CROP DISEASE AND SEVERITY DETECTION AND PREDICTION OF PESTICIDES

A project report submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE & ENGINEERING

BY

Anuj Jain (BE/25036/14)

Yash Tongia (BE/25030/14)

DEPARTMENT OF COMPUTER SCIENCE &ENGINEERING

BIRLA INSTITUTE OF TECHNOLOGY, MESRA

JAIPUR CAMPUS, JAIPUR

SP-2018

DECLARATION CERTIFICATE

This is to certify that the work presented in the project entitled "Crop Disease and Severity Detection and Prediction of Pesticides" in partial fulfillment of the requirement for the award of Degree of Bachelor of Engineering in Computer Science and Engineering of Birla Institute of Technology, Mesra, Ranchi, Extension Center Jaipur is an authentic work carried out under my supervision and guidance.

To the best of my knowledge, the content of this project does not form a basis for the award of any previous degree to anyone else.

Date:

Dr. Madhavi Sinha
(H.O.D. C.S.E.)
Birla Institute of Technology, Mesra, Ranchi
Extension Centre, Jaipur

ACKNOWLEDGMENT

We declare that the project entitled "Crop Disease and Severity Detection and Prediction of Pesticides" in partial fulfillment of the requirement for the award of Degree of Bachelor of Engineering in Computer Science and Engineering of Birla Institute of Technology, Mesra, Ranchi, Extension Centre Jaipur is an authentic work carried out by us. To the best of our knowledge, the content of this project does not form a basis for the award of any previous degree to anyone else.

Date:

1. Anuj Jain (BE/25036/14)
2. Yash Tongia (BE/25030/14)

Birla Institute of Technology, Mesra, Ranchi
Extension Centre, Jaipur

Table of Contents

1. Introduction

1.1 Introduction

- Each year, plant viruses and fungal attacks lead to crop losses of up to 30 percent and moreover excess usage of pesticides leads to spreading of fatal diseases such as cancer, autism and asthma.
- That is why it is important to detect plant disease in its early stages. The existing methods include naked eye observations by experts. This includes constant monitoring which is very costly and a tedious exercise for large farms.
- The farmers in India are less educated to understand the vulnerability and severe consequences of the diseases. They either ignore or recklessly use pesticides to get rid of the problem. This not only leads to spreading of fatal diseases and soil contamination but is also not economically beneficial for them.
- Thus, it is very important for them to know the disease and their proper treatment for better crop production and environmental balance.

1.2 Intended Audience and Reading Suggestions

The intended audience is anyone who is interested in implementing and knowing more about disease and severity detection of crop and prediction of pesticides like farmers and agriculturist. Also, the document will be utilized to evaluate the model's design and features.

1.3 Scope of a Project

The disease detection model in plants will help users to identify disease present in plant and will suggest the user amount and the type of fertilizer a user must add to cope up with the disease present in a particular type of plant.

Given a large set of data set, disease detection in plants is done by comparing different information of plant with that of given plant in dataset, disease is detected on the basis amount and type of difference between data set values and actual values.

The disease and severity detection of crop and prediction of pesticides will use various algorithms to various type of suggestion about crop condition. It:

- Provide access to a very large information collection
- Provide User-friendly interface
- Supports advanced search 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 are non-technical persons.

This system aims to automate the disease and severity detection of crop and prediction of pesticides.

1.4 Methodology

The project consists of four phases which are: -

- Image Acquisition and Preprocessing.
- Computing the features using color co-occurrence methodology.
- Classification of the disease.
- Prediction of the amount of pesticides.

Phase 1:

- Image acquisition is the very first step that requires capturing an image with the help of a digital camera.
- Preprocessing of input image to improve the quality of image and to remove the undesired distortion from the image.
- Clipping of the leaf image is performed to get the interested image region and then image smoothing is done using the smoothing filter. To increase the contrast Image enhancement is also done. Mostly green colored pixels, in this step, are masked. In this, we computed a threshold value that is used for these pixels.
- Then in the following way mostly green pixels are masked: if pixel intensity of the green component is less than the pre-computed threshold value, then zero value is assigned to the red, green and blue components of the this pixel.

Phase 2:

- Computing the features using color co-occurrence methodology. For feature extraction the method used is color co-occurrence method It is the methodology in which both the texture and color of an image are considered, to come to the unique features, which shows that image.

Phase 3:

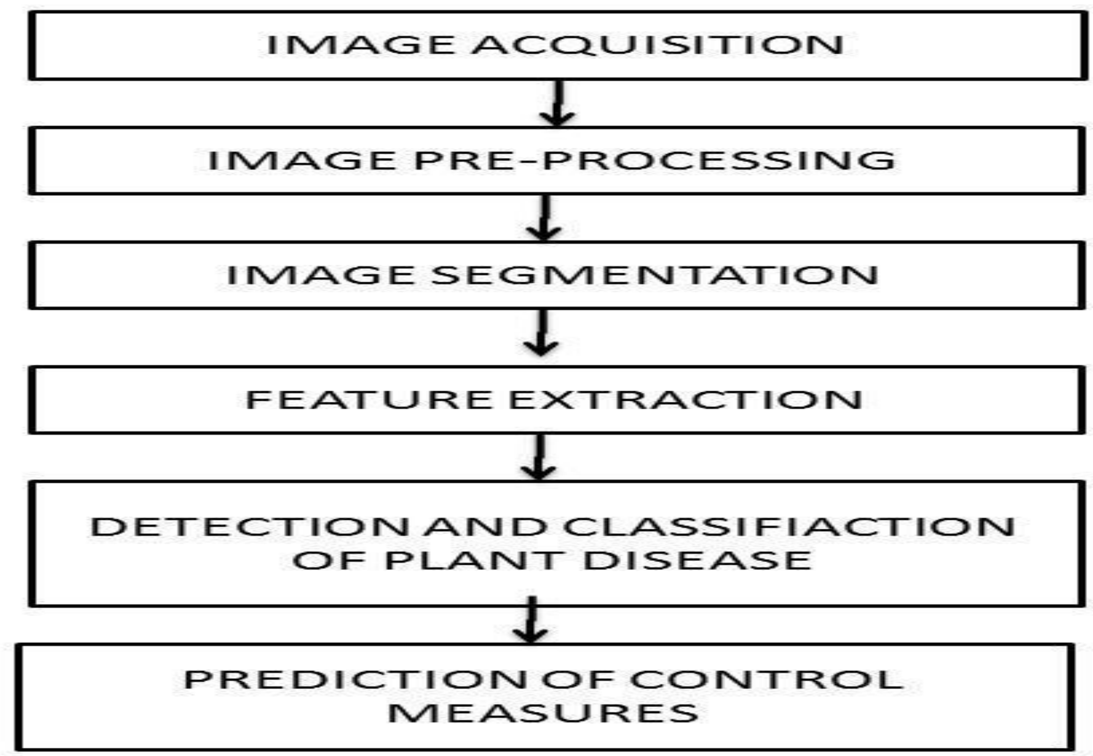
- In this phase classification, extraction and comparison of co-occurrence features for the crop leaves with the corresponding feature values stored in data set.

- The following algorithms are used for the computations:

    1. Support Vector Machine (SVM).
    2. K-Nearest Neighbours

Phase 4:

- Upon the classification and severity detection. The software tool recommends the proposed measures which should be taken in order to combat the disease.



is affected using its image. The underlying objectives are explained as follows:

- To apply image processing techniques to obtain affected portion of the crop and extraction of consequential feature values.
- To perform comparison of extracted values with sample values to identify and classify the disease using various classifier algorithms.
- To integrate and compare results of various classifier algorithms.

- To predict the necessary control measures to cure the disease without any environmental and economic damage.

1.6 Expected Outcome

- The software tool determines the affected region of the image and extracts the consequential values and compares them with sample data to determine and classify the disease associated with the crop and predicts the necessary control measures.
- Further it also compares and analyze the outcome of various classifier algorithms.

2. Software Requirements Specification

2.1. Overall Description

2.1.1 Product Perspective

- Each year, plant viruses and fungal attacks lead to crop losses of up to 30 percent and moreover excess usage of pesticides leads to spreading of fatal diseases such as cancer, autism and asthma.
- The farmers in India are less educated to understand the vulnerability and severe consequences of the diseases.
- They either ignore or recklessly use pesticides to get rid of the problem. This not only leads to spreading of fatal diseases and soil contamination but is also not economically beneficial for them.
- Thus it is very important for them to know the disease and their proper treatment for better crop production and environmental balance.
- The current system involves monitoring from naked eyes which requires time, effort and presence of specialist to detect any crop failure. Thus this tool which processes an image and detects the diseases associated with it is of utmost importance.
- This project will make use of the trained data set and will compare the received information.

2.1.2 Product Functions

- The main feature of this system is to analyze the disease with which the crop is affected. Therefore, to achieve this we need data for various crops for which the software is designed. This dataset includes the images and information about the healthy and unhealthy crops.
- The information received whether in text/image format is pre-processed as per requirement and then given for the extraction of the features of the received information. This processed data is compared and the disease is classified using various classification algorithms such as KNN, SVM.
- Finally, after the classification of the disease the control measures are listed.

2.1.3 Operating Environment

- Windows Operating System or Linux
- MATLAB

2.1.4 Hardware Requirements

- 2 GB RAM
- Intel i3 (minimum)
- 20 GB HDD space

2.2. Functional Requirements

2.2.1 User Interface

- User can use his username and password to log onto the system. If the user is new then he/she can register after specifying the required details.
- User gives the input in the form of an image which is to be processed (image can be obtained by using either upload option or from camera of the device) for detection of disease.
- User can also search information about a particular crop and the diseases associated with it.

2.2.2 System Interface

- The system will provide user name and id and also maintain user session logs.
- For a new user, the system will display images of diseased and healthy leaves of some crops user can select the crop type from the given images.
- For existing user, the system will display crops previously searched or as detailed in the sign-up form.
- There will also be a logout button to redirect back to the login/signup page.

2.2.3 Disease Detection

Disease detection of plant will take place based on information provided by the user and information present in the database by making comparison between the two-information based on several algorithms. This information about the plant in the database will be gained.

2.3 Nonfunctional Requirements

2.3.1 Performance

The system should return correct result with very less percentage from displaying other Data. Also, the web crawling should be done properly to extract all possible information about a given crop.

2.3.2 Availability

For the system to be available to users all the time, online server is must which provide 24 hour service daily.

2.3.3 Security

The password of all the users should be stored properly and not exposed to anyone.

2.3.4 Maintainability

The code will be well documented. Particular care will be taken to design the software modularly to ensure that maintenance is easy.

2.3.5 Scalability

The system will detect the presence of disease in the crop, with the type of disease that exists and if so will also recommend type and amount of pesticides to be used to make crop disease free.

3. Software Design Specification

3.1 System architecture description

This section is the main focus of the first version of SDS, the high level design. This should give a good view of exact organization of the system as per the requirements.

3.1.1 Overview of Modules

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

- **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 form of RGB (Red, Green, Blue)

value. All pixel will be assigned a value by the image segmentation.

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

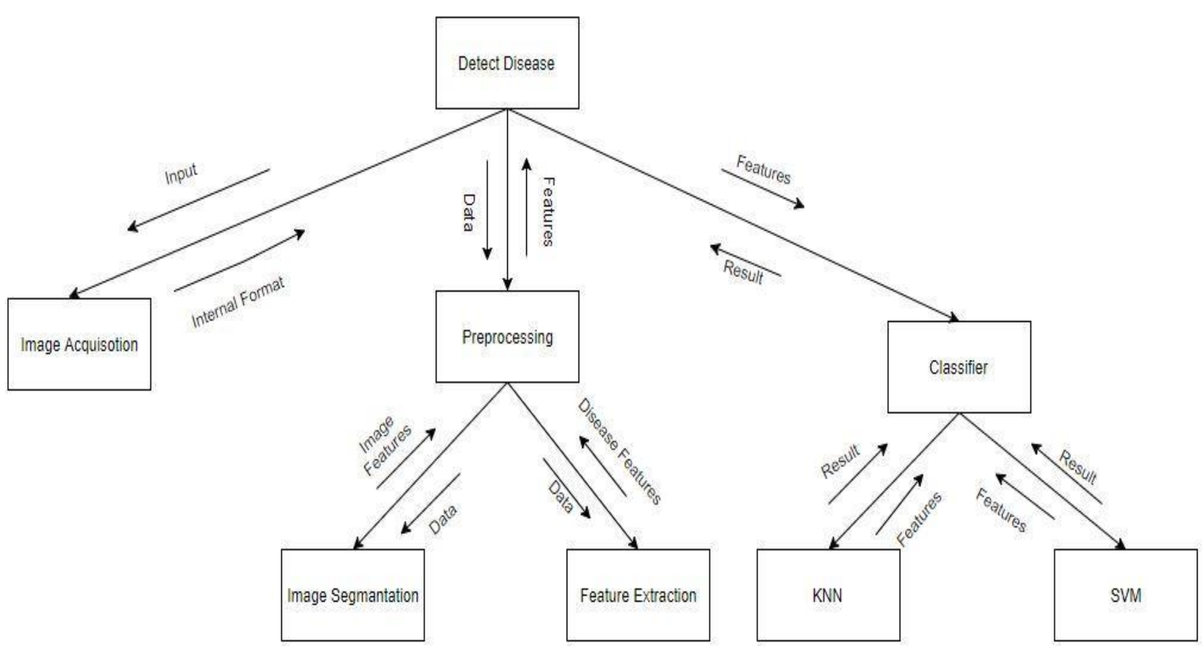
3.1.2 Structure and relationships

The functionalities available to the user will be based on the type of login details entered, we have two types of user a general user and the other one the administrator. The user can use the functionalities such as details submission, disease detection and disease enquiring. On the other hand, the administrator can add on new features such as modifying the dataset used and providing other utilities.

3.1.3 User interface issues

The GUI can be divided into 2 parts. One is for the user, and the other is for the Administrator of the system. The user side interface will be optimized to provide the user with an easy to use interface.

3.2 Structured Chart



3.3 Detailed description of components

3.3.1 Image Acquisition

Identification	Input Image
Type	Module
Input	An image in which we have to detect the disease segments.
Function	The main function of this module is to allow the user to upload the image in the system.
Output	The image provided by the user.

3.3.2 Image Pre-Processing and Segmentation

Identification	Preprocessing
Type	Module
Input	The input is the image provided by the user.
Function	The main function of this module is to pre-process the given input image. A matrix will be formed in which the value of each pixel of image will be stored in form of RGB (Red, Green, Blue) value.
Dependencies	The pixel value will be dependent on the image format.
Processing	An input image will be provided by the user. The image segmentation will be performed and the image is segmented into pixels and the value of each pixel (in form of RGB) will be stored in form of matrix.
Output	The output is the matrix in which the value of each pixel will be stored.

3.3.3 Feature Extraction

Identification	Feature Extraction
Type	Module
Input	The input is the segmented image provided by the previous module which is the processed form of the image obtained from the user.
Function	The main function of this module is to extract feature of the given input image. A matrix will be formed in which the features to be

	judged are stored which will be compared to the pre-stored values.
Dependencies	This unit depends on the segmented image.
Processing	The image is obtained from the user is segmented and given to this unit for feature extraction. The various parameters are calculated and stored to represent the color changes and textures.
Output	The output is the matrix in which the values associated with image are stored.

3.3.4 Image Classification

Identification	Classification
Type	Module
Input	The features extracted from the segmented image.
Purpose	To obtain the disease affected.
Function	The main function of this module is apply the classification algorithms on the extracted features and use predefined dataset to detect the disease with which the plant is affected.
Resources	The dataset is required for this process.
Output	The output from this unit is the integrated output from the various classifier algorithms.

3.3.5 Output

Identification	Output
Type	Module
Purpose	The purpose of this module is to provide the output which are the pesticides associated with the disease.
Processing	Depending upon the disease detected the pesticides associated and their amount is calculated according to the intensity of the disease detected.
Output	The output is pesticide type and quantity according to the disease detected.

4. Implementation

4.1 Preprocessing

Dataset containing images pertaining to various diseases and healthy leaves of pomegranate is taken to evaluate features associated with every image.

The features evaluated per image are stored in Taining_Data.m. The attributes of the image are numerical attributes representing these values:

1. Mean
2.Standard Deviation
3.Entropy
4.RMS (Root Mean Square)
5.Variance
6.Smoothness
7. Kurtosis
8. Skewness
9. IDM(Inverse Difference Moment)
10. Contrast
11. Correlation
12. Energy
13. Homogeneity

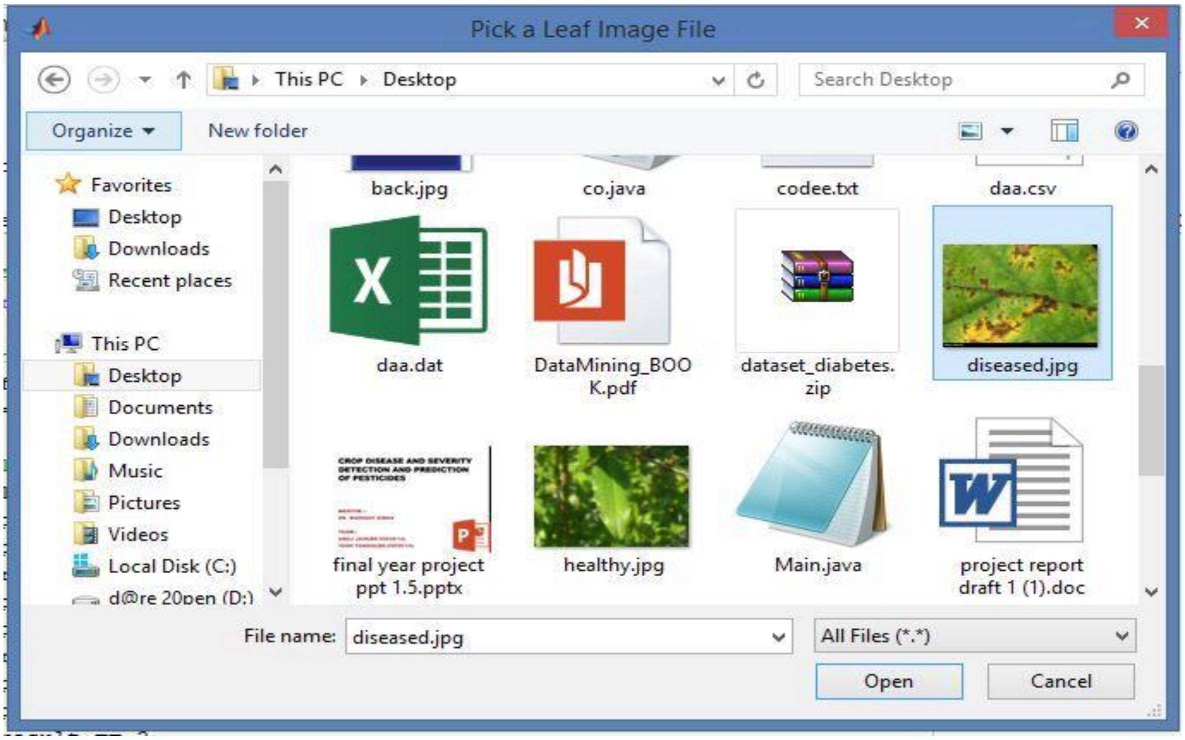
Train_Feat | Train_Label

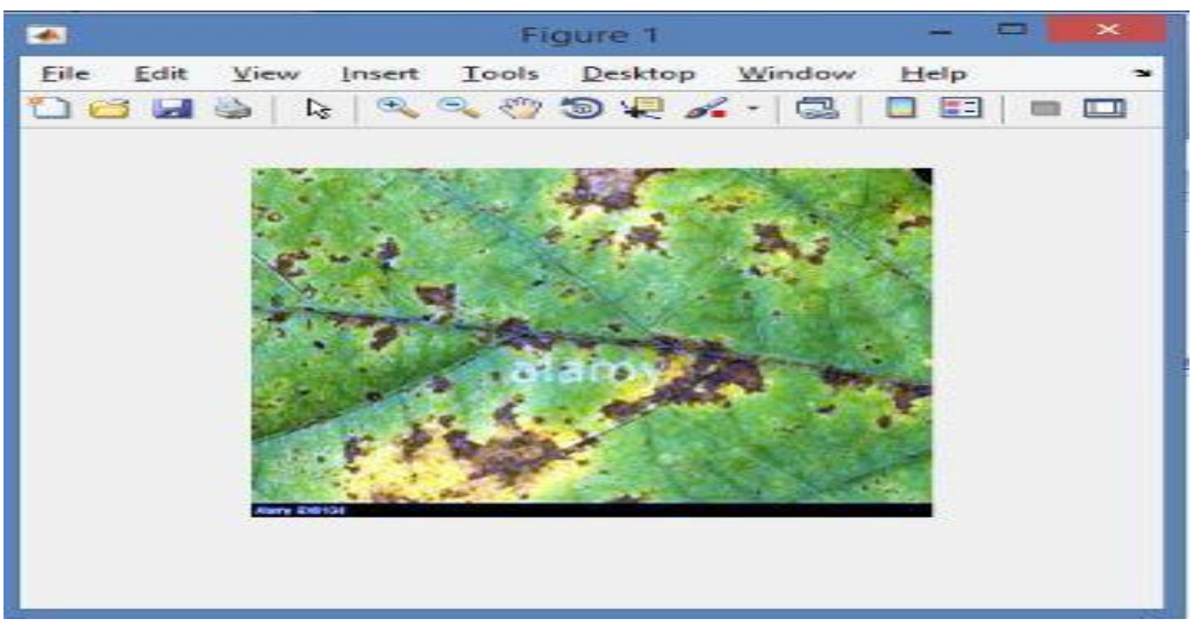
125x13 double

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0789	0.9783	0.7626	0.9749	14.8439	47.8117	1.7099	5.5748	2.1507e+03	1.0000	15.5978	3.6320	255
2	0.4668	0.8657	0.7967	0.9592	14.1501	48.1396	1.3658	4.3136	1.6322e+03	1.0000	15.7654	3.6744	255
3	0.3676	0.9102	0.7573	0.9625	16.4441	51.4194	1.6679	5.3404	2.3050e+03	1.0000	13.7926	3.4025	255
4	0.5412	0.7510	0.5382	0.9222	17.9717	37.6635	2.5829	7.4037	1.3068e+03	1.0000	10.4951	2.5883	255
5	0.5128	0.7103	0.8947	0.9717	17.1185	35.5205	2.8432	10.4505	1.1622e+03	1.0000	27.6033	4.6820	255
6	0.6976	0.8739	0.4873	0.9104	31.5604	56.4596	2.9830	8.1140	2.8443e+03	1.0000	4.4008	1.6129	255
7	0.4886	0.9580	0.2687	0.9403	71.8528	83.0729	5.1204	11.4616	5.6827e+03	1.0000	1.8270	0.6497	255
8	0.4309	0.8966	0.7660	0.9656	17.4376	52.4639	1.8789	5.7289	2.0524e+03	1.0000	12.8361	3.2736	255
9	0.5761	0.9092	0.7104	0.9584	23.8136	60.2088	1.6734	5.4362	3.2303e+03	1.0000	6.9582	2.3349	255
10	0.7462	0.9098	0.5279	0.9007	40.0473	73.8575	2.9119	7.5330	4.4652e+03	1.0000	3.9932	1.5873	255
11	0.8894	0.8263	0.8185	0.9651	16.4181	55.6534	1.3002	4.3228	2.8415e+03	1.0000	12.3204	3.3010	255
12	0.4140	0.9702	0.4106	0.9730	76.6184	97.9821	3.8439	9.3642	6.3323e+03	1.0000	1.5171	0.5990	255
13	0.0863	0.9491	0.8848	0.9934	8.5755	35.4527	0.7176	2.5151	1.1104e+03	1.0000	18.7524	4.1059	255
14	1.0451	0.8167	0.6192	0.9209	26.9492	60.8740	2.4818	6.9560	3.4713e+03	1.0000	6.6288	2.2075	255
15	0.4131	0.8459	0.8424	0.9766	10.6032	41.4724	1.1945	3.8289	1.5944e+03	1.0000	22.3461	4.4208	255
16	1.0066	0.7952	0.7960	0.9543	16.5970	54.7251	1.2976	4.6009	2.7719e+03	1.0000	12.3711	3.2826	255
17	1.3198	0.8648	0.4802	0.9194	44.1780	76.8477	3.3220	8.6895	5.4942e+03	1.0000	3.4116	1.4309	255
18	0.2745	0.8710	0.8596	0.9791	9.7420	38.4817	0.8864	3.5148	1.3888e+03	1.0000	19.3455	4.1094	175
19	0.5655	0.8692	0.6730	0.9549	21.6041	53.8695	2.3224	6.2055	2.4847e+03	1.0000	9.9314	2.7350	255
20	0.2625	0.8792	0.7989	0.9765	13.2757	42.4137	1.2711	4.3011	1.6140e+03	1.0000	12.9619	3.2818	255

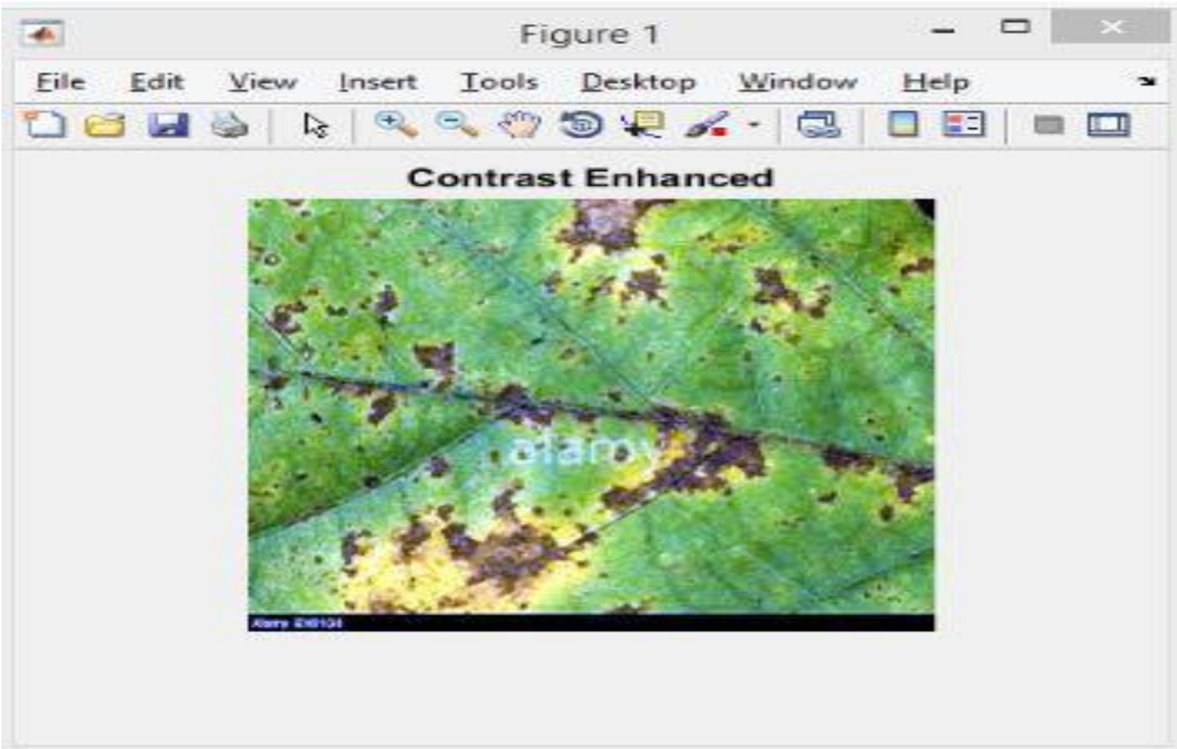
4.2 Image Acquisition

1. Image to be classified is taken from user.

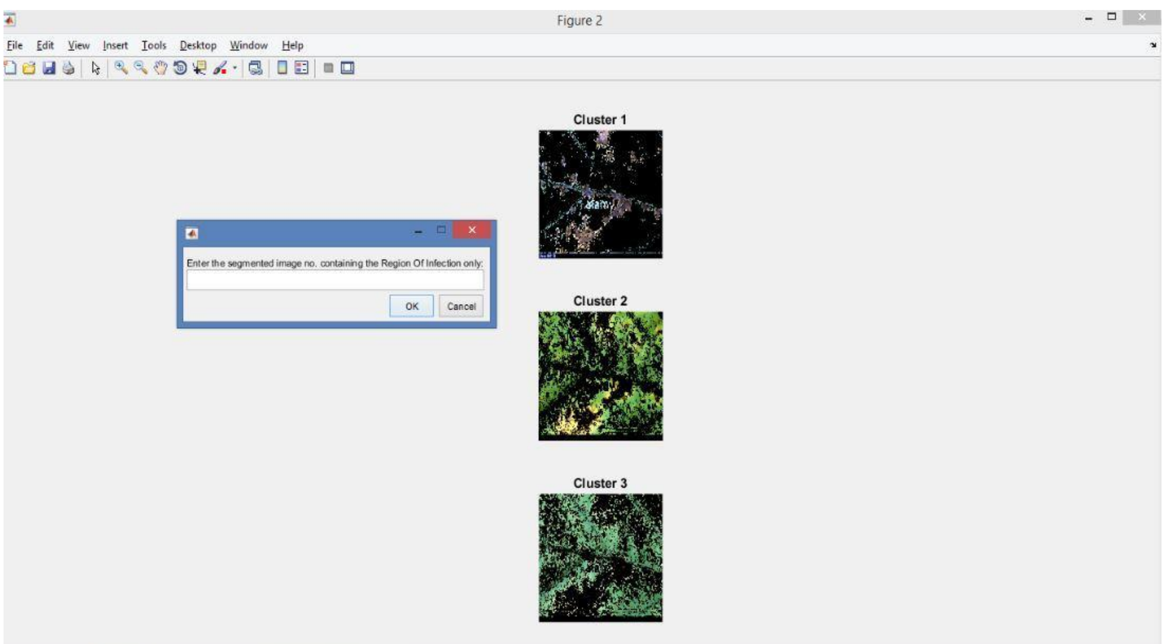




2. Image is Contrast Intensified.



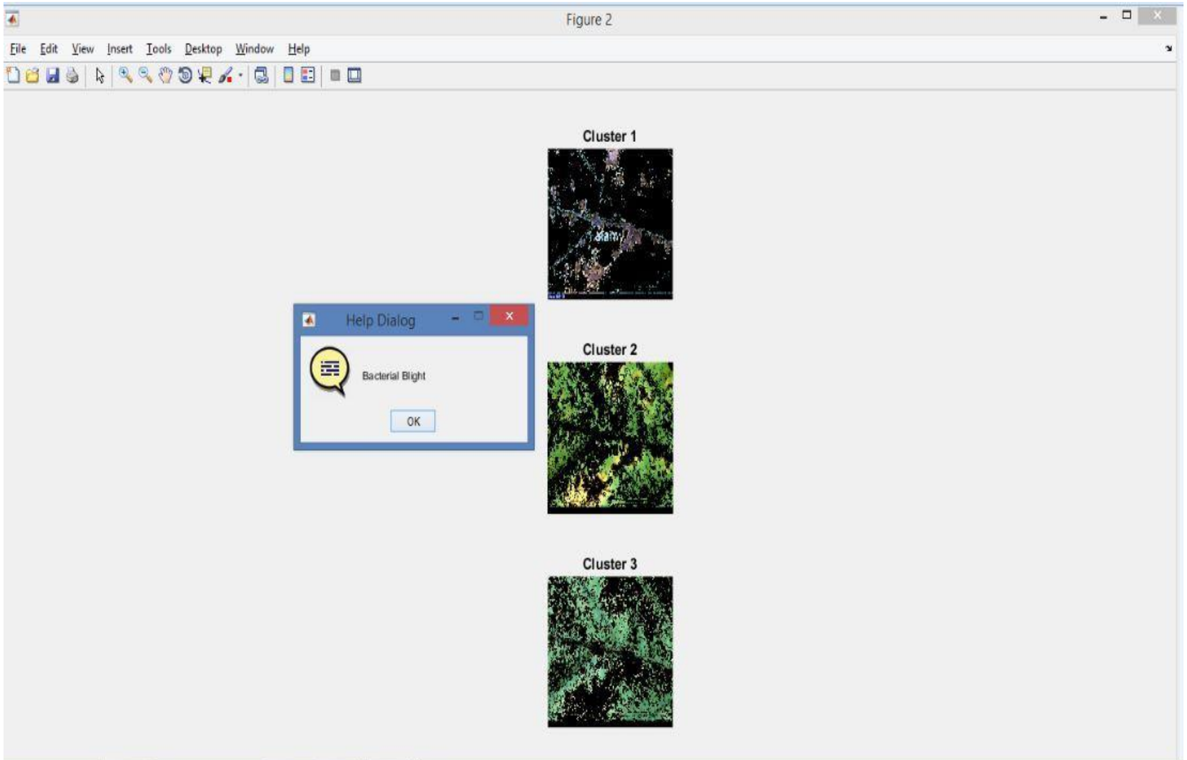
3. Image is segmented according to 3 colors (RGB). The segmented image is selected from 3 segmented images which best depicts the Region of Infection.



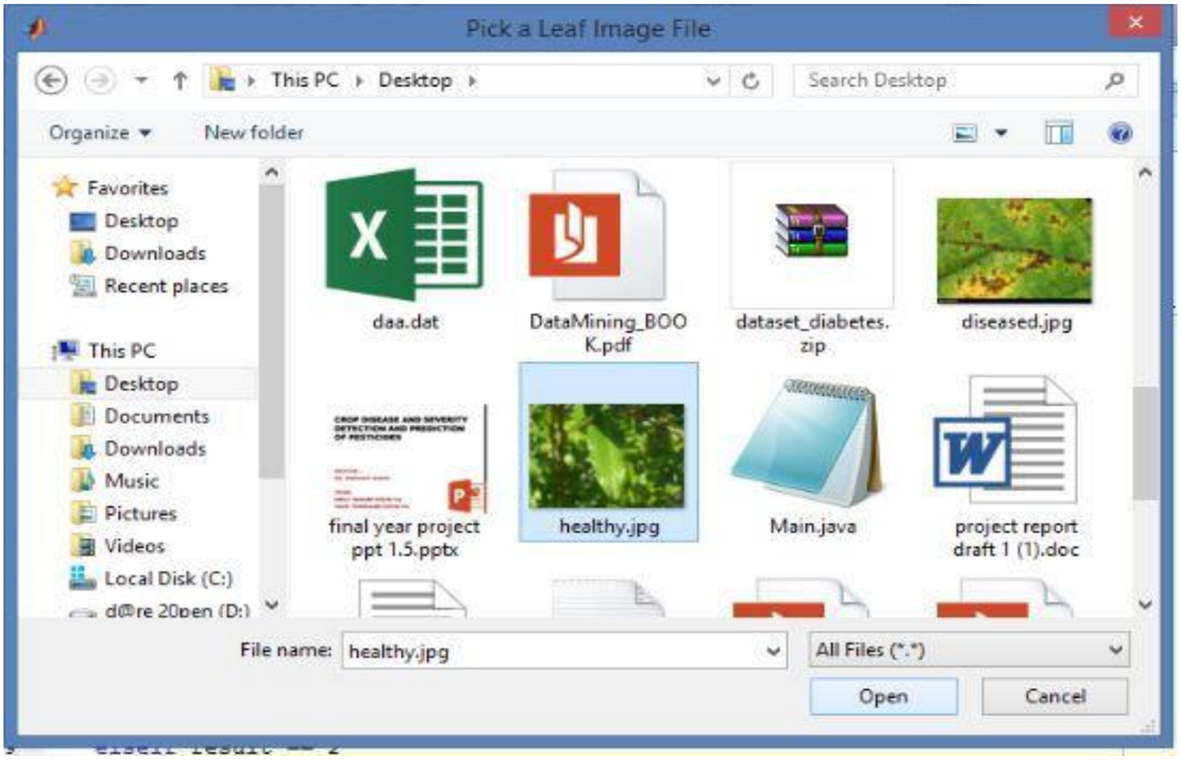
4.The features of the segmented image are calculated and presented to the classifier.

4.3 Classification using SVM

The multiSVM function detects the number of classes from Training_Data.mat.



Upon evaluation of number of classes, the multiSVM classifies the image as one of the classes. The multiSVM function returns the class number and using that the disease is displayed. Similarly, when we select a healthy image as our input then the no disease is detected and we have a healthy leaf displayed.


Pick a Leaf Image File
This PC ▸ Desktop ▸
Search Desktop
Organize ▾
New folder
Favorites
Desktop
Downloads
Recent places
This PC
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
daa.dat
DataMining_BOOK.pdf
dataset_diabetes.zip
diseased.jpg
final year project ppt 1.5.pptx
healthy.jpg
Main.java
project report draft 1 (1).doc
File name:
healthy.jpg
All Files (*.*)
Open
Cancel

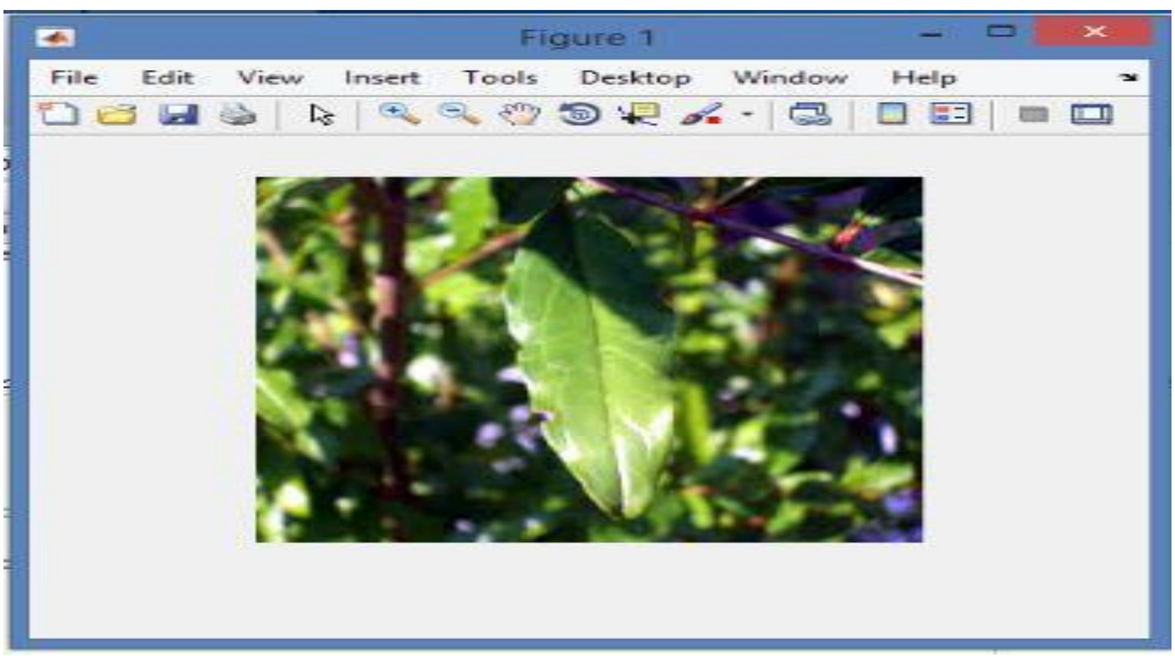
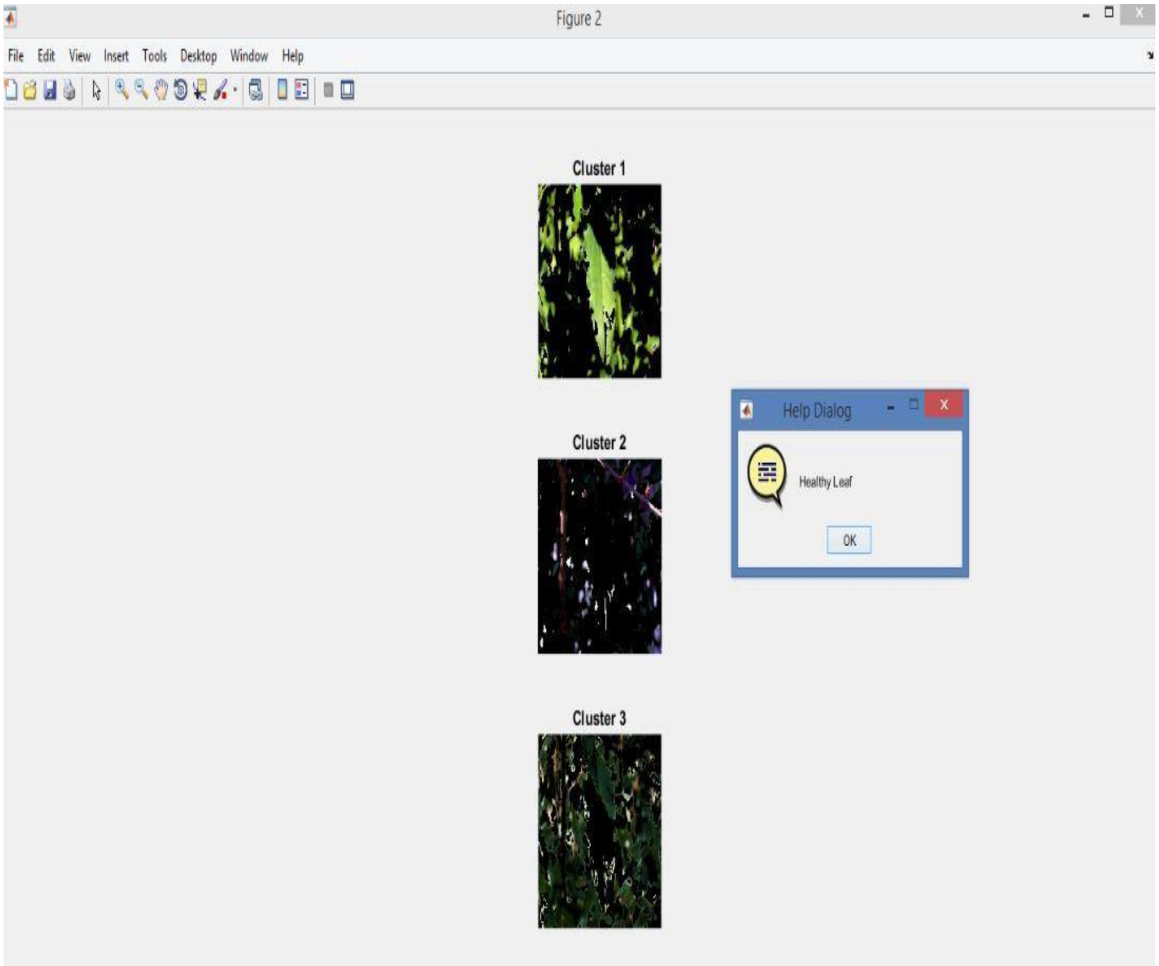
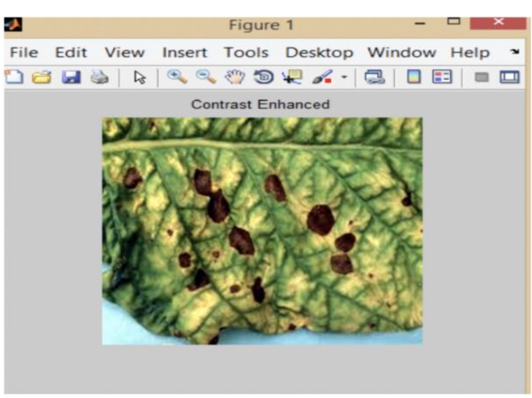


Figure 1
File
Edit
View
Insert
Tools
Desktop
Window
Help

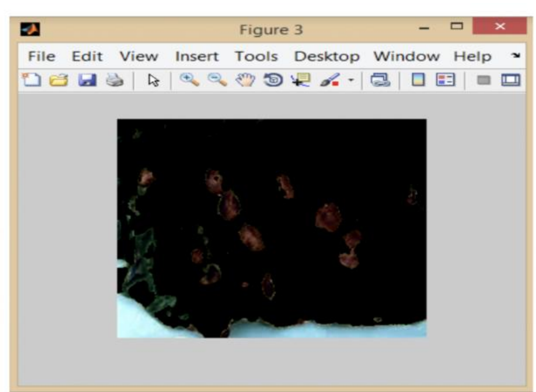


4.4 Results for various Diseases from both KNN and SVM

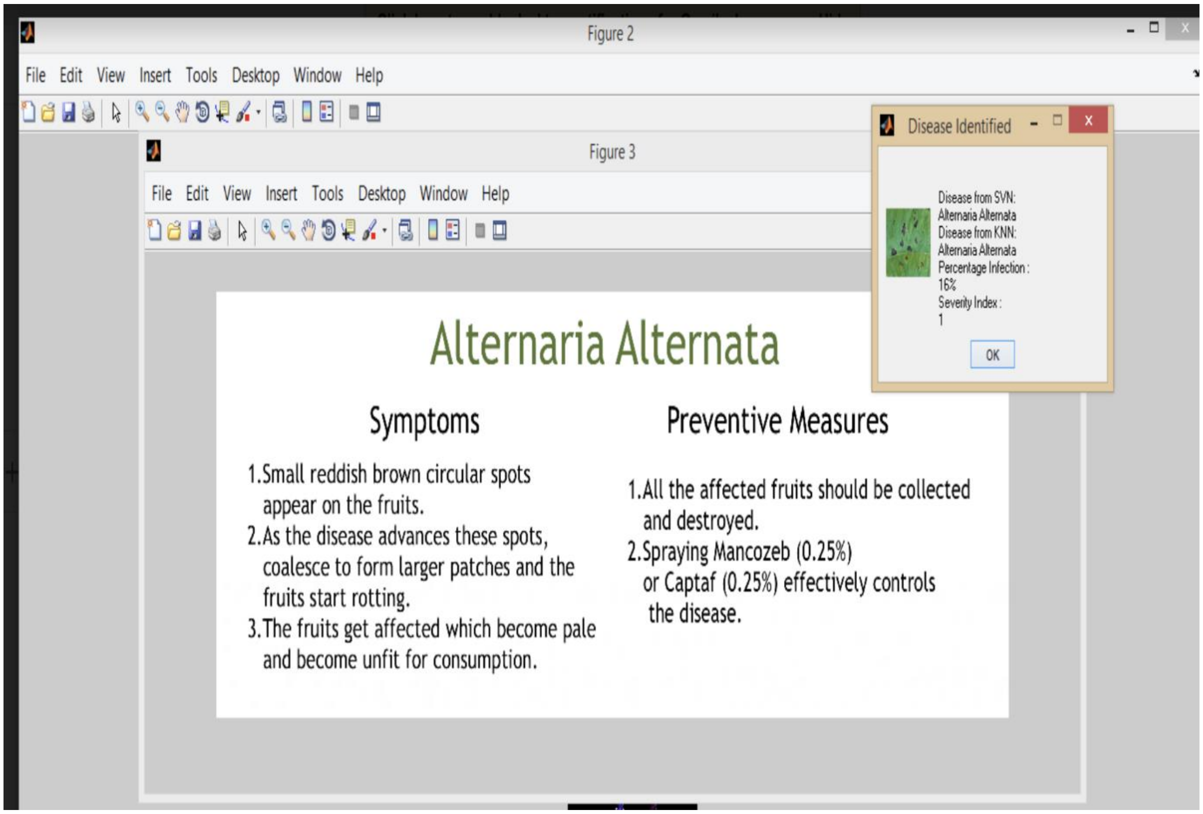
4.4.1 Alternaria Alternata



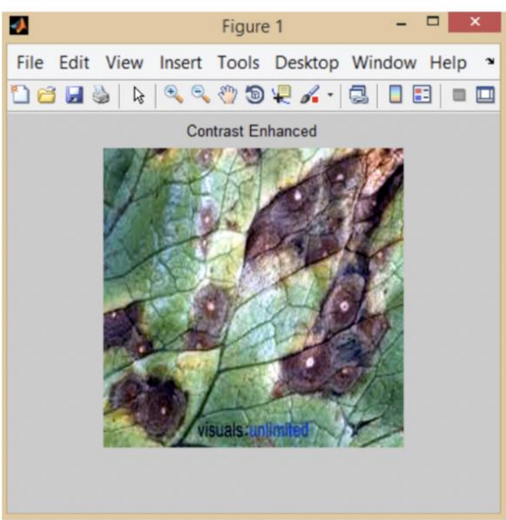
Contrasted Image



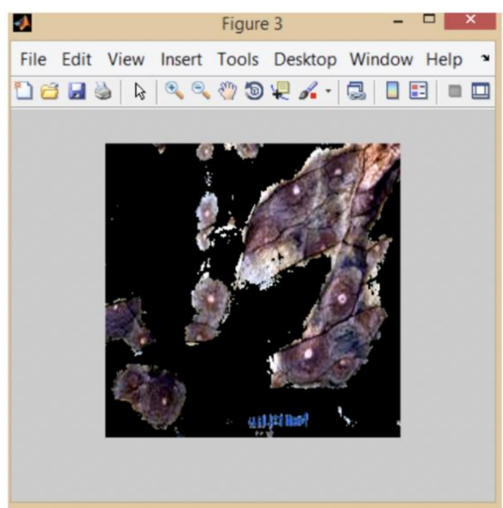
Segmented Image



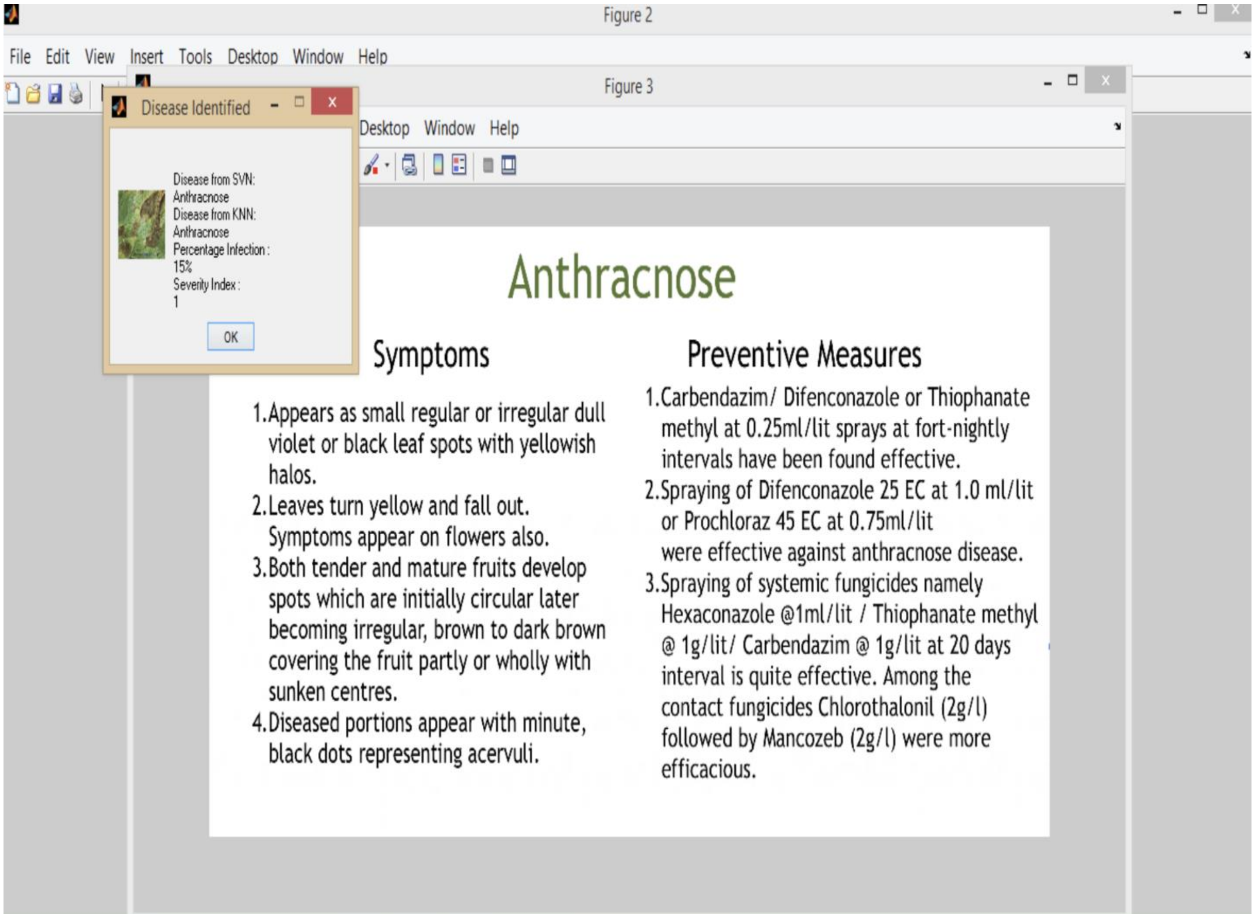
4.4.2 Anthracnose



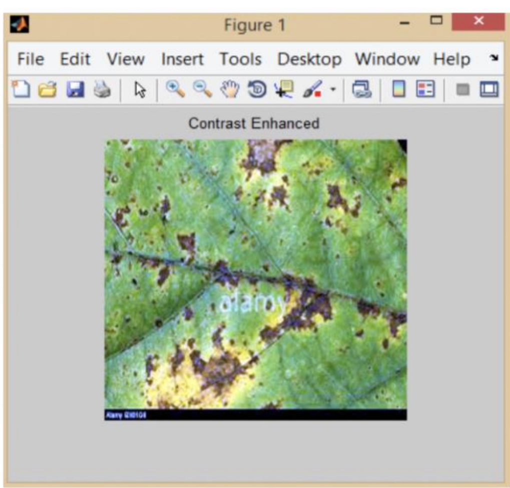
Contrasted Image



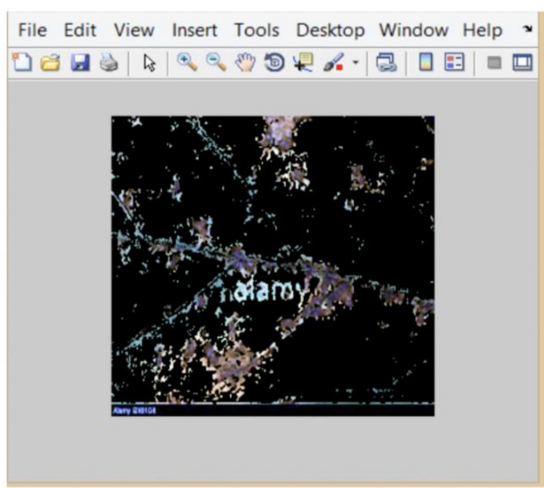
Segmented Image



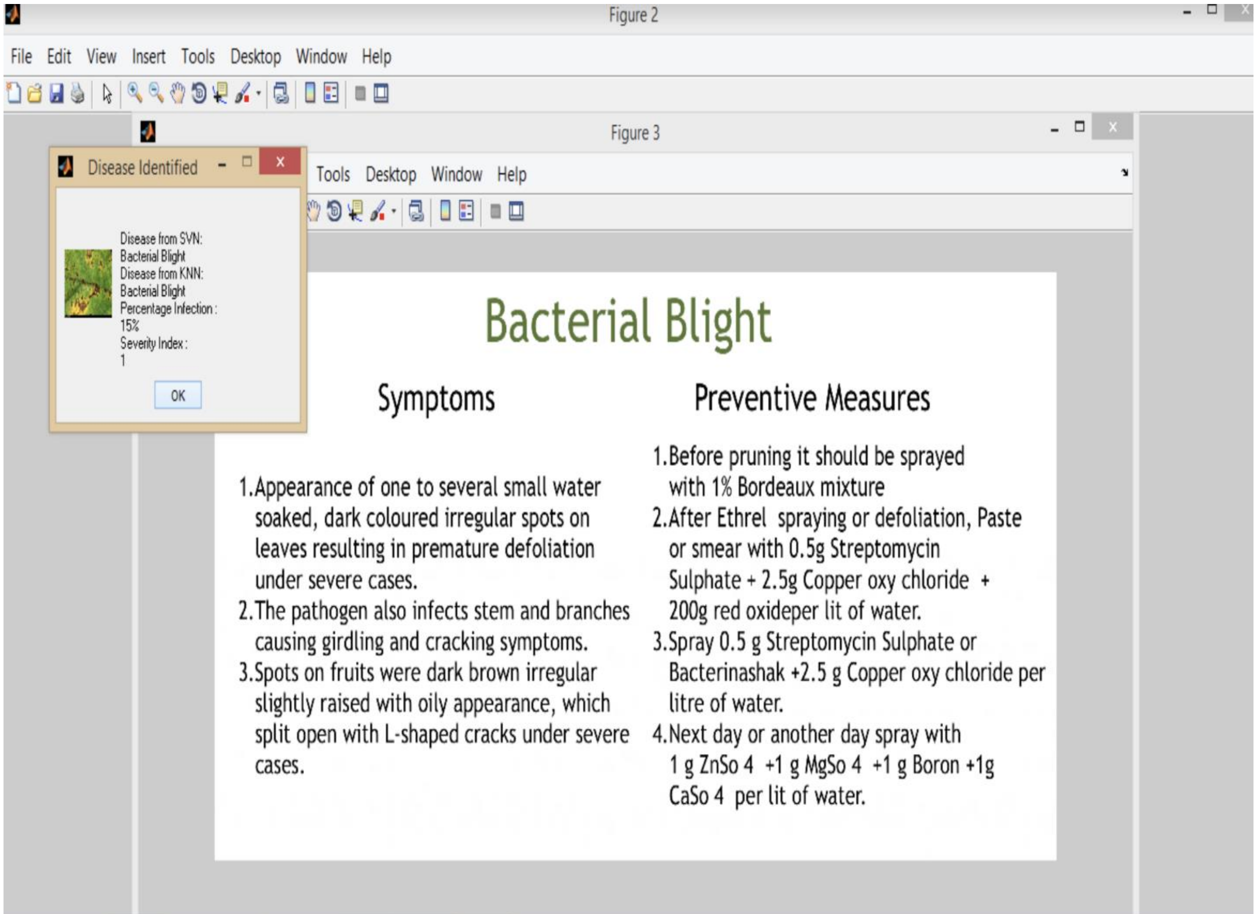
4.4.1 Bacterial Blight



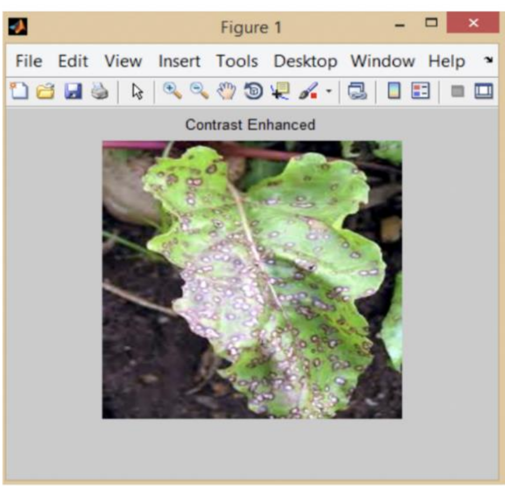
Contrasted Image



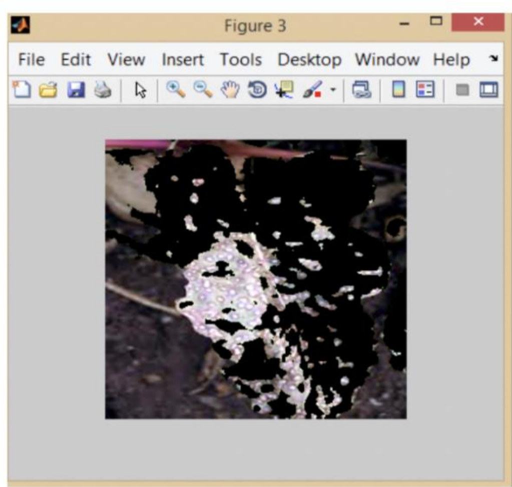
Segmented Image



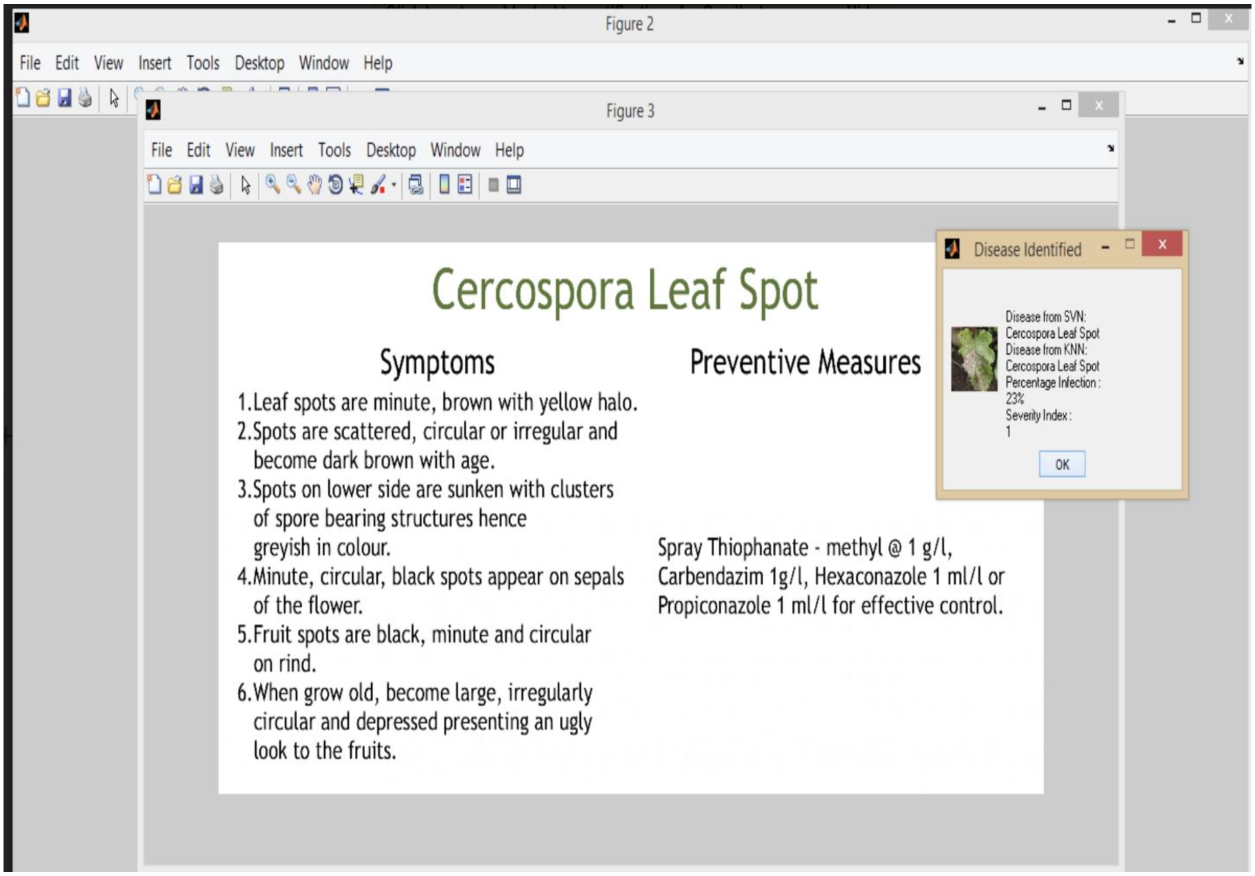
4.4.1 Cercospora Leaf Spot



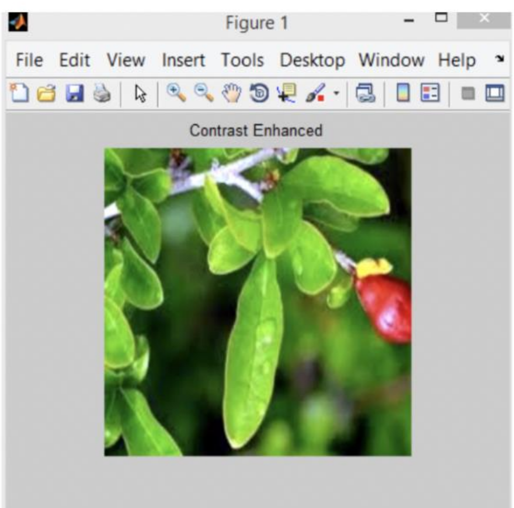
Contrasted Image



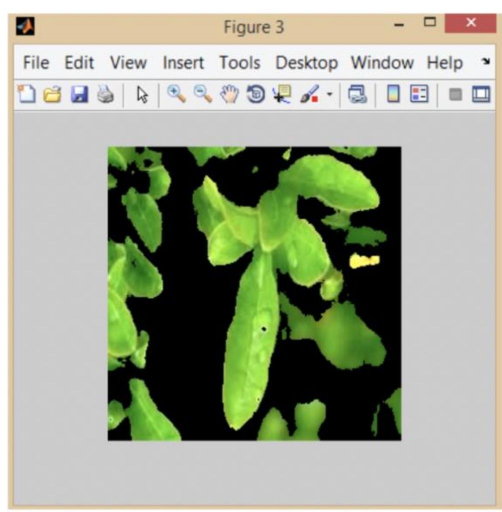
Segmented Image



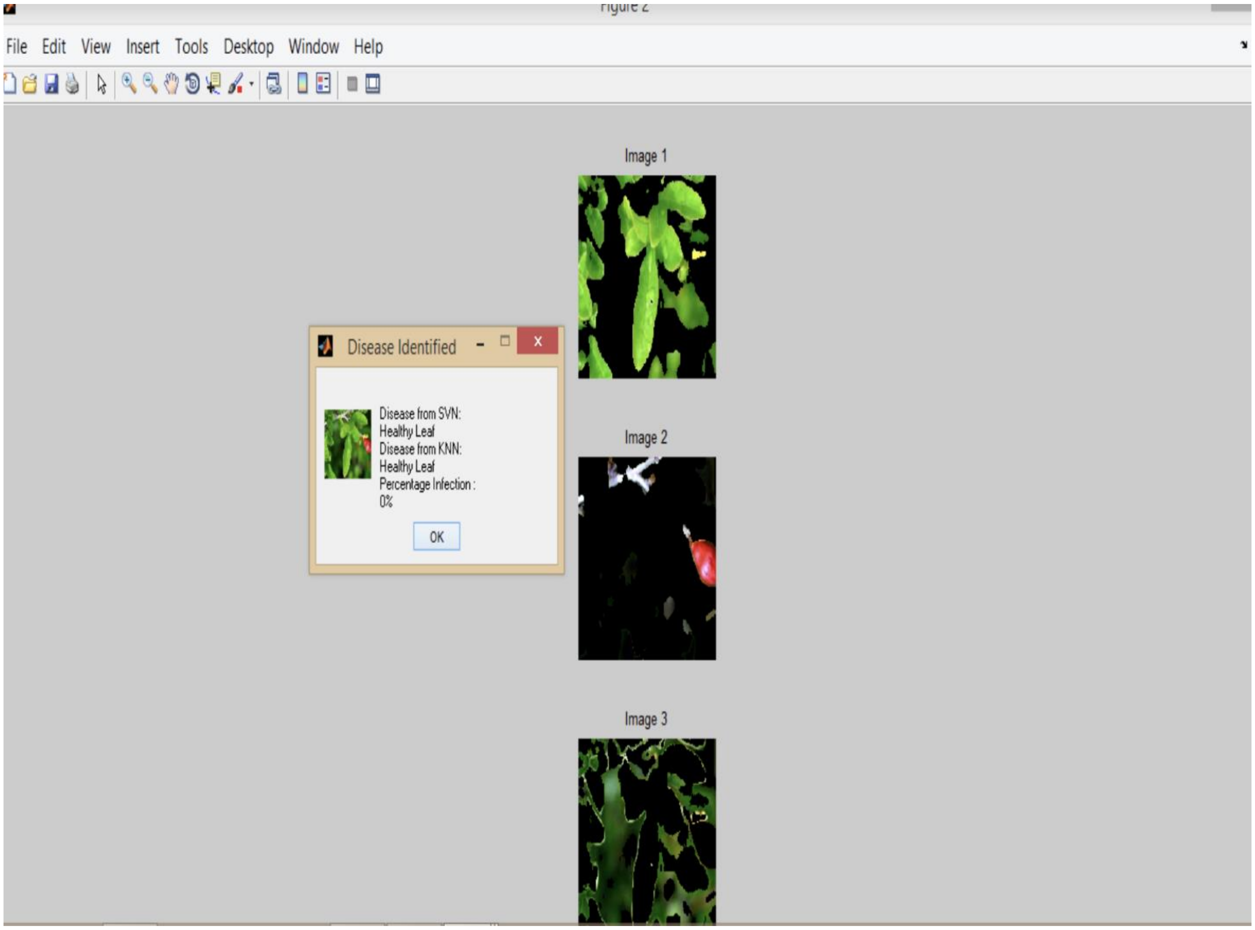
4.4.1 Healthy Leaf



Contrasted Image



Segmented Image



5. Workflow and Functions Used

5.1 Detection.

This code segment takes an image input,applies image segmentation and pre proceedings described above calculates the features of the image and send the images to the classification algorithms.

Functions performed by this detection module:

1. It takes the image from the user,convert it into standard size, enhances contrast using color transformations.
2. On the contrasted image K means clustering is applied to get color segmented images which define the disease affected image.
3. From these segmented images the image which best represents the leaf affected is presented to the classification part.
4. The image features are evaluated and presented to the classifiers which in turn assigns the class label of disease to the image.
5. Now,the affected area is analysed to calculate percentage of infection in the leaf.
6. Then the results and visualised and preventive measures according to the disease are displayed.

5.2 SVM Classifier

The following function takes training data,training labels and test data to assign label to test data.

The code determines the type of disease using 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.

5.3 KNN Classifier

The following function takes nearest neighbours, training data,training labels and test data to assign label to test data.

The code determines the type of disease using 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.

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.

6. References

- [1]www.catalog.data.gov/dataset?tags=plant+disease
- [2]www.reddit.com/r/datasets/comments/5uljlp/plant_leaf_disease_datasets/
- [3]www.sciencedirect.com/science/article/pii/S2214317316300154
- [4]https://pdfs.semanticscholar.org/0a10/96ca51b4b8067865211b42c0bec4c9969f34.pdf
- [5]http://www.krishisewa.com/articles/disease-management/398-pomegranate-diseases.html
- [6] https://www.ijarcce.com/upload/2015/october-15/IJARCCE%20116.pdf