# A Project Report
**on**
# "Plant leaf disease identification and classification using image processing technique"

submitted to the

**Savitribai Phule Pune University**

In partial fulfillment for the award of the Degree of

Bachelor of Engineering

in

Electronics and Telecommunication

by


**Kshitij Gajbhiye**

**Vishwas Meshram**

**Daulatrao Suryawanshi**


Under the guidance of

**Prof. Nilesh Kuchekar**


Department Of Electronics and Telecommunication

**Pune Vidyarthi Griha's**
**College Of Engineering & Technology**

44, Vidyanagari, Parvati, Pune-411009

Maharashtra, India


**2019-2020**

## CERTIFICATE

This is to certify that the project based seminar report entitled **Plant leaf disease identification and classification using image processing technique** being submitted by

**Kshitij Gajbhiye**
**Vishwas Meshram**
**Daulatrao Suryawanshi**

is a record of bonafide work carried out by her under the supervision and guidance of **Prof.Nilesh Kuchekar** in partial fulfillment of the requirement for **Electronics and Telecommunication Engineering - 2015 course** of Savitribai Phule Pune University, Pune in the academic year **2019-2020**

This project report has not been earlier submitted to any other Institute or University for the award of any degree or diploma.

**Prof. Nilesh Khuchekar**
**Internal Guide**

**Dr. Y.B.Thakre**
**HOD Electronics and Telecommunication.**

**(Name & Signature)**
**External Examiner**

**Dr. K.J.Kulkarni**
**Principal PVG's COET,PUNE**

Date:

Place:

# Acknowledgement

The completion of our project has valuable contribution from various people with whom we have interacted. This project, which is a complete fulfillment of the requirement for the degree of Bachelor of Engineering, would not have been possible without the constant support of the teaching staff and mutual co-ordination of several people to whom we owe our sincere gratitude. Firstly and most importantly, we would sincerely like to thank our mentor Prof.Nilesh Kuchekar.
He has supported us throughout the project development cycle through his constant inspiration,vision and motivation. He rendered his valuable knowledge in assisting us for publishing the manuscript in a renowned journal. We would like to thank our HOD Dr. K.J Kulkarni who extended every facility to us for making and completing this project smoothly. We would also like to thank our fellow peers and friends who constantly challenge us while also inspiring us, and provide us with their selfless help and invaluable advices. Lastly, we thank the teaching and non – teaching staff of our college as well as the library staff for providing us timely essential information in the form of books and for their assistance in the laboratories.

## Contents

# ABSTRACT

Economy of a nation depends upon the agricultural productivity, Identification of plant diseases is important for preventing losses in productivity and improving the quality of agricultural product.Traditional methods require human resource for visually observing changes in the plant leaf patterns and diagnose the disease. Traditional method consumes more time. In Big agricultural lands, detection of plant diseases at an early stage will reduce the loss in productivity, by using automated techniques.

Image processing algorithms are developed to detect plant diseases by identifying the color feature of the affected leaf area. Database is created which consists of trained dataset and test dataset, the test dataset is compared against the trained dataset using image processing techniques. The proposed system consists of image acquisition in which the cotton leaf images are taken using a digital camera and resized to 256x256 image size. In the pre-processing stage contrast is enhanced to make the texture of the image more significant for further analysis. Image segmentation is done to segment the diseased part of the leaf from the healthy part using k-means clustering algorithm. Feature extraction is performed to extract features such as Contrast, Energy, Homogeneity and Correlation. SVM classification is done to classify whether the image is healthy or diseased, also the percentage of disease spread is calculated, and further ANN classifier is used to perform the same function based on the performance of both classifiers we can conclude which classifier gives accurate results.

# Abbrevations and Acronyms

| Abbrevations | Acronyms |
|---|---|
| WPa | White pixels of affected area |
| WPu | White pixels of unaffected area |
| BP1 | Black pixels of unaffected area |
| BP2 | Black pixels of affected area |
| Pl | T otal pixels of leaf area |
| Pa | Percentage of area affected |

## List of Tables

# List of Figures

# Chapter 1

## 1 Introduction

### 1.1 Introduction of the proposed system

A Country's economic development depends upon the fertility of the agricultural land and productivity, Majority of the population are depended on the agriculture. Due to changes in the environmental conditions such as temperature, rain fall, soil fertility etc. The crops can be easily infected by bacteria, viruses and fungi. Farmers can use suitable pesticides and herbicides for preventing diseases and increasing the quality of the product. Detection of plant diseases at an early stage will be beneficial since the disease can be controlled. In few countries the farmers don't have enough knowledge or the required facility to contact the experts. Existing methods for detection of plant diseases is visual observation of the affected leaf patterns by the experts, But it requires a large expert team. In such situations an automated plant disease monitoring system will be very useful. By comparing the plants leaf's in the farm land with the plant leaf disease symptoms(samples) in the database through automation will be cheaper.

### 1.2 Motivation

The motivation for doing this project was basically because of our mutual interest in digital image processing domain. The topic in this domain was basically chosen to build a system which can be readily available to the farmers and agricultural communities. By choosing this project, it will give us deep insight on how the different image processing algorithms work in MATLAB and how the new techniques provide better accuracy and results

## 1.3 Objectives

- To obtain data set of healthy and diseased cotton leaf images (200 images).
- To segment(affected and unaffected part) and reduce the noise in the leaf image.
- To extract features from the affected area to train the classifier.
- To classify the type of disease.
- To measure the percentage of disease spread and the amount of pesticide required to stop the disease from growing further.
- Comparing two classifiers to see which works efficiently and which gives better results.

## 1.4 Expected Outcomes

- The proposed system will be beneficial for farmers to identify the type of disease and exact amount of pesticide required to cure it.
- This system is also beneficial for agricultural students and various biological authorities to examine diseases on cotton plants.
- Multinational pesticide companies can make use of this system to promote their pesticides for disease control.
- The database of this system can be made readily available for further research work in cotton crops.

## 1.5 Organisation of the report

- Chapter 2 includes Literature survey of the project wherein various methods of segmentation and classification techniques used are discussed.
- Chapter 3 includes the block diagram of the proposed system. Various blocks are discussed and their method of implementation is proposed.
- Chapter 4 includes the expected results of each block explained in chapter 3.

- Chapter 5 includes result analysis of the developed system, conclusion and future scope of the project.
- Chapter 6 includes the IEEE papers and a few references from Research Gate which were a prerequisite to analyze our system before the development phase.

# Chapter 2

## 2 Literature Survey

### 2.1 Introduction

This chapter highlights the study of different methods carried out for segmentation of the image and types of classifiers. It also describes the traditional methods used for detection of disease.

### 2.2 Types of segmentation algorithm and classifiers

In recent years, image processing techniques are used in various fields such as agriculture, biomedical, space research etc. The identification of plant infection using traditional methods are naked eye observation and microscopic evaluation by experts. For achieving this, continuous monitoring of plant as well as large team of experts is required, which becomes costly in case of large farms. At the same time it becomes time consuming to contact experts. Image processing techniques are used for fast and accurate detection of various plant diseases. The steps involved in detection of the leaf infection includes Image Acquisition, Image Pre-Processing, Segmentation, Feature Extraction, Classification. The accuracy of the system depends upon the type of method used for disease detection. The existing system for detection and classification of plant disease includes, Piyush Chaudhary[1] introduced a method for detection of heterosporium leaf spot disease in iris leaf In this system otsus threshold method is used to segment the disease spot from the leaf, Kyamelia Roy[2] has studied image segmentation methods using canny edge detection and marker based segmentation. Pranjali B. Padol [3] has used k means clustering algorithm with svm classifier to detect and classify downy mildew and powdery mildew on grape leaf. The proposed system uses k means clustering algorithm for image segmentation and Ann classifier to classify the types of disease such as Anthracnose, Bacterial Blight and Cercospora leaf spot on a cotton leaf.

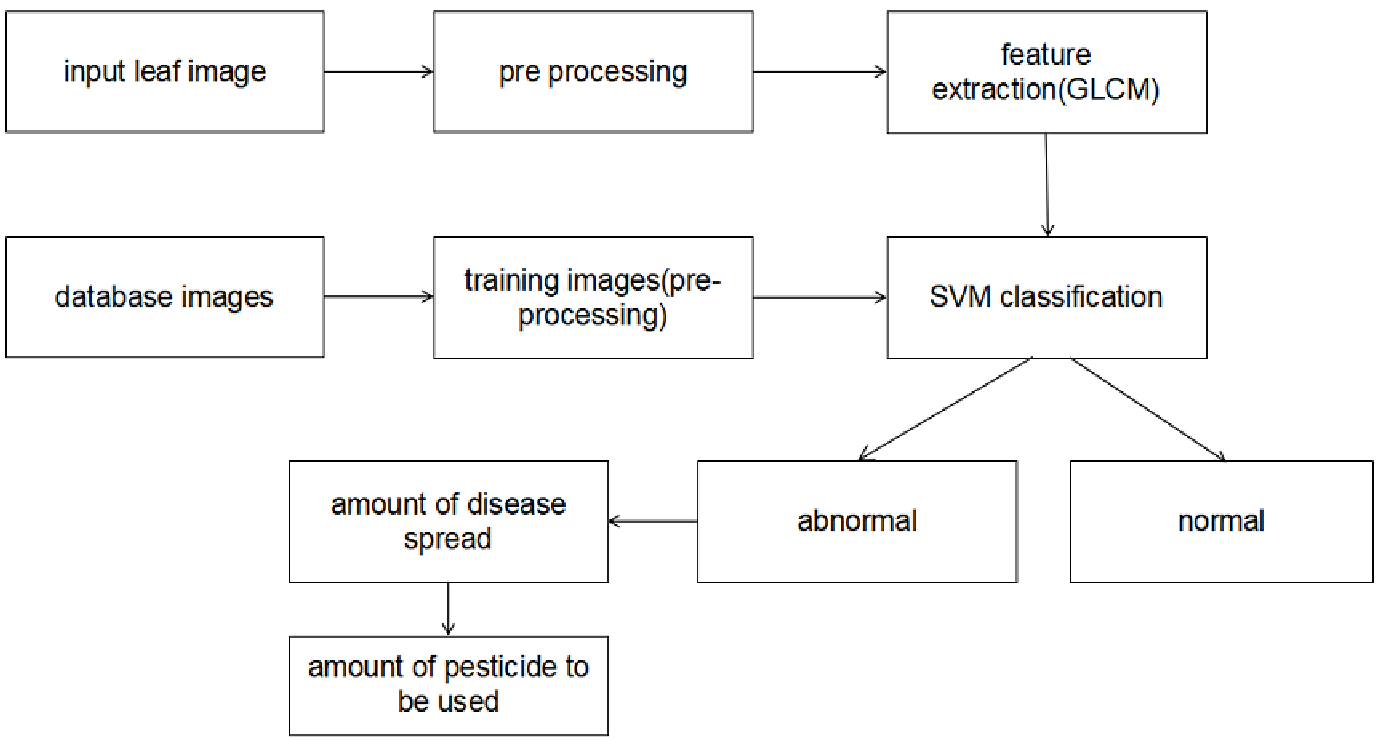
# Chapter 3

## 3 Design of proposed system

### 3.1 Introduction

This chapter illustrates the proposed system architecture, the system architecture consists of block diagram and the description of each block.

### 3.2 Block Diagram of the proposed system




(a) fig 3.2.1

### 3.3 Working of the project

1. Image Acquisition

    - In the first stage a dataset containing 200-250 images of cotton leaves are taken using a digital camera.

- These images are acquired by visiting agricultural fields, agricultural colleges and botanical gardens to create a database.

2. Image Pre - processing

   - Image pre-processing improves the quality of the image by removing noise.
   - In this stage resizing of the image to 256*256 is done, Image enhancement is performed, gray scale and HSI of the image is generated for further computation

3. Image Segmentation

   - Image Segmentation is a technique of partitioning an image into multiple segments. The goal of Image segmentation is to simplify or to change the representation of an image into something easier to analyze.
   - It is the process of partition the image carrying similar characteristics(Pixel Intensity) and assigning label to them.
   - The number of partitions depends upon the user's demand.

## 3.4 Types of segmentation techniques

- Otsus Thresholding -In This Technique Histogram Equalisation of the image is taken to find the mid point between the higher and lower pixel intensity values,If the histogram equalisation of the image is linear throughout the entire pixels ,mid point is not formed and the required segmented results cant be obtained.

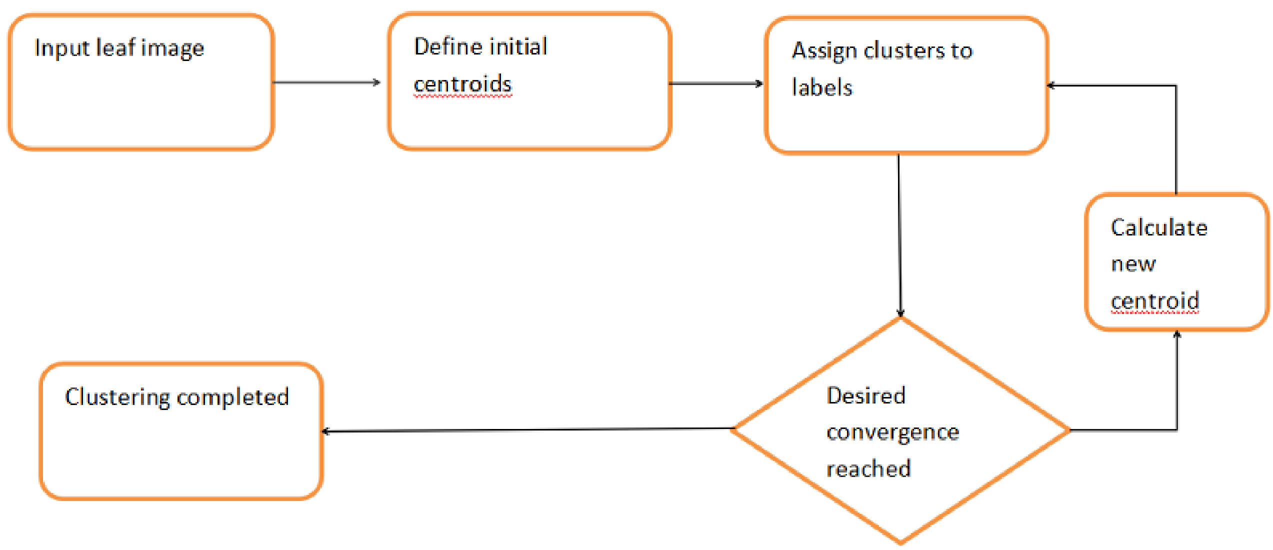
- Edge Detection method - It is a Discontinuity Detection based approach,In this technique abrupt changes in intensity values are detected and segmentation is obtained, If too many edges occur in the image this technique is not suitable to provide desired results.

- Color Based Segmentation - In this Technique based on color variations the image is segmented, It is not suitable for an image having multiple diseases.

- k means clustering - In this technique the image is segmented based on the concept of centroids and the image is segmented when the required convergence is reached,Selection of initial centroid is random and selection of optimal number of clusters is difficult.

## 3.5 k means clustering

Clustering is a method in which large set of data is grouped into smaller sets of clusters of similar data. This is done by using the Euclidean distance metric, User should select the value of k. k means the number of clusters/groups.




(a) fig 3.5.1 k-means segmentation

## 3.6 Properties of k-means clustering algorithm:

- There is k number of clusters.
- There is one centroid in each of the given cluster
- The clusters will never overlap with each other
- Each element of a single cluster is nearer to its own cluster than any other cluster.

## 3.7 Feature Extration

[5]Gray Level Co-occurrence matrix (GLCM) is a statistical method of investigating texture which considers the relationship among pixels. In the proposed technique four features are extracted which include energy, contrast, correlation.

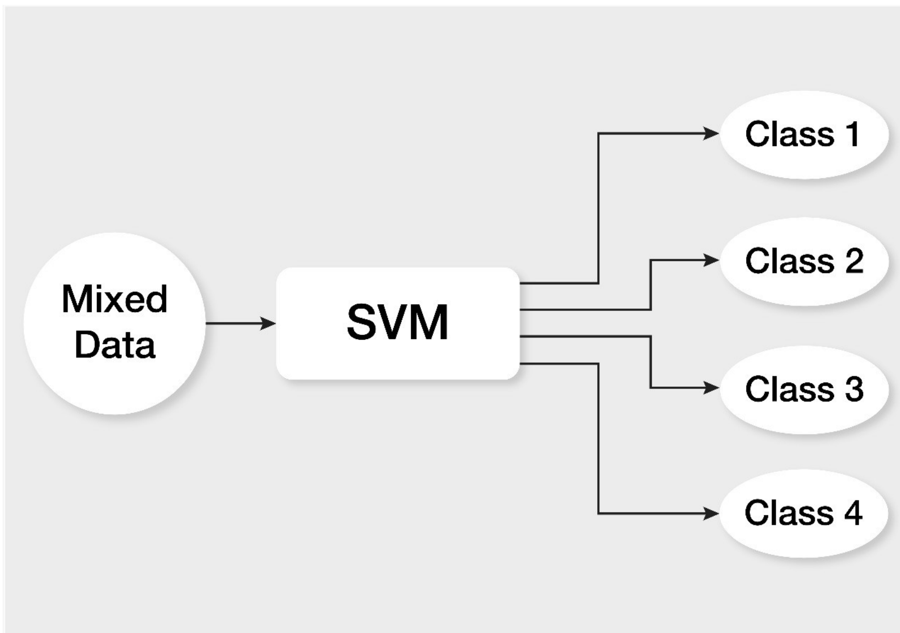
| Feature | Description | Formula (N=no. of distinct gray levels P(I,j)=pixel intensity in the matrix i*j) |
|---|---|---|
| contrast | It measures the intensity value of the pixel and its neighbouring pixel over the entire image, If the image is constant(Similar Intensity) contrast=0 | $\sum_{i,j=0}^{N-1} P(i,j) * (i-j)^2$ |
| Energy | It is the summation of square of pixel intensity values of an image.Energy range is between 0 and 1.Energy is 1 for a constant image | $\sum_{i,j=0}^{N-1} P(i,j)^2$ |
| Homogeneity | It measures the similarity between the pixels. Its range lie between 0 and 1 and Homogeneity is 1 for a diagonal GLCM. | $\sum_{i,j=0}^{N-1} \frac{P(i,j)^2}{1+(i-j)^2}$ |
| Correlation | It measures how correlated a pixel is to its neighbouring pixel, it ranges between -1 and 1. $\mu$ is mean value of all pixels and $\sigma$ represents the variance | $\sum_{i,j=0}^{N-1} P(i,j)[(i-\mu)(j-\mu)/\sigma^2]$ |

(a) Table 3.7.1 Feature Extraction

# Classifier

## 3.8 SVM Classifier

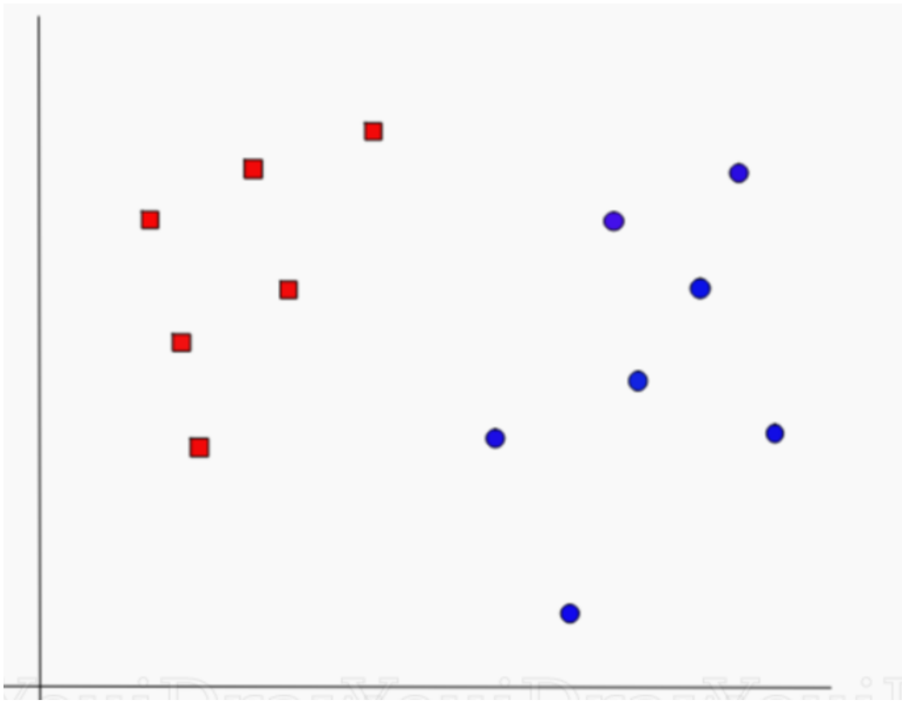
- The classification technique is used for both training and testing process. This is the last stage of the system, The features extracted from training leaves are compared with features extracted from testing leaves. Then the image is classified in a specific class based on the matched features.




(a) fig 3.8.1 SVM(17)

1. Linearly Separable Data

    - SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems, and work well for many practical problems and give solutions. The idea of SVM is simple, The algorithm creates a line or a hyperplane which separates the data into classes.
    - SVM is an algorithm that takes the data as an input and outputs a line that separates those classes if possible.

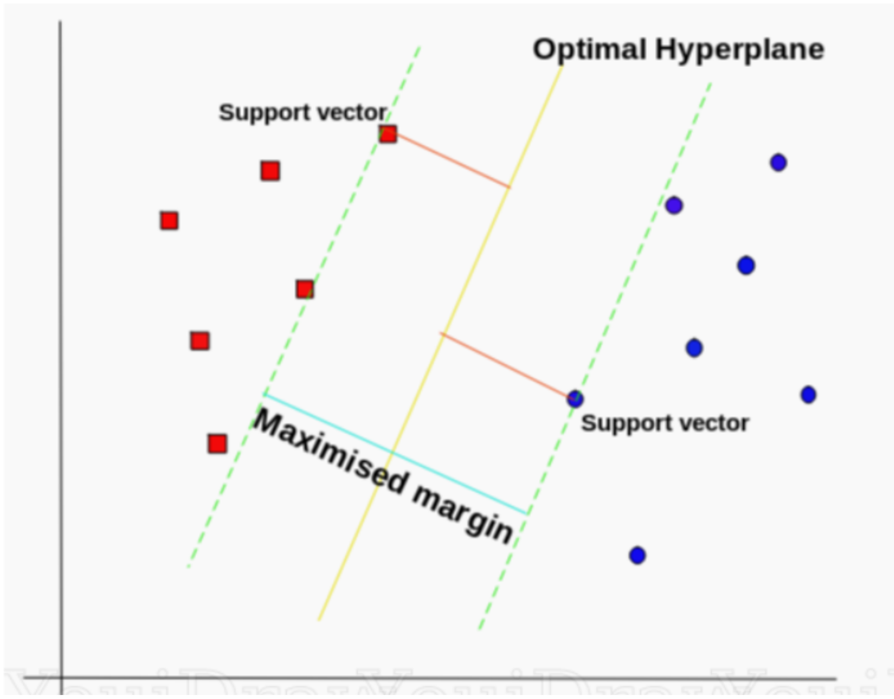


(a) fig 3.8.2 Linearly Separable Data(17)

- We have two candidates here, the green colored line and the yellow colored line.The green line in the image above is quite close to the red class. Though it classifies the current datasets it is not a generalized line and in machine learning our goal is to get a more generalized separator.
- According to the SVM algorithm we find the points closest to the line from both the classes.These points are called support vectors. Now, we compute the distance between the line and the support vectors. This distance is called the margin. Our goal is to maximize the margin. The hyperplane for which the margin is maximum is the optimal hyperplane.
- Thus SVM tries to make a decision boundary in such a way that the separation between the two classes(that street) is as wide as possible.

2. Non Linearly Separable Data

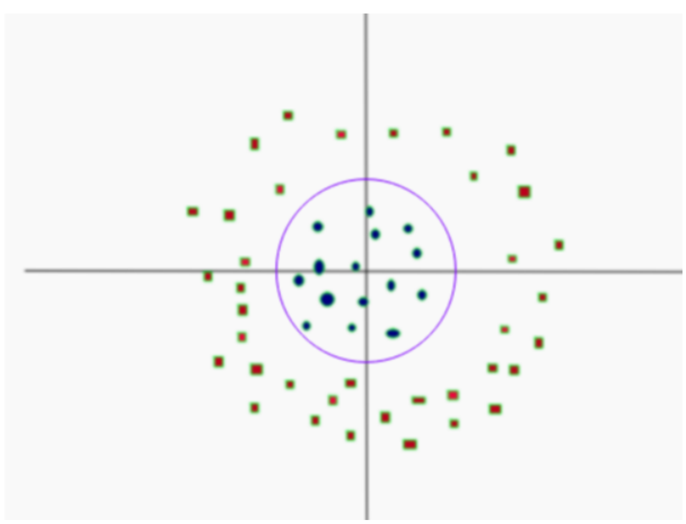
- This data is clearly not linearly separable.But, this data can be converted to linearly separable data in higher dimension. Lets add one more dimension and call it z-axis. Let the co-ordinates on z-axis be




(a) fig 3.8.3 Linearly Separable Data(17)

governed by the constraint,z = x+y. So, basically z co-ordinate is the square of distance of the point from origin.

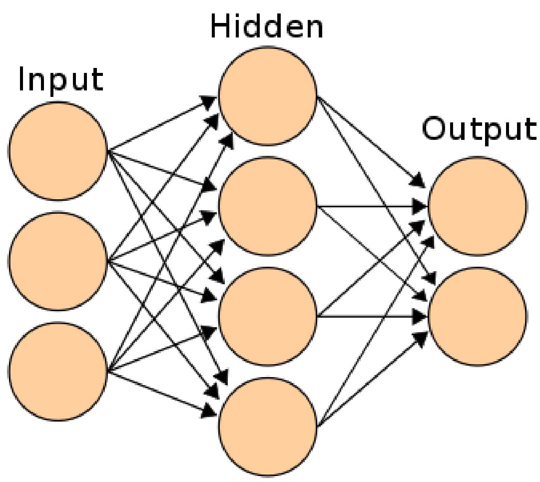
- Now the data is clearly linearly separable. Let the purple line separating the data in higher dimension be z=k, where k is a constant. Since, z=x+y we get x + y = k; which is an equation of a circle. So, we can project this linear separator in higher dimension.



(a) fig 3.8.4 Non Linearly Separable Data(17)

## 3.9 ANN Algorithm

- ANN is a tool which works as a human brain, Human brain has neurons which shoots signal to other neurons.
- Artificial Neural Networks can be best described as the biologically inspired simulations that are performed on the computer to do a certain specific set of tasks like clustering, classification, pattern recognition etc




(a) fig 3.9.1 ANN(18)

- Input layer: The Input layers contain those artificial neurons (termed as units) which are to receive input from the outside world. This is where the actual learning on the network happens, or recognition happens else it will process.
- Output layer: The output layers contain units that respond to the information that is fed into the system and also whether it learned any task or not.
- Hidden layer: The hidden layers are mentioned hidden in between input layers and the output layers. The only job of a hidden layer is to transform the input into something meaningful that the output layer/unit can use in some way..
- Most of the artificial neural networks are all interconnected, which means that each of the hidden layers is individually connected to the neurons

in its input layer and also to its output layer leaving nothing to hang in the air.

- This makes it possible for a complete learning process and also learning occurs to the maximum when the weights inside the artificial neural network get updated after each iteration.
- Artificial Neural Networks can be best viewed as weighted directed graphs, where the nodes are formed by the artificial neurons and the connection between the neuron outputs and neuron inputs can be represented by the directed edges with weights.
- The Artificial Neural Network receives the input signal from the external world in the form of a pattern and image in the form of a vector. These inputs are then mathematically designated by the notations x(n) for every n number of inputs.
- Each of the input is then multiplied by its corresponding weights (these weights are the details used by the artificial neural networks to solve a certain problem). In general terms, these weights typically represent the strength of the interconnection amongst neurons inside the artificial neural network. All the weighted inputs are summed up inside the computing unit(another artificial neuron)
- If the weighted sum equates to zero, a bias is added to make the output non-zero or else to scale up to the system's response.
- Bias has the weight and the input to it is always equal to 1.
- Here the sum of weighted inputs can be in the range of 0 to positive infinity. To keep the response in the limits of the desired value, a certain threshold value is benchmarked. And then the sum of weighted inputs is passed through the activation function.
- The activation function, in general, is the set of transfer functions used to get the desired output of it. There are various flavors of the activation function, but mainly either linear or non-linear sets of functions. Some of the most commonly used set of activation functions are the Binary, Sigmoidal (linear)

- The output of the binary activation function is either a 0 or a 1. To attain this, there is a threshold value set up. If the net weighted input of the neuron is greater than 1 then the final output of the activation function is returned as 1 or else the output is returned as 0.
- Every neuron is connected with other neuron through a connection link. Each connection link is associated with a weight that has information about the input signal, the weight usually excites the signal that is being communicated.
- The neurons are passed through activation functions which creates a threshold. ANN tool also has a back propagation process, this process is used to reduce the errors.

## 3.10 Calculation of The Diseased Part

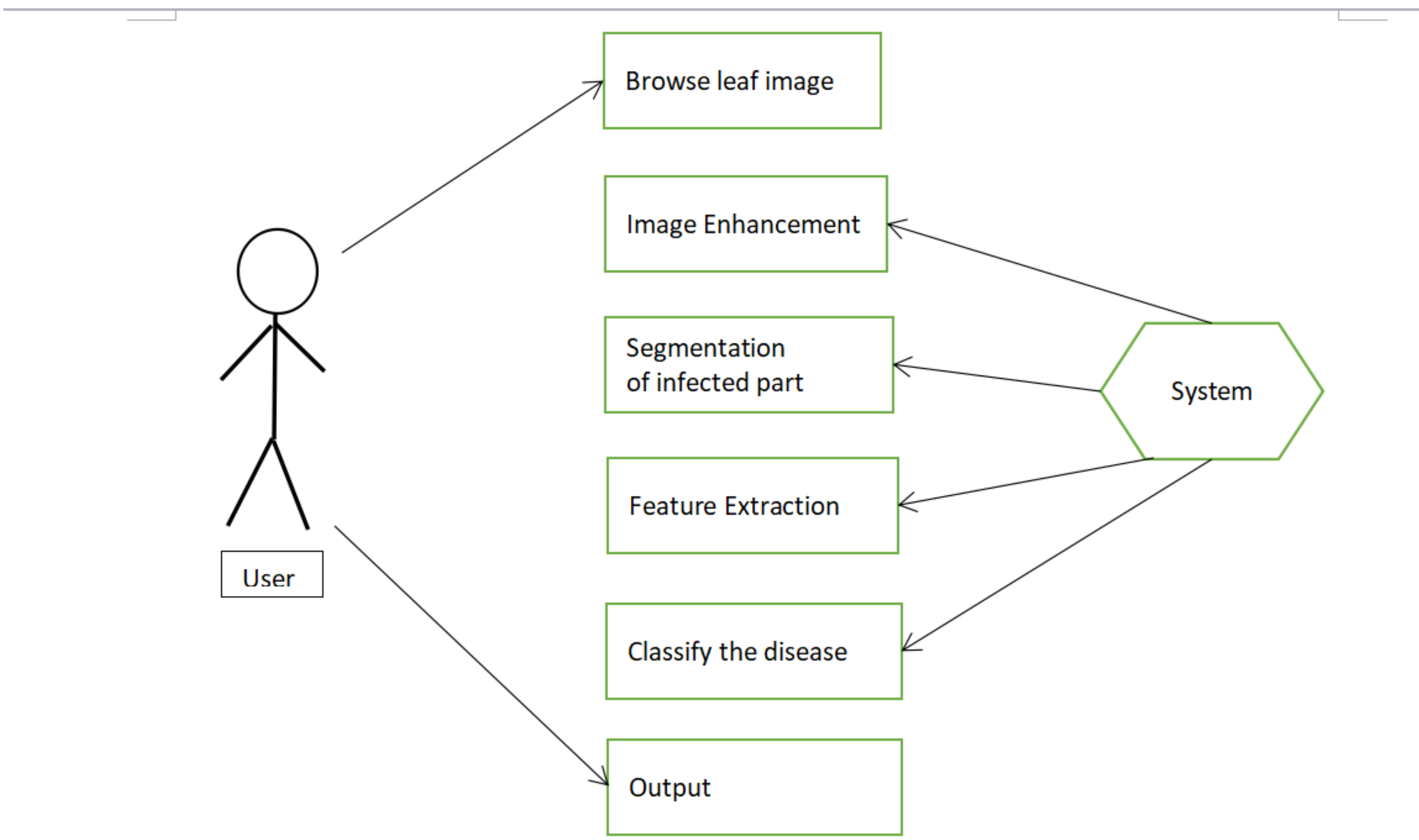
As the segmentation has been achieved in earlier stage, and the label has been assigned to affected and unaffected part, we now take the binary transform of these segmented image to calculate the percentage of disease spread.

- To calculate the WPu and BP1 of the unaffected part of the leaf and to calculate total pixel we add both the values.
- To calculate the WPa and BP 2 of the affected part of the leaf and to calculate total pixel we add both the values.
- Total pixel of leaf area is obtained by Pl=WPa + WPu.
- Total percentage of affected pixels can be obtained by following formula Pa= (WPa/Pl) *100.

## 3.11 Proposed Design of Graphical User Interface

- The GUI or Graphical user interface gives the user ease of operation when handling the system.
- The graphical user interface will allow the user to browse through the 50 set of testing images.
- The GUI will allow the user to perform image Pre-processing, segmentation, plotting of the extracted features on X-Y scale and classification of

the type of the disease along with amount of disease spread in percentage.




(a) fig 3.11.1 GUI

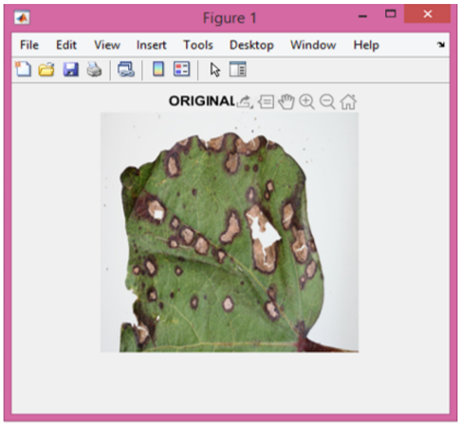
# Chapter 4

## 4 Simulation and testing

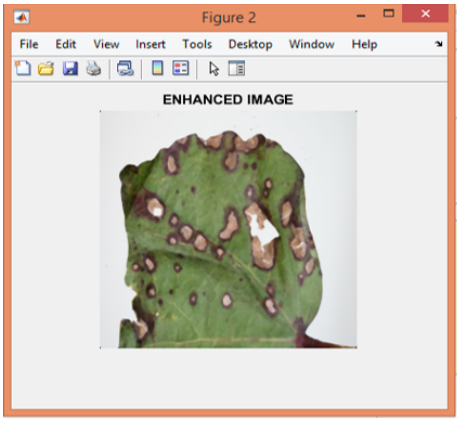
### 4.1 Introduction

This section comprises of observations and output of system design stages which includes Image pre-processing, image segmentation, feature extraction and classification.

### 4.2 Image Pre-processing

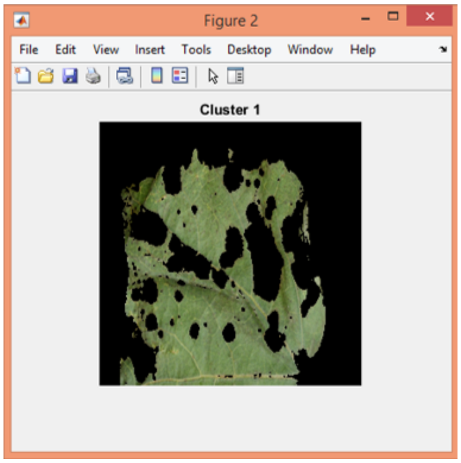



(a) fig 4.2.1 Resized Image

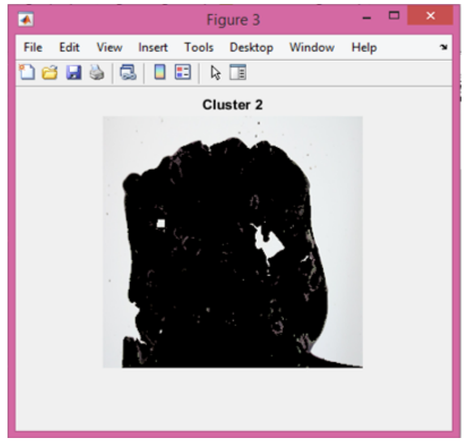



(a) fig 4.2.2 median filter on resized image

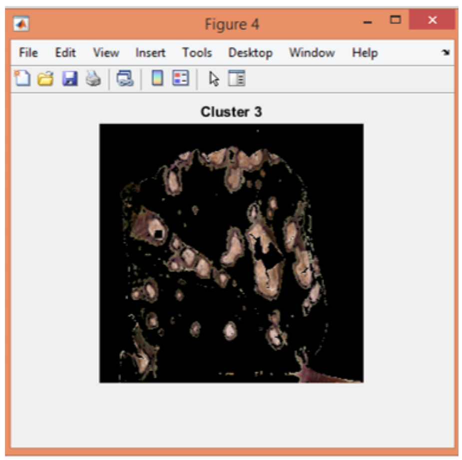
## 4.3 Segmentation Of the Diseased part




(a) fig 4.3.1 cluster 1

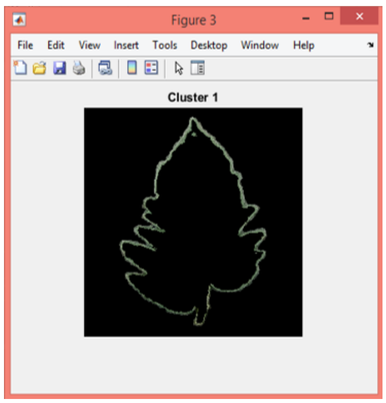



(b) fig 4.3.2 cluster 2

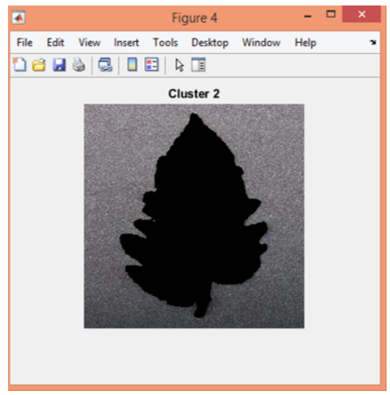



(c) fig 4.3.3 cluster 3

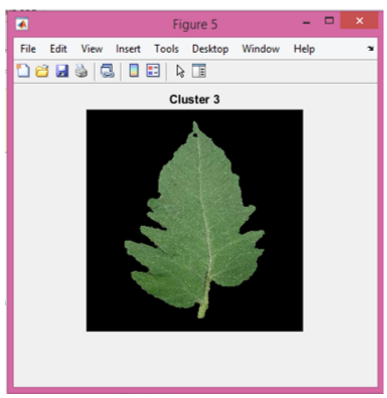
Figure 1: Image segmentation using k-means on diseased leaf




(a) fig 4.3.4 cluster 1



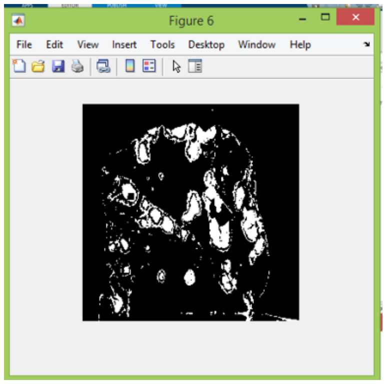

(b) fig 4.3.5 cluster 2



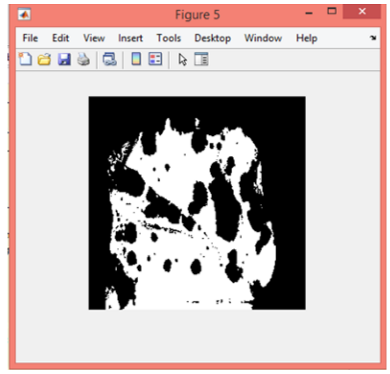

(c) fig 4.3.6 cluster 3

Figure 2: Image segmentation using k-means on healthy leaf

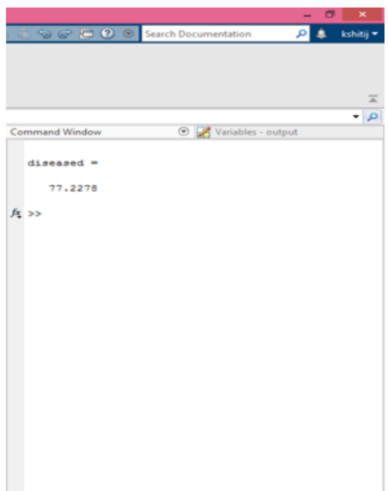
## 4.4 Calculation of the Diseased Part




(a) fig 4.4.1 infected part

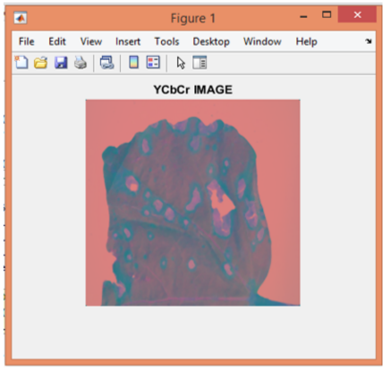



(b) fig 4.4.2 healthy part

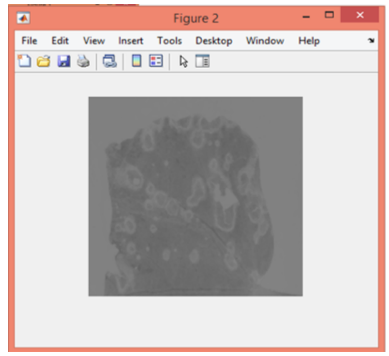



(c) fig 4.4.3 Percentage of disease spread

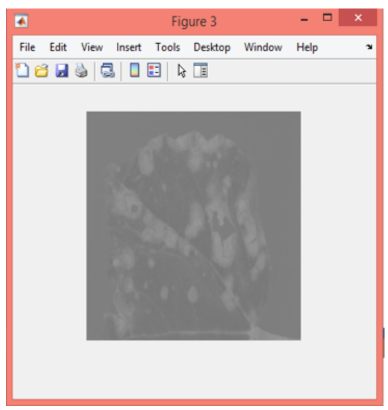
## 4.5 Image Segmentation of The Diseased Part Using Color Image Segmentation



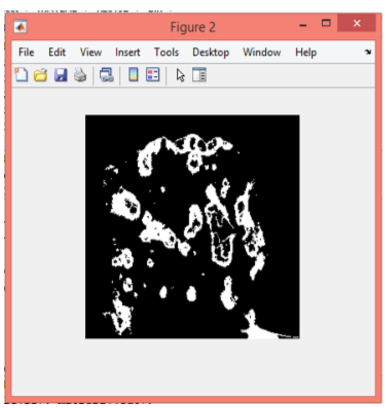

(a) fig 4.5.1 YCbCr image



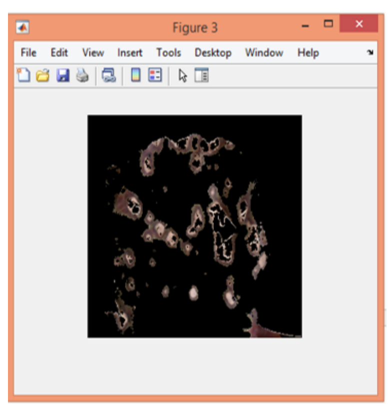

(b) fig 4.5.2 Cb plane




(c) fig 4.5.3 Cr plane

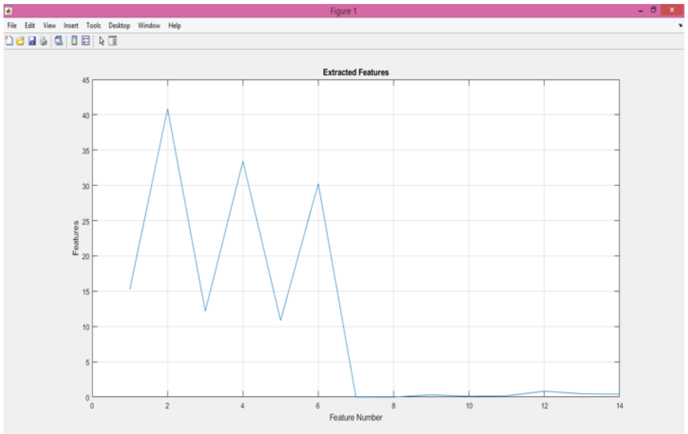



(a) fig 4.5.4 Binary image of segmented part

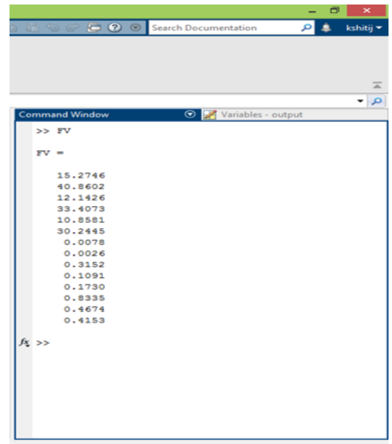



(b) fig 4.5.5 Colour image of segmented part

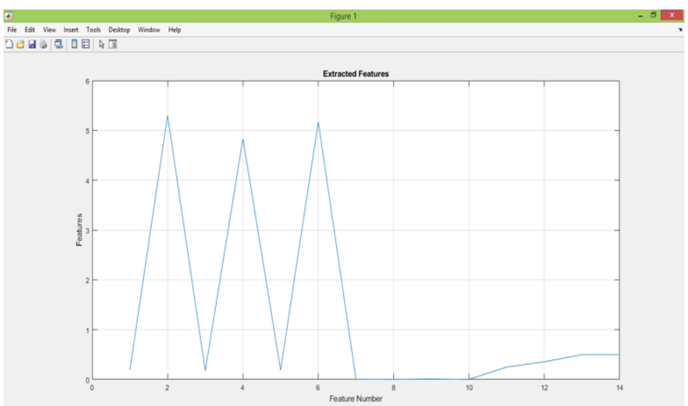
## 4.6 Extracted Features



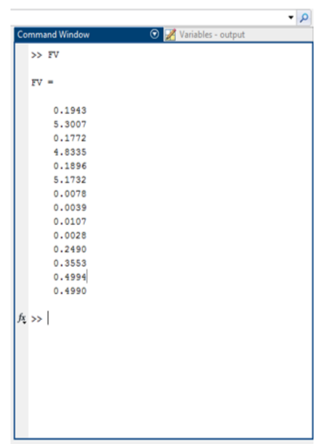
(a) fig 4.6.1 Plotting the extracted features on X-Y scale



(b) fig 4.6.2 Values of extracted features

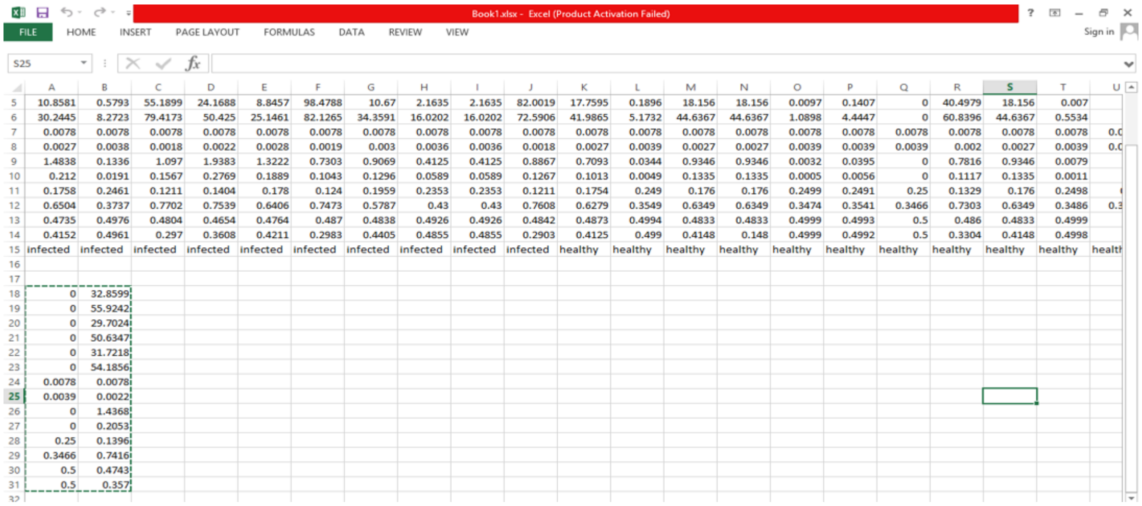


(a) fig 4.6.3 Plotting the extracted features on X-Y scale



(b) fig 4.6.4 Values of extracted features

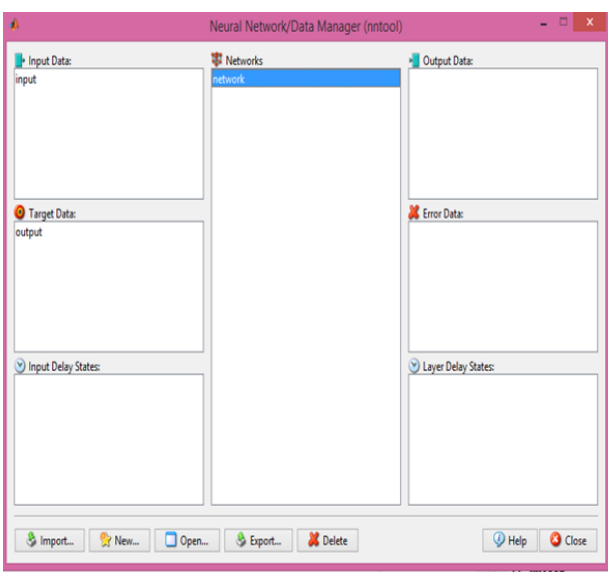
## 4.7 Experiment



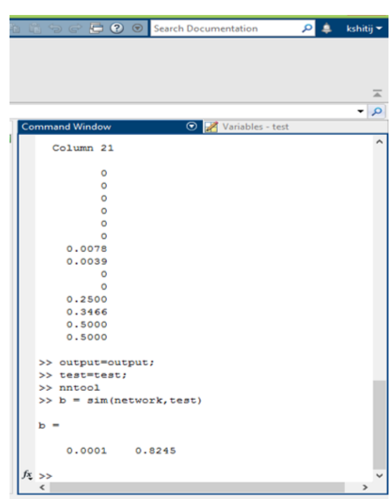
| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 10.8581 | 0.5793 | 55.1899 | 24.1688 | 8.8457 | 98.4788 | 10.67 | 2.1635 | 2.1635 | 82.0019 |
| 6 | 30.2445 | 8.2723 | 79.4173 | 50.425 | 25.1461 | 82.1265 | 34.3591 | 16.0202 | 16.0202 | 72.5906 |
| 7 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 |
| 8 | 0.0027 | 0.0038 | 0.0018 | 0.0022 | 0.0028 | 0.0019 | 0.003 | 0.0036 | 0.0036 | 0.0018 |
| 9 | 1.4838 | 0.1336 | 1.097 | 1.9383 | 1.3222 | 0.7303 | 0.9069 | 0.4125 | 0.4125 | 0.8867 |
| 10 | 0.212 | 0.0191 | 0.1567 | 0.2769 | 0.1889 | 0.1043 | 0.1296 | 0.0589 | 0.0589 | 0.1267 |
| 11 | 0.1758 | 0.2461 | 0.1211 | 0.1404 | 0.178 | 0.124 | 0.1959 | 0.2353 | 0.2353 | 0.1211 |
| 12 | 0.6504 | 0.3737 | 0.7702 | 0.7539 | 0.6406 | 0.7473 | 0.5787 | 0.43 | 0.43 | 0.7608 |
| 13 | 0.4735 | 0.4976 | 0.4804 | 0.4654 | 0.4764 | 0.487 | 0.4838 | 0.4926 | 0.4926 | 0.4842 |
| 14 | 0.4152 | 0.4961 | 0.297 | 0.3608 | 0.4211 | 0.2983 | 0.4405 | 0.4855 | 0.4855 | 0.2903 |
| 15 | infected | infected | infected | infected | infected | infected | infected | infected | infected | infected |
| 16 | | | | | | | | | | |
| 17 | | | | | | | | | | |
| 18 | 0 | 32.8599 | | | | | | | | |
| 19 | 0 | 55.9242 | | | | | | | | |
| 20 | 0 | 29.7024 | | | | | | | | |
| 21 | 0 | 50.6347 | | | | | | | | |
| 22 | 0 | 31.7218 | | | | | | | | |
| 23 | 0 | 54.1856 | | | | | | | | |
| 24 | 0.0078 | 0.0078 | | | | | | | | |
| 25 | 0.0039 | 0.0022 | | | | | | | | |
| 26 | 0 | 1.4368 | | | | | | | | |
| 27 | 0 | 0.2053 | | | | | | | | |
| 28 | 0.25 | 0.1396 | | | | | | | | |
| 29 | 0.3466 | 0.7416 | | | | | | | | |
| 30 | 0.5 | 0.4743 | | | | | | | | |
| 31 | 0.5 | 0.357 | | | | | | | | |

| | K | L | M | N | O | P | Q | R | S | T | U |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 17.7595 | 0.1896 | 18.156 | 18.156 | 0.0097 | 0.1407 | 0 | 40.4979 | 18.156 | 0.007 | |
| 6 | 41.9865 | 5.1732 | 44.6367 | 44.6367 | 1.0898 | 4.4447 | 0 | 60.8396 | 44.6367 | 0.5534 | |
| 7 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0078 | 0.0 |
| 8 | 0.0027 | 0.0039 | 0.0027 | 0.0027 | 0.0039 | 0.0039 | 0.0039 | 0.002 | 0.0027 | 0.0039 | 0.0 |
| 9 | 0.7093 | 0.0344 | 0.9346 | 0.9346 | 0.0032 | 0.0395 | 0 | 0.7816 | 0.9346 | 0.0079 | |
| 10 | 0.1013 | 0.0049 | 0.1335 | 0.1335 | 0.0005 | 0.0056 | 0 | 0.1117 | 0.1335 | 0.0011 | |
| 11 | 0.1754 | 0.249 | 0.176 | 0.176 | 0.2499 | 0.2491 | 0.25 | 0.1329 | 0.176 | 0.2498 | |
| 12 | 0.6279 | 0.3549 | 0.6349 | 0.6349 | 0.3474 | 0.3541 | 0.3466 | 0.7303 | 0.6349 | 0.3486 | 0.3 |
| 13 | 0.4873 | 0.4994 | 0.4833 | 0.4833 | 0.4999 | 0.4993 | 0.5 | 0.486 | 0.4833 | 0.4999 | |
| 14 | 0.4125 | 0.499 | 0.4148 | 0.148 | 0.4999 | 0.4992 | 0.5 | 0.3304 | 0.4148 | 0.4998 | |
| 15 | healthy | healthy | healthy | healthy | healthy | healthy | healthy | healthy | healthy | healthy | healtl |

(a) fig 4.7.1 Values of extracted features of 10 healthy and 10 infected leaf images stored in a CSV file with one healthy and one infected features of leaf image for testing

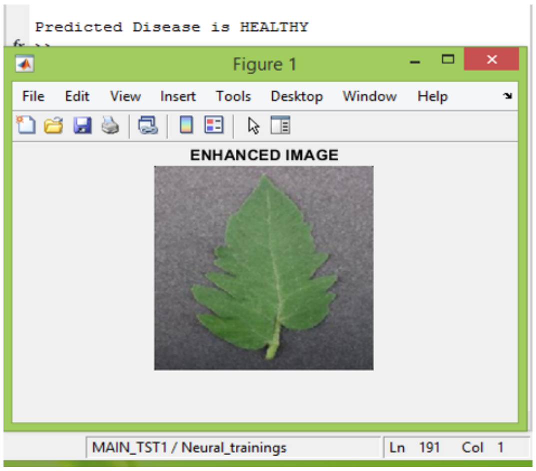



(a) fig 4.7.2 Neural network nntool

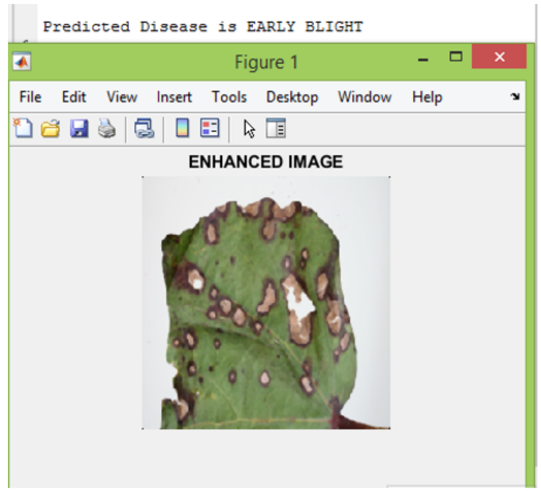



(b) fig 4.7.3 Result of the nueral network

## 4.8 Classification of the type of Disease




(a) fig 4.8.1 Predicted disease is healthy




(b) fig 4.8.2 Predicted disease is early blight


(a) fig 4.8.3 GUI template

# Chapter 5

## 5 Results analysis and conclusion

### 5.1 Introduction

This section contains the result, analysis, conclusion and future scope of the designed system.

### 5.2 Result analysis

- The pre processing stage consists of converting the trained data set to a size 250 * 250. to reduce the high frequency components in the image , median filter was applied to the image.
- Image segmentation was carried using k- means segmentation algorithm, 3 centroids(clusters) were taken to segment the image. Due to the drawbacks of k means algorithm everytime the centroids altered as builtin k means function was used to segment the image.an alternative to k means is colour image segmentation using YCbCr color modelwhich is used for further analysis.
- The features for instance Contrast,Dissimilarities,Energy,Entropy,Homogeinity were extracted from the infected part of the image.
- The extracted features were used to classify the type of disease.

### 5.3 Conclusion

Plant Disease Detection and classification can be categorized into four different sections: Image pre-processing,Image Segmentation,Feature Extraction and Image Classification.K means Segmentaion along with Color Image Segmentation using YCbCr color model was implemented.It can be concluded that the algorithms and the extracted features used in the proposed system are all meant to increase the efficiency of the system to achieve better results.k means segmentation gives the promising results.For the segmentation of the diseased part as compared to Color Image Segmentation.as an experiment 10 healthy and 10 infected tomato leaf images were used to train the nntool present in the matlab, 5 testing images were taken to classify if the plant was healthy or diseased.Of the two classification methods used,It was found Artificial Neural Network to have a higher classification accuracy as compared to Support Vector Machines.This System can be trained to classify different diseases found in different plants,Since the extracted features will have different values for different diseases.

## 5.4 Future scope

In the future one may use a more diverse dataset to increase the classification accuracy,another approach to improve the result is to extract more features from the infected part which will drastically improve the accuracy of the neural network, This System can be implemented to detect other kinds of diseases on different plants as well.
The project can be further extended to determine the amount of pesticide required to cure the disease which would most likely require a dataset on pesticides,The system can be implemented in real time by integrating the MATLAB software with a camera which will be continously taking snapshot of the plant with an interval of more than a minute give time for the neural network to classify the disease.

# Chapter 6

## 6 References


1 Piyush Chaudhary, Anand K. Chaudhari, Dr. A. N. Cheeran and Sharda Godara “Color Transform Based Approach for Disease Spot Detection on Plant Leaf” ,International Journal of Computer Science and Telecommunications [Volume 3, Issue 6, June 2012]

2 Kyamelia Roy,etc,”Segmentation Techniques for Rotten Fruit detection ”,2019 IEEE

3 Pranjali B. Padol, Prof. Anjali A. Yadav,”SVM Classifier Based Grape Leaf Disease Detection”,2016 Conference on Advances in Signal Processing (CASP) Cummins College of Engineering for Women, Pune.

4 Dilpreet Kaur, Yadwinder Kaur,”Various Image Segmentation Techniques: A Review ”,IJCSMC, Vol. 3, Issue. 5, May 2014.

5 R.Meena Prakash,G.P.Saraswathy, G.Ramalakshmi,”Detection of Leaf Diseases and Classification using Digital Image Processing”,2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

6 Pooja V,rahul Das,Kanchana V, “ Identification of plant leaf diseases using image processing techniques “,department of computer science,amrita school of arts and sciences,amrita university mysuru campus,karnataka,india.

7 Namrata R. Bhimte, V.R.Thool, “Disease detection of cotton leaf spot using image processing and svm classifier”,conference on IEEE 2018

8 K.Narsimha Reddy, B.Polaiah,N.Madhu, ”A literature survey :leaf disease detection using image processing techniques”,vardhaman college of engineering department of electronics and communication engineering,telangana,india JECE)(may-june 2017)

9 Aditya Parikh,Mehul S. Raval,, Chandrasinh Parmar and Sanjay Chaudhary,”Disease Detection and Severity Estimation in Cotton Plant from Unconstrained Images” 2016 IEEE International Conference on Data Science and Advanced Analytics

10 Bhumika S.Prajapati, Vipul K.Dabhi, Harshadkumar B.Prajapati,”A Survey on Detection and Classification of Cotton Leaf Diseases ” International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) - 2016

11 Bharat Mishra,Mamta Lambert,Sumit Nema,Swapnil Nema,”Recent Technologies of Leaf Disease Detection using Image Processing Approach – A Review ”,2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

12 Chaitali G. Dhaware ,Mrs. K.H. Wanjale,”A Modern Approach for Plant Leaf Disease Classification which Depends on Leaf Image Processing ”,2017 International Conference on Computer Communication and Informatics (ICCCI -2017), Jan. 05 – 07, 2017, Coimbatore, INDIA

13 Nikhil Shah1, Sarika Jain, ”Detection of Disease in Cotton Leaf using Artificial Neural Network”, Amity Institute of Information Technology, Amity University Uttar Pradesh ,2019 IEEE

14 Anand R,Veni S,Aravinth J, ”An Application of image processing techniques for Detection of Diseases on Brinjal Leaves Using K-Means Clustering Method” ,Department of Electronics and Communication Engineering Amrita School of Engineering, Coimbatore,2016 Fifth Interna-

tional Conference On Recent Trends In Information Technology.

15 https://data-flair.training/blogs/svm-support-vector-machine-tutorial

16 https://www.analyticsvidhya.com/blog/2014/10/ann-work-simplified

17 Support vector machine: https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989

18 Artificial neural network: https://mindmajix.com/artificial-neural-network-and-how-it-works

# Chapter 7

## 7 Appendix

### 7.1 Minimum system requirements

1 **Software requirements**

- MATLAB 2019b.
- Database: database collected from kaggle.
- Latex software.

2 **Hardware Requirements:**

- Intel Dual Core Dual processor or advanced version
- Minimum 4GB of RAM
- Minimum 20GB of Hard disk space

### 7.2 Technical Reference Manual

Installing MATLAB is required for the implementation of this system.Licensed version of the software is obtained from the official MATLAB website. We need to follow the instructions given in the manual for installing and running the software smoothly.
All images loaded must be of desired format (JPG),and must not exceed 250 * 250 dimension.

## 7.3 Plagiarism Report

SmallSEOTools

### PLAGIARISM SCAN REPORT

| | | | |
|---|---|---|---|
| Words | 938 | Date | June 14,2020 |
| Characters | 6152 | Exclude Url | |

| | | | |
|---|---|---|---|
| 0% Plagiarism | 100% Unique | 0 Plagiarized Sentences | 47 Unique Sentences |

Content Checked For Plagiarism

SmallSEOTools

### PLAGIARISM SCAN REPORT

| | | | |
|---|---|---|---|
| Words | 611 | Date | June 15,2020 |
| Characters | 3573 | Exclude Url | |

| | | | |
|---|---|---|---|
| 3% Plagiarism | 97% Unique | 1 Plagiarized Sentences | 28 Unique Sentences |