አዲስ አበባ ዩኒቨርሲቲ

Addis Ababa University

2.

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

AUTOMATIC FLOWER DISEASE IDENTIFICATION USING IMAGE PROCESSING

Getahun Tigistu

ADVISOR:

Yaregal Assabie(PhD)

APPROVED BY

EXAMINING BOARD:

1. Dr. Yaregal Assabie, Advisor ______________________
2. ______________________ ______________________
3. ______________________ ______________________

ACKNOWLEDGEMENT

First of all, I would like to express my deepest gratitude to my advisor Dr. Yaregal Assabie for his motivate and constructive guidance right from the moments of problem formulation to the completion of the work. Many thanks and appreciations go to him for the discussions with him always made me think that things are possible.

I would also like to thank Mr.Nebiat Fikru, Mr. Lijalem Kelemu, Mr.serekberhan F/Mariayam and Mr.Addis Hailu for sharing of valuable ideas and encouragement. Without their support, I may not have the opportunity to finish my thesis.

I would also like to thank the Ethiopian horticulture agency for supplying samples of flower images for different disease category. I would like to extend my appreciation and thanks to Joe and Herburg flower producing company.

I would like to thank my wife Heran Yehenew and my daughter Eldana Getahun for they have showed me their love and support despite the fact that I abandoned them focusing on my work.

Finally, I would like to thank my father and mother and all the rest of my families, friends and peers who, in one or the other way brought me up to a success in my academic endeavor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
BPNN	Back Propagation Neural Network
CCM	Color Co-occurrence Method
HIS	Hue, Intensity, Saturation
JPEG	Joint Photographer Expert Group
MLP	Multilayer Preceptron
MSE	Mean Squared Error
RGB	Red, Green, Blue
YCbCr	Luminance; Chroma: Blue; Chroma: Red

ABSTRACT

Currently, the cultivation of flowers is becoming popular. However, during the cultivation process there may be a number of challenges that affect it, one of which is flower disease. Most flower diseases are caused by insects, fungi, and bacteria. Identification of these diseases need experienced experts in this area. Thus, developing a system that automatically identifies flower diseases can help to support the experienced experts.

In view of this, an image processing based system for automatic identification of flower disease is proposed. The proposed system consists of two main phases. In the first phase normal and diseased flower image are used to create a knowledge base. During the creation of the knowledge base, images are pre-processed and segmented to identify the region of interest. Then, seven different texture features of images are extracted using Gabor texture feature extraction. Finally, an artificial neural network is trained using seven input features extracted from the individual image and eight output vectors that represent eight different classes of disease to represent the knowledge base. In the second phase, the knowledge base is used to identify the disease of a flower.

In order to create the knowledge base and to test the effectiveness of the developed system, we have used 40 flower images for each of the eight different classes of flower disease and we have a total of 320 flower images. From those images 85% of the Dataset is used for training and 15% of the data set is used for testing. The experimental result demonstrates that the proposed technique is effective technique for the identification of flower disease. The developed system can successfully identify the examined flower with an accuracy of 83.3%.

Keywords: Gabor Feature Extraction, Artificial Neural Network, Texture Feature

`CHAPTER ONE: INTRODUCTION

1.1 Background

Ethiopia is one of the countries in Africa which have a huge potential for the development of different varieties of horticultural crop. The country is endowed with natural resources in different agro-ecological zones which are suitable for the cultivation of horticultural products. Accordingly, large varieties of flower are being grown currently in various areas of the country.

Even though the revival of the horticulture industry in Ethiopia started only about a decade ago, the industry has scored significant positive developments. Currently, more than 120 foreign and local companies are engaged in the cultivation of horticultural export products. The majority of the companies operating in the sector are owned by foreign investors in the form of sole proprietorships. Out of the total number of horticulture producers and exporters, about 80% are engaged in the floriculture business, whereas the remaining 20% are involved in the development of vegetable and fruit [45].

In the cultivation of flowers there are a number of risks. One of the risks is disease that attacks them. In plant science, there are a number of diseases that attack a flower species. Plant pathology or physiopathology is the science dealing with plant diseases and their control. Plant pathologists study plant diseases caused by fungi, bacteria, viruses, nematodes, and parasitic plants. They also study plant disorders caused by nutrient imbalances, air pollution, and other unfavorable growing conditions [15].

As diseases of a plant are inevitable, detecting disease plays a major role in the field of Agriculture. Plant disease is one of the crucial causes that degrades quantity and reduces quality of the agricultural products [40].

To identify disease that cause the problem of a flower, it is usually necessary to look at the flower closely; examine the flower, leaves, stem and sometimes the roots; and do some detective work to determine possible causes. This identification of flower disease can be done through experts on this filed. However, getting experts for the identification and investigation of flower disease is difficult (expensive) for flower growers that are far from the place where experts are found [15]. In order to minimize and properly identify plant disease, it is possible to develop a system that identifies plant disease by using an image processing. Image processing can be used, to detect diseased plant, to recognize affected areas of plant, to identify shape and size of fruit and etc. [11].

Many flower disease produce symptoms which are the main indicators in filed diagnosis. The development of proper methodology for flower disease identification is quite useful. In general, in this research paper we are going to develop an automatic flower disease identification mechanism using a digital image processing

1.2. Motivation

After a few years Ethiopia flower sector will outgrows Kenya flower sector. Currently, Ethiopia is now Africa's second largest flower exporter after Kenya, with its export earnings growing by 500% over the past year [43]. This has left Kenya stunned, given that five years ago, the Horn of Africa country was doing less than $20 million of exports compared with the East African giant's $300 million. It is estimated that, this year, Ethiopia will close its books at $120 million, slightly less than half of Kenya's earnings. According to the recent data, in Ethiopia a total of 1442 hectares of land is covered by more than 80 flower growers who come from the Netherlands, India and Israel as well as domestic investors [45]. In fact, biological pest control has the great advantage in assuring the safety of employee, protecting the environment, and also to reduce cost while increasing quality [14]. So developing a system that identifies flower disease in its early stage will help the flower growers to use the biological pest control mechanism effectively.

1.3. Statement of the Problem

In the agrarian economy, the major source of economy is the cultivation of plants. Currently, flower plantation becomes the major investment in Ethiopia. However, in the cultivation process there are a number of challenges that affect it, from those challenges disease is one of them. The identification of flower disease needs a special attention that requires experienced experts in the area. However, in areas that cultivates flowers and far from expertise may suffer in the identification of disease that attacks their plantations. This is due to the fact that plant disease detection and identification requires continuous monitoring of flowers by experts [15]. Despite the importance of the subject of identifying plant disease using digital image processing and although this has been studied for at least 30 years, the advances achieved to be a little small [41]. Some facts lead to this conclusion:

Methods are too specific: many of the methods that are being proposed not only are able to deal with only one species of plant, but also those plants need to deal with at certain growth stage in order for the algorithm to be effective.

Operation conditions are too strict: In previously developed plant disease identification and detection methods, the desire images are captured under a certain very strict conditions like lighting, angle of capture, distant between objects and capture device.

In the current flower plantation, diseases are identified and detected through a traditional process like people with experience and experts. However, identifying and detecting flower diseases in such a way are difficult and expensive. Thus, developing a system that uses a computer Vision will help the farmers to control and respond effectively for the disease they encounter. In this research work, we develop an automatic flower disease identification method that helps the agricultural society by using image processing techniques.

1.4. Objectives

General Objective

The general objective of this research is to develop an automatic flower disease identification method using image processing.

Specific Objective

In order to achieve the general objective, the following specific objective are identified

- Review works that are related to flower and leaf disease identification and detection mechanisms.
- Explore the different types of flower disease in different development stage of the flower.
- Select an appropriate methodology (tool) to analyze the images of flower.
- Design architecture of the proposed system.
- Select algorithms that are best suited for image segmentation and feature extraction process.
- Develop the prototype.
- Test and evaluate the system.

1.5. Scope and Limitation of the Study

The main focus of this research is to develop an automatic flower disease detection and identification mechanism. After the identification and detection of the disease, there may be a way to recommend the appropriate medicine for that disease, however recommending the appropriate treatment for the identified disease, estimating the severity of the detected disease is beyond the scope of this thesis work.

1.6. Methodology

Literature Review

To understand the subject matter, literatures related to the work that we are going to develop will be reviewed. This will help us to summarize all related works that are relevant to this work. Since this work is on developing a system that identifies flower disease, we will look at related topics that deal about Agro-industrial disease identification and detection mechanisms trough a computer vision.

Data Collection

We will collect flower images from two flower plantation in Ethiopia, namely Herburg Roses and Joe Flowers and other source of flower image repository. The images were collected by using a digital camera in certain conditions.

Tools

For image preprocessing and analysis of flower images, MATLAB for windows was used. This tool has a great capability on array based data processing. Thus, for the purposes of pre-processing like enhancement and segmentation of an image and creating the GUI, MATLAB will be used.

For neural network pattern recognition, a Neural Network Toolbox provided by MATLAB will be used. It provides function and application for modeling complex nonlinear systems that are not easily modeled with a closed-form equation. The Neural Network Toolbox supports supervised learning with feed forward, radial basis, and dynamic network [30]. It also supports unsupervised learning with self-organizing maps and competitive layers.

1.7. Application of Results

The development of such a system has a great advantage for the flower agro-industry. Some of the application areas of this work are:

- Early stage flower disease identification: since we are using a real time flower disease identification system the development of this application will increase the effectiveness of disease controlling mechanism. This application also helps the farmers to take an effective protection approach by identifying the disease in its early stage

- Minimize the needs of experts: since we are developing an automatic system that uses a machine vision to identify flower disease, it will eliminate the needs of experts in that area.

1.8. Organization of the Thesis

The rest of this thesis report is organized as follows. Chapter two discusses the different issues related to flower diseases, digital image processing and the related subject areas as literature review. Chapter three is devoted to discuss related works done on plant disease identification and other topics related to weed and product classification. Chapter four gives a detailed description of the architecture and design issues of our system. The main components of the system, their functional operation and the specific components are discussed in this Chapter. Chapter 5 presents the implementation of the proposed system architecture and experimental results. Chapter 6 concludes the thesis by recommending some feature work. It also shows some research directions that can be used in the future to improve the flower disease identification system.

CHAPTER TWO: LITERATURE REVIEW

2.1. Introduction

This chapter focuses on diseases that attack flower, digital image processing techniques and approaches for real world applications. In digital image processing, each and every systems of machine vision starts from image acquisition. After the images are captured there are a number of processes that we follow to reach the desired goal of a machine vision system. In this chapter, we explain the science of a machine vision system.

2.2. Flower Diseases

A plant becomes diseased when it is continuously disturbed by some causal agent that results in an abnormal physiological process that disrupts the plant's normal structure, growth, function, or other activities. Plant disease can be broadly classified according to the nature of their primary causal agent, either infectious or noninfectious [44]. Infectious plant diseases are caused by a pathogenic organism such as a fungus, bacterium, mycoplasma, virus, viroid, nematode, or parasitic flowering plant. An infectious agent is capable of producing within or on its host and spreading from one susceptible host to another.

In nature, plants may be affected by more than one disease-causing agent at a time. A plant that must contained with a nutrient deficiency or an imbalance between soil moisture and oxygen is often more susceptible to infection by a pathogen; a plant infected by one pathogen is often prone to invasion by secondary pathogens.

Flowers are one of the plant species that are attacked by different type of disease. Flower diseases are caused by fungi, bacteria and pests. Some of the disease that attack the flower species are powdery mildew, Aphid, Japanese beetles, Rosettle, Goldenrode-soldier, crown gall, Rust, Soldier beetles, Gray mold, Botrytis blight and Black spot [46].

In Ethiopia in the year 2011and 2012, these flower diseases affected different flower plantation area as it is shown in Table 2.1 [44].

Table2. 1: Examples of flower diseases that attack Ethiopian horticulture

No	Name of company	Damage causing pest	Crop attacked	Covered Area
1	Minaye flowers	Spider mite	Rose	40
2	Tinaw	Spider mite	Rose	13.6
3	Rainbowcolor	Soldier beetles	Rose	15.3
4	Minaya flowers	Mossyrose gall wasp	Rose	27
5	AQ Roses	Powdery mildew	Pepper	Unknown
6	Joe Flowers	Rose rust	Rose	34
7	Joe Flowers	Aphid	Rose	21
8	Herburg	Aphid	Rose	24
9	Joe Flowers	Mossyrose gall wasp	Rose	29.5
10	Joe Flowers	Black spot	Rose	unknown

Some of the above flower diseases are explained as follows:

Rose Aphid: (Greenfly) (Order Hemiptera Family Aphididae) Macrosiphumrosae Macrosiphum rosae – Likely to be found on new shoots and buds, aphids are soft bodied insects 1-2mm long. Often green but occasionally light-brown, and sometimes with wings, they may cover (in a colony) the complete growing tip of the plant. Aphids are most active in spring and summer and multiply at a prodigious rate feeding on the sap of the plant by piercing the plant cells via a proboscis. In large quantities they may seriously retard the growth of the plant and ruin buds. Some of the symptoms are:

- Green or pink aphids, cluster on the flower buds, shoot tips and young foliage
- White cast aphid skins are often seen on infested flower buds and leaves
- Flower buds and foliage covered in a sticky honeydew that aphids excrete
- Black sooty moulds may grow on the honeydew

Japanese Beetles: the Japanese beetles are small bug and has iridescent, copper and green colorings. Roses can be a favorite food for these pests, although some varieties are more resistant than others. They can be quite destructive; chewing holes in leaves and flowers.

Black Spot: Black or brown spots appear on leaves, turning them yellow and causing them falloff. This is the most-common rose disease. If left untreated, it will infect the whole plant. Some of the symptoms are:

- Typically, a rapidly enlarging purplish or black patch appears on the upper leaf surface, with diffuse and radiating strands of the fungus sometimes just visible
- Leaf tissues may turn yellow color does not appear, but infected leaves often drop, even though other parts are as yet unaffected
- At other times, the yellow color does not appear, but infected leaves still drop
- Sometimes, the spots remain relatively small and the leaf does not drop
- Small, black, scabby lesions may also appear on young stems

Mossyrose gall wasp: Flower galls come in a fascinating variety of strange forms, textures and colors. Some are irregular, bumpy, or warty; others are smooth and spherical. Some galls sport thick growths of fuzz, hair or spines. One of my favorites is the aptly named mossy rose gall - it occurs on rose and looks like moss. Galls result from an intricate interaction between two living organisms. The gall-maker (insect disease, mite) causes the flower to change the course of normal growth and modifies growing tissue into a special swelling that surrounds the gall-maker.

Powdery mildew: Rose powdery mildew is a disease of roses caused by the fungus *Podosphaerapannosa*. The conspicuous white growth can affect all aerial parts of the plant, producing microscopic spores that spread the disease. High humidity is favorable for infection, and plants growing in areas where air movement is poor or the soil is dry can be severely affected. Some of the symptoms are:

- Older lesions turn brown and appeared shriveled
- Mycelium of fungus forms mats and appears as white, grayish white or tan colored patches on leaves, buds, stems or young fruit
- Fruiting bodies (cleistothecia) appear as small black or brown specks on the mycelial mats
- Infected leaves often appear chlorotic due to decreased photosynthesis
- Infected fruit and flowers are often aborted or malformed
- early signs include small chlorotic spots or blistering on leaves or flowers

Rose Rust: Rust is a fungal disease that seems to be more prevalent in West Coast gardens; it is rarely indicated east of the Rocky Mountains. Rust appears as its name implies, as red-orange spots (raised looking like warts) on undersides of leaves and yellow blotches on top surfaces. Long, narrow rust spots or streaks may also form on young canes. It left on roses, these spots will develop into large groupings of rust 'warts' and in autumn they will turn black. Ultimately this disease will cause the entire defoliation of the rose plant.

- Rose rust most commonly appears in spring and fall, but can appear in the summer months as well
- Rose rust fungus appears as small orange or rust colored spots on the leaves and will grow to bigger markings as the infection advances. The spots on the canes of the rose bush are orange or rust colored but become black in the fall and winter
- Rose rust fungus appears as small orange or rust colored spots on the leaves and will grow to bigger markings as the infection advances. The spots on the canes of the rose bush are orange or rust colored but become black in the fall and winter

Rose-rosettle: Rose rosette disease, also known as witches'-broom of rose, is a virus or virus-like disease, such as a phytoplasm, that is spread by a very small, eriophyid mite. The disease is limited to plants in the genus *Rosa.*

- Increased growth/rapid elongation of shoots
- Abnormal red discoloration of shoots and foliage
- Witches broom (prolific clustering of small shoots)
- Shortening of internodes(shorter stem length between leaves)
- Distorted or dwarfed leaves
- Deformed buds and flowers

Botrytis Blight: During periods of cool and wet weather, Botrytis blight frequently develops on roses. The disease may affect flowers which may not open and may become covered with grayish brown fungal growth.

The botrytis blight fungus is sort of grayish brown and looks fuzzy or wooly. The botrytis blight fungus seems to attack mostly hybrid tea rose bushes, attacking the leaves and canes of the subject rose bush. It will prevent the blooms from opening and many times causes the bloom petals to turn brown and shrivel up.

2.3. Digital Image Processing

An image may be defined as a two-dimensional function, $f(x, y)$, where x and y are spatial (plane) coordinates, and the amplitude of at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, we call the image a digital image [7]. The field of digital image processing refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, and pixels. Pixel is the term most widely used to denote the elements of a digital image. Vision is the most advanced of our senses, so it is not surprising that images play the single most important role in human perception. However, unlike humans, who are limited to the visual band of the electromagnetic (EM) spectrum, imaging machines cover almost the entire EM spectrum, ranging from gamma to radio waves. They can operate on images generated by sources that humans are not accustomed to associating with images. These include ultra-sound, electron microscopy, and computer-generated images. Thus, digital image processing encompasses a wide and varied field of applications. There is no general agreement among authors regarding where image processing stops and other related areas, such as image analysis and computer vision, start. Sometimes a distinction is made by defining image processing as a discipline in which both the input and output of a process are images. We believe this to be a limiting and somewhat artificial boundary [6]. For example, under this definition, even the trivial task of computing the average intensity of an image (which yields a single number) would not be considered an image processing operation. On the other hand, there are fields such as computer vision whose ultimate goal is to use computers to emulate human vision, including learning and being able to make inferences and take actions based on visual inputs. This area itself is a branch of artificial intelligence (AI) whose objective is to emulate human intelligence.

There are no clear-cut boundaries in the continuum from image processing at one end to computer vision at the other. However, one useful paradigm is to consider three types of computerized processes in this continuum: low-, mid-, and high-level processes.

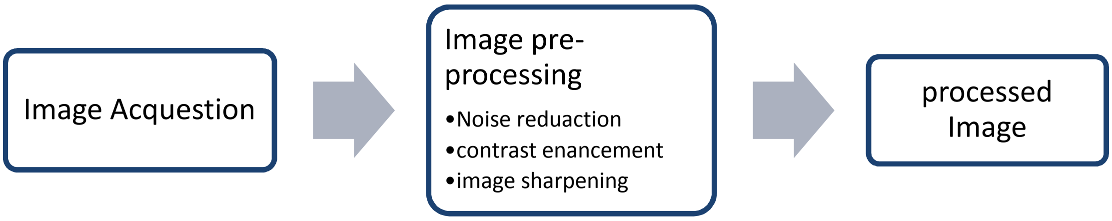
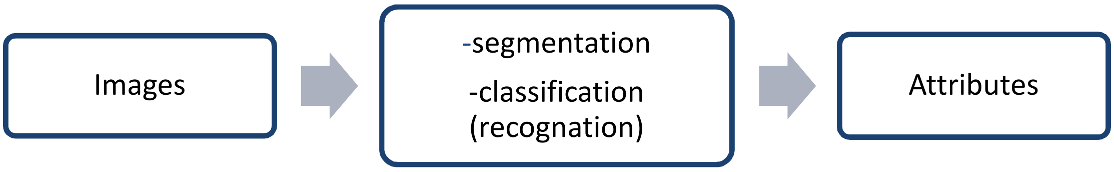


Image Acquestion
Image pre-processing
•Noise reduaction
•contrast enancement
•image sharpening
processed Image


Images
-segmentation
-classification (recognation)
Attributes

simple illustration to clarify these concepts, consider the area of automated analysis of text. The processes of acquiring an image of the area containing the text, pre-processing that image, extracting (segmenting) the individual characters, describing the characters in a form suitable for computer processing, and recognizing those individual characters are in the scope of what we call digital image processing.

Today, there is almost no area of technical endeavor that is not impacted in some way by a digital image processing [7]. The application areas of a digital image processing are so varied. One of the simplest ways to develop an understanding of the extent of image processing application is to categorize images according to their source. The principal energy source for images in use today is the electromagnetic energy spectrum. Other important source of energy includes acoustic, ultrasonic, and electronic (in the form of electron beams used in electron microscopy). Thus, imaging techniques based on this source of energy includes gamma-ray imaging (nuclear medicine and astronomical observation), X-ray imaging(medical diagnosis and astronomy), imaging in the ultraviolet band(lithography, industrial inspection, microscopy, lasers, biological imaging and astronomical observation), imaging in the visible and infrared bands(light microscopy, astronomy, remote sensing, industry, and law enforcement), imaging in the microwave band(radar system) and imaging in the radio band(medicine and astronomy).

2.4. Fundamental Steps of Digital Image Processing

Basically different image processing applications may follow different steps. However, there are some fundamental steps that every image processing applications pass through. These steps are shown in Figure 2.3 [31].

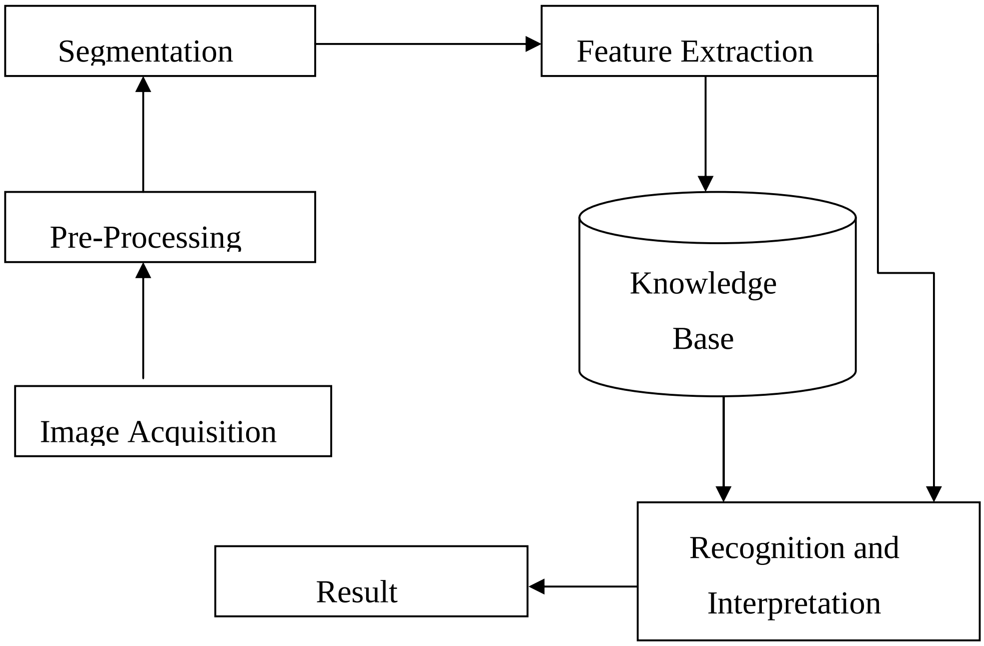


Figure2.3: Fundamental steps of digital Image processing

2.4.1. Image Acquisition

Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source, usually a hardware-based source, so it can be passed through whatever process need to occur afterward. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the result of whatever hardware was used to generate it, which can be very important in some fields to have a consistent baseline from which to work. One of the ultimate goals of this process is to have a source of input that operates with in such controlled and measured guidelines that the same image can, if necessary, be nearly perfectly reproduced under the same conditions so anomalous factors are easier to locate and eliminate [20].

Depending on the field of work, a major factor involved in image acquisition in image processing sometimes is the initial setup and log-term maintenance of the hardware used to capture the image. The actual hardware device can be anything from desktop scanner to a massive optical telescope.

2.4.2. Image Pre-processing

Image pre-processing is a common name for operations with image at the lowest level of abstraction both input and output are intensity of image. Most image-processing techniques involve removing low-frequency background noise, normalizing the intensity of the individual particles images, removing reflections, and masking portions of images. Image preprocessing is the technique of enhancing data images prior to computational processing. Image processing usually refers to digital image processing, however optical and analog image processing also are possible. Some of image pre-processing tasks are:

Image Re-sampling

It is the process of transforming a sampled image from one coordinate system to another (changing the pixel dimension of an image) [1]. Re-sampling also affects the display size of an image. When we resample an image it can lose some information from the image(down sampling) or increase the number of pixels in the given image, new pixels are added based on color values of existing pixels

Image Enhancement

Image enhancement is one of the simplest and most appealing areas of digital image processing. Basically the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. A familiar example of enhancement is when we increase the contrast of an image because it looks better. It is important to keep in mind that enhancement is a very subjective area of image processing.

Image Restoration

Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration technique tends to be based on mathematical or probabilistic models of image degradation.

2.4.3. Segmentation

Segmentation refers to the operation of partitioning an image into component parts, or into separate objects (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Segmentation algorithms are based on the two basic properties of intensity values- discontinuity (to partition an image based on sharp changes in intensity) and similarity (to partition an image into regions that are similar according to a set of predefined criteria) among the pixels. Some of the well-established segmentation techniques are:

Segmentation by Thresholding

Because of its intuitive properties and simplicity of implementation, image thresholding enjoys a central position in application of image segmentation. In this method the pixels are partitioned depending on their intensity value. By selecting an adequate threshold value T, the gray level image can be converted to binary image. The binary image should contain all of the essential information about the position and shape of the objects of interest (foreground) [2, 3].

Thresholding may be viewed as an operation that involves tests against a function T of the form

$$\boldsymbol{T}[\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})] \tag{2. 1}$$

Where f(x,y) is the gray level of point (x, y) and p(x, y) denotes some local property of this point- for example, the average gray level of a neighborhood centered on (x, y). A threshold image g(x, y) is defined in equation 2.2.

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} \boldsymbol{1}, \boldsymbol{if}\ \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) > T \\ \boldsymbol{0}, \boldsymbol{if}\ \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) \leq \boldsymbol{T} \end{cases} \tag{2. 2}$$

Where T is the threshold, *g(x, y)*=1 for image elements as the objects, and g(x, y)=0 for image elements as the background. In this case T depends only on f(x, y) and this tresholding is called global tresholding. When the pixel values of the component and that of Background are fairly consistent in their respective values over the entire image, global thresholding could be used. A simple iterative algorithm for threshold *(T)* selection in a given image is shown below.

Step1: Choose an initial threshold $T \leftarrow T_0$

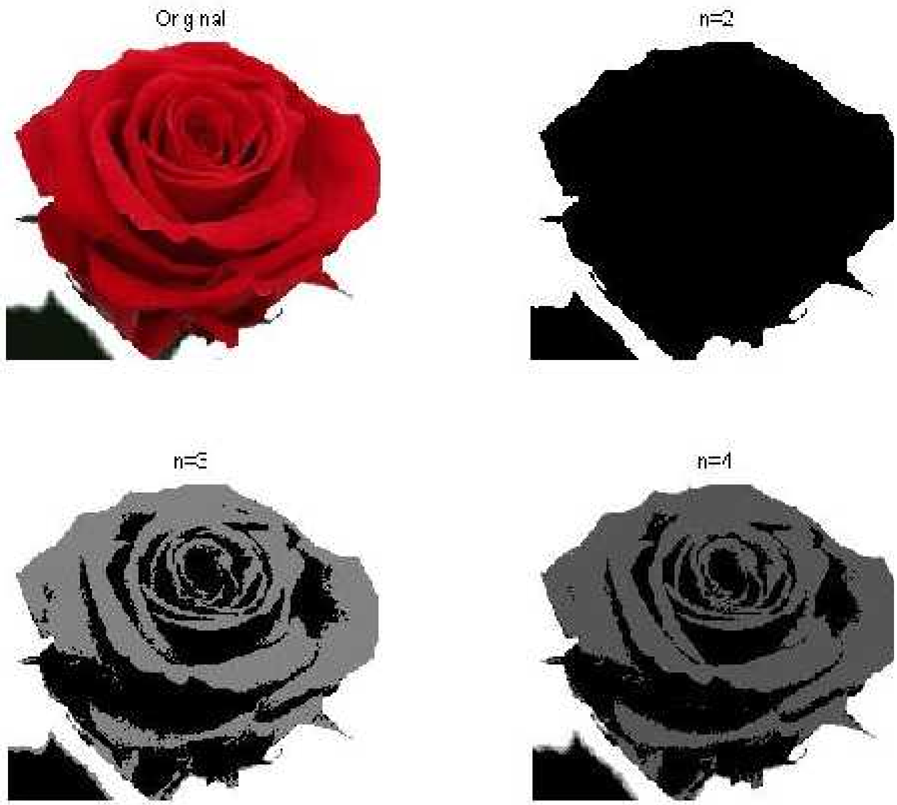
Step2: Partition the image using *T* in two regions- background and foreground (object)

Step3: Compute means gray values μ_1 and μ_2 of background and object region respectively.

Step4: Compute the new threshold $T \leftarrow \frac{\mu_1+\mu_2}{2}$

Step5: Repeat Steps 2 to 4 until there is no change of *T*

If T depends on both *f(x, y)* and *p(x, y)*, or if it depends on the spatial coordinates x and y the thresholding is called dynamic or adaptive. So if an image contains different pixel intensities in each region adaptive thresholding selects an individual threshold for each pixel based on the range of intensity values in its local neighborhood. Figure 2.4 shows results of segmentation by Otsu thresholding using different threshold value.


Original
n=2
n=3
n=4

at location *(x, y)* is the gradient of the image. Since an image *f(x, y)* is a two dimensional function its gradient is a vector

$$\nabla f = grad(f) = \begin{bmatrix} Gx \\ gy \end{bmatrix} = \begin{bmatrix} \frac{df}{dx} \\ \frac{df}{dy} \end{bmatrix} \tag{2. 3}$$

The vector has the important geometrical property that it point in the greatest of change of f at location (x, y). The magnitude of the gradient is the value of the rate of change in the direction of the gradient vector and calculated according to the equation 2.4.

$$\left.\begin{aligned} G[f(x,y)] &= \sqrt{G_x^2 + G_y^2} \\ G[f(x,y)] &= |G_X| + |G_y| \\ G[f(x,y)] &= \max\{|G_x|, |G_y|\} \end{aligned}\right\} \tag{2. 4}$$

The direction of the gradient with respect to the x-axis calculated according to the equation 2.5:

$$\theta(x,y) = \tan^{-1}\left({G_y}/{G_x} \right) \tag{2. 5}$$

Gradient operators compute the change in gray level intensities and also the direction in which the change occurs. This is calculated by the difference in values of the neighboring pixels. Gradient operators require two masks, one to obtain the x-direction gradient and the other to obtain the y-direction gradient. These two gradients are combined to obtain a vector quantity whose magnitude represents the strength of the edge gradient at a point in the image and whose angle represents the gradient angle. A number of edge detector based on a single derivative have been developed by various researchers. Amongst them the most important are Robert Operator edge detector, Soble operator-based edge detector, Prewitt. In each of these operator-based edge detection strategies, we compute the gradient magnitude by using the above formula.

Robert Operator-Based Edge detector

The Robert cross operator is a simple gradient operator based on 2×2 gradient operator. This operator provide the simplest approximation of the gradient magnitude given by equation 2.6

$$\boldsymbol{G}[\boldsymbol{f}(\boldsymbol{i},\boldsymbol{j})] = [\boldsymbol{f}(\boldsymbol{i},\boldsymbol{j}) - \boldsymbol{f}(\boldsymbol{i}+\mathbf{1},\boldsymbol{j}+\mathbf{1})][\boldsymbol{f}(\boldsymbol{i}+\mathbf{1},\boldsymbol{j}) - \boldsymbol{f}(\boldsymbol{i},\boldsymbol{j}+\mathbf{1})] \qquad (2.\ 6)$$

The convolution mask for the Robert's operator is shown below.

1	**1**
-1	**-1**

Soble Operator-Based Edge Detector

The Soble edge detector [7] is a nonlinear edge enhancement technique. It is another simple techniques for enhancing edges.(i, j) enumerated in the counterclockwise direction as follows:

a_3	a_2	a_1
a_4	(i,j)	a_0
a_5	a_6	a_7

The Soble edge magnitude image $m\epsilon R^x$ is given by equation 2.7.

$$\boldsymbol{n} = (\boldsymbol{i},\boldsymbol{j}) = (\boldsymbol{u}^2 + \boldsymbol{v}^2)^{\frac{1}{2}}, \qquad (2.\ 7)$$

Where

$$u = (a_5 + 2a_6 + a_7) - (a_1 + a_2 + a_3)$$

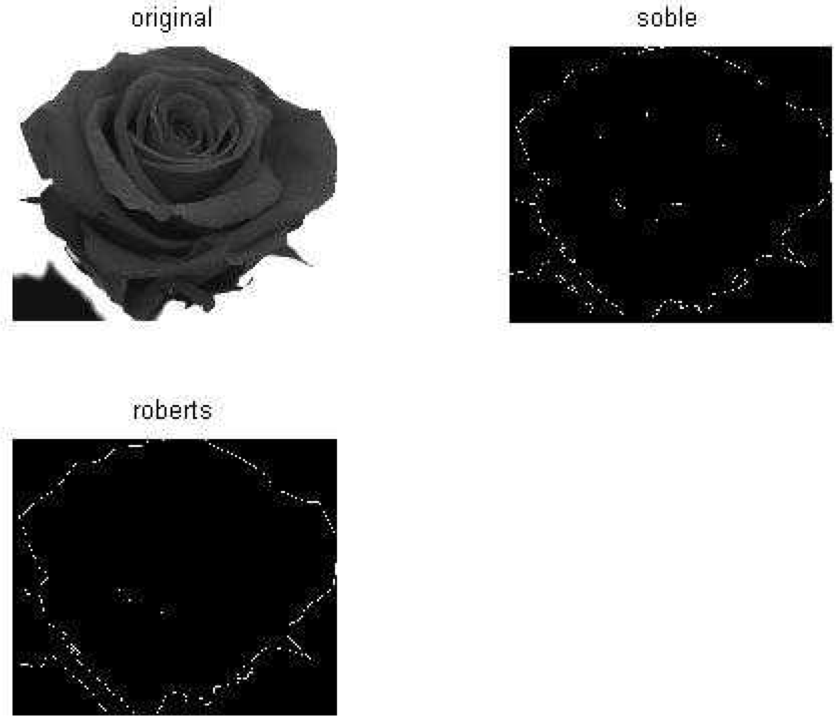
and

$$v = (2a_0 + a_1 + a_7) - (a_3 + 2a_1 + a_5)$$

The convolution mask for the Soble operator is defined by the two kernels as shown below. The two masks are separately applied on the input image to yield two gradient components G_x and G_y in the horizontal and vertical orientations respectively [7, 8, and 9].

1	2	1
0	0	0
-1	-2	-1

1	0	-1
2	0	-2
1	0	-1



2.4.4. Feature Extraction

The segmentation techniques yield data in the form of pixels along a boundary or pixels in region. After an image is segmented into regions; the resulting aggregate of segmented pixels is represented and described for further computer processing. In the process of image representation regions can be represented in terms of its external characteristics (boundary) or in terms of its internal characteristics (pixels comprising the region). If our primary focus is on shape characteristic, we can use external representation, on the other hand if our primary focus is on region properties like color and texture we can use internal representations. Sometimes, we can also use both representation techniques. Thus, based on these region representation techniques, we make the data useful to computer.

It is standard practice to use schemes that compact the segmented data into representation that facilitate the computation of descriptors. We have different external image representation techniques such as chain code, polygonal approximations, signatures, boundary segments and skeletons and we have also some internal representation techniques like color and texture.

Chain Codes

Chain codes are used to represent a boundary by a connected sequence of straight-line segments of specified length and direction (coded by using numbering schema). Typically this representation is based on 4-or-8 connectivity of segments [7].The chain code of boundary depends on the starting point. However, the code can be normalized with respect to starting point by a straightforward procedure. The chain code can be simply treated as a circular sequence of direction numbers and redefine the starting point so that the resulting sequence of numbers forms an integer of minimum magnitude.

Polygon Approximation

A digital boundary can be approximated with arbitrary accuracy by a polygon [7, 21]. For a closed curve, the approximation is exact when the number of segments in the polygon is equal to the number of points in the boundary so that each pair of adjacent points defines a segment in the polygon. In practice, the goal of polygonal approximation is to capture the "essence" of the boundary shape with the fewest possible polygonal segments. This problem in general is not trivial and can quickly turn into a time-consuming iterative search. However, several polygonal approximation techniques of modest complexity and processing requirement are well suited for image processing application.

After the representation of the image, different features of that image are extracted using different feature extraction techniques.

Image Features

A functional definition of features is something that can be measured from an image. This feature is a distinguishing primitive characteristics or attribute of an image. Rather than analyzing the image in its original form, extracting the features of the image will minimize the time to analyze it. To make the process of the image analysis simple and less time consuming, some quantitative information is extracted from the image to be analyzed. Extracting the region of interest from an image also minimize the computational cost of object recognition [22,23].

Several image features have been used to represent an image for object recognition/identification systems. Most popular among them are color, texture, shape and topology of an image.

Color Features

Color is one of the most widely used visual features in object recognition/identification. While we perceive only a limited number of gray levels, our eyes is able to distinguish thousands of colors and a computer can represent even millions of distinguishable colors in practice. Color has been successfully applied to recognize images, because it has very strong correlations with the underlying objects in an image. Moreover, color feature is robust to background complication, scaling, orientation, perspective, and size of an image.

Texture Features

Texture is a very interesting image features that has been used for characterization of images, with application in object recognition/identification. There is no single formal definition of texture in the literature. However a major characteristic of texture is the repetition of pattern or patterns over a region in an image. The elements of patterns are sometimes called textons. The size, shape, color, and orientation of the textons can vary over the region. The difference between two textures can be in the degree of variation of the textons. It can also be due to spatial statistical distribution of the textons in the image.

Texture contains important information regarding underlying structural arrangement of the surfaces in an image.

A variety of techniques have been used for measuring textural similarity. In 1973, Haralick et al. proposed co-occurrence matrix representation of texture feature to mathematically represent gray level spatial dependency of texture in an image [24]. In this method the co-occurrence matrix is constructed based on the orientation and distance between image pixels.

Another texture feature extraction method is a wavelet transform, which is based on multi-resolution decomposition of the images and representing texture in different scales. Form this wavelet filters, Gabor filter were found to be very effective in texture analysis. Gabor features are simply the response of Gabor filters. To understand the concept of Gabor filtering, we must first understand Gabor wavelets. Each Gabor wavelet is formed form two components; a complex sinusoidal carrier placed under a Gaussian envelops [24, 27, 29].

Thus, apart from the Gaussian envelop in each 2D Gabor wavelet, sinusoidal carrier has a frequency and orientation of its own and the system is similar to those of human visual system, known as the visual cortex. Gabor filtering has been found to be appropriate for discrimination and representation of texture features in particular [28].

Since Gabor filter for texture feature extraction is reported to yield a good result [24], we use here a Gabor wavelet based texture feature extraction technique. We use the following family of two-dimensional Gabor kernel as it shown in equation 2.8[25, 26], :

$$\begin{cases} \boldsymbol{\text{И}}_{(x,y,\vartheta,\lambda,\varphi,\sigma,\gamma)} = \boldsymbol{exp}\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \\ x' = x\cos(\theta) + y\sin(\theta) \\ y' = -x\sin(\theta) + y\cos(\theta) \end{cases} \tag{2. 8}$$

Where *(x, y)* specify the position of a light impulse in the visual field and μ, φ, γ, λ, σ is parameters of the wavelet. We have chosen the same parameter used by Mohammed Haghighat [27] as shown in Table 2.2.

Table2. 2: Parameters Used for Gabor feature extraction

Parameter	Symbol	Values
Orientation	θ	$\{0, \frac{\pi}{8}, \frac{2\pi}{8}, \frac{3\pi}{8}, \frac{4\pi}{8}, \frac{5\pi}{8}, \frac{6\pi}{8}, \frac{7\pi}{8}\}$
Wavelength	λ	$\{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$
Phase	φ	$\{0, \frac{\pi}{2}\}$
Gaussian Radius	σ	$\sigma=\lambda$
Aspect Ratio	γ	1

In summary, as it is reported in different research papers Gabor feature extraction is an effective method for texture feature extraction [24, 27, 28]. Thus, for this research work we have used Gabor texture feature extraction technique to extract the texture feature of flower image.

2.4.5. Recognition and Interpretation

What we call object recognition may also be called pattern recognition. A pattern is an arrangement of descriptors. These descriptors may have more forms, but they are primarily vectors and strings. Pattern recognition by machine involves techniques for assigning patterns to their respective classes- automatically with as little human intervention. [11].

A pattern recognition system based on any pattern recognition method mainly includes three mutual-associate and differentiated processes. One is data building; the other two are pattern analysis and pattern classification. Data building converts original information into vector which can be dealt with by computer. Pattern analysis task is to process the data (vector), such as feature selection, feature extraction, data-dimensional compress and so on. The aim of pattern classification is to utilize the information acquired from pattern analysis to discipline the computer in order to accomplish the classification.

The basic thing in pattern recognition is the design of a classifier, a mechanism which takes features of objects as its input and which results in a classification or label or value indicating to which class the object belongs. The performance of a classifier is ultimately limited by the overlap of the classes in feature space. Basically, classifiers can be categorized into one of the following categories, Decision theoretic and Structural method.

Decision-Theoretic Method

Decision-theoretic approaches to recognition are based on the use of decision (or discrimination) functions. Let $x=(x_1, x_2,\ldots,x_n)^T$ represent an n-dimensional pattern vector. For W pattern classes $w_1, w_2, \ldots, w_w$, the basic problem in decision-theoretic pattern recognition is to fined W decision function $d_1(x), d_2(x), \ldots, d_w(x)$ with the property that, if a pattern x belongs to class w_i, then

$$\boldsymbol{d_1(x) > d_j} \qquad \boldsymbol{j = 1, 2, \ldots W\,; j \neq i} \tag{2. 9}$$

In other words, an unknown pattern x is said to belong to the i^{th} pattern class if, upon substitution of x into all decision functions, $d_1(x)$ yields the largest numerical value [7].

The decision boundary separating class w_i from w_j is given by values of x for which $d_i(x)=d_j(x)$ or, equivalently, by values of x for which

$$d_i(x) - d_j(x) = 0$$

Common practice is to identify the decision boundary between two classes by the single function $d_{ij}(x) = d_i(x)-d_j(x) =0$. Thus $d_{ij}(x)>0$ for pattern of class w_i and $d_{ij}(x) <0$ for pattern class $w_{j.}$ there are different decision-theoretic approaches for pattern classification. Here we will discuss two of them namely, matching and Artificial neural network.

Matching

Recognition technique based on matching represent each class by a prototype pattern vector; an unknown pattern is assigned to the class to which it is closest in terms of a predefined metric. The simplest approach is the minimum- distance classifier, which, as its name implies, computes the (Euclidean) distance between the unknown and each of the prototype vectors. It chooses the smallest distance to make a decision. And the other approach is based on correlation, which can be formulated directly in terms of images and is quite intuitive.

Artificial Neural network

Artificial Neural networks (ANN) are highly distributed interconnection of adaptive nonlinear processing elements. In the other word, they are large set of interconnected neurons, which execute in parallel to perform the task of learning. Hence, ANN resembles human brain in two respects. The first property is that knowledge is acquired by the network through a learning process. The other is interneuron connection strengths known as weights are used to store the knowledge, i.e., the weights on the connection encode the knowledge of networks [12, 13].

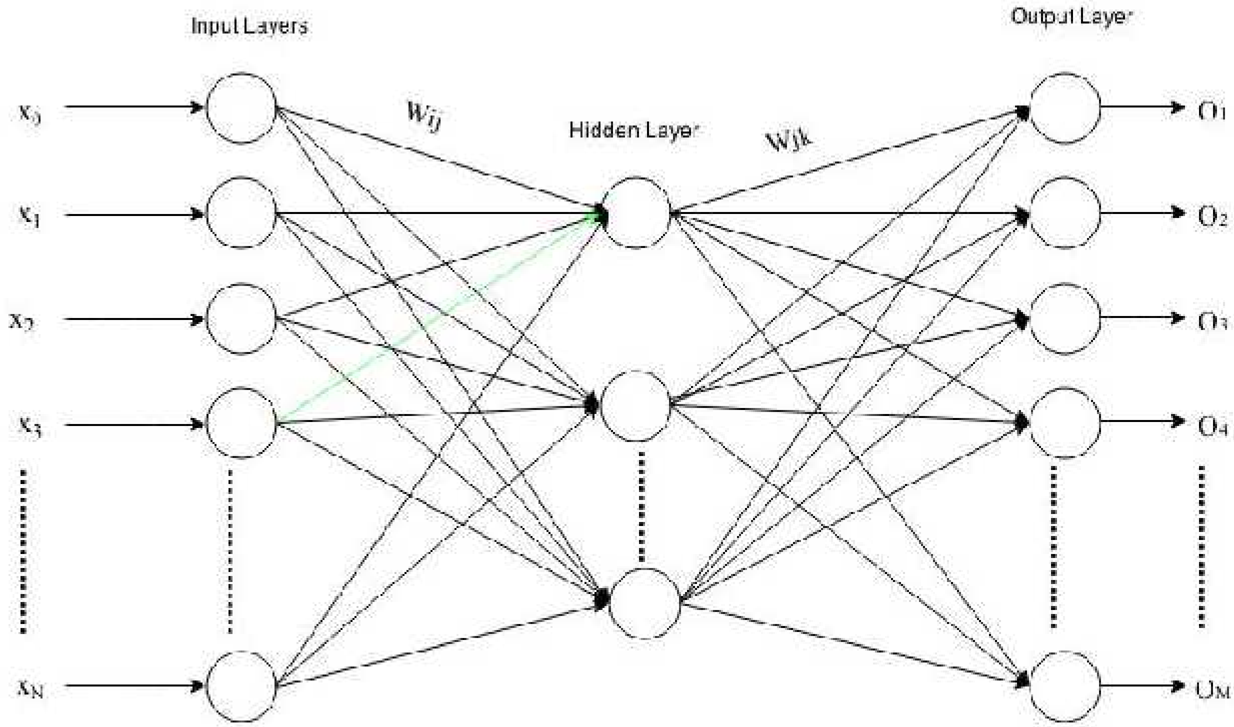
In connection with this, the ANN features of distributed processing, adaptation and nonlinearity are the hallmark of biological information processing system. Therefore, ANNs are working with the same basic principles as biological brains. That is, ANNs mimic biological brains [13].

Distributed computation of ANN has the advantage of reliability, fault tolerance, high throughput (division of computation tasks) and cooperative computing. The adaptation is the ability to change a system's parameters according to some rule (normally, minimization of an error function). Adaptation enables the system to search for optimal performances. The ANN property of nonlinearity is also important in dynamic range control for unconstrained variables and produces more powerful computation schemes when compared to linear processing. However, it complicates theoretical analysis tremendously [12].

Unlike more analytical based information processing methods, neural computation effectively explores the information contained within input data, without further assumptions. Statistical methods are based on assumptions about input data ensembles (i.e. a prior probabilities, probability density functions, etc.). Neural networks, on the other hand build relationships in the input data sets through the iterative presentation of the data and the intrinsic mapping characteristics of neural topologies, normally referred to as learning.

There are two basic phases in neural network operation. They are training or learning phase and testing-recall or retrieval phase. In the learning phase, data is repeatedly presented to the network, while weights are updated to obtain a desired response. In testing phase, the trained network with frozen weights is applied to data that it has never seen.

Although there exists many models and representations of ANNs, each one of these network possesses four tuple attributes $<Nc, W, \sigma, \delta, >$, where Nc is a finite set of highly interconnected neurons with outputs $n_1, n_2, \ldots\ldots n_k$;W denotes a finite set of weights which represents the strength w_{ij} of the interconnection between neurons ni and nj; σ is a propagation rule which shows how the input signals to a neurons ni propagates through it. A typical propagation rule may be $\delta(i) = \sum n_i w_{ij}$ and δ is an activation function which is usually a nonlinear function like sigmoid function [12]. The most popular neural network model is the multilayer perceptron (MLP), which is an extension of the signal preceptron



Architecture of Back-propagation Neural network set of labeled pattern samples belonging to each specific class, there is a desired output. The actual response of the neurons at the output layer will deviate from the desired output, which may result in an error at the output layer. The error at the output layer is used to compute the error at the hidden layer immediately preceding the output layer and the process continues [12].

In view of the above, the net input to the j^{th} hidden neuron is expressed in equation 2.10

$$I_j^h = \sum_{n=1}^{N} x_n w_{ij}^h + \theta_j^h \qquad (2.\ 10)$$

The output of the j^{th} hidden layer neuron is calculated using equation 2.11

$$O_j = f_j^h(I_j^h) = \frac{1}{1+e^{-I_j^h}} \qquad (2.\ 11)$$

Where *x1,x2….xn* is the input pattern vector, weight w_{ij} represents the weight between the hidden layer and the input layer, and θ_j^h is the bias term associated with each neuron in the hidden layer. These calculations are known as forwarded pass. In the output layer the desired or target output is set as Tk and the actual output obtained from the network is Ok. The error (Tk-Ok) between the desired signal and the actual signal is propagated backward during the backward pass. The equation governing the backward pass is used to correct the weight. Thus the network learns the desired mapping function by back propagating the error. The average error *E* is a function of weight as shown below.

$$E(W_{jk}) = \frac{1}{2}\sum_{k=1}^{M}(Tk - Ok)^2 \qquad (2.\ 12)$$

To minimize the error *E* we have to find the root of the partial derivatives

$$\sum_{k=1}^{M}\frac{\partial E}{\partial w_{jk}} = 0 \qquad (2.\ 83)$$

Hence, from this we can obtain the value of updated weights as follows

$$W_{jk}^{(new)} = W_{jk}^{(old)} + \eta\delta_j O_j,$$

Whereη is the learning rate of the hidden layer neurons

In summary, artificial neural networks can be regarded as an extension of many classification techniques, which have been developed over several decades. These networks are inspired by the concept of the biological nervous system, and have provedto be robust in dealing with the ambiguous data and the kind of problems that require the interpolation of large amounts of data. Instead of sequentially performing a program of instructions, neural networks explore many hypotheses simultaneously using massive parallelism. Neural networks have the potential for solving problems in which some inputs and corresponding output values are known, but the relationship between the inputs and outputs is not well understood or is difficult to translate into a mathematical function [12, 13]. These conditions are commonly found in tasks involving grading and classification of agricultural products and plant disease identification [42].

CHAPTER THREE: RELATED WORKS

3.1. Introduction

In this Chapter, related works of different researchers in the area of plant disease identification using machine vision system were reviewed. In fact there are no any researches that are done regarding the identification of flower disease using image processing. Thus, papers related to plant leaf, stem and root disease identification, weed and crop detection and identification, and agricultural product classification will be assessed and reviewed. In order to describe each research clearly we will divide the chapter in to different sections based on the area of research.

3.2. Plant Leaf Disease Identification

Based on a computer image processing different research regarding plant leaf disease detection and identification was developed [15]. In this research paper the application of k-mean clustering and neural networks have been formulated for clustering and classification of disease that affect plant leaves. In this work images are segmented in order to isolate the necessary component of the leaf. After the segmentation of the image they added two main steps. The first steps they identity the mostly green colored pixels. Next these pixels are masked based on specific threshold values that are computed using Otsu's method, than those mostly green pixels are masked. The other steps are the pixels with zero red, green and blue values and the pixels on the boundaries of the infected cluster (objects) were completely removed. After this color and texture features are extracted by using the color Co-occurrence Method (CCM) (both color and texture features are selected). Finally, based on the selected feature values (color and texture) the classification and detection of the disease was done by using artificial neural network. One of the basic steps for the system is the identification of green colored pixels and masking them to identify the region of interest. However, identification of plant leaf diseases rather than green colored plants is not possible. The system also leads to a miss- classification for normal plants rather than green color.

Panagiotts Tzional and stelios E.Papadakis also developed a research to classify plant leaves based on morphological features and a fuzzy surface selection technique [16]. In this research several morphological and geometrical features that represent the leaves are extracted. Then the necessary features are selected by using a fuzzy surface approach.

Based on this selected features a neural network classifier was designed for classification of the leaf.

Plant leaf disease spot identification is also done by using different color model for feature extractions [18]. Images are color transformed from RGB image to one of the color space named by YCbCr, HIS and CIELAB color space. The color transformed images are passed through median filter to remove unnecessary spots. In last step Otsu threshold is applied on RGB image, 'A' component of CIELAB color space, 'H' component of HIS color space and 'Cr' component of YCbCr color space is used to detect the disease spot.

In general, most of the researches done in plant leaf disease identification and detection are color dependent. However the identification of flower disease based on color is difficult. This is due to the fact that, there are diseases that are caused by bacteria and fungi and changes the color of one flower to other colored flower. In addition to this, a flower at different stage has different color, thus selecting color features for flower disease identification is not suitable.

3.3. Fruit/Food Grading

Bredon J. Woodford , Nikola k –kasabov and c – Howared wearing in paper titled ''Fruit image analysis using Wavelets'' [14] proposed wavelet based image processing technique and neural network to develop a method of a line identification of pest damage pip fruit in orchards. There pests that are prevalent in orchards were selected as the candidates for this research; the leaf roller, coding moth, and apple leaf curling midge. Fast wavelet transform with special set of Doubenchies wavelet was used to extract the important features. To retrieve the related images, the search is done in two steps. The first steps match the image by comparing the standard deviations for the three color components. In the second steps, a weighted version of the Euclidean distance between the feature coefficient of an images selected in the first step and those of the querying image is calculated and the images with the smallest distance are selected and as matching images to the query.

Image processing concept for grading of bakery products, fruits, vegetables and grains were considered in [33]. Fruits characterized by color, size and shape, its condition in pre and post harvesting, damages were attributes for grading. Vegetable specifically roots, tomatoes, mushrooms were also compared with its attributes for grading purpose. Socio economic limitations were also discussed. Similar kinds of approaches for grading of grain, vegetables, were reviewed by other researchers [34]. The methods or techniques in image processing such as image segmentation, shape analysis and morphology, texture analysis, noise elimination, 3D vision, invariance, pattern recognition and image modality were applied for grading these categories.

Identification and classification of normal and infected agricultural/horticultural produce based on combined color and texture feature also designed by other researchers [17]. In this research color feathers are extracted by using HIS (Hue, Intensity, and Saturation) from the RGB(Red, Green, and Blue) image. In order to extract texture features, color co-occurrence matrix was used. Finally a multilayered back propagation neural network (BPNN) was used as a classifier of different produce.

Corn verity identification on the basis of the color, shape and geometric features using discrimination analysis and neural network was proposed in [32]. To avoid illumination and manmade disturbances image were captured with flatbed scanner. Features were extracted using morphological feature and color features. Morphological feature analysis was used for classification. Mean and standard deviation of this color component calculated to extract 28 color features for identification. To reduce computational burden and to enhance the performance of classification stepwise discrimination analysis was used. Mahalanobis distance method with back propagation neural network was used to train and classify the corns. Since the researcher uses the shape and color feature of an image it will make it difficult to identify flower diseases at different stage.

3.4. Weed Detection

Weed infestation rate (WIR) in synthetic images using Gabor filtering and results of wavelets were compared in [36]. Daubechies, Symlet, Coiflet, Meyer, Biorthogonal, R-biorthogonal wavelets were used for image decomposition and reconstruction. Results of these and global confusion matrix were used to classify crop and weed into true and false categories. Daubechies 25 wavelet and the Meyer wavelets give better result than Gabor filtering in the cost of average time in both syntactic and real images.

To classify weeds in to broad and narrow leaves erosion and dilation segmentation was used [37]. The RGB images was decomposed in to R, G, B components and this components are converted in to binary image which discriminates bright pixel as weed and dark as background. In order to eliminate the irrelevant details from the image erosion by structured element was used and then dilation was applied, for storing the result in the form of tables. Success rate of classification was remarkable with less algorithm execution time.

Real-time image processing for crop/weed discrimination in maize filed also presented in other research paper [42]. The system consists of two independent subsystems. A fast image processing delivering results in real-time (Fast Image Processing, FIP). The other subsystem is a slower component that is used to correct the first subsystem's mistakes'. They have used an Otsu tresholding technique for segmentation of an image and the system was tested on real-time images and the experimental result shows that, the system detects an average of 95% of weed and 80% of crops under different illumination conditions.

A digital image analysis technique based on morphological and color features was also developed to classify different verities of Ethiopian coffee based on their growing region [22]. For their research, sample coffees were taken from six coffee growing regions (Bale, Harar, Jimma, Limu, Sidamo and Welega) which are popular and widely planted in Ethiopia. For the classification analysis, ten morphological and six color features were extracted from each coffee bean images. According to the researcher, the processing type of the coffee (washed or unwashed) has been also defined during the analysis. After this the researcher compared classification approaches of Naïve Bayes and Neural Network classifiers on each classification parameters of morphology, color and the combination of the two. To evaluate the classification accuracy, from the total of 4844 data sets 80% were used for training and the remaining 20% was used for testing. Finally, It was found that the classification performance of neural networks classifier was better than Naïve Bayes classifier. And the overall classification accuracy was 77.4%. However, in our research work we will use the texture features of flower image to identify the normal and diseased flower image. In addition to this, identifying flower disease using color features is not possible, due to the fact that one disease can change the color of one flower to other colored flower.

Table 3. 1 Summary of related works

Category	Research Title	Algorithm	Authors	Limitation
Plant leaf disease identification	Fast and Accurate Detection and classification of Plant Disease	K-mean clastering and Neural Network	Al.Hiary and Bani-Ahmad [15]	✓ Color dependent ✓ Stage dependent
	Plant leaves classification based on morphological features and a fuzzy surface selection technique	Fuzzy surface selection and Neural Network	Panagiotis Tzionas and Stelios E.Papadakis [16]	✓ Stage dependent ✓ Shape feature dependent
	Color Transform Based Approach for Disease spot Detection on Plant Leaf	Color Transform Approach	Piyush Chaudhary and Anand K. Chaudhari [18]	✓ Color dependent
Fruit/Food Grading	Fruit Image Analysis Using Wavelets	Doubenchies wavlet and Euclidian distance	Brendom Woodford and Nikola Kasabov [14]	✓ Shape dependent ✓ Color dependent
	Identification and classification of Normal and affected Agricultural/horticultural products	Color Co-occurrence Method and ANN as a classifier	Basvaraj Anami and J.D Pujari [17]	✓ Color dependent ✓ Shape dependent
	Detection of fungus infected corn kernels using near infrared reflectance spectroscopy and color imaging	Linear discrimination analysis and ANN	J.G. Tallada and D.T. Wicklow [32]	✓ Color dependent ✓ Shape dependent
Weed Detection	Wavelet transform to discriminate between crop and weed in perspective agronomic images	Wavelet Transform	J.Bossua and Ch.Geea [36]	✓ It is done for the identification of weed in certain growing stage ✓ Stage dependent
	Real-time image processing for crop/weed discrimination	Point matching	Xavier Artizzu Angela Ribeiro [42]	✓ used video data ✓ used a simple matching technique
	Image analysis for Ethiopian coffee classification	CCM , Support vector machine and ANN	Habtamu Minassie [22]	✓ Color and shape dependent

Generally, all the above researches are to detect plant leaf that has green color, to identify weeds, and to classify agricultural products. If we use them to identify diseases that appeared in multi-colored flowers and that change the color of one flower to other flower, it is insufficient. It is also difficult to assign a single color for one flower species, because flowers at different stage have different color and shape.

CHAPTER FOUR: DESIGN AND IMPLEMENTATION OF FLOWER DISEASE IDENTIFICATION

4.1. Introduction

Identification of an object and classifying it in to its appropriate classes becomes a wide area of research. The identification of an object passes through a series of steps/procedures that will be applied to differentiated items, in which each new item must be categorized in to one of a predefined classes based on the basis of observed attribute or features. In this chapter, we will explain the design and system architecture of flower disease identification.

4.2. System Architecture

The overall concept that is the framework for any vision related algorithm for image identification is almost the same. First, the digital images are acquired from the environment using a digital camera. Then image pre-processing techniques are applied to the acquired images to extract useful features that are necessary for further analysis. After this, several analytical discriminating techniques are used to classify the images according to the specific problem at hand. Figure 4.1 depicts the basic procedure of the proposed vision-based identification of flower disease. Thus, the identification of flower disease starts with acquiring images of flowers using a digital camera. In order to remove noises that occur during the image acquisition step, we have applied an image processing techniques like image enhancement and segmentation. Then, features that are best suited to represent the image are extracted from the image using an image analysis technique. Based on the extracted features the training and testing data that are used to identify are extracted. Finally, appropriate machine learning pattern identifier is selected to classify an image in to its class of disease. Figure 4.1 shows the architecture of the proposed vision based identification of flower disease.

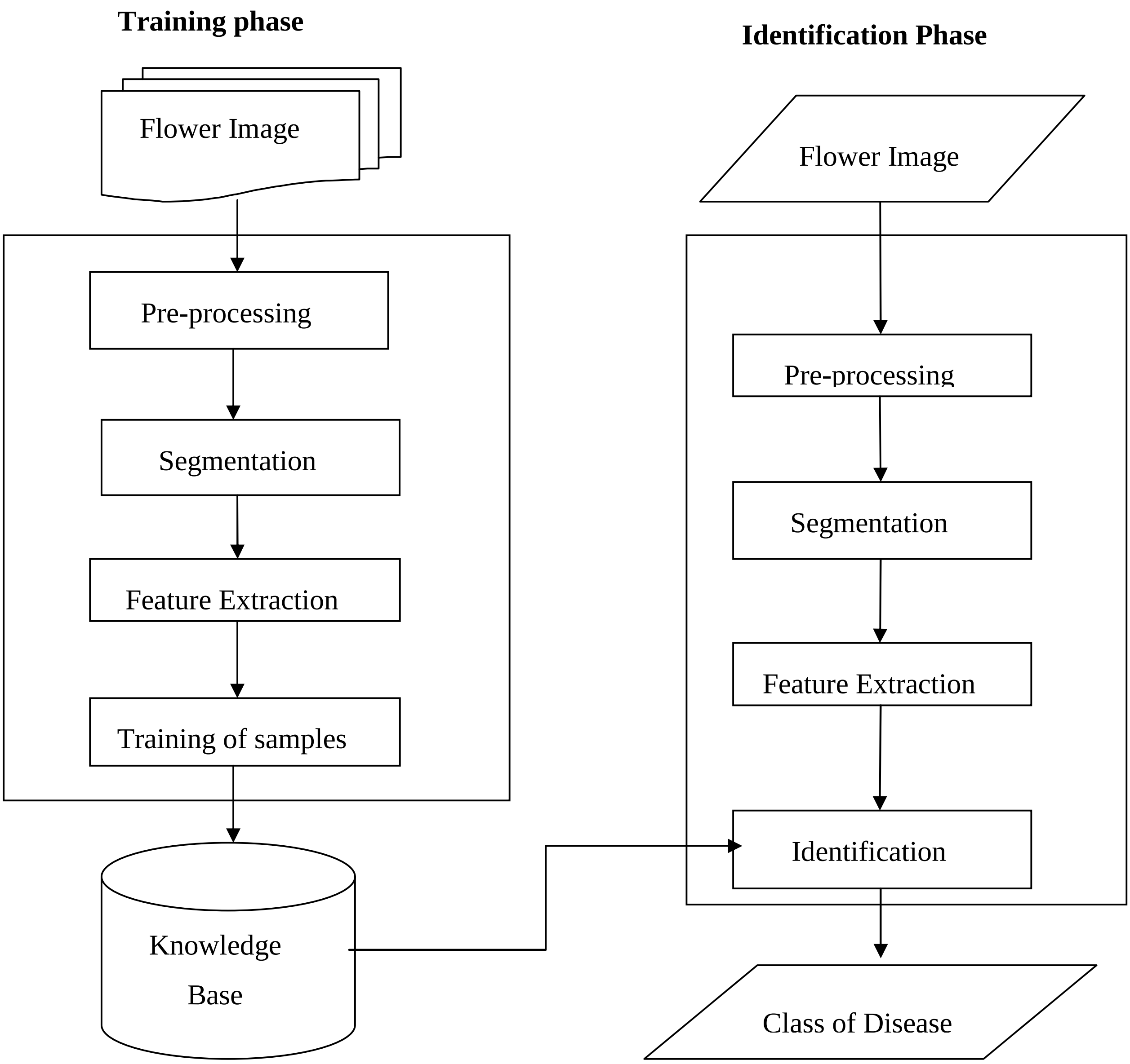


Figure4. 1: Proposed System Architecture

The first part of the proposed system architecture is the training phase. In this phase, flower images are captured using a digital camera. Then, those flower images are pre-processed using an image processing techniques. After preprocessing a useful image features are extracted using feature extraction technique and then an artificial neural network is trained using the extracted features to create the knowledge base. Each step of the training phase is explained as follows.

original image

Result of image after pre-processing

$$\sigma_{\omega}^{2}(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \qquad (4.1)$$

Weights ω_i are the probabilities of the two classes separated by a threshold t and σ_i^2 variance of the classes. Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance as it is shown in equation 4.2:

$$\sigma_b^2\ (t) = \sigma^2 - \sigma_{\omega}^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \qquad (4.2)$$

The class probability $\omega_1(t)$ is computed from the histogram as t:

$$\omega_1(t) = \sum_0^t p(i) \qquad (4.3)$$

While the class mean $\mu_1(t)$ is:

$$\mu_1(t) = \frac{\sum_0^t p(i)x(i)]}{\omega_1} \qquad (4.4)$$

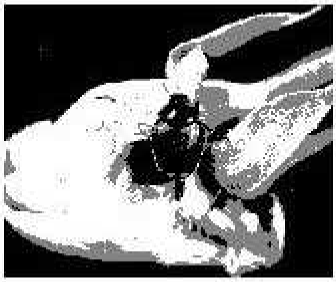
The class probabilities and class means can be computed iteratively. This idea yields an effective algorithm

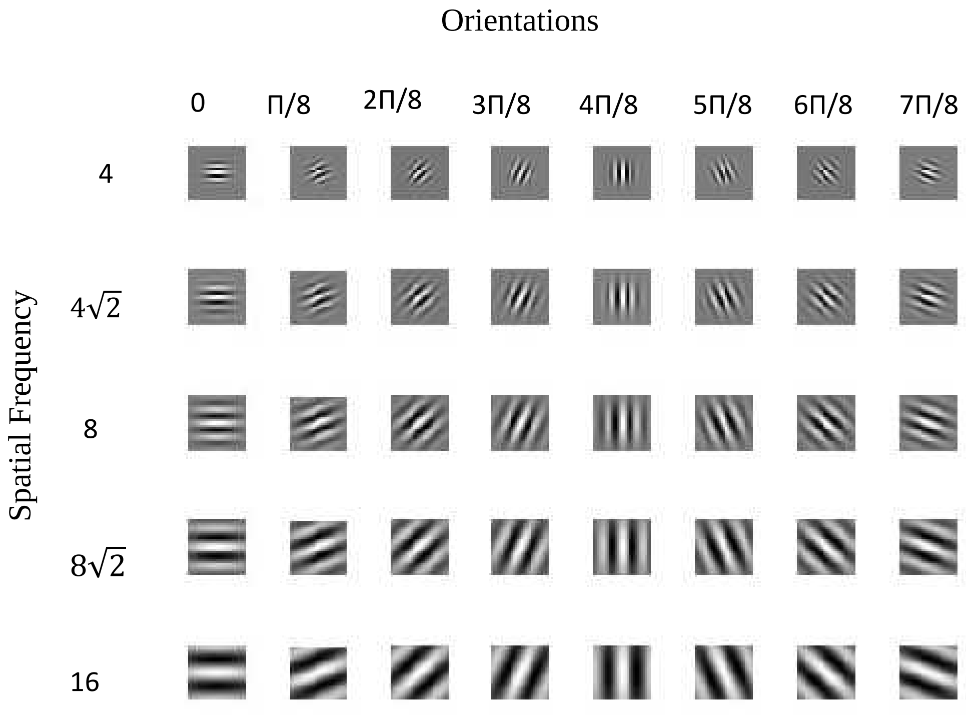
1. Compute histogram and probabilities of each intensity level
2. Set up initial $\omega_1(0)$ and $\mu_1(0)$
3. Step through all possible threshold *t=1*….maximum intensity
   a) Update ω_1 and μ_1
   b) Compute $\sigma_b^2\ (t)$
4. Desired threshold corresponds to the maximum $\sigma_b^2\ (t)$
5. You can compute two maximum (and two corresponding thresholds). $\sigma_{b1}^2\ (t)$ is the greater max and $\sigma_{b2}^2\ (t)$ is the greater or equal maximum
6. Desired threshold = $\frac{threshold1 + threshold2}{2}$

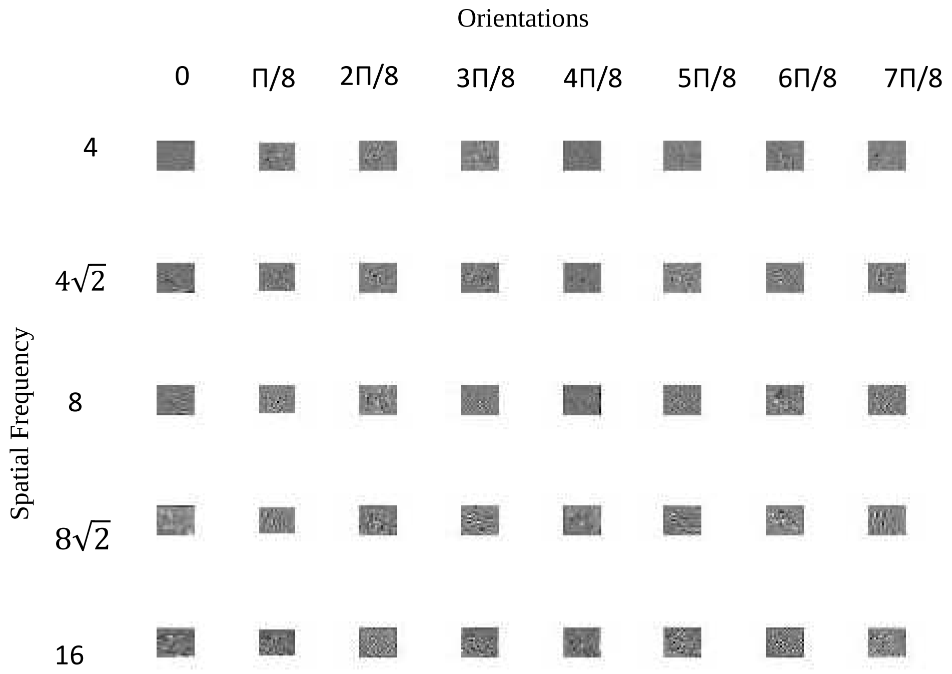
Segmentation results for a single flower image attacked by Japaese-beetles are shown in Figure 4.3.

Original image

segmented image




Orientations
0
Π/8
2Π/8
3Π/8
4Π/8
5Π/8
6Π/8
7Π/8
Spatial Frequency
4
4√2
8
8√2
16



re ca

$\bar{x}$ of lata set ca be

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Square of Deviations: The Square of deviation *DEVSQ* is computed by using equation 4.6:

$$DEVSQ(x_1 \dots x_n) = \sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{4. 6}$$

Absolute deviation: The absolute deviation *Adev* is computed by using equation 4.7

$$Adev(x_1 \dots x_n) = \frac{1}{n}\sum_{i=1}^{n}|x_i - \overline{x}| \tag{4. 7}$$

Variance: The variance *var* is computed by using equation 4.8

$$var(x_1 \dots x_n) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{4. 8}$$

Standard Deviation: the standard deviation *Sdv* is calculated as the square root of variance by using equation 4.19:

$$Sdv(x_1 \dots . x_n) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{4.9}$$

Skeweness: characterizes the degree of asymmetry of a distribution around its mean. It characterizes the shape of the distributions and calculated by using equation 4.10:

$$skew(x_1 \dots x_n) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i - \overline{x}}{\sigma}\right]^3 \tag{4.10}$$

Kurtosis: it is non-dimensional quantities that are used to measure the relative peakedness or flatness of a distribution. It can be calculated by using equation 4.11:

$$kurt(x_1 \dots x_n) = \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i - \overline{x}}{\sigma}\right]^4\right\} - 3 \tag{4. 5}$$

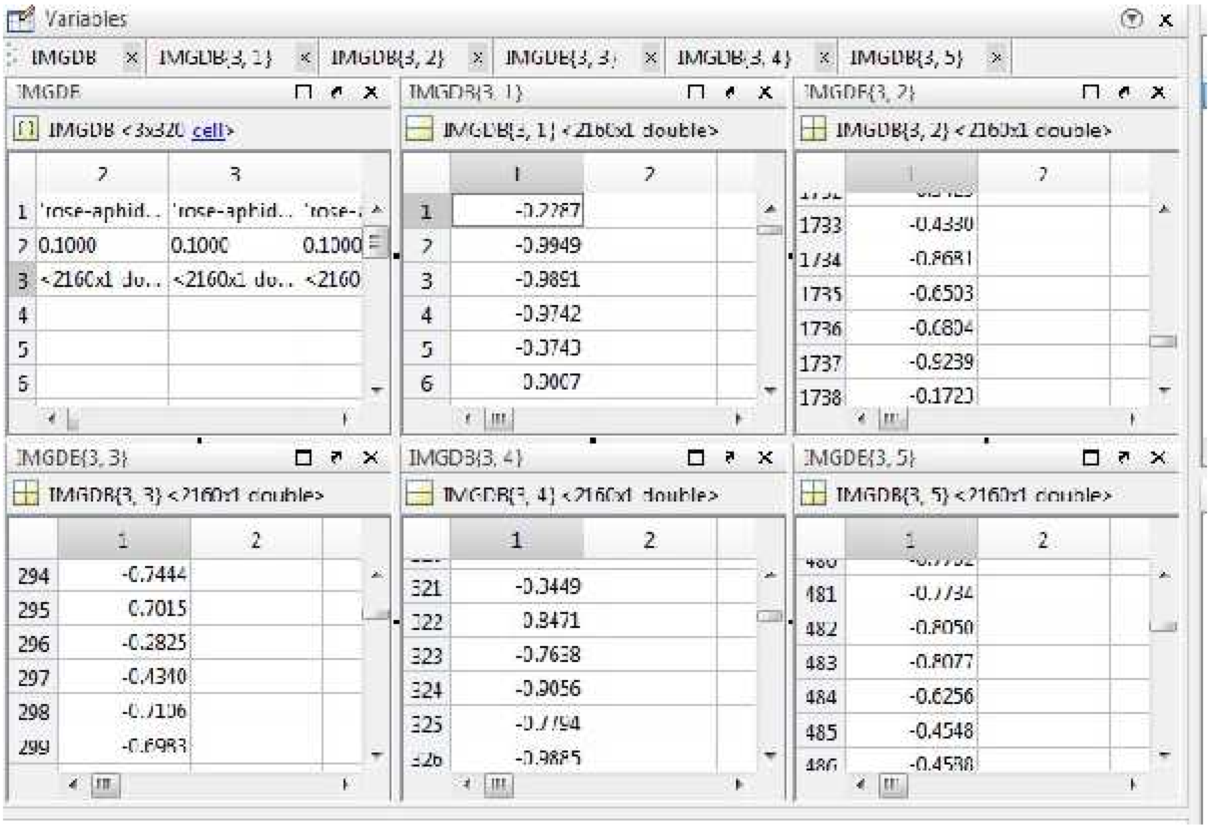
Class of disease   Input Vector   OutputVector

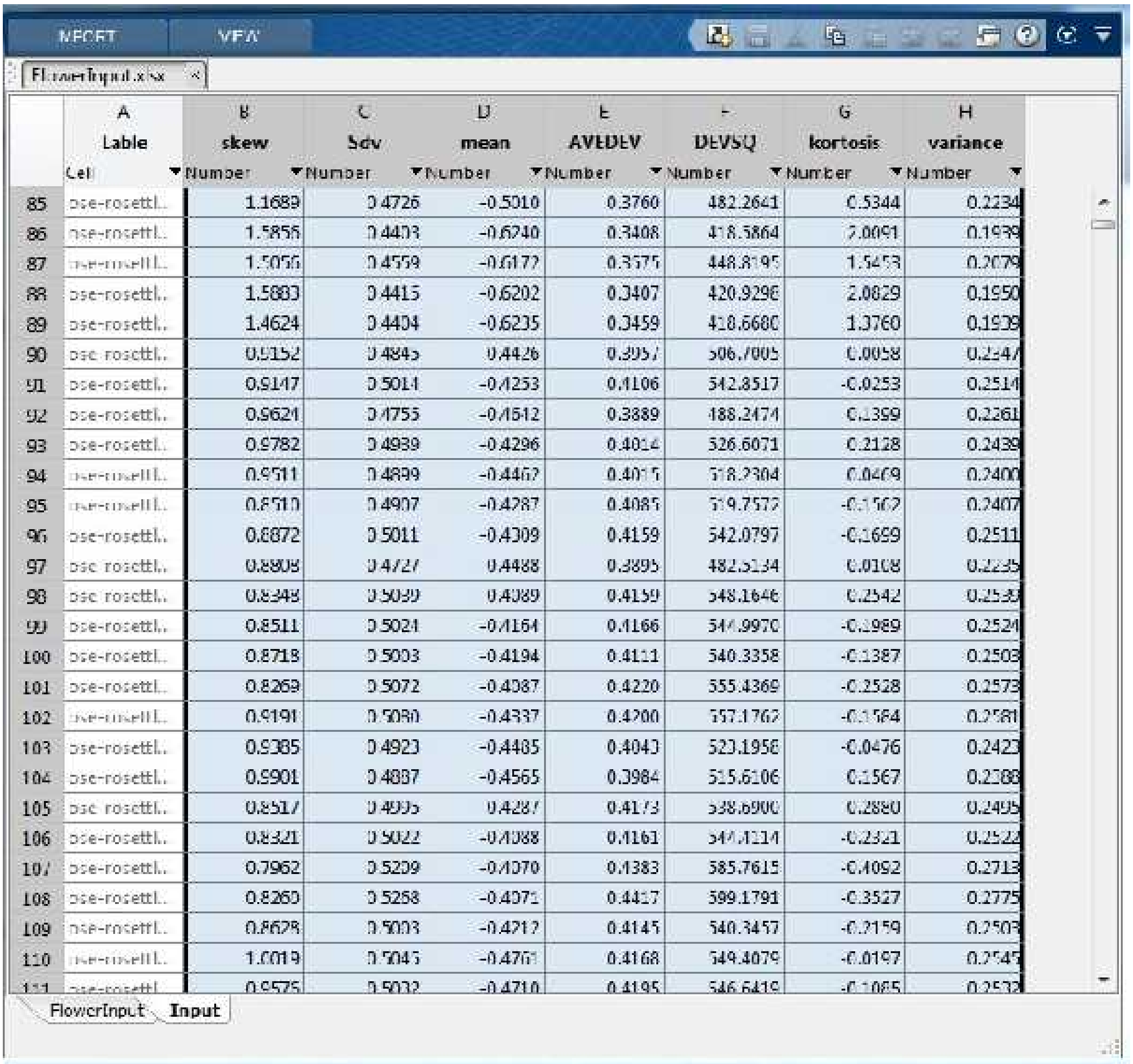
$$\begin{bmatrix} \text{Rose-aphid} \\ \text{Rose –japanese beetles} \\ \text{Rose-rosettle} \\ \text{Rose-goldenrod-soldier} \\ \text{Rose-mossyrose gall wasp} \\ \text{Rose-normal} \\ \text{Rose-rose rust} \\ \text{Rose –soldier beetles} \end{bmatrix} \rightarrow \begin{bmatrix} \text{Skew} \\ \text{Sdv} \\ \text{Mean} \\ \text{ADEV} \\ \text{DEVSQ} \\ \text{Kurtosis} \\ \text{Variance} \end{bmatrix} = \begin{bmatrix} 10000000 \\ 01000000 \\ 00100000 \\ 00010000 \\ 00001000 \\ 00000100 \\ 00000010 \\ 00000001 \end{bmatrix}$$

Figure 4.7: Vector representation of input and output features

As it is described in Figure 4.7 we have seven different inputs and 8 different output classes. The output vectors are represented by using the binary number 1 and 0. Since we have eight different groups of flower image, we have used eight bit binary number. Each bit refers whether that feature belongs to a group represented at that bit position. As shown in Figure 4.7, rose-aphid, rose-japanese beetles, rose-rosettle, rose-goldenrose, rose-mossyrose gall wasp, rose-normal, and rose-soldier beetles are represented by first, second, third, fourth, fifth, sixth, seventh and eight bits respectively. The bit value indicates whether the feature data set is categorized under that group or not. If the value is one, the feature data set is the member of that group if it is not then the value will be zero.

As described in Section 4.5, we have extracted the texture features of an image using Gabor feature extraction and from this texture features, we have calculated seven different measures of central tendency and dispersion of the extracted Gabor texture features. Texture features of an image that are extracted by using Gabor feature extraction are sown in the Figure 4.8.


Variables
IMGDB
IMGDB{3, 1}
IMGDB{3, 2}
IMGDB{3, 3}
IMGDB{3, 4}
IMGDB{3, 5}
IMGDB <3x320 cell>
'rose-aphid... 'rose-aphid... 'rose-...
0.1000 0.1000 0.1000
<2160x1 do... <2160x1 do... <2160
IMGDB{3, 1} <2160x1 double>
1 -0.2287
2 -0.9949
3 -0.9891
4 -0.9742
5 -0.0743
6 0.0007
IMGDB{3, 2} <2160x1 double>
1733 -0.4330
1734 -0.8681
1735 -0.6503
1736 -0.6804
1737 -0.9239
1738 -0.1720
IMGDB{3, 3} <2160x1 double>
294 -0.7444
295 0.7015
296 -0.2825
297 -0.4340
298 -0.7106
299 -0.6983
IMGDB{3, 4} <2160x1 double>
321 -0.0449
322 0.8471
323 -0.7638
324 -0.9056
325 -0.7794
326 -0.9885
IMGDB{3, 5} <2160x1 double>
481 -0.7734
482 -0.8050
483 -0.8077
484 -0.6256
485 -0.4548
486 -0.4538



IMPORT VIEW

FlowerInput.xlsx

	A	B	C	D	E	F	G	H
	Lable	skew	Sdv	mean	AVEDEV	DEVSQ	kortosis	variance
	Cell	Number	Number	Number	Number	Number	Number	Number
85	ose-rosettl...	1.1689	0.4726	-0.5010	0.3760	482.2641	0.5344	0.2234
86	ose-rosettl...	1.5855	0.4403	-0.6240	0.3408	418.5864	2.0091	0.1939
87	ose-rosettl...	1.5056	0.4559	-0.6172	0.3575	448.8195	1.5453	0.2079
88	ose-rosettl...	1.5883	0.4415	-0.6202	0.3407	420.9298	2.0829	0.1950
89	ose-rosettl...	1.4624	0.4404	-0.6235	0.3459	418.6680	1.3760	0.1939
90	ose-rosettl...	0.9152	0.4845	0.4426	0.3957	506.7005	0.0058	0.2347
91	ose-rosettl...	0.9147	0.5014	-0.4253	0.4106	542.8517	-0.0253	0.2514
92	ose-rosettl...	0.9624	0.4755	-0.4512	0.3889	488.2474	0.1399	0.2261
93	ose-rosettl...	0.9782	0.4939	-0.4296	0.4014	526.6071	0.2128	0.2439
94	ose-rosettl...	0.9511	0.4899	-0.4467	0.4015	518.2304	0.0469	0.2400
95	ose-rosettl...	0.8510	0.4907	-0.4287	0.4085	519.7572	-0.1562	0.2407
96	ose-rosettl...	0.8872	0.5011	-0.4309	0.4159	542.0797	-0.1699	0.2511
97	ose-rosettl...	0.8808	0.4727	0.4488	0.3895	482.5134	0.0108	0.2235
98	ose-rosettl...	0.8348	0.5039	0.4089	0.4159	548.1646	0.2542	0.2530
99	ose-rosettl...	0.8511	0.5024	-0.4164	0.4166	544.9970	-0.2989	0.2524
100	ose-rosettl...	0.8718	0.5003	-0.4194	0.4111	540.3358	-0.1387	0.2503
101	ose-rosettl...	0.8269	0.5072	-0.4087	0.4220	555.4369	-0.2528	0.2573
102	ose-rosettl...	0.9191	0.5080	-0.4337	0.4200	557.1762	-0.1584	0.2581
103	ose-rosettl...	0.9385	0.4923	-0.4485	0.4040	523.1958	-0.0476	0.2423
104	ose-rosettl...	0.9901	0.4887	-0.4565	0.3984	515.6106	0.1567	0.2388
105	ose-rosettl...	0.8517	0.4995	0.4287	0.4173	538.6900	0.2880	0.2495
106	ose-rosettl...	0.8321	0.5022	-0.4088	0.4161	544.4114	-0.2321	0.2522
107	ose-rosettl...	0.7962	0.5209	-0.4070	0.4383	585.7615	-0.4092	0.2713
108	ose-rosettl...	0.8260	0.5268	-0.4071	0.4417	599.1791	-0.3527	0.2775
109	ose-rosettl...	0.8628	0.5003	-0.4212	0.4145	540.3457	-0.2159	0.2503
110	ose-rosettl...	1.0019	0.5043	-0.4761	0.4168	549.4079	-0.0197	0.2545
111	ose-rosettl...	0.9575	0.5032	-0.4710	0.4195	546.6419	-0.1065	0.2532

FlowerInput Input

Since we have a labeled data, we use a supervised machine learning method to identify the class of the image; we clearly stated the number of input and outputs. For the identification of the image we have selected a neural network that is able to generalize among images, that is adaptable so that it can easily learn to recognize new images. We have chosen neural networks for this task because of their favorable properties that make them an excellent choice for object classification. The most important of these properties are generalization, expandability, representing multiple samples, and memory saving.

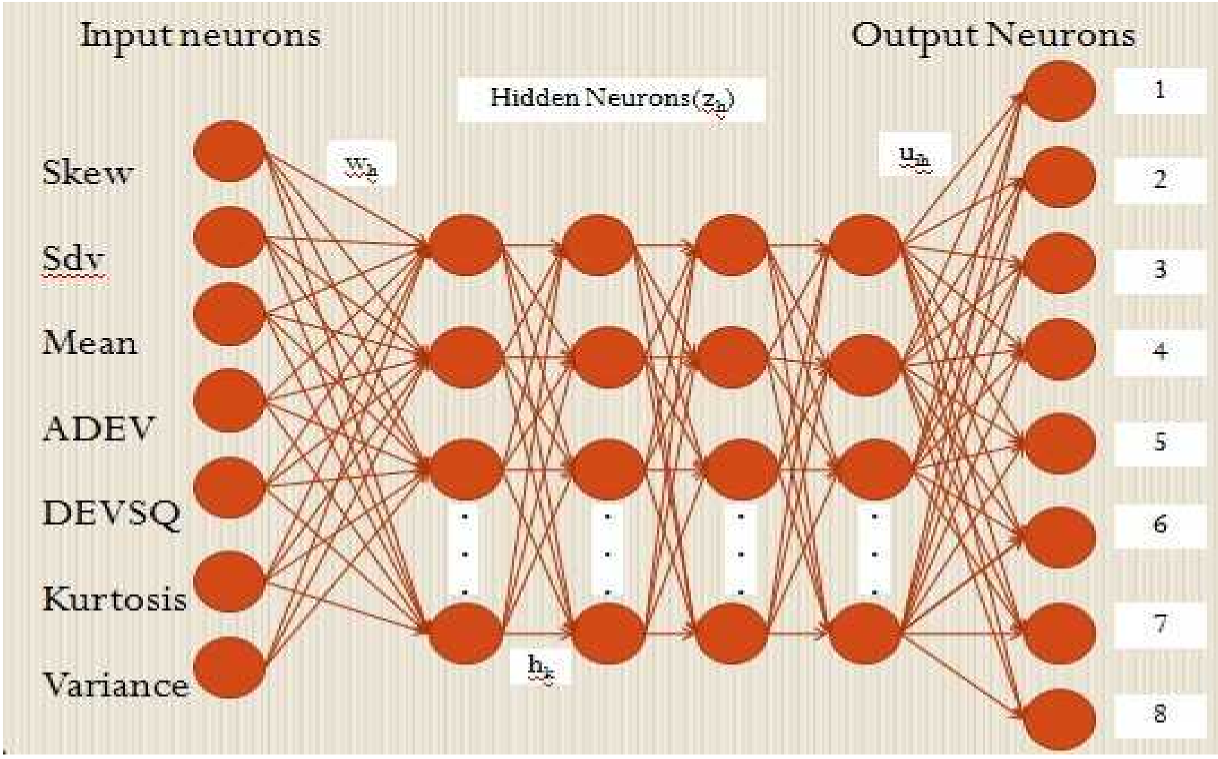
Generalization: Small distortion can be handled easily, a necessity for object identification and recognition.

Expandability: Another great benefit of artificial neural network is that, it can easily be expandable. In order for it to learn new images there is no need to start all over and redefine distance measures and distributions.

Representing multiple samples:A class image can be easily represented by multiple samples images under multiple conditions. Because a neural network incorporates in its structure what it learns, the recognition/identification of an image becomes a single step. In this way there is no need for multiple comparisons as in many conventional systems. The network determines in one single step to what class the object belongs.

Memory saving: An advantage stemming from the previously mentioned characteristics is that there is no need to store the standard images to be saved for comparison. Once a network is trained properly it contains the necessary information and the image data becomes expandable and can be removed from memory.

In our study, feed forward multi-layer perceptron architecture is used. In this architecture the output from one layer of neurons feed forward into the next layer of neurons. There are never any backward connections, and connections never skip a layer [38, 39]. Typically the layers are fully connected, meaning that all units at one layer are connected with all units at the next layer as it is shown in Figure 4.10. Feed-forward multilayer neural network consists of a layer of input units, one or more of hidden units, and one output layer of units.



s i accordin g

$$v \qquad {}_{h=0}^{H} u_{ih} z_h$$

Hidden Layer Unit: Layers other than input and output layers are called hidden layers. The connection coming out of an input unit have weighted associated with them. A weight going to hidden unit z_h from input unit x_j would be labeled w_{hj}. The bias input node, x_0, is connected to all the hidden units, with weights w_{h0}. In the training, these bias weights, w_{h0} are treated like all other weights, and are updated according to the back propagation algorithm. We have different techniques to fix the number of hidden neurons. In review of methods to fix the number of hidden neurons in neural network the following method has a better result and we have used it to decide the number of hidden layers by using equation 4.13[38].

$$z_h = \frac{4d^2+3}{d^2-8} \tag{4. 7}$$

Where z_h is number of hidden layers and d is number of input layer

Each hidden node calculates the weighted sum of its input and applies a thresholding function to determine the output of the hidden node. The weighted sum of the inputs for hidden node is calculated using equation 4.14:

$$k = \sum_{j=0}^{i} w_{hj}x_j \tag{4. 8}$$

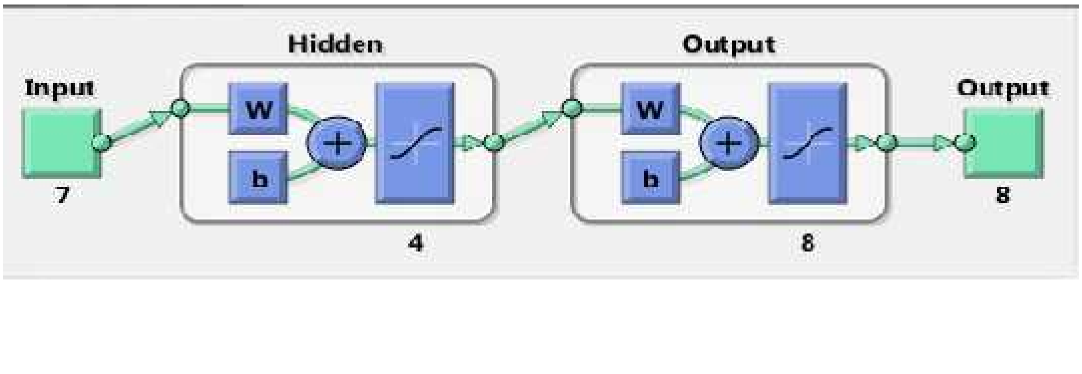
Where k is the weighted sum of the input for hidden node *h*

Learning in ANN indicates the methods used to determine the adaptation of weights between the connections of two neurons. A MLP feed-forward network uses back propagation learning algorithm for adapting its weights to a sequence of training samples during a learning phase.

The general idea with the back propagation algorithm is to use a gradient decent to update the weights so as to minimize the squared error between the network output values and the target output values. The update rule derived by talking the partial derivative of the error function with respect to the weights to determine each weight's contribution to the error. Then, each weight is adjusted, using gradient decent, according to its contribution to the error. This process occurs iteratively for each layer of the network, starting with the set of weights, and working back towards the input layer. A back-propagation learning algorithm involves two phases: forward and backward uation 4.

$$e_{i=d_i-y_i}$$

$$F\ x\quad \frac{1}{1+e^{-x}}$$



4.7. Identification

As we explained in the previous section, in the first phase flower images are used to create the knowledge base and this knowledge base are used for identification purpose. In the identification phase, flower images that are different from the samples that we have used for training are captured. This flower images contains more than forty eight samples of eight different classes of flower images. Then, this flower images are pre-processed using an image pre-processing technique that are used in the training phase. In addition to image pre-processing those flower images are segmented using Otsu's method of image segmentation to identify the region of interest. After we identify the region of interest, useful features are extracted in order to reduce the computational cost of the system. Since we have used texture features of an image for creation of the knowledge base, here also we have extracted the texture features of the image using Gabor texture feature extraction technique. From each flower image, we have extracted seven different features of the image. After we get the necessary texture features of the image, we use the network that is built during the training phase to identify the image in to its class of disease.

In summary, in this work an image processing based system for automatic identification of flower disease is proposed and implemented. The proposed system consists of two main phases. In the first phase normal and diseased flower image are used to create a knowledge base. During the creation of the knowledge base, images are pre-processed and segmented to identify the region of interest. Then, seven different texture features of images are extracted using Gabor texture feature extraction. Finally, an artificial neural network is trained using seven input features extracted from the individual image and eight output vectors that represent eight different classes of disease to represent the knowledge base. In the second phase, the knowledge base is used to identify the disease of a flower.

CAHPTER FIVE: EXPERIMENT

5.1. Introduction

In image processing, the first step is the acquisition of an image. During the acquisition of the image, we have to consider different environmental factors like lightening, camera resolution in order to minimize the tasks in image pre-processing. The geometry of the viewing situation, i.e., the relative position of the sources and camera with respect to the objects of interest, usually also has a major impact on the contrast between the object and their background. Thus, as described in Section 2.4.1 image acquisition is the critical step in machine vision system.

5.2. Dataset

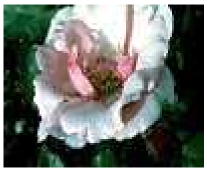
In this study images are taken from two flower producers (Herburg Roses and Joe Flowers) and from other image repository. During the image acquisition both normal and diseased flower images are taken. Seven disease of flower (especially rose) namely, Japanese beetle (popillia japonica), goldenrod solider beetle (Chauliognathus pensylvanicus), soldier beetles (chauliognathu sspp), rose rosettle (Emaravirus RRD), mossyrose gall wasp (Diplolepis rosae), rose rusts (Phragmidium spp), and rose aphid (Macrosiphum rosae), are selected. We have selected those diseases because the diseases occurred in Ethiopian floriculture industry most frequently.

In both places the images are taken in certain parameters. In order to take the image Sony digital camera was used to capture the images with the following parameter.

- The distance between the images and the camera was approximated to 30 cm
- For all images, the pictures are taken after the sunset.
- All the images are in JPEG (Joint Photographer Expert Group) file format
- All the captured images were taken in the same controlled environment in-order to avoid external effects of sunlight and other environmental conditions,
- All images have a size of 81x100 pixels

Rose-goldenrod solider beetle

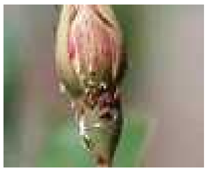
Rose-normal



Rose-soldier beetles

Rose-japanese beetle

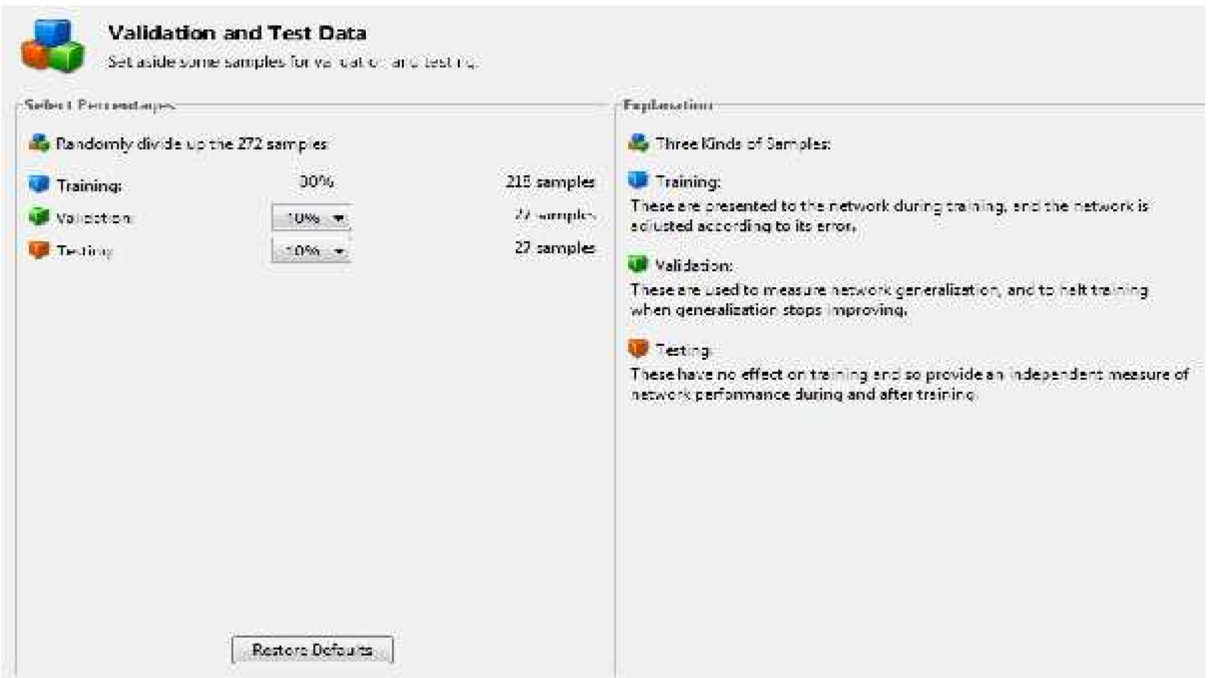
Rose-rust disease

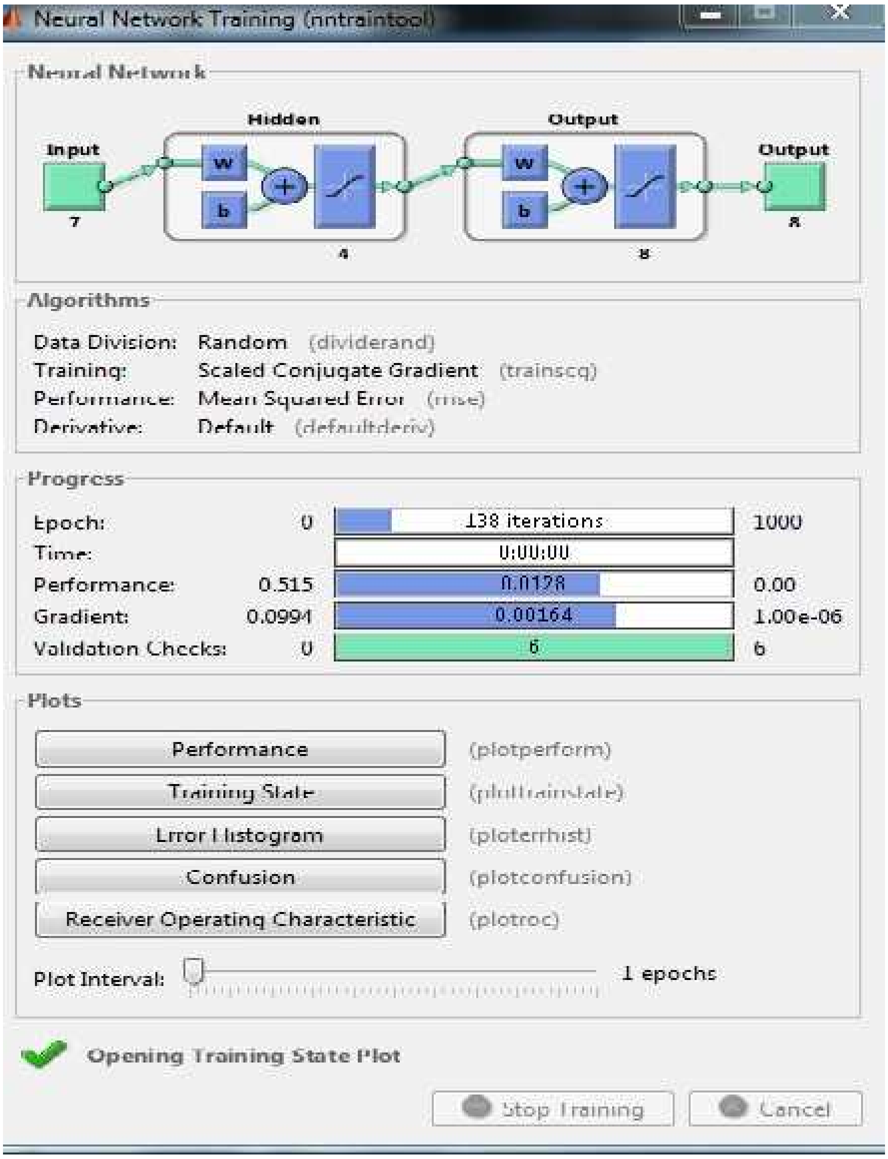


Rose-aphid

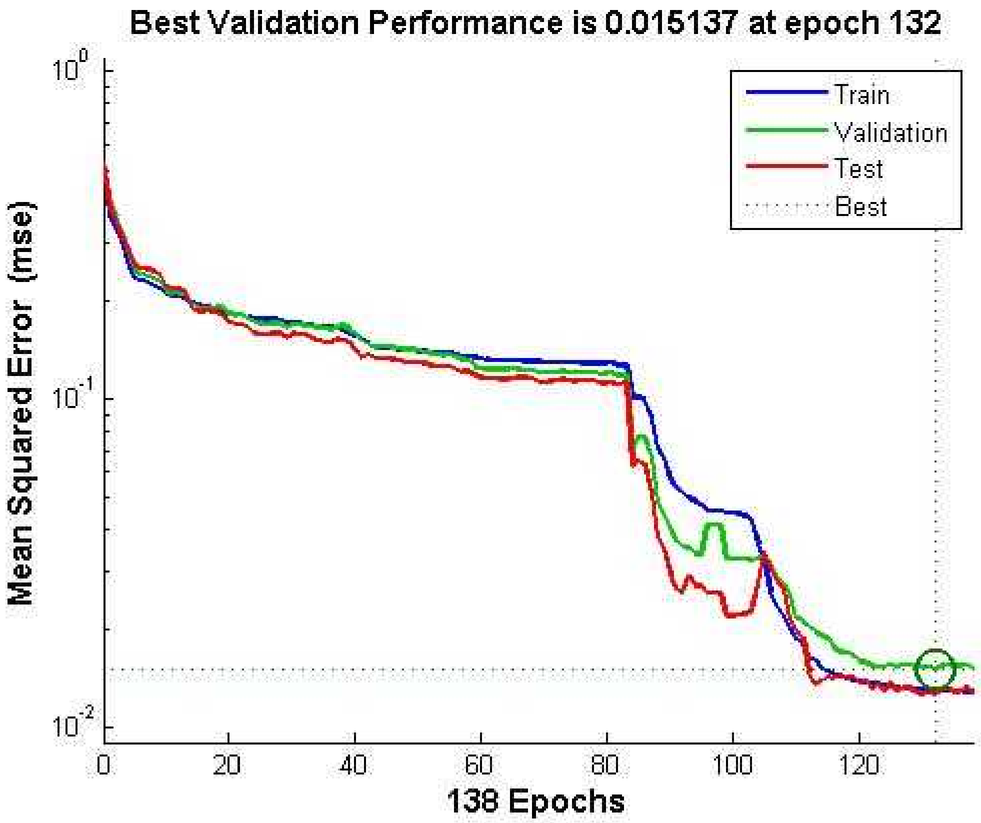
Rose-mossyrose gall wasp

Rose-rosettle disease

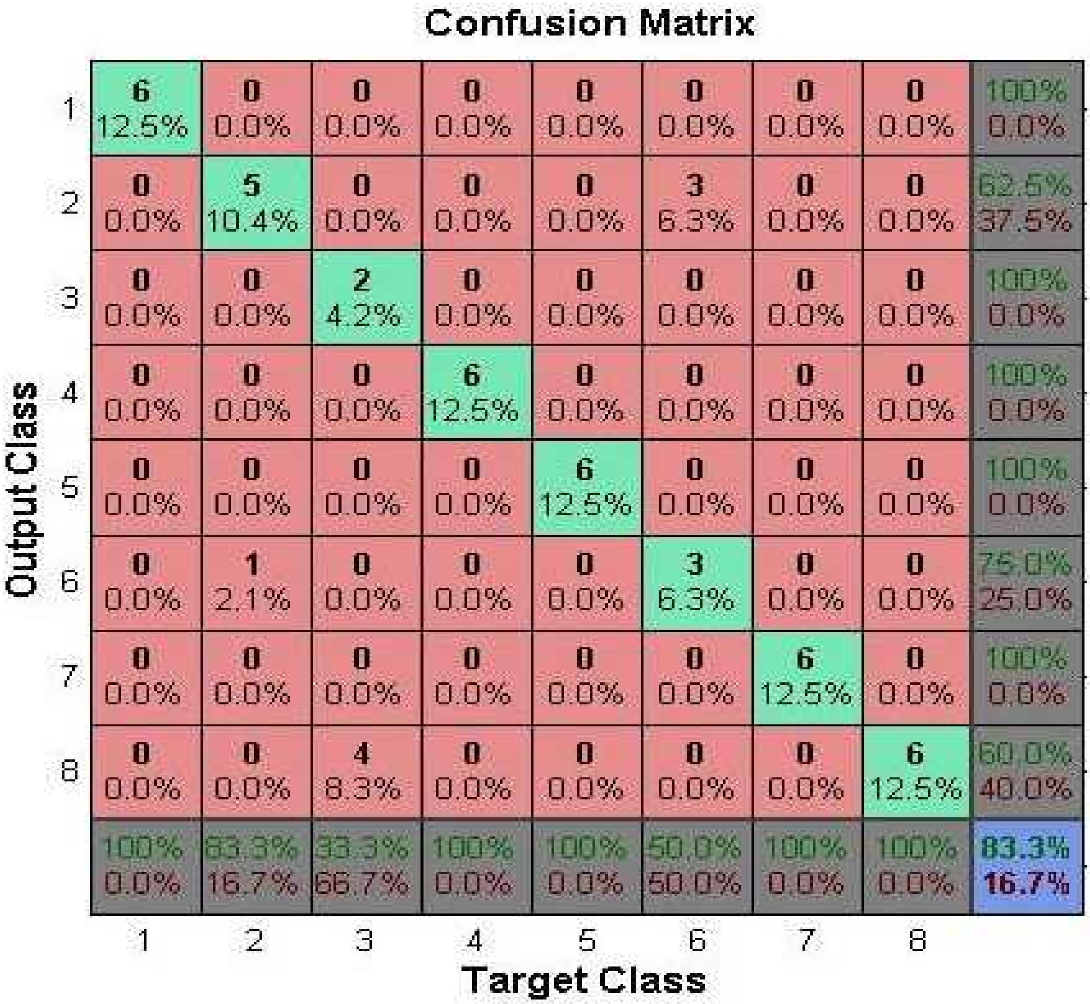

Validation and Test Data
Randomly divide up the 272 samples:
Training:
22 samples
Restore Defaults
Three Kinds of Samples:
Training:
These are presented to the network during training, and the network is adjusted according to its error.
Validation:
These are used to measure network generalization, and to halt training when generalization stops improving.
Testing:
These have no effect on training and so provide an independent measure of network performance during and after training.



Neural Network Training (nntraintool)
Neural Network
Hidden
Output
Input
Output
7
4
8
8
Algorithms
Data Division: Random (dividerand)
Training: Scaled Conjugate Gradient (trainscg)
Performance: Mean Squared Error (mse)
Derivative: Default (defaultderiv)
Progress
Epoch: 0 138 iterations 1000
Time: 0:00:00
Performance: 0.515 0.0128 0.00
Gradient: 0.0994 0.00164 1.00e-06
Validation Checks: 0 6 6
Plots
Performance (plotperform)
Training State (plottrainstate)
Error Histogram (ploterrhist)
Confusion (plotconfusion)
Receiver Operating Characteristic (plotroc)
Plot Interval: 1 epochs
Opening Training State Plot
Stop Training
Cancel



Best Validation Performance is 0.015137 at epoch 132
Train
Validation
Test
Best
Mean Squared Error (mse)
10^0
10^{-1}
10^{-2}
0
20
40
60
80
100
120
138 Epochs


Confusion Matrix
Output Class
Target Class
6 12.5%
5 10.4%
3 6.3%
2 4.2%
6 12.5%
1 2.1%
4 8.3%
62.5% 37.5%
75.0% 25.0%
60.0% 40.0%
83.3% 16.7%
33.3% 66.7%
50.0% 50.0%
100% 0.0%
83.3% 16.7%

From the above confusion matrix 1, 2, 3, 4, 5, 6, 7, 8 represents, rose-aphid, rose-japanese-beetle, rose-rosettle-disease, rose-goldenrod-solider, rose-mossay rose-gell-wasp, rose-normal, rose-rust, and rose-solider-beetles, respectively. And the confusion matrix is represented in Table as it is shown in Table 5.2:

Table5. 2: Summery Result of the identification

Target Class / Output class	1	2	3	4	5	6	7	8
1	6	0	0	0	0	0	0	0
2	0	5	0	0	0	1	0	0
3	0	0	2	0	0	0	0	4
4	0	0	0	6	0	0	0	0
5	0	0	0	0	6	0	0	0
6	0	3	0	0	0	3	0	0
7	0	0	0	0	0	0	6	0
8	0	0	0	0	0	0	0	6
Total	6	8	2	6	6	4	6	10
Percent correct (%)	100	62.5	100	100	100	75	100	60

$$x = \frac{u}{T} \times 100 \tag{5.1}$$

$$y = \frac{v}{T} \times 100 \tag{5.2}$$

Where *T* is total test set, *x* is correctly identified, *y* is incorrectly identified, *u* is the summation of images that have the same output class and target class in Table 5.2. *v* is the summation of images that have different input and output class.

Using equation 5.1 and 5.2, the summery result of the identification showed that from the total test set of 48 samples, 40 (83.3%) were identified correctly and 8 (16.7%) were not identified in its correct disease class.

5.3.3. Analysis of the Result

The identification accuracy of the Neural Network for rose-aphid, rose-japanese-beetle, rose-rosettle-disease, rose-goldenrod-solider, rose-mossay rose-gell-wasp, rose-normal, rose-rust, rose-solider-beetles are 100%, 62.5%, 100%, 100%, 100 %, 75%, 100% and 60% respectively.

Flowers that are affected by rose-japanse-beetle were identified as normal rose (16.7%). This shows that some flower images that are affected by rose-japanese-beetle and normal rose has a slight similarity between there texture features.

Flowers that are affected by rose-rosettle-disease were identified as rose-soldier-beetles (66.7%). This shows that there is a strong similarity between the texture features of flower images affected by rose-rosettle-disease and rose-soldier-beetles.

Normal rose images also identified as images that are affected by rose-japanese-beetles (50%). This is due to the strong similarity between the texture features of normal rose and roses that are affected by rose-japanese-beetles. The other flower images were identified into its correct disease category.

In general, from the total test set that we have used to test the system 83.3 % are identified in its correct disease class and the remaining 16.7% are misclassified.

CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS

6.1. Conclusion

According to the Ethiopian Horticulture Sector, biological pest control is an important practice leading to reduction or switching entirely from the use of chemical pesticides. In fact, biological pest control has the great advantage in assuring the safety of employee, protecting the environment, and also to reduce cost while increasing quality. In order to achieve the biological pest controlling mechanism we have to identify the disease in its early stage. Thus, developing an automatic system that identifies the disease of flowers in its early stage is unquestionable.

In line with this, to identify different flower diseases, we have selected a digital image processing technique which is a recent research area in computer science and information technology. Digital image processing is a means of processing digital images using a digital computer. Every digital image processing application follows some fundamental steps like image acquisition, preprocessing, representation and interpretation, and recognition and interpretation. In this research, we have used a digital image processing to develop a method for automatic identification of flower diseases. The method contains two main phases: training phase and identification phase

In the first phase flower images are captured. Then, this flower images are preprocessed in order to remove noises, lightening effects and others. After preprocessing those flower images are segmented using Otsu's method to identify the region of interest and useful features are extracted. For this research we have selected the texture feature of flower image and extracted using Gabor feature extraction. This feature extraction method uses a wavelet transform approach to extract the texture feature of the given image. For this study, a set of Gabor filters with five spatial frequencies and eight distinct orientations are used. This will make forty different Gabor kernels. Using those Gabor kernels we have extracted the texture feature of the image and we have represented those texture features using seven different statistical data representation techniques. Finally, those seven texture features are used to create the knowledge base which is used to train an artificial neural network.

In the identification phase flower images that are different from images that we use in training phase, are captured. Then, like the first phase images are preprocessed, segmented and useful texture features of those images are extracted from the image using the aforementioned technique.

To test the identification accuracy of the system an independent data set was used. This data set contains texture features of a normal and diseased flower image that are extracted using Gabor feature extraction. The experimental result shows that flower disease identification using texture features are efficient to classify the disease of a flower in its class of disease.

In general, identification of flower disease can be done automatically using an image analysis technique. Using the test data the eight class of disease, rose-aphid, rose-japanese-beetle, rose-rosettle-disease, rose-goldenrod-solider, rose-mossay rose-gell-wasp, rose-normal, rose-rust, rose-solider-beetles are identified 100%, 62.5%, 100%, 100%, 100 %, 75%, 100% and 60% respectively, and the overall performance was 83.3%

6.2. Recommendation

The horticulture industry in Ethiopia started only about a decade ago; the industry has scored significant positive developments. Currently, more than 120 foreign and local companies are engaged in the cultivation of horticultural export products. The identification and classification of different agricultural products using image processing applications are reported by different researchers. In Ethiopia no researches have been conducted in the identification of agricultural disease to support the sector. Hence, this research work may encourage different researchers to work on this area.

Image analysis for the identification of flower disease can be further investigated. The work can also be seen in depth and researched by the different structure of flower image. The following recommendations are made for further research and improvement.

- ✓ In this work we have built a system that identifies what type of disease attack the flower. However, the system does not estimate the severity of the identified disease. Thus, automatically estimating the severity of the identified disease can be one research direction.

- ✓ After the identification of the flower disease it is better to recommend appropriate treatment. Thus, automatically recommending the appropriate treatment for the identified disease is also other research direction.

REFERENCES

[1] Al-Amri, SalemSaleh, and Namdeo V. Kalyankar, "Image segmentation by using threshold techniques." arXiv preprint arXiv: 1005-4020 (2010).

[2] Otsu, Nobuyuki, "A threshold selection method from gray-level histograms." Automatica 11, no. 285-296 (1975): 23-27.

[3] U.kAnasuya Devi," Pixel-level image processing methodology pre-processing technique for an OCR system for the Brahmi script".

[4] Jemil A.M Saif et.al, "Image segmentation using edge detection and thresholding", *The 13th international Arab congruence on IT*, December 10 – 13, 2012.

[5] Saluja, Sonam, Aradhana Kumari Singh, and Sonu Agrawal, "A study of edge-detection methods." *International Journal of Advanced Research in Computer and Communication Engineering* (IJARCCE) 2 (2013): 994-999.

[6] Pitas, Ioannis, *Digital image processing algorithms and applications*. John Wiley & Sons, 2000.

[7] Rafael C. Gonzalez, Richard Eugene Woods, and Steven L. Eddins. *Digital image processing using MATLAB*. Pearson Education India, 2004.

[8] Wilson, Joseph N., and Gerhard X. Ritter. Handbook of computer vision algorithms in image algebra. CRC press, 2000.

[9] Bhabhor, C. H., and A. B. Upadhyay, "Performance Analysis for Image Segmentation of Various EDGE detection techniques." ICRDETS-2013 Nov 12 (2013).

[10] BARBU, Tudor, "An Approach to Image Objects Recognition And Interpretation." *Proceedings of the Romanian Academy: Mathematics, physics, technical sciences, information science* 4, no. 3 (2003): 217.

[11] Weber, Markus, Max Welling, and Pietro Perona. "Towards automatic discovery of object categories." In Computer Vision and Pattern Recognition, 2000. *Proceedings IEEE Conference* on, vol. 2, pp. 101-108. IEEE, 2000.

[12] Acharya, Tinku, and Ajoy K. Ray. *Image processing: principles and applications*. John Wiley & Sons, 2005.

[13] Coolen, Anthony CC, Reimer Kühn, and Peter Sollich, *Theory of neural information processing systems*. *Oxford University Press*, 2005.

[14] Woodford, Brendon J., Nikola K. Kasabov, and C. Howard Wearing. "Fruit image analysis using wavelets." In *Proceedings of the ICONIP/ANZIIS/ANNES*, vol. 99, pp. 88-91. 1999.

[15] H.Al.Hiary,S.Bani.Ahmed et.al, “Fast and Accurate Detection and classification of plant diseases”, *International Journal of Computer Applications*(0975-8887),Volume 17 No 1, March 2011.

[16] Tzionas, Panagiotis, Stelios E. Papadakis, and DimitrisManolakis. "Plant leaves classification based on morphological features and a fuzzy surface selection technique." *In Fifth International Conference on Technology and Automation*, Thessaloniki, Greece, pp. 365-370. 2005.

[17] Anami, Basvaraj S., J. D. Pujari, and Rajesh Yakkundimath. "Identification and classification of normal and affected agriculture/horticulture produce based on combined color and texture feature extraction." IJCAES I (III) (2011): 356-360.

[18] Chaudhary, Piyush, Anand K. Chaudhari, and ANCheeran&ShardaGodara, "Color transforms based approach for disease spot detection on plant leaf."*International Journal of Computer Science and Telecommunications* 1, no. 6 (2012): 65-70.

[19] He, Qinghai, Benxue Ma, DuanyangQu, Qiang Zhang, XinminHou, and Jing Zhao, "Cotton pests and diseases detection based on image processing."Telkomnika *Indonesian Journal of Electrical Engineering* 11, no. 6 (2013): 3445-3450.

[20] Arivazhagan, S., R. Newlin Shebiah, S. Ananthi, and S. Vishnu Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features." *Agricultural Engineering International*: CIGR Journal 15, no. 1 (2013): 211-217.

[21] R.parker, ”*Algorithm for image processing and computer vision*”,2nd edition, p 67-82

[22] HabtamuMinassie,”Image analysis for Ethiopian coffee classification”. Unpublished *master’s thesis*, school of graduate study, Addis Ababa university, 2008

[23] Collen,A.c.c,R.kuhun and P.sollich,”*Theory of natural information processing system*”, *oxford university press*,2005.

[24] Chellappa, Rama, Charles L. Wilson, and SaadSirohey, "Human and machine recognition of faces: A survey." *Proceedings of the IEEE* 83, no. 5 (1995): 705-741.

[25] Ziga Gosar, ”seminar on surface defect defection with neural networks”, *Seminar in computer science*, *University of Lijubljani*, *Faculty of Mechanical Engineering*, 2011/2012

[26] Wikipedia, Gabor filter, http://en.wikipiedia.org/wiki/gaborfilter, visited in November, 2014.

[27] Mohammad Haghighat, Saman Zonouz, and Mohamed Abdel-Mottaleb, "Identification using encrypted biometrics." In Computer Analysis of Images and Patterns, pp. 440-448. *Springer Berlin Heidelberg*, 2013.

[28] Jemaa, Yousra Ben, and Sana Khanfir, "Automatic local Gabor features extraction for face recognition." arXiv preprint arXiv:0907.4984 (2009).

[29] BENČO, Miroslav, and Robert Hudec. "Novel method for color textures features extraction based on GLCM." *Radio engineering* 4, no. 16 (2007): 64-67.

[30] Neural Network Toolbox, 2014, from http://www.mathworks.com/products/neural-network. , visited on June 12.

[31] Solomon, Chris, and Toby Breckon. Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons, 2011.

[32] Tallada, J. G., D. T. Wicklow, T. C. Pearson, and P. R. Armstrong, "Detection of fungus infected corn kernels using near infrared reflectance spectroscopy and color imaging." Trans. ASABE 54, no. 3 (2011): 1151-1158.

[33] Raji, A. O., and A. O. Alamutu. "Prospects of computer vision automated sorting systems in agricultural process operations in Nigeria." (2005).

[34] Narendra, V. G., and K. S. Hareesha, "Prospects of computer vision automated grading and sorting systems in agricultural and food products for quality evaluation." *International Journal of Computer Applications* 1, no. 4 (2010): 1-9.

[35] Davies, E. R. "The application of machine vision to food and agriculture: a review." *The Imaging Science Journal* 57, no. 4 (2009): 197-217.

[36] J.Bossua, Ch.Geea,G.jones,F.Trechetetb, 2009,"wavelet transform to discriminate between crop and weed in perspective agronomic images", *Science Direct-computers and electronics in agriculture* (2009),133-143.

[37] Siddiqi, Muhammad Hameed, Irshad Ahmad, and Suziah Bt Sulaiman, "Weed recognition based on erosion and dilation segmentation algorithm." *In Education Technology and Computer*, 2009. ICETC'09. *International Conference on*, pp. 224-228. IEEE, 2009.

[38] Sheela, K. Gnana, and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks." Mathematical Problems in Engineering2013 (2013).

[39] Lynne E.parker, "Notes on multilayer feed forward neural network" *project in Machine learning, spring* 2006.

[40] Anand.H.Kulkarni and AshwinPatilR.K, "Appling image processing technique to detect plant disease" vol.2 Issue.5 Sep-Oct 2012

[41] Barbedo, Jayme Garcia Arnal, "Digital image processing techniques for detecting, quantifying and classifying plant diseases." *Springer Plus 2,* no. 660 (2013): 1-12.

[42] Burgos-Artizzu, Xavier P., Angela Ribeiro, Maria Guijarro, and Gonzola Pajares, "Real-time image processing for crop/weed discrimination in maize fields." *Computers and Electronics in Agriculture* 75, no. 2(2011): 337-346.

[43] Ethiopian flower export, from http://ethiopianflowerexport.com/ , visited in September 2014.

[44] Jain, Anil K. *Fundamentals of digital image processing*. Vol. 3. Englewood Cliffs: prentice-Hall, 1989.

[45] Ethiopian horticulture sector Statistical Bulletin, issue 01, 2012, from http://www.ehda.gov.et/downloads.html, visited in October 2014.

[46] Insect and diseases Management, from http://leon.ifas.ufl.edu/COMHORT/Insect_and_Diseases_Management. , visited in September, 2014

APPENDIX A

TEXTURE FEATURE EXTRACTION MATLAB CODE

```
//Creating The Gabore Filter Bank
functiongaborArray = gaborFilterBank(u,v,m,n)

if (nargin ~= 4)    % Check correct number of arguments
error('There should be four inputs.')
end
% Create u*v gabor filters each being an m*n matrix

gaborArray = cell(u,v);
fmax = 0.25;
gama = sqrt(2);
eta = sqrt(2);

for i = 1:u

fu = fmax/((sqrt(2))^(i-1));
alpha = fu/gama;
beta = fu/eta;

for j = 1:v
tetav = ((j-1)/v)*pi;
gFilter = zeros(m,n);

for x = 1:m
for y = 1:n
xprime = (x-((m+1)/2))*cos(tetav)+(y-((n+1)/2))*sin(tetav);
yprime = -(x-((m+1)/2))*sin(tetav)+(y-((n+1)/2))*cos(tetav);
gFilter(x,y) = (fu^2/(pi*gama*eta))*exp(-
((alpha^2)*(xprime^2)+(beta^2)*(yprime^2)))*exp(1i*2*pi*fu*xprime);
end
end
gaborArray{i,j} = gFilter;

end
end
```

//Extracts the Gabor Features of the Image

```
Function featureVector = gaborFeatures(img,gaborArray,d1,d2)
if (nargin ~= 4)    % Check correct number of arguments
error('Use correct number of input arguments!')
end

if size(img,3) == 3
% % Check if the input image is grayscale
img = rgb2gray(img);
end
img = double(img);
[u,v] = size(gaborArray);
gaborResult = cell(u,v);
for i = 1:u
for j = 1:v
gaborResult{i,j} = conv2(img,gaborArray{i,j},'same');
% J{u,v} = filter2(G{u,v},I);
end
end
% Extract feature vector from input image
[n,m] = size(img);
s =((n*m)/(d1*d2));
 l = (s*u*v);
featureVector =zeros(l,1);
c = 0;
for i = 1:u
for j = 1:v

        c = c+1;
gaborAbs = abs(gaborResult{i,j});
gaborAbs = downsample(gaborAbs,d1);
gaborAbs = downsample(gaborAbs.',d2);
gaborAbs = reshape(gaborAbs.',[],1);

gaborAbs = (gaborAbs-mean(gaborAbs))/std(gaborAbs,1);

%          featureVector(((c-1)*s+1):(c*s)) = gaborAbs;
end
```

Declaration

I, the undersigned, declare that this thesis is my work and has not been presented for a degree in any other university, and that all source of materials used for the thesis have been duly acknowledge.

Declared by:

Name:________________________________

Signature:_____________________________

Date:_________________________________

Confirmed by advisor:

Name:________________________________

Signature:_____________________________

Date:_________________________________

Place and date of submission: Addis Ababa, February, 2015