MACHINE VISION BASED CARROT DISEASE RECOGNITION

BY

PAPIYA HOSSAIN LIMA
ID: 142-15-4049

AND

SAIFUDDIN SOURAV

ID: 142-15-4021

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

ANUP MAJUMDER
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Md. Tarek Habib
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

MAY 2018

APPROVAL

This Research titled "**Machine Vision based Carrot Disease Recognition",** submitted by **Papiya Hossain Lima** and **Saifuddin Sourav** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 6^{th} May 2018.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Zahid Hasan — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Anup Majumder, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Anup Majumder
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Md. Tarek Habib
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Papiya Hossain Lima
ID: 142-15-4049
Department of CSE
Daffodil International University

Saifuddin Sourav
ID: 142-15-4021
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty Allah for His divine blessing makes us possible to complete this project successfully.

We fell grateful to and wish our profound our indebtedness to **Anup Majumder, Lecturer,** Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of Research in Machine Learning influenced us to carry out this research. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this research.

We would like to express our heartiest gratitude to **Professor Dr. Syed Akhter Hossain**, Head**,** Department of CSE, and **Md. Tarek Habib,** Assistant Professor**,** Department of CSE, Daffodil International University, Dhaka and for their kind help to finish our research and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Most of the people in Bangladesh lives on agriculture producing rice, jute, vegetables and so many crops. Fruits and vegetables have a large part of the agricultural sector which is still produced and supervised manually. Sometimes the diseases are not seen in the initial stage, which can severely affect our GDP. To ensure the freshness of fruits and vegetables modern image processing tools can help a lot. Experts can detect the defected fruits and vegetables by watching them with their eyes but the process is too long and not suitable for all the stores, farms, supermarkets or the exporters all around. There comes the blessings of new computer vision technologies with image processing techniques that can do a lot of works in a second. In this paper an automated approach is developed to detect defects of fruits and vegetables and recognize diseases by using machine vision based image processing techniques which is implemented in MATLAB including a machine learning algorithm with the Multiclass SVM classifier. There are many algorithms that can detect defects of fruits and vegetables hence, we separated the defected parts of the carrots using K-means clustering and then classified it with multiclass support vector machine classifier. Here, a supervised machine learning concept is implemented to recognize various carrot diseases. As the domain of this research model, carrot diseases are classified and 96% of accuracy is achieved which can certainly help in our agricultural science along with proper maintenance.

TABLE OF CONTENT

CONTENTS **PAGE**

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Introduction

Most people in this densely populated country Bangladesh are living on agriculture, contributing 19.6 percent to the national GDP and providing employment for 63 percent of the population. Mainly producing rice, jute, tea also with fruits and vegetables people are living by here. But it is very pathetic that our country is not well digitized yet. There are so many features available provided by the modern technologies that can make a great change in our agricultural sector. If we say about developed countries like U.S.A, England, China, Malaysia etc. they are practicing many advanced technologies in their farming where in an agriculture based country like Bangladesh has not given much attention to the modern technologies that can contribute an enormous amount of productivity in farming. There are noble and prominent research possibility and opportunity are available in our agricultural science that are not being risen properly. The neighbor country India is using many features of advanced image processing technologies that helps their farmers to get a secured farming environment. Fruits and vegetables like Apple, Guava, Pomegranate and Tomato etc. are supervised by automated system that recognizes problems in a very short time. Our country also started to use many digital techniques to reduce the risks of loss and produce more profit. Producing profit from harvest is not very easy to make if disease of the crop attacks unexpectedly. The entire harvest can be destroyed without proper maintenance which can be a matter of fluctuation in our targeted GDP also to the livelihood of the farmers. To avoid this kind of trouble a proper maintained system should be there which can help the farmers to detect the diseases before it is too late so that they can take proper step for the future harvesting. In this research an automated system is developed which can certainly help farmers to detect diseases by capturing the image with a camera and then they should input the image to the processing system. The system analyzes the image and extract the features from it. The image is segmented through k-means clustering and then classified by multi SVM classifier to detect the predicted disease name.

1.2 Motivation

Bangladesh, for the first time in history, has fulfilled the eligibility criteria set by the United Nations to be recognized as a developing country, crossing over from the list of least developed countries (LDCs) in this year. Agriculture is one of the major sources of income from the early period of Bangladesh. But still we are depending on the manual supervision system which follows very slow process of managing the problems by naked eye while for the safe harvesting a quick and easy approach should be made to ensure the constant production of crops. Production system or marketing system can go down if the diseases cannot be discriminate in time. Diseases come ghastly and leave an awful situation to the farmers. A solution f this huge problem can be solved by developing a system that can recognize and detect the disease in crops in a second.

Digital Image processing techniques are being vastly used to detect diseases worldwide. To research on this topic make it fast and more accurate, various carrot diseases are used. The diseases are-

- Black Rot
- Growth Crack
- Root Fly
- Root Knot
- Violet Root Knot, and
- Healthy Carrots

These six categories are used to analyze the prediction system. Some precautions can be made if the disease are remarked at early age. For example, if "Black Rot" disease is detected in a land then it is advised not to harvest more carrots on that land until the risks are gone and seeds for further production should not be collected from the affected plants.

To build the system many images are collected to train the system in MATLAB using k-means clustering and multiclass support vector machine classifier.

1.3 Objectives

The objectives of this research to find out and analyze a technique which can help various farmers, wholesaler or retailers to detect the defected fruits or vegetables from the farms or store and thus make a suitable profits. After identifying the diseases, farmers can apply necessary steps to avoid big damage of harvesting. The proposed system is user friendly, fast to execute and can be implemented easily in any system.

- The system can reduce traditional problems of farmers of Bangladesh.
- Diseases of fruits and vegetables can bring a huge loss to the farmer’s as well as our country’s economic growth.
- Early detection of the diseases can help them to minimize the danger.
- This approach can be used to make profit to producers.
- Also detection of particular problem will lead them to get advices to grow crops in a more secured way.
- To explore features that suited the fruit disease recognition problem.
- To use SVM in the Multi-class scenario.

1.4 Expected Outcome

Our farmers are mostly lives in remote areas where problem solving is often a matter of time when they need to wait for the actual agriculturist to show their crops or they need to go to the agriculturist physically with the disease affected crops. We thought of a procedure which will be available to them to identify the disease by capturing images of the particular affected fruit or vegetable and give the image as input. The problem will be identified by using k-means segmentation and SVM classifier.

K-means is a technique which serves the partitioning of the observations with the mean of the nearest clusters. Using k-means and multiclass support vector machine images of carrots are processed and implemented in MATLAB to get the final output.

- Using this approach defected fruits and vegetables can be detected.
- Diseases are classified.
- An automated system can be designed to avoid unexpected rotten fruits or vegetables
- Good profit can be made by producers or wholesalers.
- Reduce time, loss of profit and unorganized production.

1.5 Report Layout

Chapter 1: Introduction

In this chapter we have discussed about the introduction, motivation of the work, objectives and expected outcome of the research work and the report layout.

Chapter 2: Background

We discussed about the background circumstances of our work. We also delivered the literature review, research summery, scope of the problem and challenges of the system.

Chapter 3: Research Methodology

This chapter is all about the procedure used to build the system. This section has the methods and steps, data collection procedure, some statistical analysis of the proposed system.

Chapter 4: Experimental results and discussion

In this chapter all the experimental result that has been achieved by the proposed system is discussed along with the performance analysis and a summary of the result is covered.

Chapter 5: Conclusions

This chapter contains the conclusion part and the ideas of implication of further study on this topic.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

In this chapter, we discuss on several research work done by researchers in the area of image categorization, fruits recognition, fruit diseases identification.

Farmers in our country are not aware of the technological growth that happened in the developing countries. They also need to have the opportunity to cultivate their crops in a better way that reduces the chance of loss on their profit. Detection of diseases or defects in fruits and vegetables with computer technology needs an image processing technique and a platform on which the experiments can be done fruitfully. The proposed system is able to detect defected carrots and also can recognize the specific disease of carrot from images. Many researcher has experimented on many techniques that detects the diseases of fruits and vegetables using various techniques.

Image processing is a technique that acquires an image and analyze it, enhance or collect useful information from images and finally it output the result in an explainable or apprehensible format. The image may be analyzed to find patterns that aren't visible by the human eye. People can take decisions after getting the output, sometimes the decisions also can be made by the machine itself.

In this research MATLAB is used to train and analyze the data from numerous images of carrots. MATLAB is a platform where the digital image processing algorithms are implemented.

Earlier, many people have worked with image processing to ensure the use of technology in agriculture. But it is not being used in our country adequately. Several image processing techniques are being developed to ensure the compactness of methodologies available.

Computer vision is a system that can describe a system what it is containing and what does it mean. In this work computer vision based system is used recognize disease from numerous carrot images. It is a supervised learning process where disease names are used to label the classes.

2.2 Literature Review

Many Image processing techniques were developed through many years of research for object detection and use of classifiers as well. In image segmentation histogram based techniques are efficient because it requires only one pass through the pixels [1]. Image acquisition is the process of collecting an image from some hardware based sources that can be used next for further processing [2].

In paper [3], size measurement, average slope difference, first derivative of the diam. profile, and connected component algorithms are used with computer vision concept to detect carrot's defected areas for grading of carrots.

In paper [4], a quantitative method is developed to estimate tip shape for classification of carrots which uses curvature profile based on Freeman-chain code. A parameter estimation technique Levenberg-marquardt method is used to extract several critical features of the carrots and its shapes. Finally a Bayes decision classifier is used to classify the tip shapes. Using this classifier 86% of the data were successfully classified.

Potatoes were inspected dynamically in a rule based control system in paper [5]. Multivariate discrimination techniques used to classify from the color and the shape separation is done by Fourier domain. The developed method used a combined approach of statistical and structural texture analysis for bruise, disease, injury and russeting scab. The multivariate discriminant technique gave an accuracy of 90% for greening potato.

F. López-García and G. Andreu-García in their paper [6] used a Multivariate Image Analysis strategy to develop an automatic detection of citrus fruit defects where T2 statistics – an applied statistics field is used to detect defective parts of citrus peel. This unsupervised method is based on Principal Component Analysis (PCA) that uses color and texture features together extracting Eigenspace from unfolding color and spatial data. The correct detection of individual fruit defects is 91.5% and damaged/sound samples classification ratio of 94.2% is achieved from this experiment.

In paper [7] authors used K-Means clustering and image is converted from RGB color space to L*a*b* color space. Color features are Color coherence vector (CCV) and Global color histogram (GCH). Complete local binary pattern (CLBP), Gabor Features, Local binary pattern (LBP) and Local ternary pattern (LTP) texture features are used. Then authors applied Random forest classifier on segmented image to classify the diseases of apples. Fruit disease is shown after classification by tagging every pixels by using this method. Finally feature level fusion by fusing more than two texture and color features are performed to improve accuracy of the proposed system.

In paper [8] authors described image processing technique for Apple fruit disease detection. Apple Bloch, Apple Rot and Apple Scab are used for fruit disease identification and the features used are global Color Histogram, Local Binary Pattern, Color Coherence Vector and Complete Local Binary Pattern. K-means Clustering in L*a*b* space is used for defect segmentation. Multi-class SVM Classifier was used to train and test images. In paper [9] authors used K-Means clustering and image is converted from RGB color space to L*a*b* color space. Color features are Color coherence vector (CCV) and Global color histogram (GCH). Complete local binary pattern (CLBP), Gabor Features, Local binary pattern (LBP) and Local ternary pattern (LTP) texture features are used. Then authors performed feature level fusion by fusing more than two texture and color features also applied Random forest classifier on segmented image. Fruit disease is shown after classification by tagging every pixels by using this method.

In paper [10] four different algorithms were tested to identify the diseases with various combination of Golden Delicious apples images using color information. The Mahalanobis distance between each pixel and both medians is computed and compared with the global model which is used next to enhance and filter images to detect various defects such as bruises, russet, scab, fungi or wounds of apples. A hybrid method is used for color segmentation which is based on seeded region growing algorithm [11]. Color segmentation is done by split and merge algorithm to partition it into some regions that was used for knowledge-based exploration of scene [12]. In paper [13] authors proposed a web based tool that has acquired an accuracy of 82% to identify pomegranate diseases. The system contained trained dataset of pomegranate fruits. In first step, feature extracting with color, morphology and CCV is done then clustered by K-means. Next, SVM classifier is used to classify the disease.

The author Monica Jhuria [14] provided an approach detecting disease on fruit. Grapes, Apple and mangoes are selected as the domain to experiment the algorithms. Morphology, color and texture feature vectors are chosen for feature extraction. Some image processing techniques were used to detect fruit defect detection on the experiment. . Back propagation is used for weight adjustment of images that are stored in learning database. In that case, morphology feature gives 90% accurate results which was the highest accuracy gained than other feature vector. In [15] a framework for the defect segmentation of fruits using images is proposed and evaluated using K-means clustering technique. Apple is used in the experiment to detect defected area of the fruit part along with the stem and calyx using k-means. In that system three or four clusters are considered as observation to detect the defects in fruits.

Many color models are used to represent the color like RGB, CMY, HSV, HSI while in paper [16] an approach is proposed which is UTSO method which is used to segment the Grayscale image into two or three classes based on color. The original image is pre-processed and then converted to L*a*b* from RGB. After that Otsu method of segmentation is applied based on the color of the particular object. Finally the segmented image is produced. This method is very effectual in segmenting images.

To detect and identify defects on fresh peach from markets a machine vision system was developed in the laboratory [17]. Scar, cuts, bruise, scale, wormhole, and brown rot were identified by segmenting the images of defected region with a hybrid classifier which described the region as a specific type. The experimental result acquired a performance of the classifier having 31% error rate for the near-infrared system and 40% for the color system.

Machine vision system is now widely used in industrial applications for fruit processing and grading. Using an hybrid algorithm in [18] a system is proposed to classify the fruit quality using k-means segmentation as the first step of the system, then calculating Euclidean distance for L*a*b* color space to produce an over segmentation result. Merge procedure of the algorithm was based on a region adjacent graph representation.

2.3 Research Summary

In this research so many images of carrots are collected from various places. Then the images are pre-processed for further processing. Features of the images are evaluated and extracted from the images and clustered by k-means clustering. Dataset is prepared and trained along with specific label for each class of the disease which is done in MATLAB through the calculation of features. When all the processing is done our system is ready to use. Test image is captured and given as input after preprocessing. The features are extracted and the selected cluster image is compared with the previous dataset. Multiclass SVM classifier is used to identify the class name or the disease name. Finally the output of the system is shown as the predicted disease name in the screen.

2.4 Scope of the Problem

The proposed framework can help farmers as well as supermarkets owner to have benefits of modern technology to prevent improper loss of their profit. Although the whole process like capturing images, entering the image as input and the implementation of the application may be a fact to a non-professional user who do not have any knowledge of this type of system. But this is a proposed framework which includes the main idea of steps how to process the data and how the algorithms should be implemented. It can be implemented in any kind of platform regardless of choice. Using this approach mobile applications or online based web applications can be developed to reach the farmers easily.

2.5 Challenges

To build such a system it needs images of carrots to be trained first and then some more images are needed for the testing purpose. Images were collected from various lands of carrots.

The big challenge with this approach is the quality of the images. Images of low resolution cannot be processed easily. Image gets blurry due to few number of pixels containing the defected area. The amount of RGB value are most important in image processing which need to be clarify first. Also the background of the image makes it difficult sometimes to detect the affected part of the carrots. If the background colors are much similar to the fruit part then the segmentation cannot be done accurately.

CHAPTER 3

Research Methodology

3.1 Introduction

To get the disease detected fruit and vegetable part accurately, image should be segmented first. Otherwise, the non-affected part of the major area can be a dispersion of this method. Fig 3.1 shows the basic steps of proposed framework of this fruit and vegetable disease classification problem. In this paper, to determine the ROI (Region of Interest) K-means clustering is used. After that, feature extractions are done from the segmented image. Multiclass SVM classifier are used to train the dataset and finally classify the disease as output.

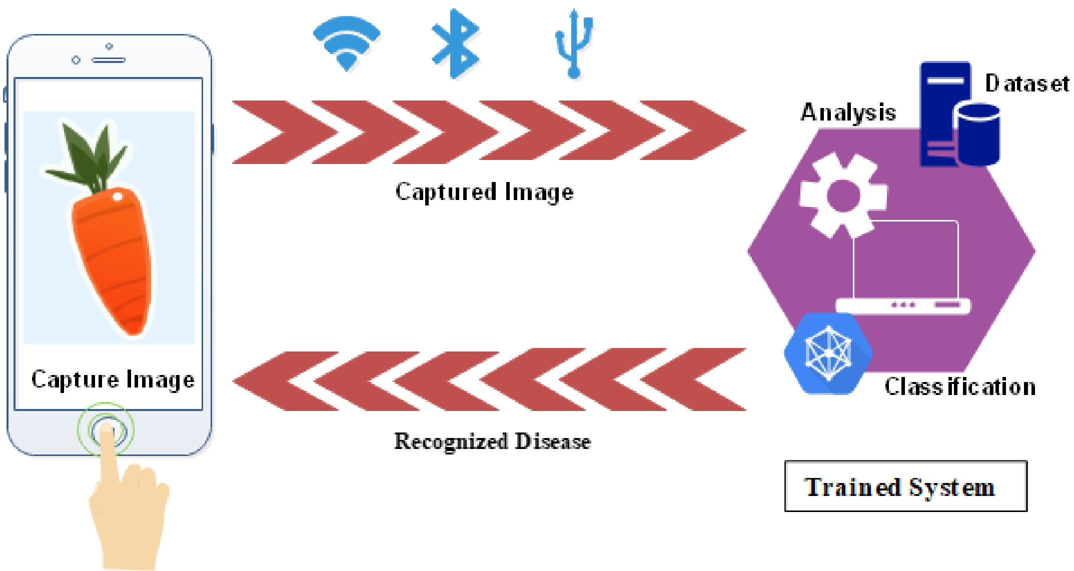

Figure 3.1: Working procedure for the proposed framework

Fig 3.1 shows how the proposed machine vision system works. To implement such a machine vision system a machine learning system is required which is described in this research and fig 3.2 shows the steps to build the framework for the system.

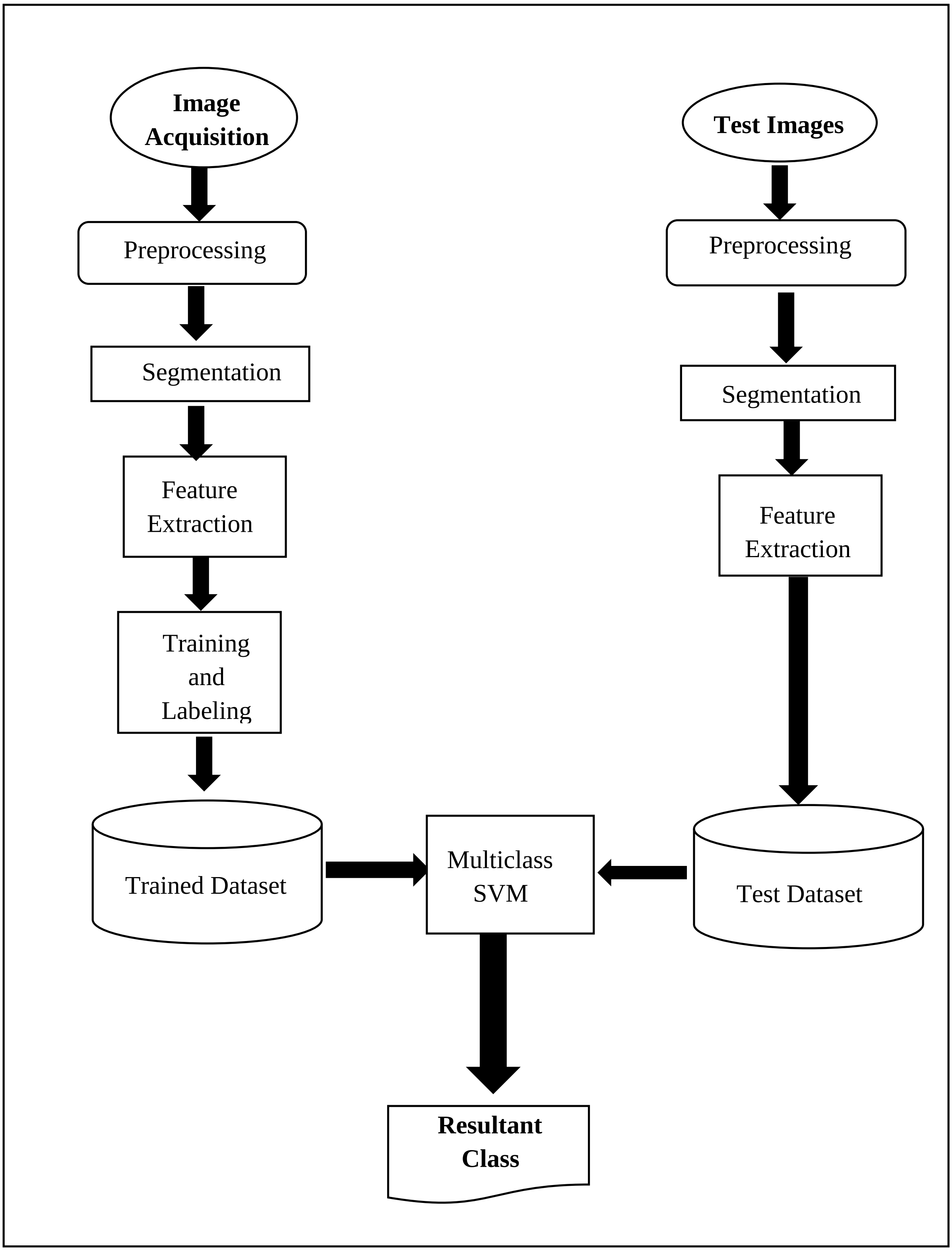

Figure 3.2: Proposed data flow diagram

3.2 Methods and steps

3.2.1 Image Acquisition

In image processing image acquisition means delivering an image from some source - usually a hardware-based source for processing. It is the foremost step of processing digital image. The raw image is received which are processed next to get desired information from it.

In this research work numerous carrot images are used which are collected and captured from various places. Used images are formed with RGB (Red, Green, and Blue) color model. RGB is in general a color model which describes images with their color values in the pixels consisting of different RGB values. The images has differences in quality and formation which needs to be processed to a common form in order to train the dataset.

Figure 3.3: Disease affected carrots (containing root fly)

3.2.2 Image Pre-processing

Before using the image for segmentation, some preprocessing of the images were done like Cropping- for cut-out the unimportant part of the image, Smoothing, enhancement to adjust the color and contrast, Rotating, Resizing to get all the images in a general form.

3.2.3 Image Segmentation

a) K-Means clustering

In this paper, K-means clustering is used to segment the images. Also boundary and spot detection algorithms are used. In boundary detection, the 8-connected pixel algorithm is used. Here, for K-means clustering Euclidean distance are used for three clusters of the ROI which are to be selected.

The algorithm for K-means clustering-

- Perform conversion from RGB to L*a*b color model.
- Identify each pixel in the image to the cluster that minimizes the distance from the center value of cluster.
- Perform color segmentation of the image.
- Finally, select the cluster containing ROIs only.

Fig 3.4 shows three cluster being produced by K-means. From those clusters one is chosen containing ROI for further processing.

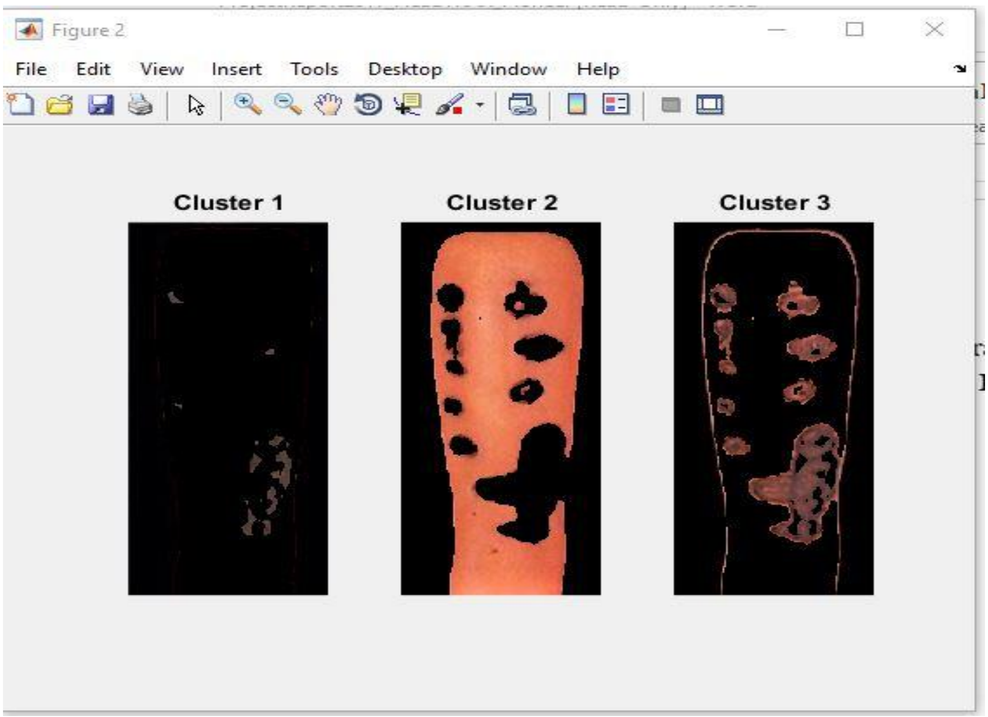

Figure 3.4: Image clustering by k-means

b) L*A*B Color Space

Color space defined by CIE, based on one channel for Luminance (lightness) L and two color channels (a and b).

In XYZ color system, Colorimetric distance between the individual colors do not correspond to perceived color differences. The CIE solved this problem in 1976 with the development of the three -dimensional L*A*b color space (or CIELAB color space) [19].

This color model is better suited to many digital image manipulation with numerous different devices than the RGB space.

AS L*a*b model is a three dimensional model, so it can only be represented accurately in three-dimensional space. The formula for RGB converting digital images from RGB space to L*a*b are-

$$\begin{bmatrix} L^* = 116\, f\left(\frac{Y}{Yn}\right) - 16 \\ a^* = 500\left[f\left(\frac{X}{Xn}\right) - f\left(\frac{Y}{Yn}\right)\right] \\ b^* = 200\left[f\left(\frac{Y}{Yn}\right) - f\left(\frac{Z}{Zn}\right)\right] \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

Where X, Xn, Y, Yn, Z, Zn are the coordinates of CIEXYZ color space.

And where,

$$f(x) = \begin{cases} x^{\frac{1}{3}} & ; \text{if } x > 0.620 \\ 7.787x + \left(\frac{16}{116}\right) & ; \text{Others} \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

The L*a*b color space includes all apprehensible colors because of its gamut exceeds those of the RGB and CMYK color models. It provides opportunity to communicate different color across the different device. That’s why it is used as an interchange format between different devices. Classically L*a*b color space is used in image editing program [19].

3.2.4 Features extraction

This is the most important part of the whole task. To identify an object, features extraction is used in various image processing applications Bhabani J. Samapti and Shesshang D. Degadwala [22] described color feature and texture feature extraction methods to extract color, shape and texture of fruits for disease classification where color features were Color Coherent Vector(CCV) and Global Color Histogram(GCH) and texture features were Complete Local Binary Pattern(CLBP), Gabor features, Local Binary Pattern(LBP) and Local Ternary Pattern(LTP). Numerous methods can be used for detection of disease part of the fruit or vegetables.

In our approach, both color and texture are taken by converting RGB to this model.

$$H = \begin{cases} \text{Theta} & \text{if } B<R \\ 360\text{-Theta} & \text{if } B>R \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R, G, B)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

$$I = \frac{1}{3}(R+G+B) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(iii)}$$

Spatial gray level Dependence Matrices (SGDM) method is a way of extracting statistical features. And GLCM functions characterized the texture of an image by calculating how often pairs of pixel with specific values and in a spatial relationship occur in an image [20]. GLCM is basically Gray Level Co-Occurrence Matrix which gives the distribution of co-occurring values in a specific area of interest. GLCM is created from a gray scale image. So before getting GLCM features first the image need to be converted to gray scale image. Then the statistical features like contrast, correlation, energy and homogeneity can be calculated by the formulas.

The statistical features from GLCM are-

- **Contrast-** measures the local variation in the gray-level co-occurrence matrix. **C**ontrast is 0 for a constant image. The formula for calculating contrast of an image is-

$$\sum_{i,j} |i-j|^2 p(i,j) \qquad \ldots\ldots\ldots\ldots\text{ (i)}$$

- **Correlation-** measure the joint probability occurrence of the special pixel pairs. The formula for calculating correlation in an image is-

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu i)p(i,j)}{\sigma_i \sigma_j} \qquad \ldots\ldots\ldots\ldots\text{ (ii)}$$

- **Energy-** provides the sum of squared elements in the GLCM. Also known as uniformly or the angular second moment. And the formula is-

$$\sum_{i,j} p(i,j)^2 \qquad \ldots\ldots\ldots\ldots\text{ (iii)}$$

- **Homogeneity-** measure the closeness of the distribution of elements in the GLCM to the GLCM diagonal. The formula for homogeneity calculation is given below-

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad \ldots\ldots\ldots\ldots\text{(iv)}$$

Another seven more spectral features-

- Mean
- Standard Deviation
- Entropy
- RMS
- Variance
- Kurtosis
- Skewness -

These were calculated in matrix in total eleven features to detect the disease defected object part from the images.

3.3 Data Collection Procedure

Carrots are produced in well-drained, fine featured soils with good water holding capacity. Gritty, loamy or ordure-based soils are preferable. For the research purpose different sort of images are collected from many farms with the help of the root level farmers. The raw images are different in size, shape and category with variation of color combination and textures. All the images are preprocessed to get them all in a suitable form of data.

3.4 Statistical Analysis

As per the analysis of data, calculation of the accuracy of the proposed system is achieved 96%. To calculate that, some more attributes are extracted with the help of confusion matrix. For example the confusion matrix of violet root knot is given below.

Table 3.1: Confusion Matrix for Violet root Knot

	Positive	**Negative**
True	8	93
False	0	2

To measure the performance of the system we need to define-

True Positive (TP) = the number of cases correctly identified actual class.

False Positive (FP) = the number of cases incorrectly identified actual class.

True negative (TN) = the number of cases correctly identified negative classes.

False negative (FN) = the number of cases incorrectly identified negative classes.

Table 3.2: Attributes collected from confusion matrices of all classes

	Healthy Carrots	**Black Rot**	**Growth Crack**	**Root Fly**	**Root Knot**	**Violet Root Knot**
TP	16	19	11	15	19	8
TN	85	82	90	86	82	93
FP	5	2	2	0	3	0
FN	0	2	4	3	2	2

From the confusion matrices for each classes sensitivity, specificity, precision, FPR and FNR are also calculated. They are-

Accuracy = 96%

Sensitivity = 87.13%

Specificity = 97.74%

Precision = 88%

FPR = 2.27%

FNR = 12.87%

CHAPTER 4

Experimental Results and Discussion

4.1 Introduction

To get the final result, first the raw images were collected and captured from numerous farms and shops. After pre-processing of data, the main trained dataset were prepared. Finally the test image is compared with the dataset using Multi-class Support Vector Machine the final outcome is shown as the detected class.

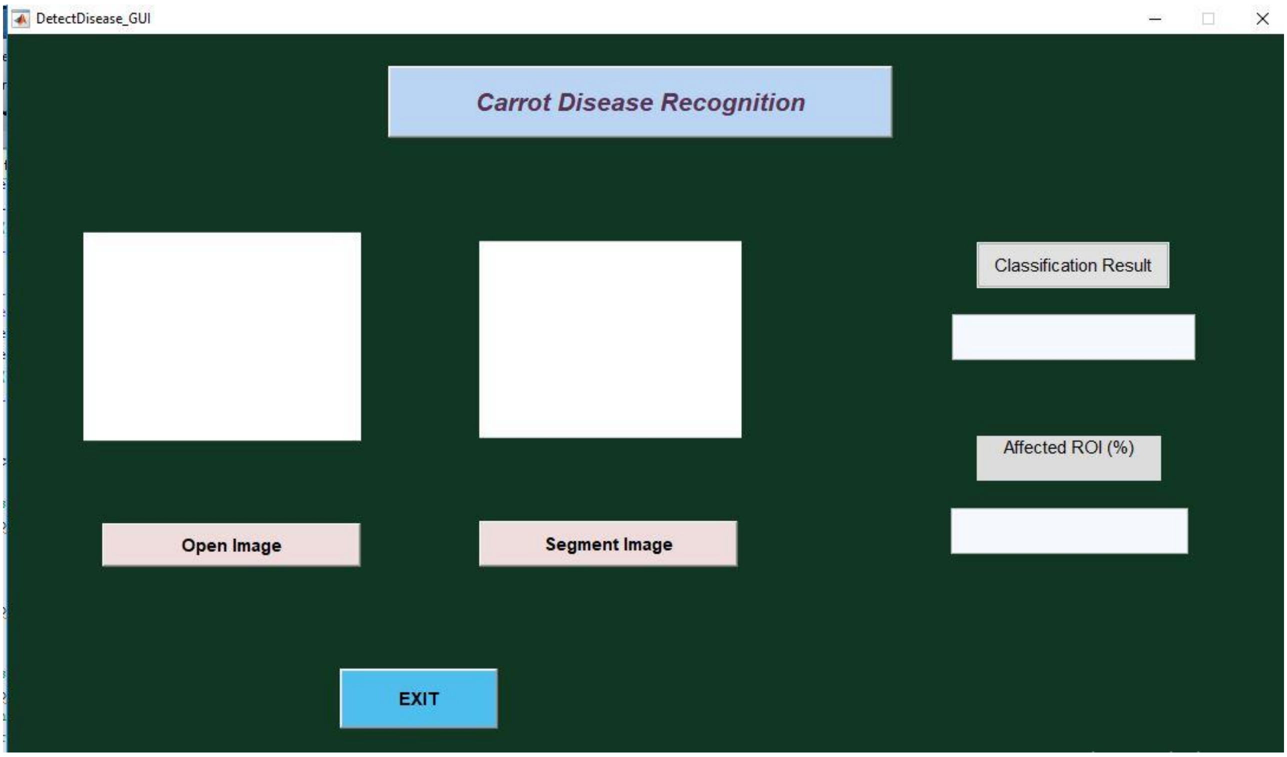

Figure 4.1: GUI of the system

4.2 Experimental Result

In this paper to demonstrate the detection of fruits and vegetables diseases the used domain is carrot. Five types of diseases of carrot were identified with images collected for six categories including detection of healthy carrots. The diseases are Black Rot (42 images), Growth Crack (28 images), Root Fly (36 images), Root Knot (42 images), Violet Root Knot (20 images) and Healthy carrots (32 images) in total 202 images are used.

Figure 4.2: Carrots images (Root fly)

After training the dataset, test image is captured and given to the system as input. Image is then segmented with K-means clustering. Predicted class is shown in the output [fig 4.2] after analyzing the data of that cluster entered containing the Region of Interest (ROI) only.

Next, some screenshots are provided from the experiments of detecting diseases for Black Rot, Growth Crack, Root fly, Root Knot and Violet root Knot respectively with the clustering process and the calculated affected part or the region of interest. The whole process are implemented in MATLAB, the experimental result is much favorable.

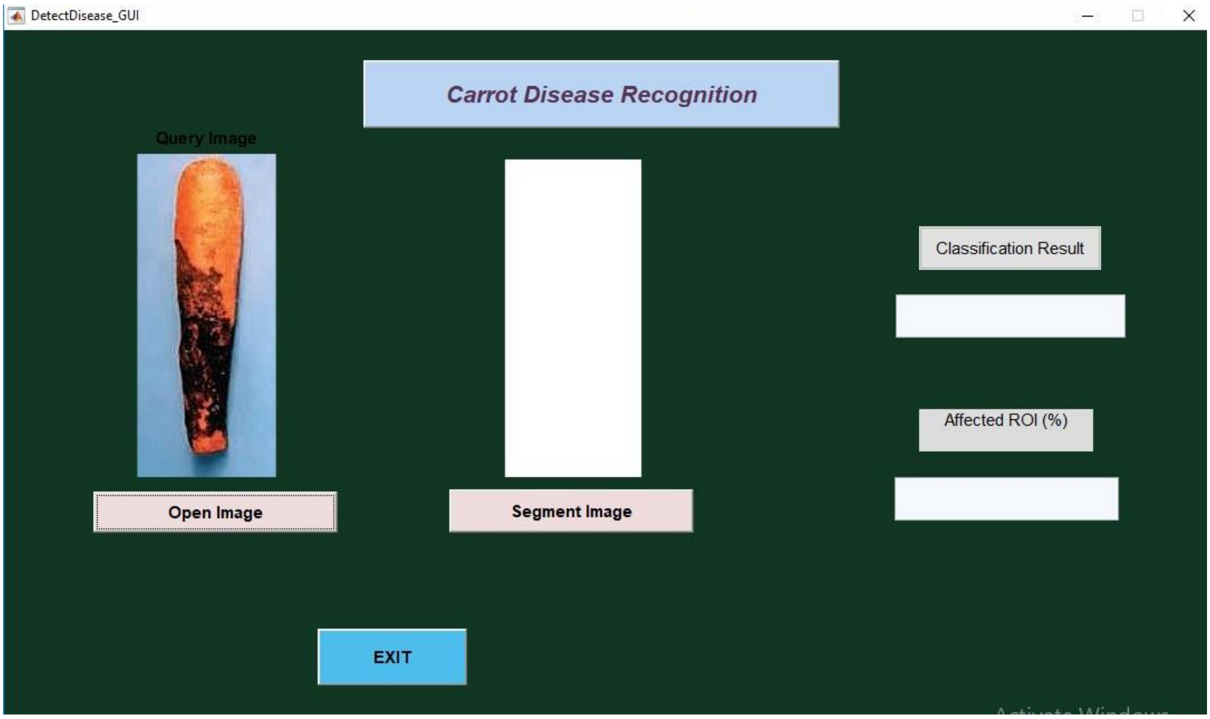

Figure 4.5: Taking input for Black Rot

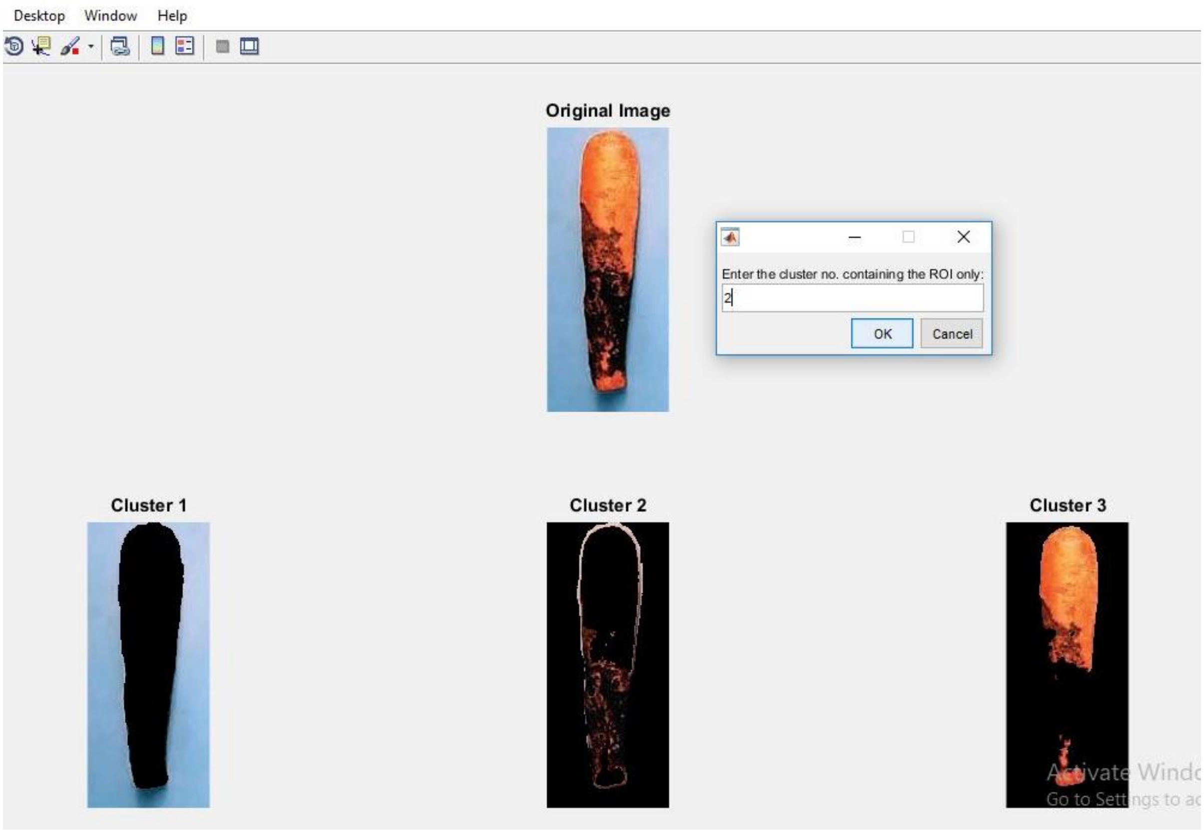

Figure 4.6: Taking suitable cluster from user (Black Rot)

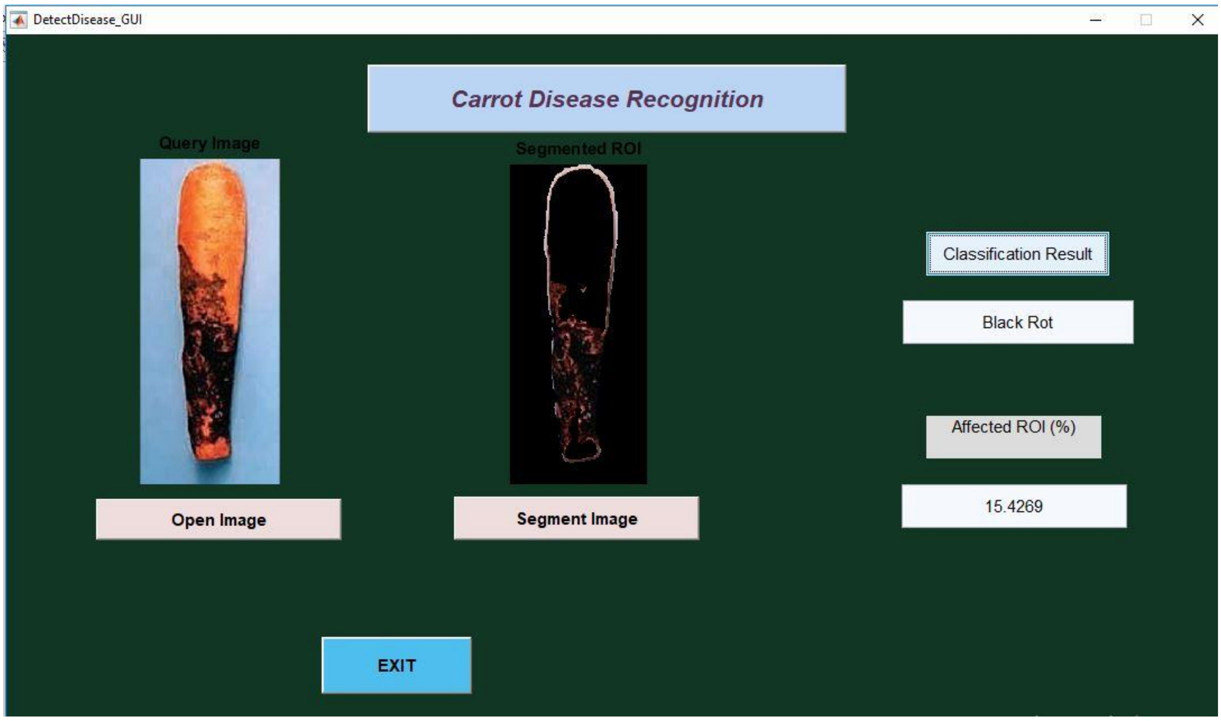

Figure 4.7: Classification result shown in the GUI (Black Rot)

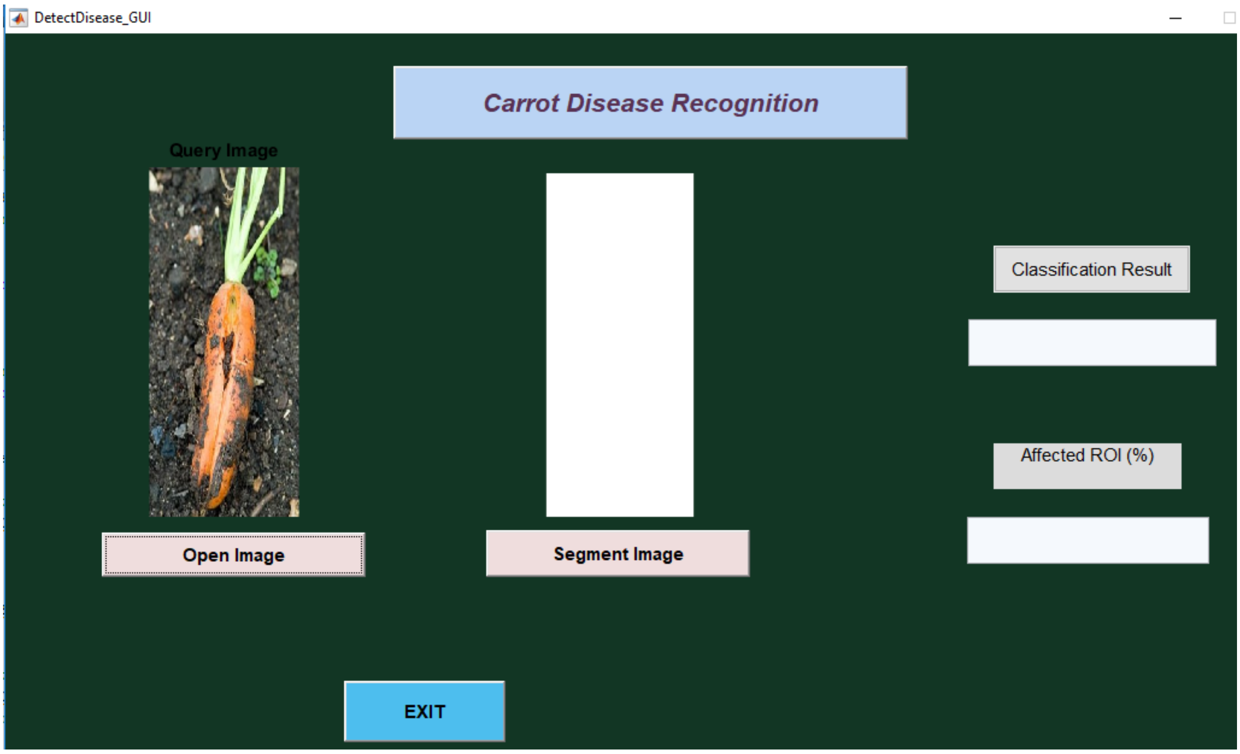

Figure 4.8: Taking input for Growth Crack

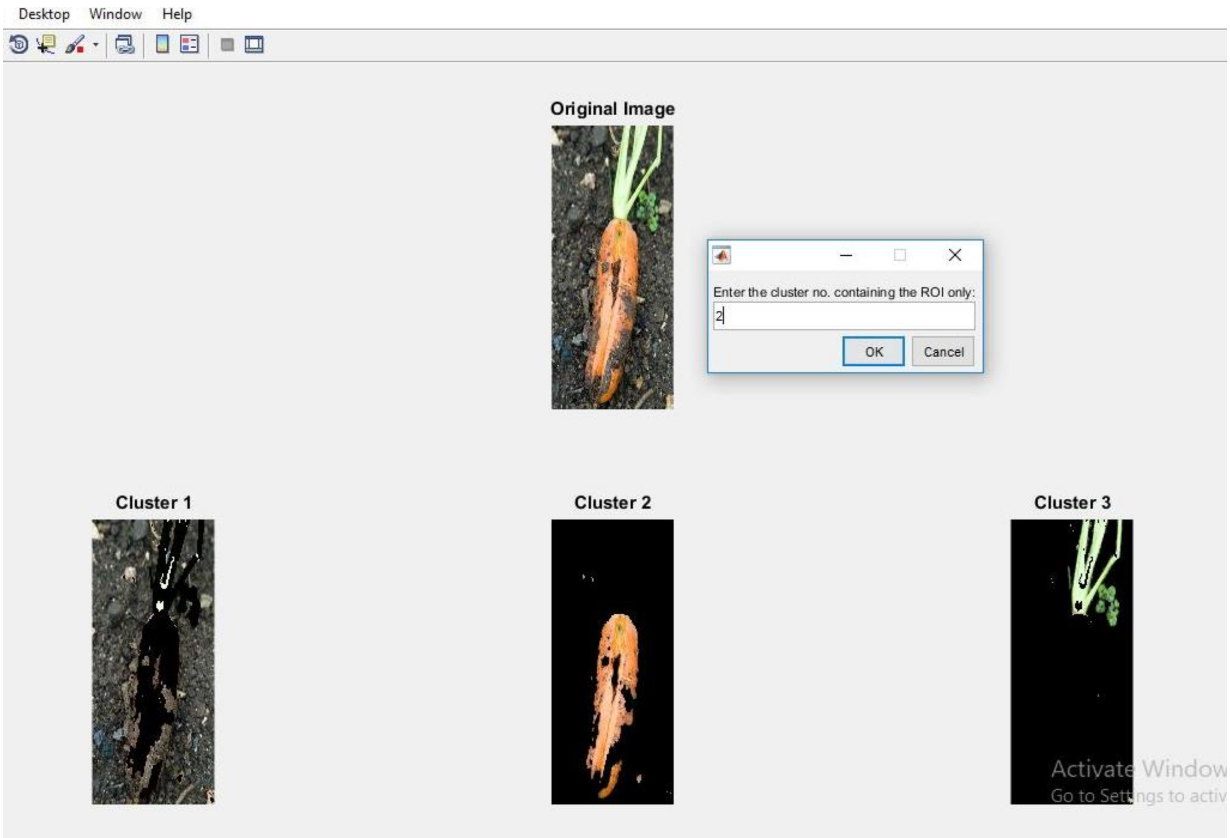

Figure 4.9: Taking suitable cluster from user (Growth Crack)

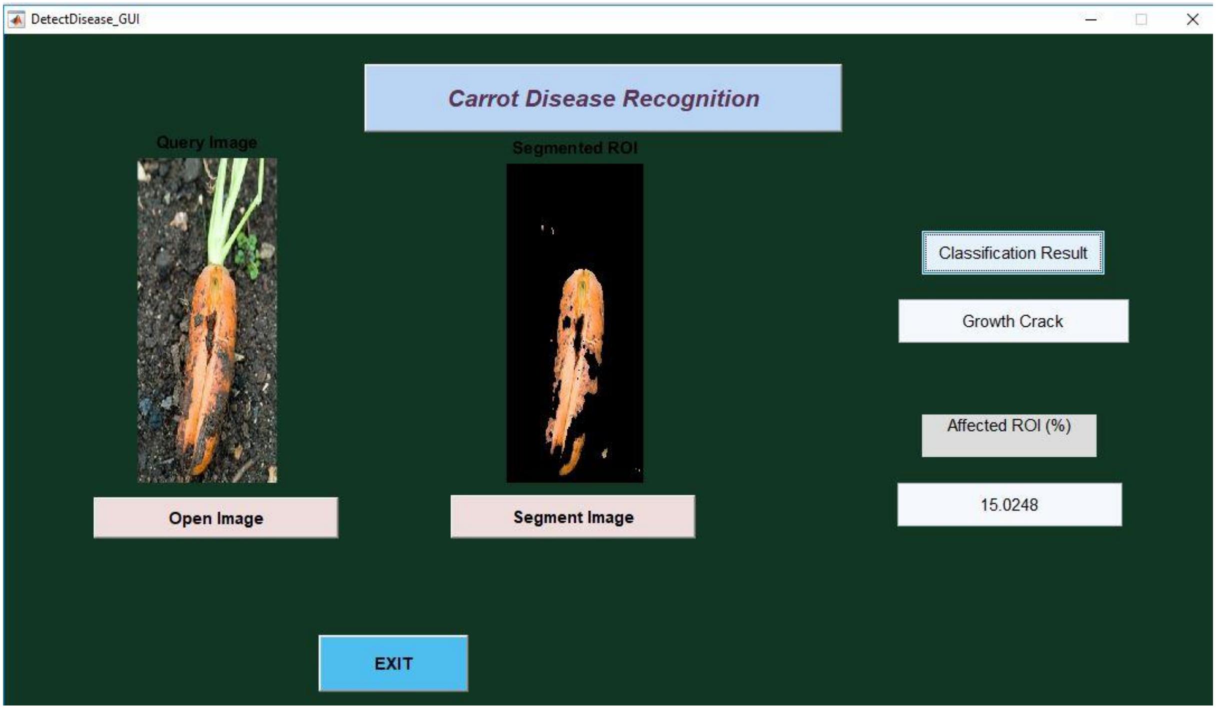

Figure 4.10: Classification result shown in the GUI (Growth Crack)

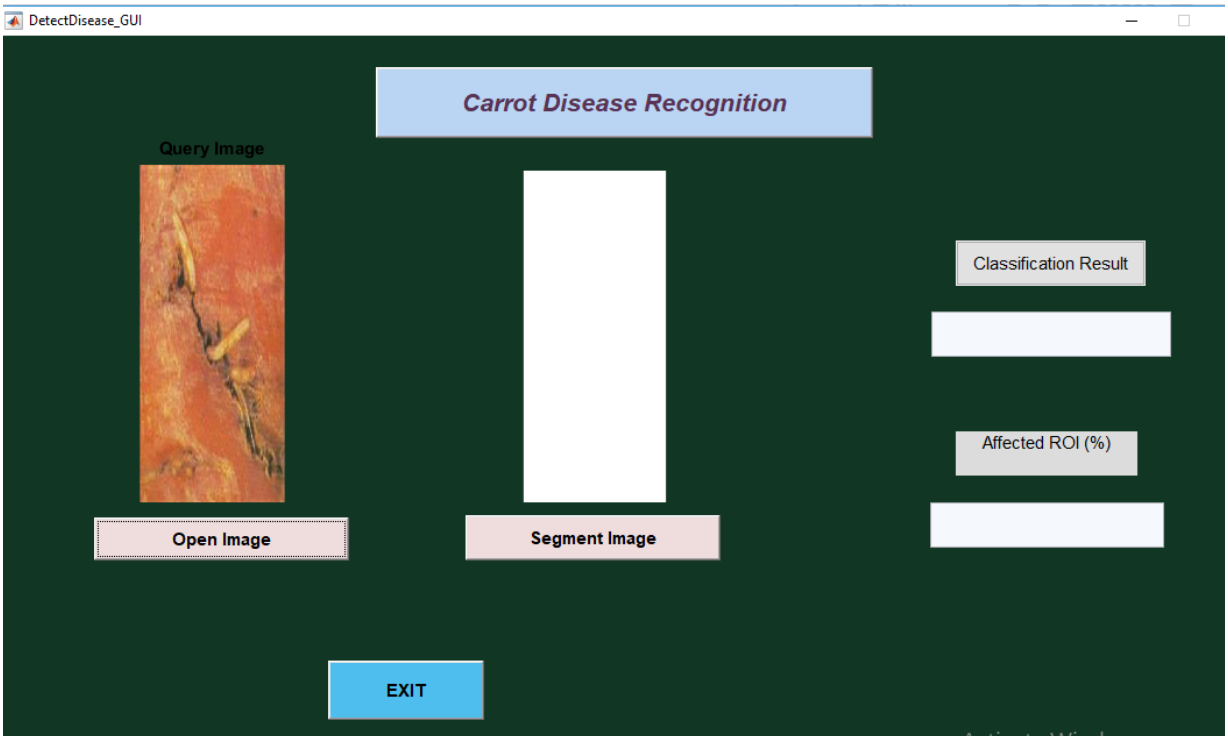

Figure 4.11: Taking input for Root Fly

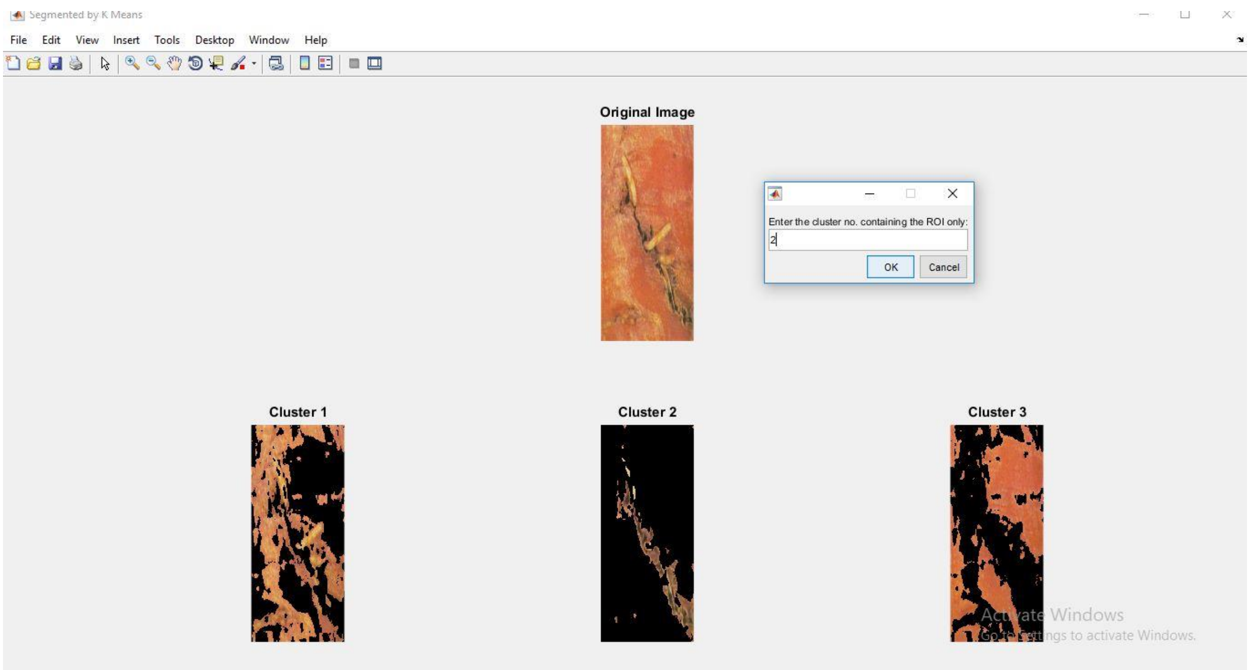

Figure 4.12: Taking suitable cluster from user (Root Fly)

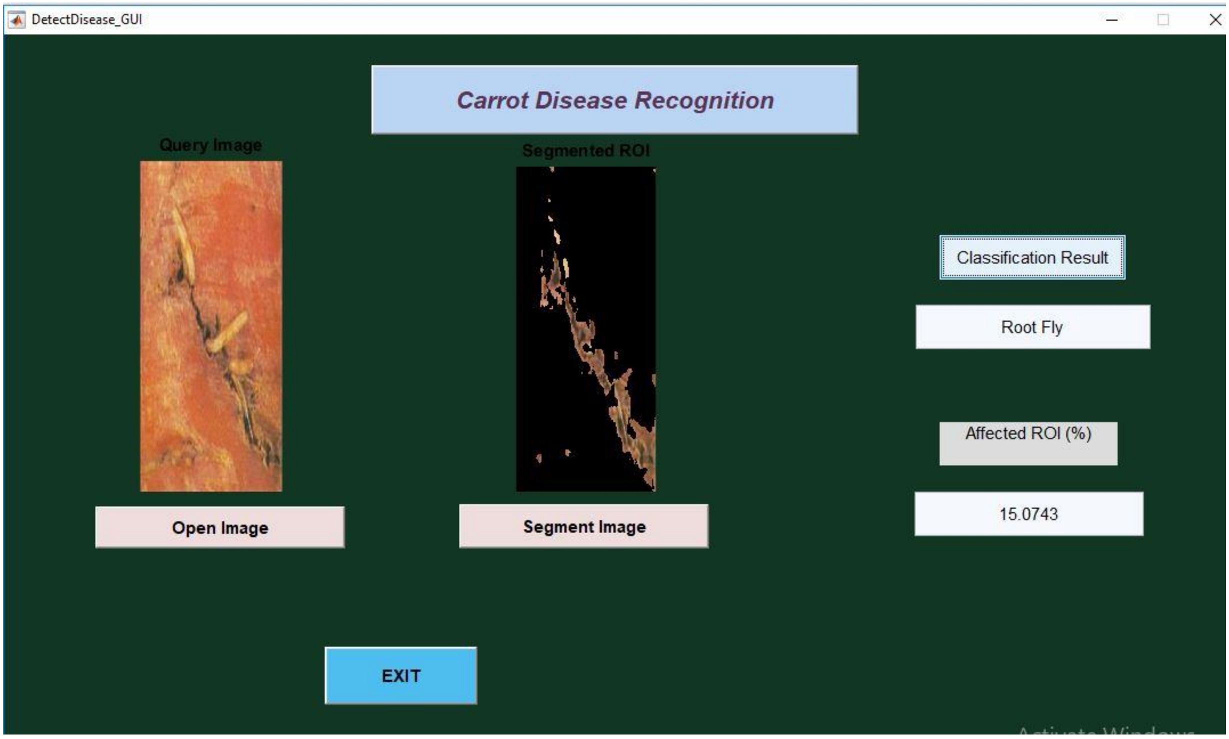

Figure 4.13: Classification result shown in the GUI (Root Fly)

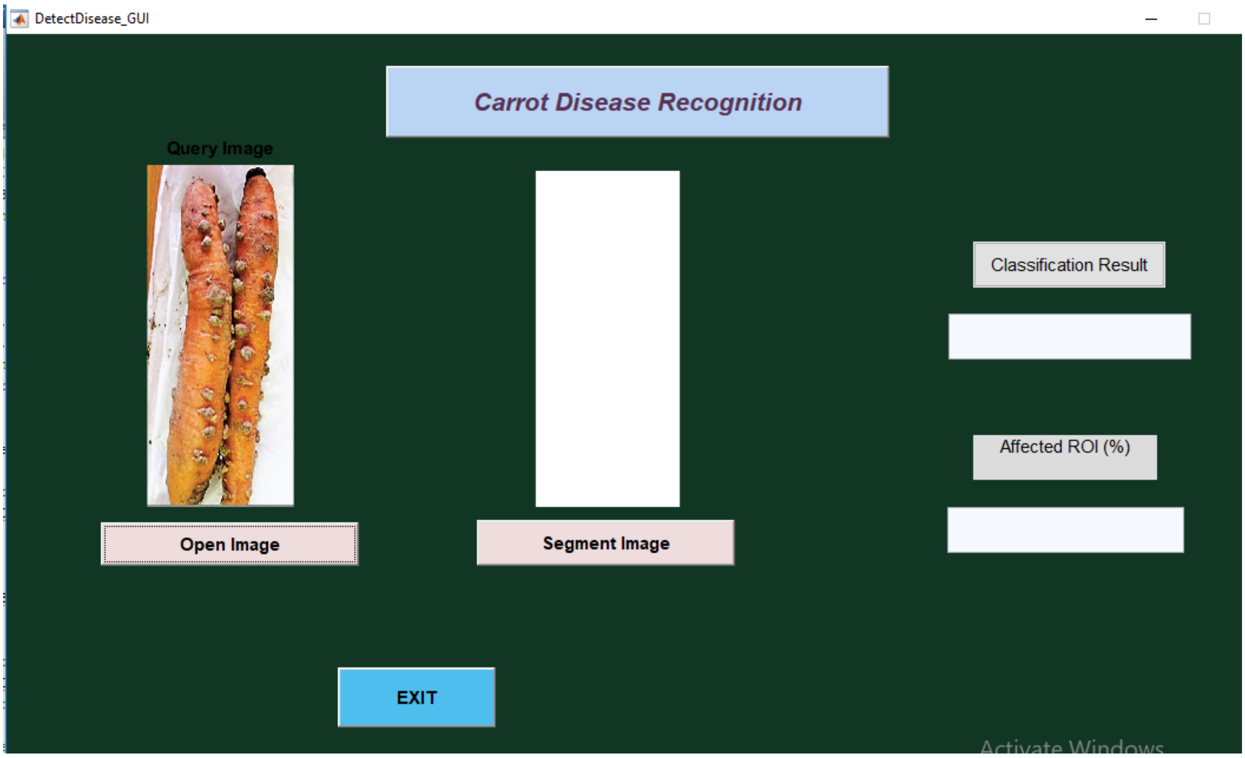

Figure 4.14: Taking input for Root Knot

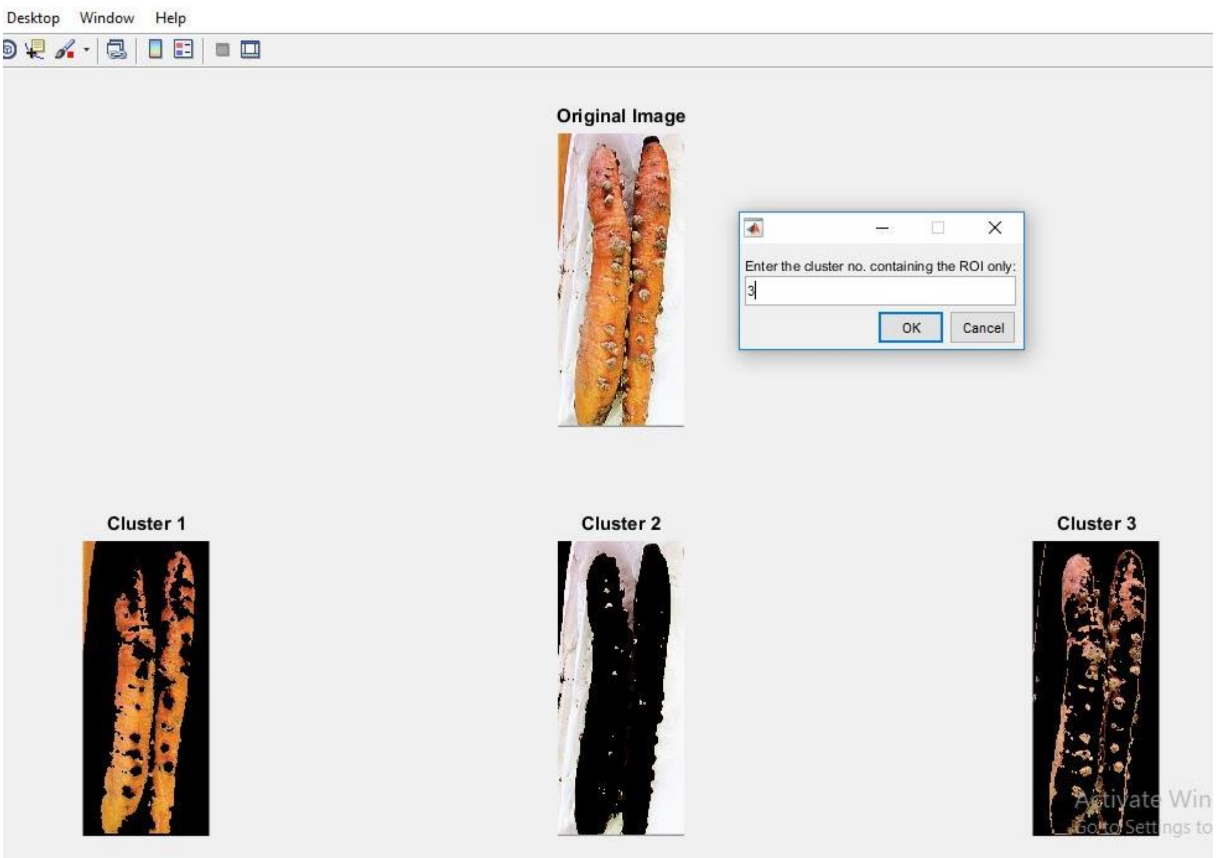

Figure 4.15: Taking suitable cluster from user (Root Knot)

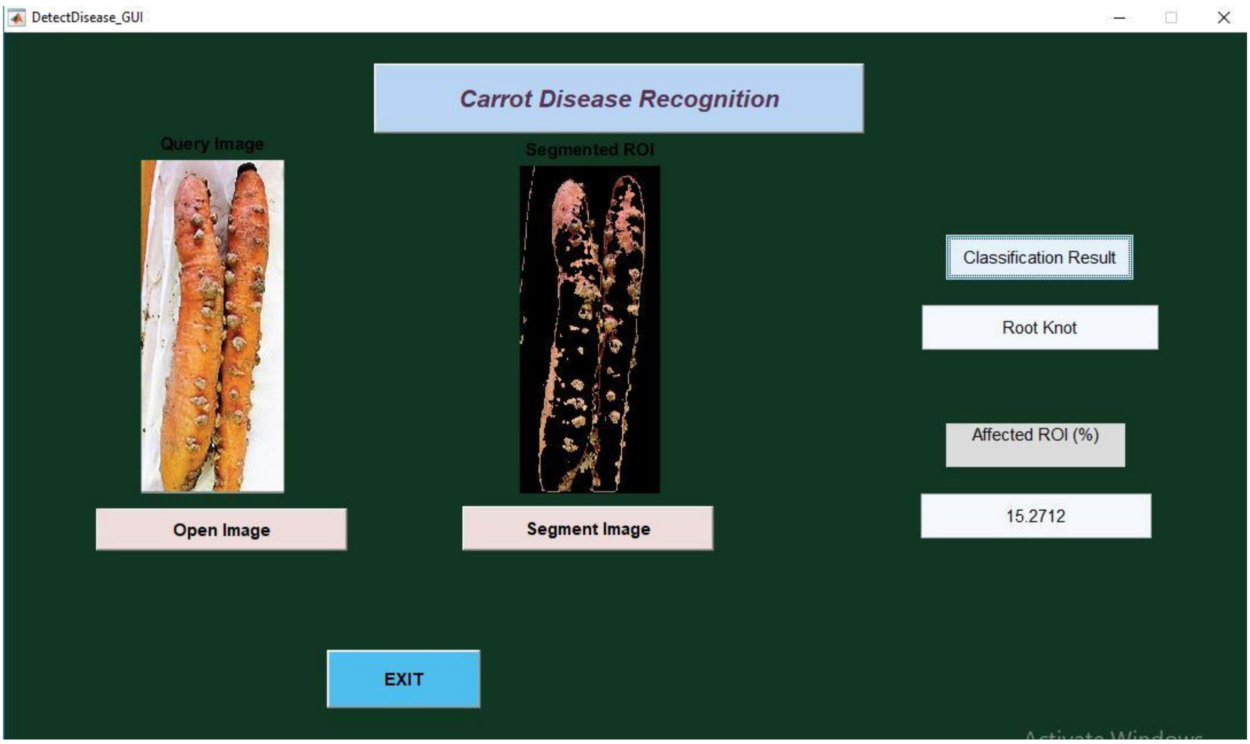

Figure 4.16: Classification result shown in the GUI (Root Knot)

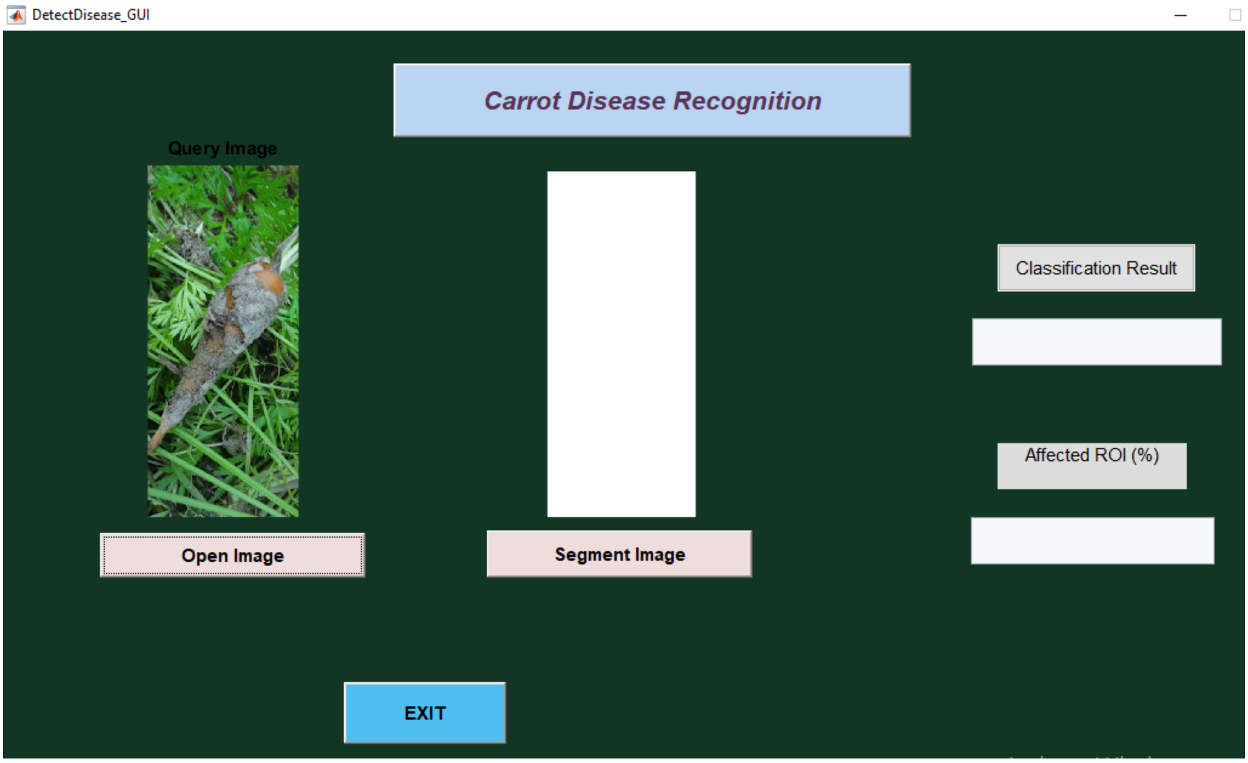

Figure 4.17: Taking input for Violet Root Knot

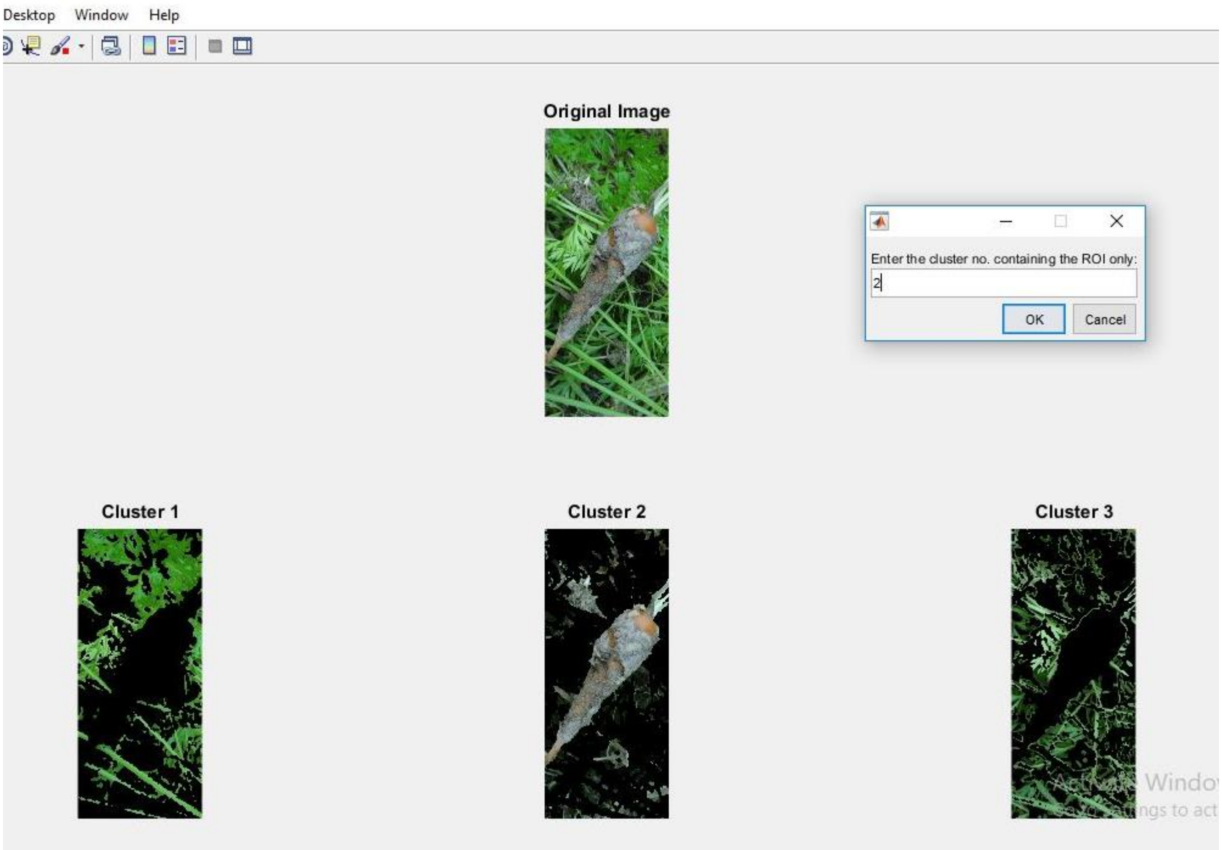

Figure 4.18: Taking suitable cluster from user (Violet Root Knot)

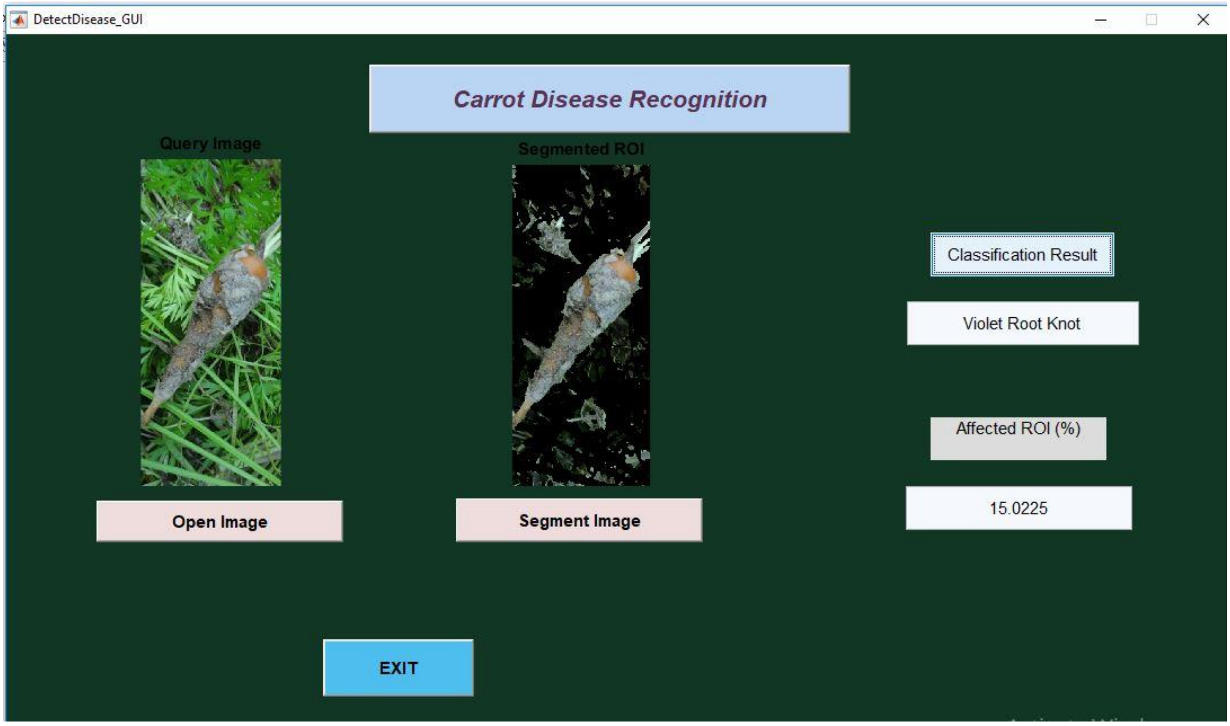

Figure 4.19: Classification result shown in the GUI (Violet Root Knot)

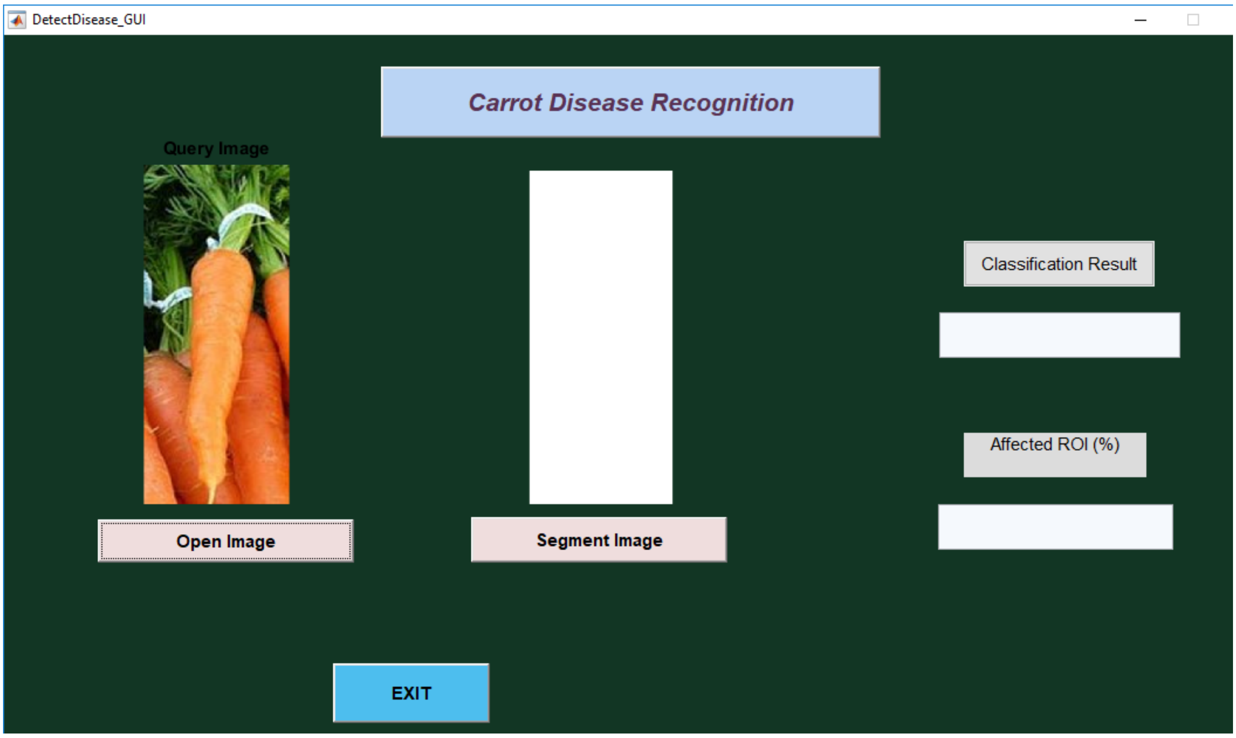

Figure 4.20: Taking input for Healthy Carrots

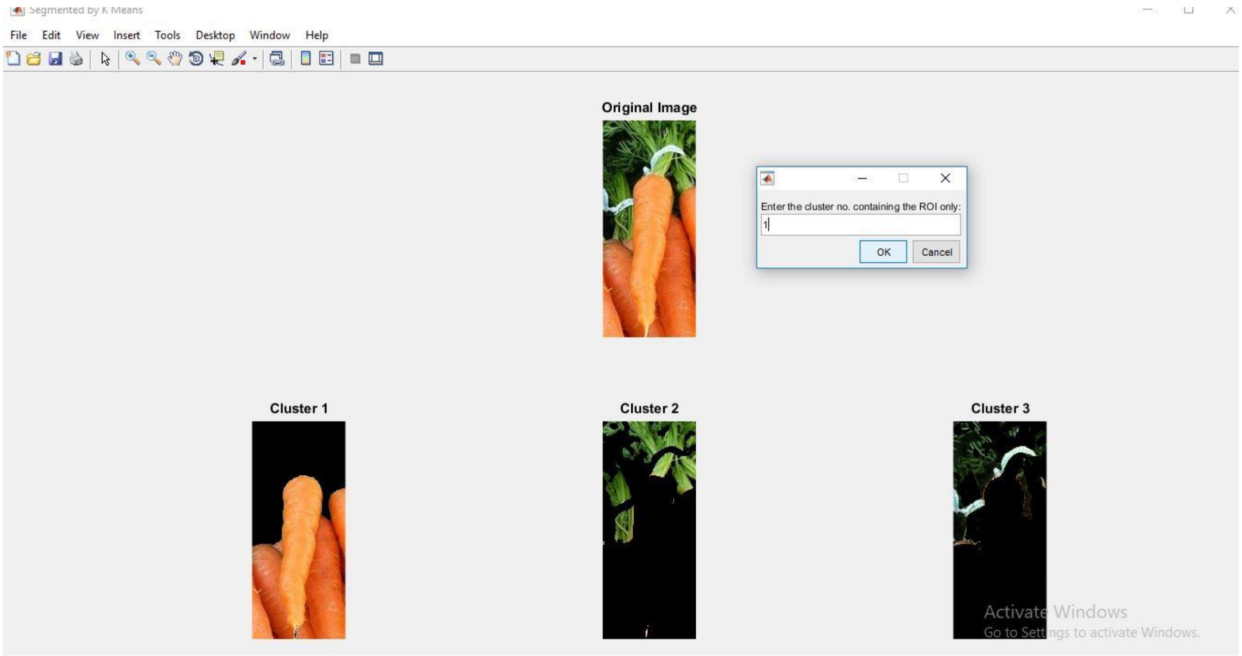

Figure 4.21: Taking suitable cluster from user (Healthy Carrots)

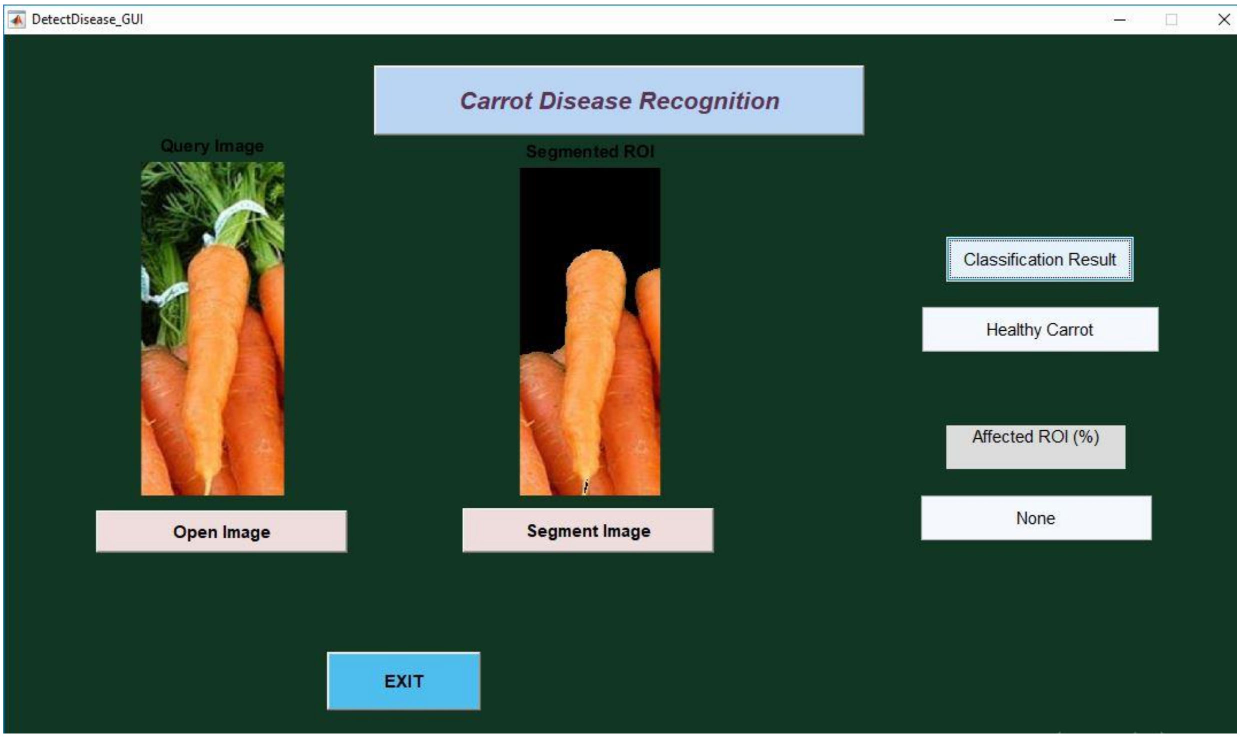

Figure 4.22: Classification result shown in the GUI (Healthy Carrots)

The accuracy of the proposed system is 96% calculated with all the other features- Sensitivity of 87.13%, Specificity of 97.74%, Precision of 88%, FPR of 2.27% and FNR of 12.87%.

4.3 Performance analysis

The extracted features of the ROI containing cluster is labeled and all the values of eleven features are stored in a $m \times n$ matrix where m is the number of images trained and n is the number of feature that is extracted from the image.

Test image features are extracted to compare with the trained dataset and Multiclass SVM algorithm is applied in the system to classify which class the test image's information belongs to.

Finally the predicted disease name is printed in the screen.

To measure the performance of the system we need to define-

True Positive (TP) = the number of cases correctly identified actual class.

False Positive (FP) = the number of cases incorrectly identified actual class.

True negative (TN) = the number of cases correctly identified negative classes.

False negative (FN) = the number of cases incorrectly identified negative classes.

The four cases of the classification result can be represented by the following 2 by 2 confusion matrix (contingency table):

Table 4.1: Confusion matrix formulation

	Positive	**Negative**
True	TP	TN
False	FP	FN

i. **Accuracy:** The accuracy of this system is its ability to differentiate the actual class and negative classes correctly. To estimate the accuracy mathematically, this can be stated as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad \dots \quad \dots \quad \dots \text{(i)}$$

ii. **Sensitivity:** The sensitivity of this system is its ability to determine the actual classes correctly. . To estimate the sensitivity mathematically, this can be stated as:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100\% \quad \dots \quad \dots \quad \dots \text{(ii)}$$

iii. **Specificity:** The specificity of this system is its ability to determine the negative cases correctly. To estimate the specificity mathematically, this can be stated as:

$$\text{Specificity} = \frac{TN}{FP+TN} \times 100\% \quad \dots \quad \dots \quad \text{(iii)}$$

iv. **Precision:** Precision is the number of relevant classes among all the positive cases. To estimate the precision mathematically, this can be stated as:

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad \dots \quad \dots \quad \dots \quad \text{(iv)}$$

v. **FPR:** False Positive Rate of the system is the number of predicted class with which the classifier makes a mistake by classifying normal state as pathological. False Positive Rate can be stated as:

$$\text{FPR} = \frac{FP}{FP+TN} \times 100\% \quad \dots \quad \dots \quad \dots \quad \text{(v)}$$

vi. **FNR:** reflects the frequency with which the classifier makes a mistake by classifying pathological state as normal. False Negative Rate can be stated as:

$$\text{FNR} = \frac{FN}{FN+TP} \times 100\% \quad \dots \quad \dots \quad \dots \quad \text{(vi)}$$

Confusion matrix for the system is shown in table 4.2

Table 4.2: Confusion Matrix for the proposed framework

Predicted Class

Actual Class

	Healthy Carrots	**Black Rot**	**Growth Crack**	**Root Fly**	**Root Knot**	**Violet Root Knot**
Healthy Carrots	32	0	0	0	0	0
Black Rot	0	38	2	0	2	0
Growth Crack	4	0	22	0	4	0
Root Fly	0	2	2	30	2	0
Root Knot	2	2	0	0	38	0
Violet Root Knot	4	0	0	0	0	16

From the above contingency table accuracy of the proposed system is 96% calculated with all the other features- Sensitivity of 87.13%, Specificity of 97.74%, Precision of 88%, FPR of 2.27% and FNR of 12.87%.

4.4 Summary

Every year our country imports many fruits and vegetables from different countries especially from India which requires a huge investment from the government. But this land is very friendly to grow crops and vegetables in a rapid productive way that it should not be necessary to import those from outside. The proper maintenance and guidance can help the farmers to make their land fruitful if major problems of their cultivation can be solved. This image processing technique can solve a leading problem of the farmers by identifying diseases in fruits and vegetables. Applying this proposed method, diseases of carrots are recognized well and also can be implemented to define other kind of fruit or vegetable diseases also that will surely bring a tremendous change in our agriculture sector.

CHAPTER 5

Conclusions and Implication for the Future Research

5.1 Conclusions

In this work carrot diseases are recognized through image processing techniques which can be used in different kind of applications to detect any fruit or vegetable diseases which can open a door for helping big markets, super shops, exporter or agricultural farms to grow a highly secured economic development. Machine vision based carrot disease recognition is done with a supervised learning process. The whole process is done with a 96% of accuracy using collected carrot images which are affected with diseases. Though there were some barriers while working, the background color and poor quality of images can distract the application to give more accurate result. Using this techniques various fruit and vegetable disease can be detected and classified with a computer vision system. Some precautions can be made by notifying the farmers about the disease which would a relief to them as if they will not be facing any unexpected disease attack into their land or they would be able to detect the problems so they can make proper decision for further harvesting on the land.

5.2 Implication of further study

To make our life easier, gradually we are getting very much dependent on modern technologies where in our country, agriculture sector is far behind from using these technologies which can be a matter of prosperity in a remarkable rate. The proposed system shows a new way to involve with the machine learning method which is able to detect disease affected part of carrots and also recognize the disease automatically. This approach can be implemented into any kind of mobile based application or web based application to reach the root level farmers easily. Farmers sends the captured image from the land directly and get the output on their hand in a second showing which disease their fruits or vegetables are having so that they can take instant decision of their next step. To prevent immense loss of their profit they were in need of a system like this from the very early period of time. In this research we made a fair achievement from our study and effort which can be applied to our agricultural technologies to get a great performance of the production system and hereby producing a good amount of profit.

REFERENCES

[1] Linda G. Shapiro and George C. Stockman (2001): “Computer Vision”, pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3

[2] L.J.Rozario, T.Rahman, and Md.S.Uddin, "Segmentation of the Region of Defects in Fruits and Vegetables," International Journal of Computer Science and Information Security, vol.14, no.5,pp.398-406, May 2016.

[3] Howarth, M.S.; Searcy, S.W., 1989: “Algorithms for grading carrots by machine vision”. Paper American Society of Agricultural Engineers (89-7502): 17 pp

[4] Howarth, M.S.; Searcy, S.W., 1992 “Estimation of Tip Shape for Carrot Classification by Machine Vision”, J. ugric. Engng Res. (1992) 53, 123-139

[5] Tao, Y.; Morrow, C. T.; Heinemann, P. H.; Sommer, J. H. “Automated machine vision inspection of potatoes”. American Society of Agricultural Engineers 1990 No.90-3531 pp.23 pp. ref.27

[6] F. López-García, G. Andreu-García, “Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach”, Computers and Electronics in Agriculture, Volume 71, Issue 2, May 2010, Pages 189-197

[7] Bhavini J. Samajpati, Sheshang D. Degadwala “Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier” IEEE International Conference on Communication and Signal Processing, pp. 978-50900396, 2016.

[8] Anand Singh Jalal, Shiv Ram Dubey “Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns” IEEE Third International Conference on Computer and Communication Technology, pp. 978-0-7695-4872, 2012.

[9] Bhavini J. Samajpati, Sheshang D. Degadwala “Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier” IEEE International Conference on Communication and Signal Processing, pp. 978-50900396, 2016.

[10] Leemans, V., Magein, H., & Destain, M. F. “Defect Segmentation on Golden Delicious Apples by using Color Machine Vision” Computers and Electronics in Agriculture, 20, pp 117-130, 1998

[11] Z. Wasik and A. Saffiotti, "Robust Color Segmentation for the RoboCup Domain," Proc. of the Int. Conf. on Pattern Recognition (ICPR), vol.2, pp.651-654, 2002.

[12] Gy. Dorko, D. Paulus, and U. Ahlrichs, ”Color segmentation for scene exploration,” Workshop Farbbildverarbeitung, Oct 200, Berlin, Germany, 2000

[13] Manisha Bhange H.A. Hingoliwala, “Smart Farming: Pomegranate Disease Detection Using Image Processing” – Procedia Computer Science 58 pp280-288, 2015

[14] Monika Jhuria, Ashwani Kumar, Rushikesh Borse, “Image Processing for Smart Farming: Detection of Disease and Fruit Grading”, IEEE Second International Conference on Image Information Processing, 2013, pp 521-526.

[15] Shiv Ram Dubey, Pushkar Dixit, Nishant Singh, Jay Prakash Gupta, "Infected Fruit Part Detection using K-Means Clustering Segmentation Technique" International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 2, Nº 2.

[16] D.Napoleon, A.Shameena, and R.Santhosh,"Color Image Segmentation using OTSU Method and Color Space," International Journal of Computer Applications (0975-8887),pp.5-8,2013.

[17] Explanation of the LAB Color Space:
https://www.aces.edu/dept/fisheries/education/pond_to_plate/documents/ExplanationoftheLABColorSpace.pdf , [Last Accessed on 01.04.2018 at 10:00pm.]

[18] Van Huy Pham and Byung Ryong Lee "An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm", Vietnam Journal of Computer Science, February 2015, Volume 2, Issue 1, pp 25–33

[19] Texture Analysis Using the Gray-Level Co-Occurrence Matrix (GLCM) https://www.mathworks.com/help/images/texture-analysis-using-the-gray-level-co-occurrence-matrix-glcm.html , [Last Accessed on 01.04.2018 at 10:00pm.]

[20] B. K. Miller and M. J. Delwich " PEACH DEFECT DETECTION WITH MACHINE VISION" https://www.researchgate.net/publication/276014187_Peach_defect_detection_with_machine_vision, [Last Accessed on 01.04.2018 at 10:00pm.]

Appendices

Appendix A: Research Reflection

The purpose of this Appendix is to provide an introduction to Research refection. The group research project was a challenging and enjoyable experience typical of the course as a whole. We have had little exposer to group work at university. So, it was a nice change to be part of an effective and dynamic team.

The experience taught us that planning and crafting responses takes a longer time in teams than on your own. The extensive effort required was ultimately a good thing. In our group we are constantly developing and refining one another's ideas. We had to go to villages and farms to collect the images, that was very enjoying and also challenging for us. We enjoyed a lot talking to the farmers who helped us a lot. This research results would help them to their future cultivation by all means.

Appendix B: Related Issues

Collecting images from this kind of urban area like Dhaka was very difficult. We had to go to villages and markets to capture the disease affected carrot images. We had to talk to the farmers a while to let them understand the problem and importance of the study. They were very friendly though to help us.

We had to learn so many new algorithms and techniques to implement our ideas and research work to be effective. Variation of the image backgrounds and quality of the images were challenging to modify and reduce the changes in results hereby.

PLAGIARISM REPORT