CSE 4019 - IMAGE PROCESSING

PROJECT REPORT

TITLE: SKIN CANCER DETECTION USING K-MEANS ALGORITHM

SHIKHAR SINGH-16BCE2316

SUBMITTED TO:

PROF. :RAJAKUMAR K

SLOT: G1+TG1

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

SKIN CANCER DETECTION USING K-MEANS ALGORITHM

A Project Report

Submitted in partial fulfillment of the requirement for the evaluation of J component for the subject CSE 4019 IMAGE PROCESSING

by

Shikhar Singh(16BCE2316)

Faculty Incharge

Prof. RAJAKUMAR K

School of Computer Science and Engineering

Vellore Institute of Technology, Vellore-632014

DECLARATION

I hereby declare that the project work entitled "**SKIN CANCER DETECTION USING K-MEANS ALGORITHM"** submitted by me for J component evalution for the subject CSE 4019-Image Processing, the work reported in this report has been submitted, having being completed to the best of my ability.

Place : Vellore **Signature of the Candidate**

Shikhar Singh(16BCE2316)

Date :

ACKNOWLEDGEMENTS

I would like to articulate my deep gratitude to my project guide Prof. **RAJAKUMAR K** who has always been source of motivation and firm support for carrying out the project. I would also like to convey our sincerest gratitude and indebtedness to all other faculty members and staff of Department of Computer Engineering, VIT University, who showed their great efforts and guidance at required times without which it would have been very difficult to carry out my project work.

Moreover, an assemblage of this nature could never have been attempted with my reference to the works of others whose details are mentioned in the references section at the last.

I acknowledge my indebtedness to all of them. Furthermore, I would like to take the name of our parents and God who directly or indirectly encouraged and motivated us during this dissertation.

CONTENTS **Page No.**

Declaration

Acknowledgement

Table of Contents

1. ABSTRACT

2. OVERALL ARCHITECTURE

2.1 Background

2.2 Work done

2.3 Segmentation using K-Means Algorithm

3. DETAILED DESCRIPTION OF MODULES

3.1 Image Pre-Processing

3.2 Segmentation

4. RESULTS

4.1 Benign Images

4.2 Malignant Images

5. APPENDICES

6. DATASET USED

7. CONCLUSION

8.REFERENCES

ABSTRACT

Skin cancer is term given to the uncontrolled growth of strange skin cells. It occurs whenever unrepaired DNA damages to skin cells triggers mutations, or any other genetic defects, that lead the skin cells to multiply readily and form malignant tumors. Image processing is a commonly used method for skin cancer detection from the appearance of affected area on the skin. The input to the system is the skin lesion image and then by applying novel image processing techniques, it analyses it to conclude about the presence of skin cancer. The Lesion Image analysis tools checks for the various Melanoma parameters Like Asymmetry, Border, Colour, Diameter, (ABCD rule) etc. by texture, size and shape analysis for image segmentation and feature stages. The extracted feature parameters are used to classify the image as Normal skin and Melanoma cancer lesion.

Artificial Neural Network (ANN) is one of the important branches of Artificial Intelligence, which has been accepted as a brand new technology in computer science for image processing. Neural Networks are currently the area of interest in medicine, particularly in the fields of radiology, urology, cardiology, oncology, etc. Neural Network plays a vital role in an exceedingly call network. It has been used to analyse Melanoma parameters Like Asymmetry, Border, Colour, Diameter, etc. which are calculated using MATLAB from skin cancer images intending to developing diagnostic algorithms that might improve

triage practices in the emergency department. Using the ABCD rules for the melanoma skin cancer, we use ANN in classification stage. Initially, we train the network with known target values. The network is well trained with 96.9% accuracy, and then the unknown values are tested for the cancer classification. This classification method proves to be more efficient for the skin cancer classification.

INTRODUCTION

In recent days, skin cancer is commonly seen as one of the most dangerous forms of the Cancers identified in Humans. Skin cancer is classified into various types such as Melanoma, Basal and Squamous Cell Carcinoma out of which Melanoma is the most unpredictable and the most common form of cancer. Melanoma could be a notably deadly variety of skin cancer, and though it justifies solely 4% of all types of skin cancers, it is responsible for 75% of all skin cancer deaths. Image processing is one of the most widely used methods for skin cancer detection. 'Dermoscopy' could be a non-invasive examination technique that supports the cause of incident light beam and oil immersion technique for the visual investigation of surface structures of the skin. The detection of melanoma using dermoscopy is higher than individual observation based detection, but its diagnostic accuracy depends on the factor of training the dermatologist. The diagnosis of melanoma is not very clear and easy to identify, especially in the early stage. Thus, automatic diagnosis tool is more effective and essential. Other than 'dermoscopy', a computerised melanoma detection using Artificial Neural Network classification has been adapted which is efficient than the conventional one for classification.

OVERALL ARCHITECTURE

Background of the problem:

Melanoma, a type of deadly skin cancer affects the region of skin which is exposed directly to UV radiation which shows a rapid death chance. In order to lower the death rate early detection methods are adapted. According to the statistical information available, it has been proved that the melanoma incidence rates showed an increase of 2% to 7% per year in between 2006-2010 and also the death rate showed an increase of 1.1% in males and 0.2% in females per year. According to the recent survey of 2014, it has been stated that the total effects of melanoma are around 76,100 and the deaths are around 9,710. In order to decrease the death rate, image processing is used for the detection of skin cancer. By using this methodology early detection of skin cancer can be achieved. It lessens the burden of Dermatologists.

A solution to this problem is using image processing to detect skin cancer.

Work done:

1. **Pre-processing part:**

Median filtering has been used for pre-processing part. It focuses on elimination of salt and pepper noises. A technique where window's centre value is replaced by the value which is the median of the 8 neighbourhood point's pixel values. It is a nonlinear filter. It is also a sliding-window spatial filter and considered as the most commonly used smoothing filter. This filter with suitable mask size eliminates the artifacts in dermoscopy images.

2. **Erosion:**

Erosion is one of the two basic operators in the area of mathematical morphology, the other being dilation. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels shrink in size, and holes within those areas become larger.

Segmentation using K-Means Algorithm:

K-Means is a least-squares partitioning method that divide a collection of objects into K groups. The algorithm iterates over two steps:

1. Compute the mean of each cluster.
2. Compute the distance of each point from each cluster by computing its distance from the corresponding cluster mean. Assign each point to the cluster it is nearest to.
3. Iterate over the above two steps till the sum of squared within group errors cannot be lowered any more.

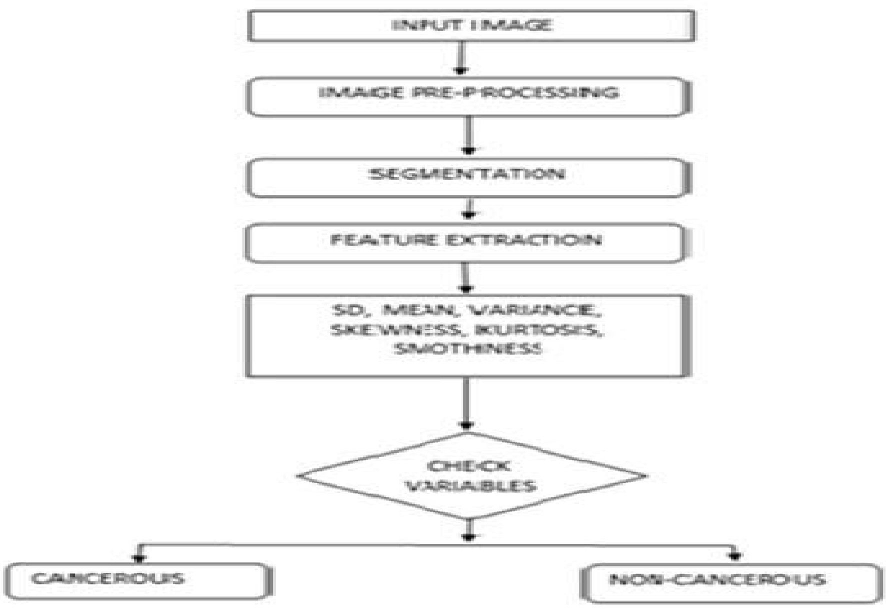
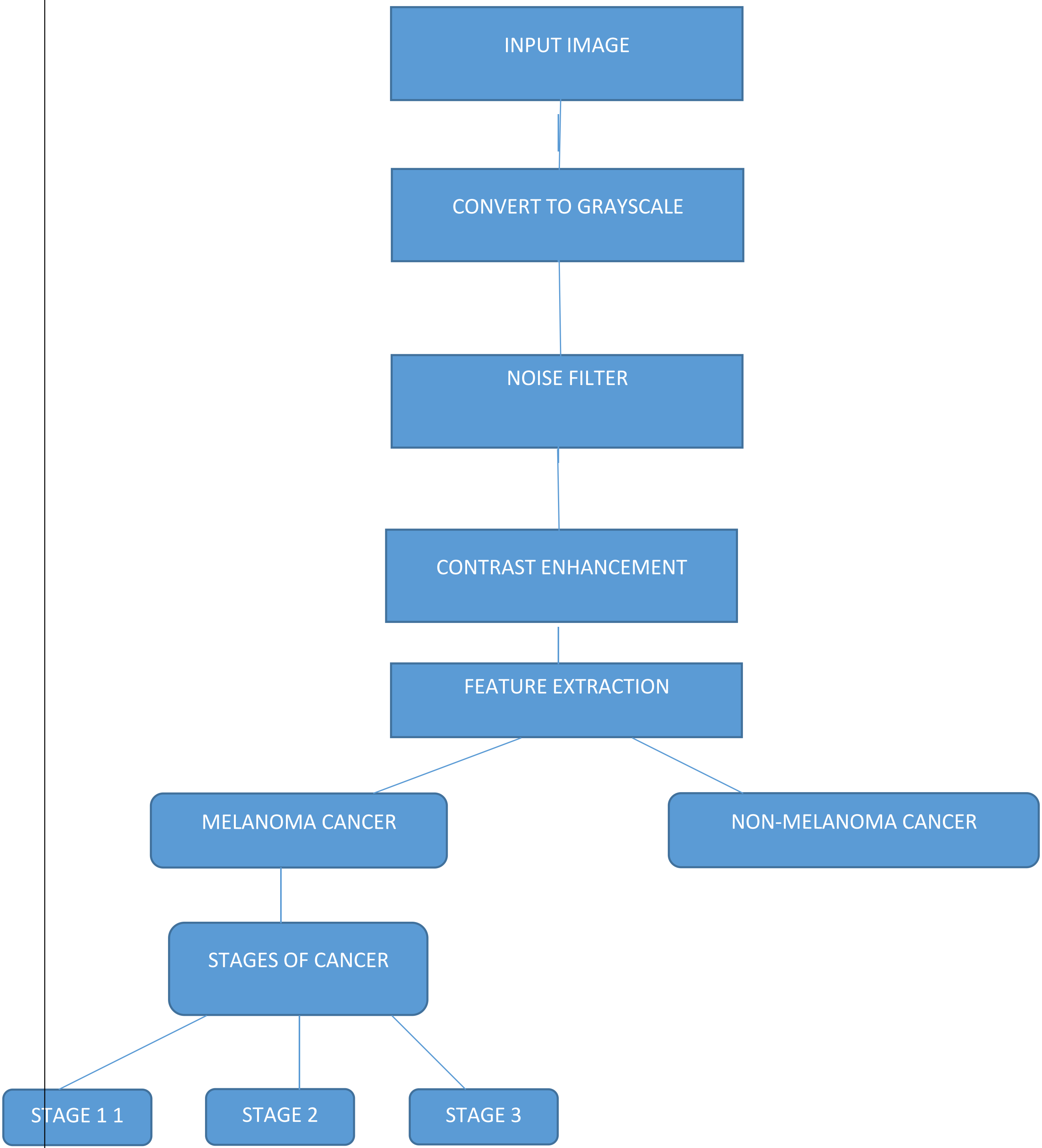


Fig 1: Flow of the System

Our flow chart is as follows:



DETAILED DESCRIPTION OF MODULES

A.)Image Pre-Processing

The image of skin lesion is given to the computer diagnostic system can be captured in any lighting condition or by using any camera. Hence, it needs to pre-process. Here, the pre-processing is the process of image resizing (scaling) and contrast and brightness modification, which is done in furtherance of compensating the non-uniform illumination in the image.

Image scaling:

Image scaling is the course of action of resizing a digital image. The size of an image is reduced or enlarged, the pixels that form the image become increasingly visible, making the image appear soft.

RGB to grayscale image:

The rgb2gray function converts the true color image RGB to the grayscale intensity image, by eliminating the saturation information.

Grayscale to Binary image:

Im2bw command converts the grayscale image to a binary image. The output image replaces all pixels in the input image with luminance exceeding the level with the value 1(white) and substitute all other pixels with the value 0 (black). If we do not define the level, then im2bw uses the value 0.5.

B.) Segmentation

Image segmentation is the course of action of segregating an image into multiple parts, which is used to identify objects or other relevant information in digital images.

Background subtraction:

Background subtraction, also known as blob detection, is an emerging technique in the fields of image processing wherein an image's foreground is extracted for further processing. Typically, an image's regions of interest are objects in its foreground.

Edge detection:

Edge detection is a significant image processing technique for catching the boundaries of objects within images. It works by detecting discontinuities in brightness.

Masking:

Masking involves setting the pixel values in an image to zero, or some other "background" value. It is used to separate the lesion from the skin image. The masked image obtained contains only the skin lesion.

Feature Extraction:

The foremost features of the Melanoma Skin Lesion are its Geometric Features. Hence, we propose to extract the Geometric Features of the segmented skin lesion. Here, we used some classic geometry features (Area, Perimeter, Greatest Diameter, Circularity Index, Irregularity Index) adopted from the segmented image containing only skin lesion, the image blob of the skin lesion is analysed to extract the geometrical features.

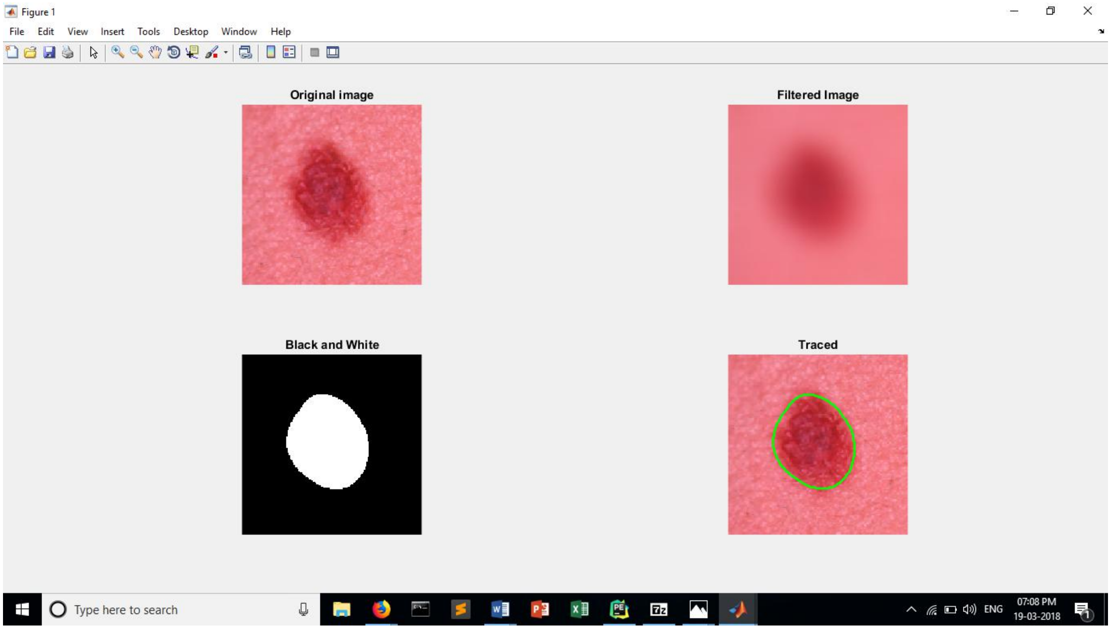
RESULTS (OUTPUT, RESULT ANALYSIS BASED ON METRICS)

SKIN CANCER DETECTION:

1.)BENIGN IMAGES

a.)Input image:

Outputs:



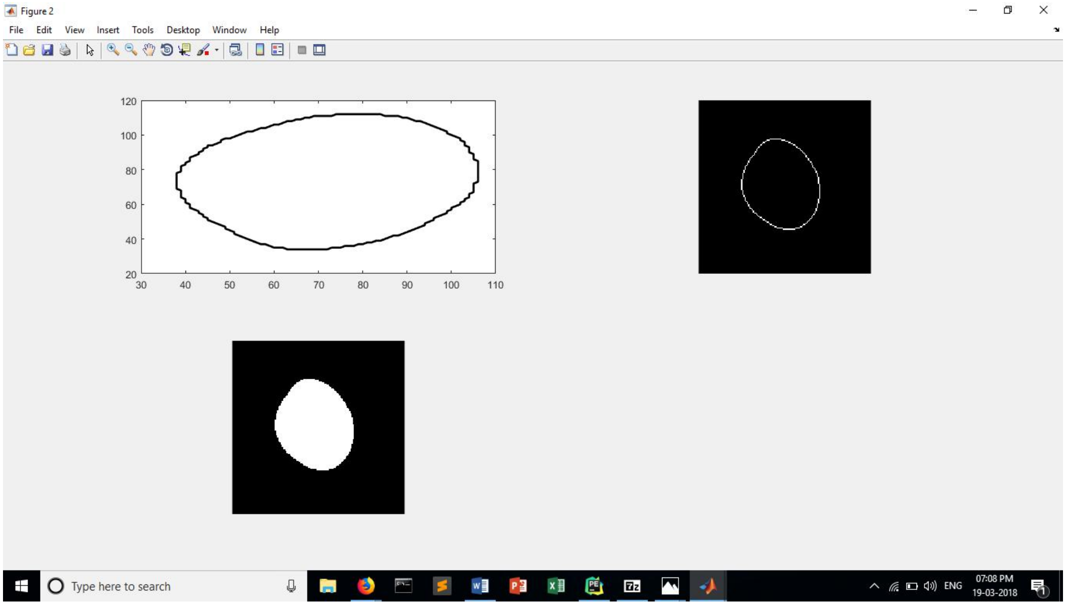


Figure 2
File Edit View Insert Tools Desktop Window Help
120
100
80
60
40
20
30 40 50 60 70 80 90 100 110
Type here to search
07:08 PM
19-03-2018

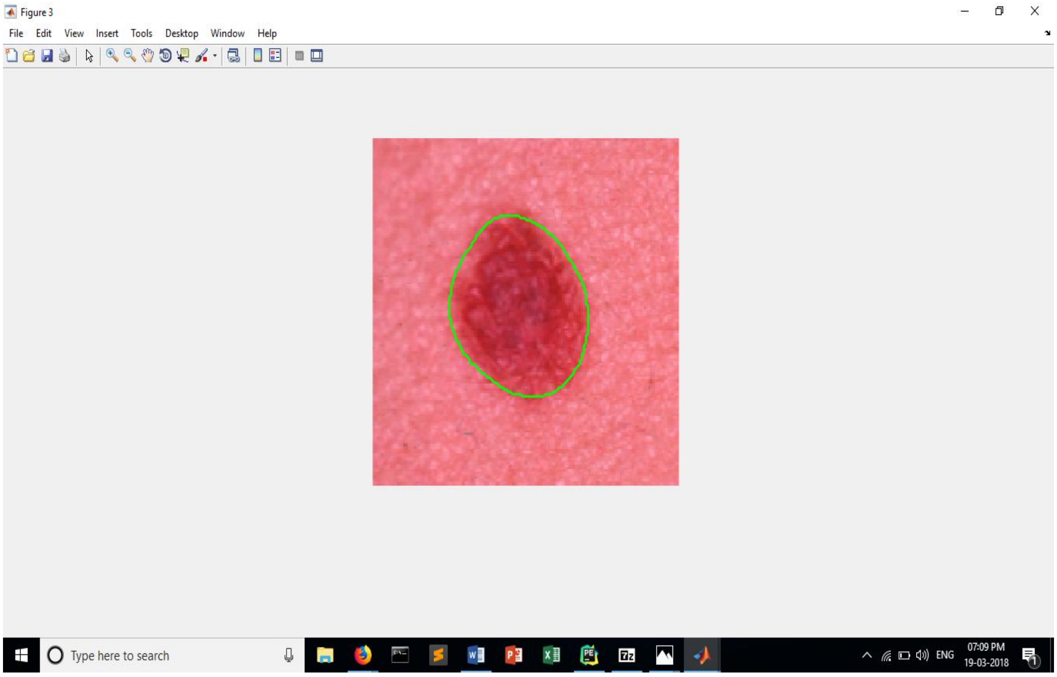
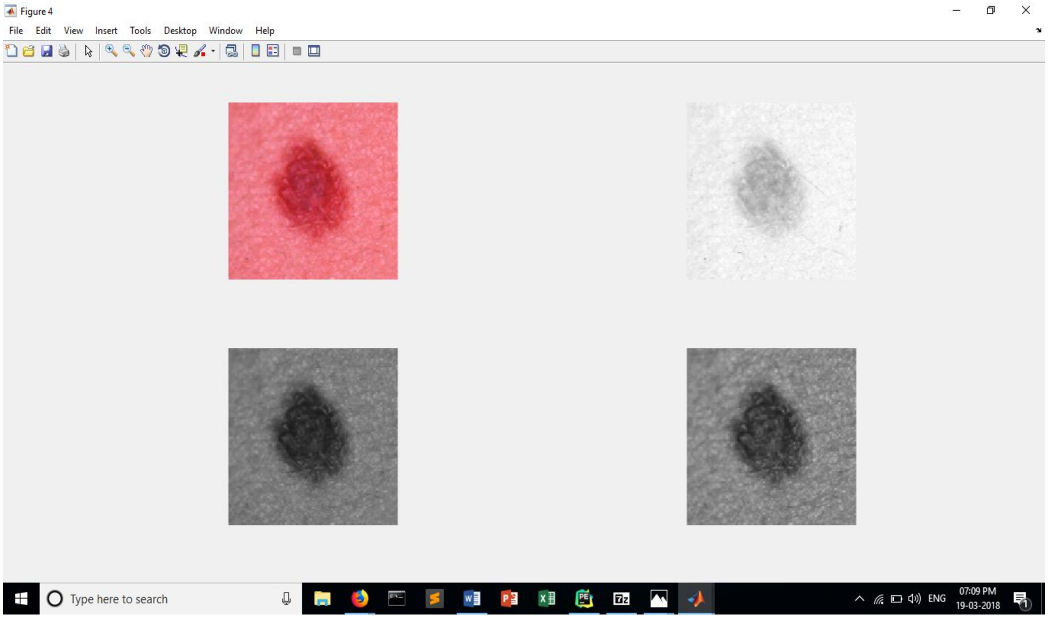


Figure 3
File Edit View Insert Tools Desktop Window Help
Type here to search
07:09 PM
19-03-2018



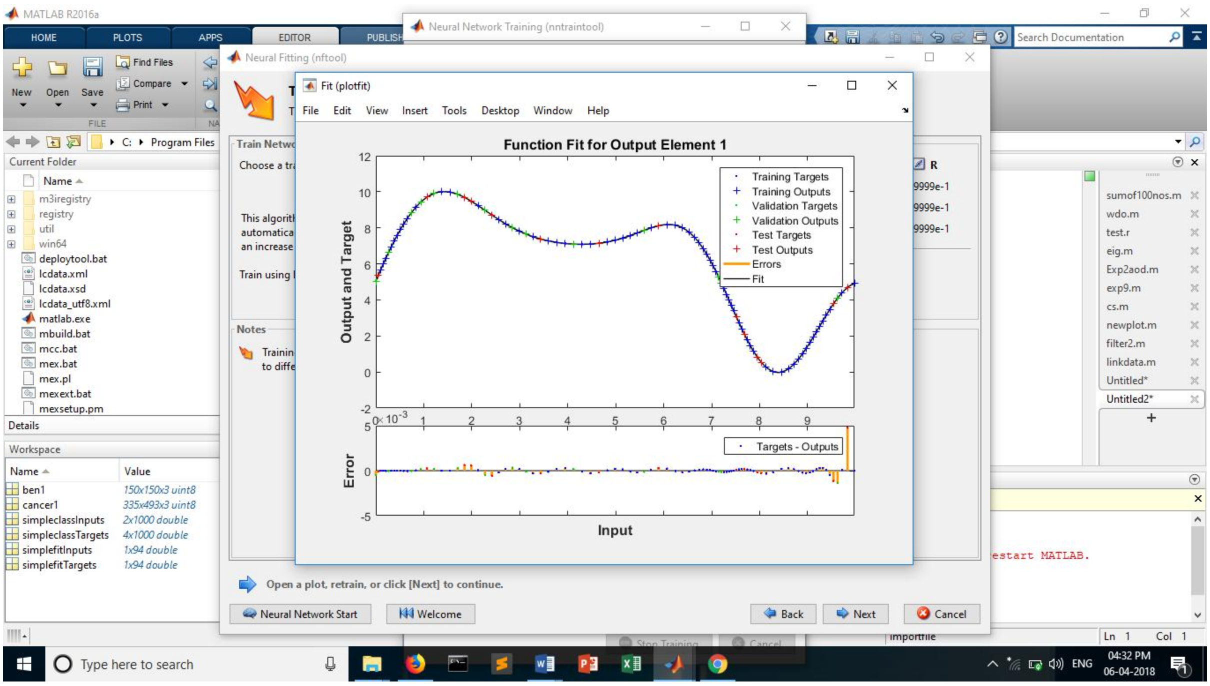
ANALYSIS:

- Asymmetry. **Not** very asymmetric
- Border irregularity. The edges are **not** ragged, notched, or blurred.

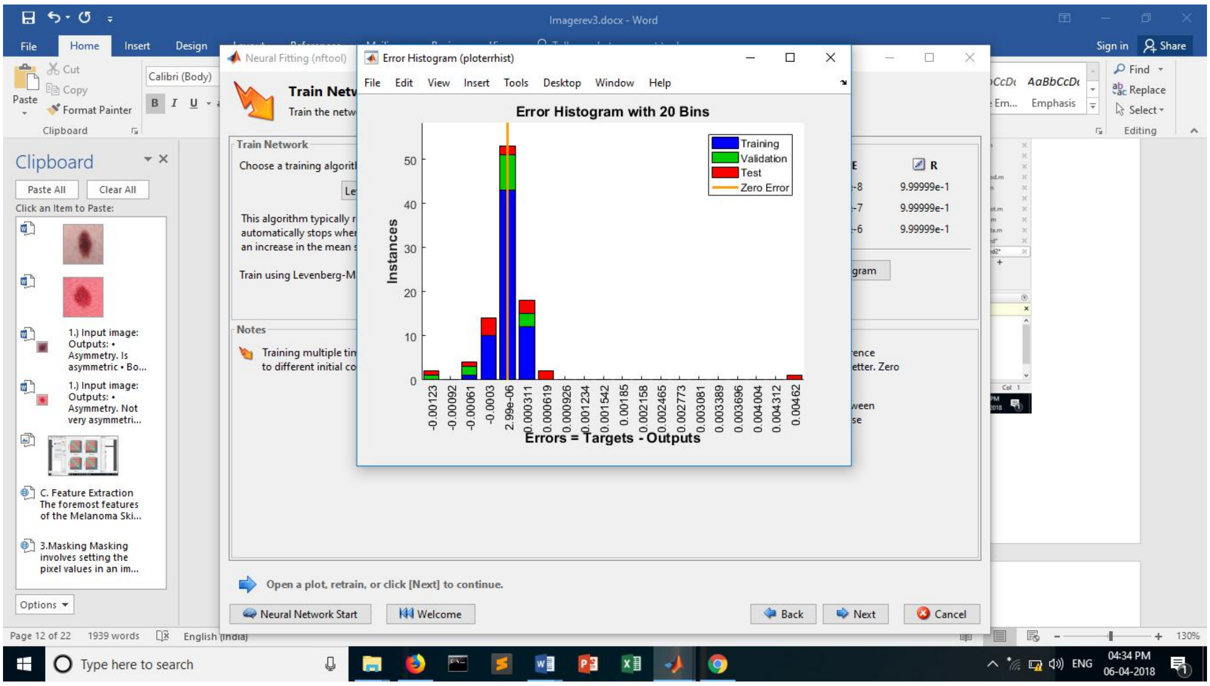
- Color. The color (pigmentation) is uniform. Shades of tan, brown, and black are **not** Present. **Not** a mottled appearance.

- Diameter. The size of the mole is **lesser** than 1/4 inch (6 mm), about the size of a pencil eraser..

PLOTFIT:



PLOT ERROR HISTOGRAM



Thus the skin lesion examined is **non-cancerous**

b.)Input image

Outputs:

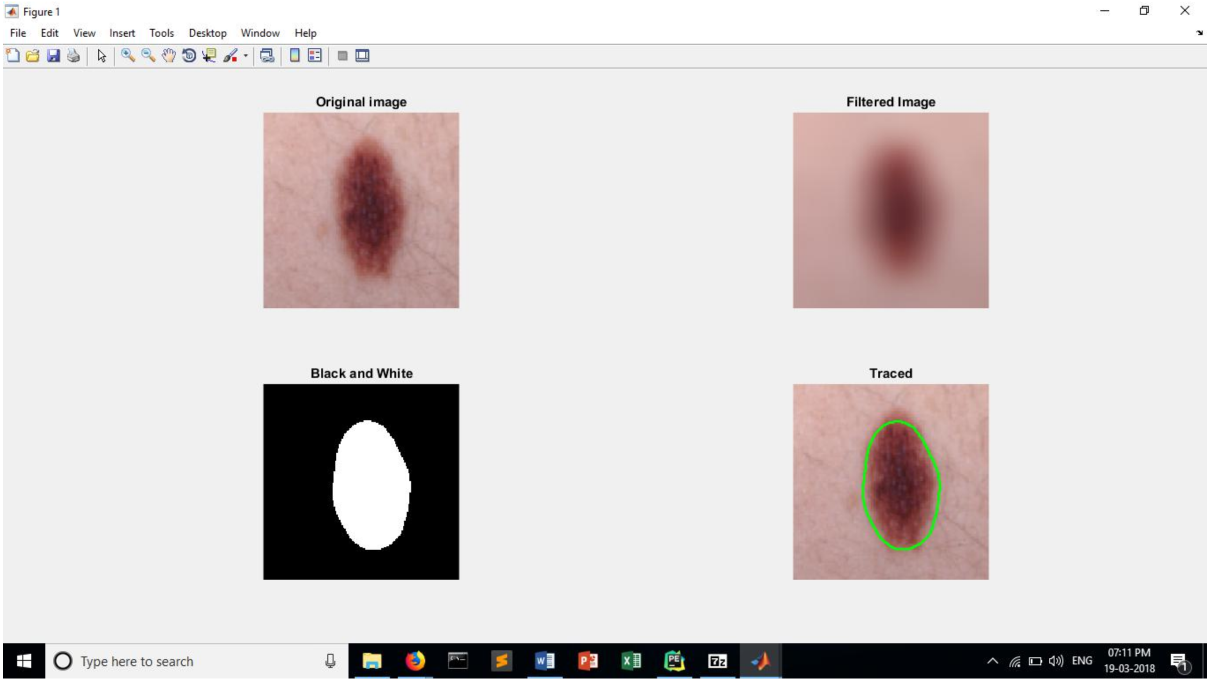


Figure 1
File Edit View Insert Tools Desktop Window Help
Original image
Filtered Image
Black and White
Traced
Type here to search
ENG
07:11 PM
19-03-2018

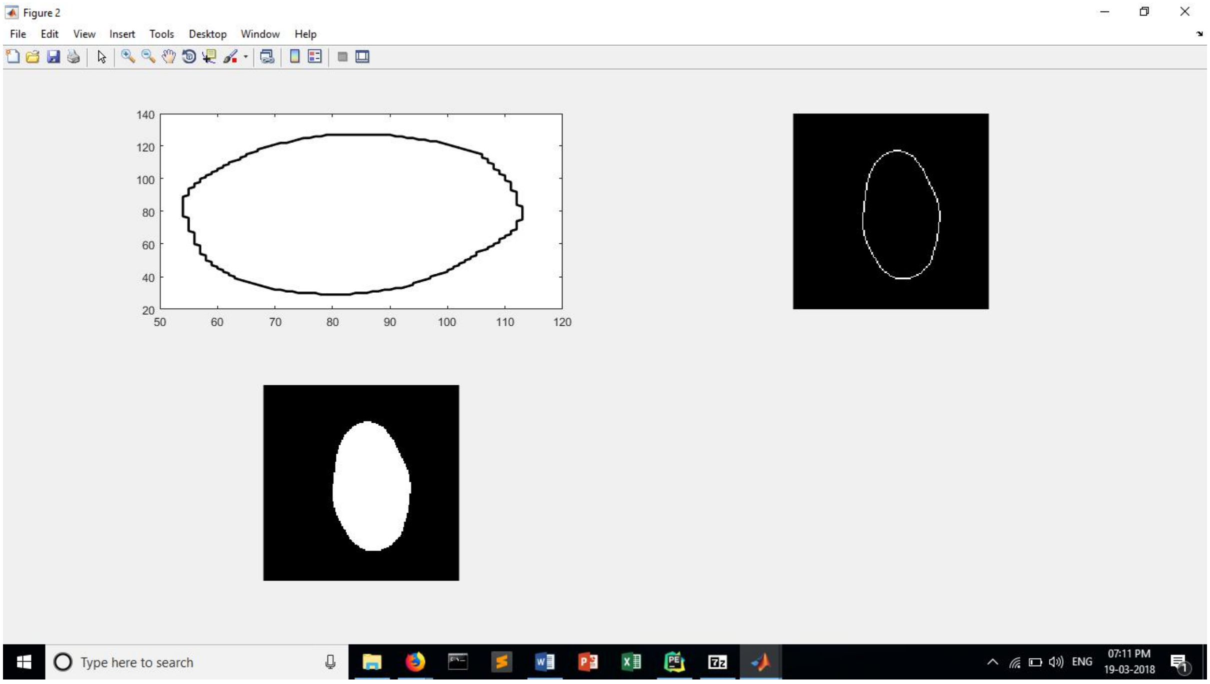
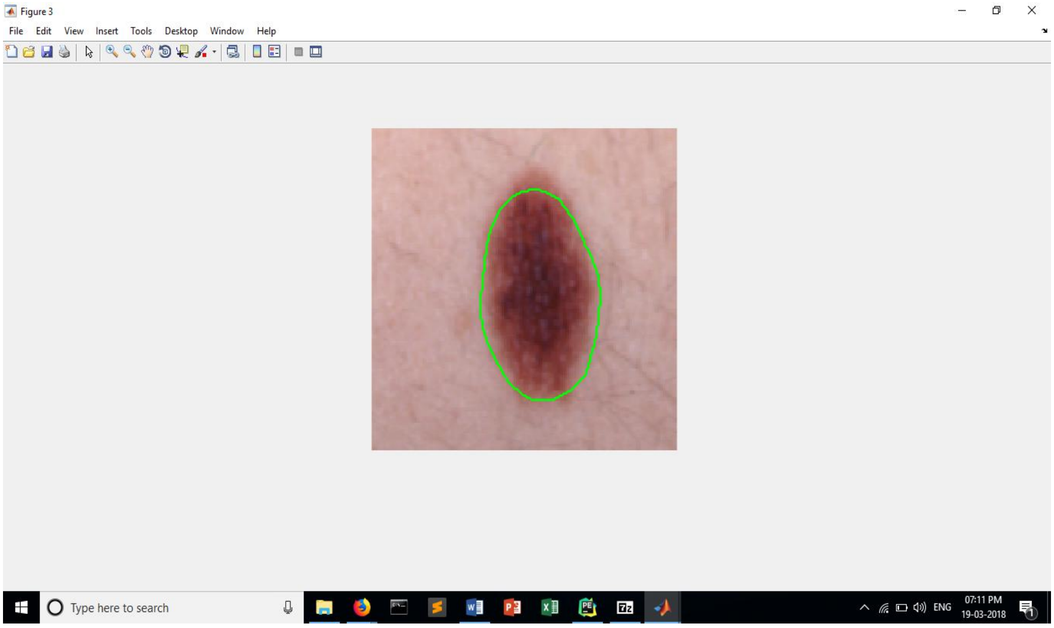
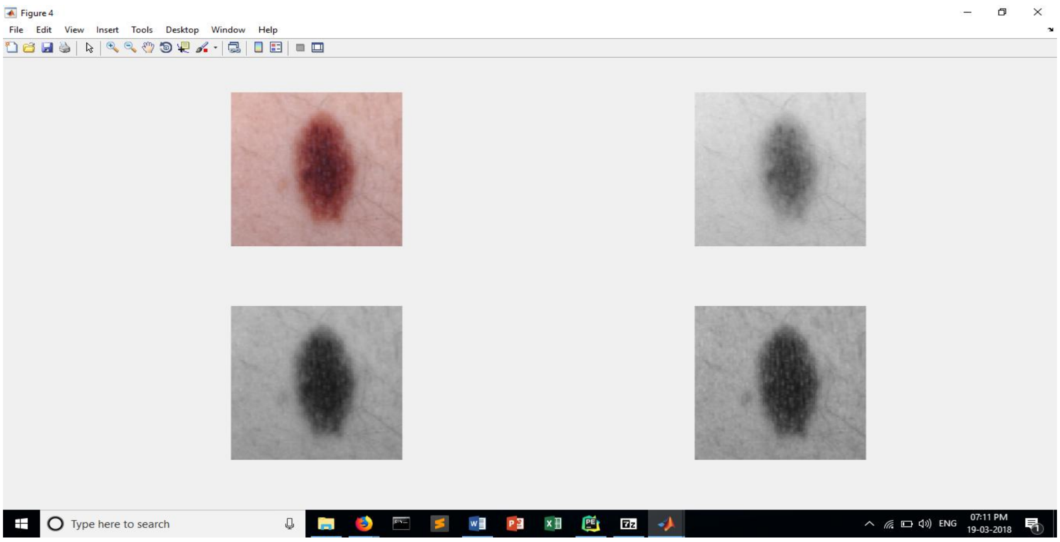


Figure 2
File Edit View Insert Tools Desktop Window Help
140
120
100
80
60
40
20
50
60
70
80
90
100
110
120
Type here to search
ENG
07:11 PM
19-03-2018





ANALYSIS:

- Asymmetry. **Not** very asymmetric
- Border irregularity. The edges are **not** ragged, notched, or blurred.

- Color. The color (pigmentation) is uniform. Shades of tan, brown, and black are **not** Present. **Not** a mottled appearance.

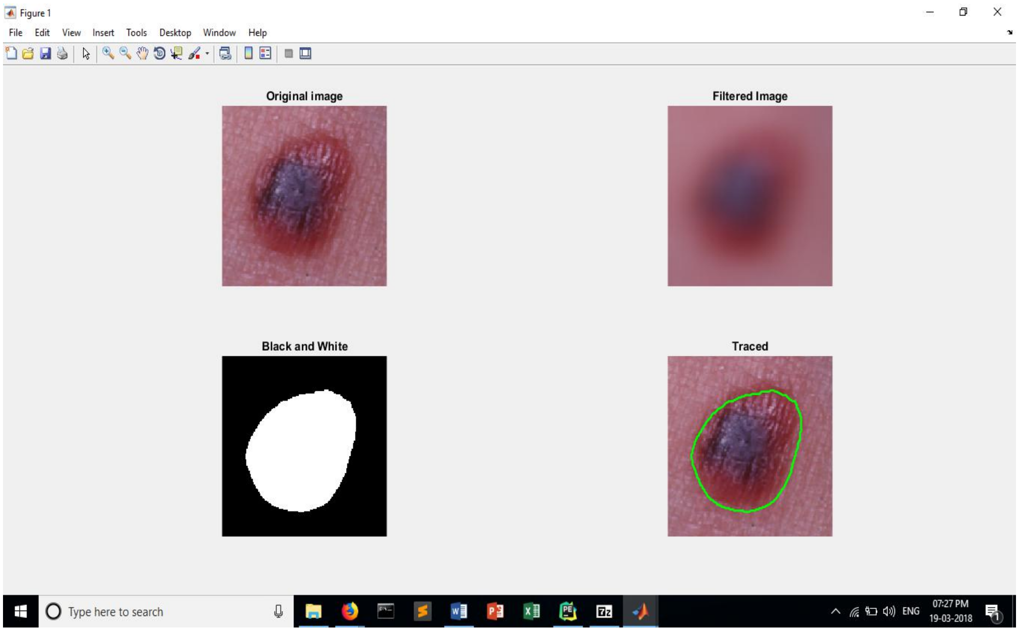
- Diameter. The size of the mole is **lesser** than 1/4 inch (6 mm), about the size of a pencil eraser..

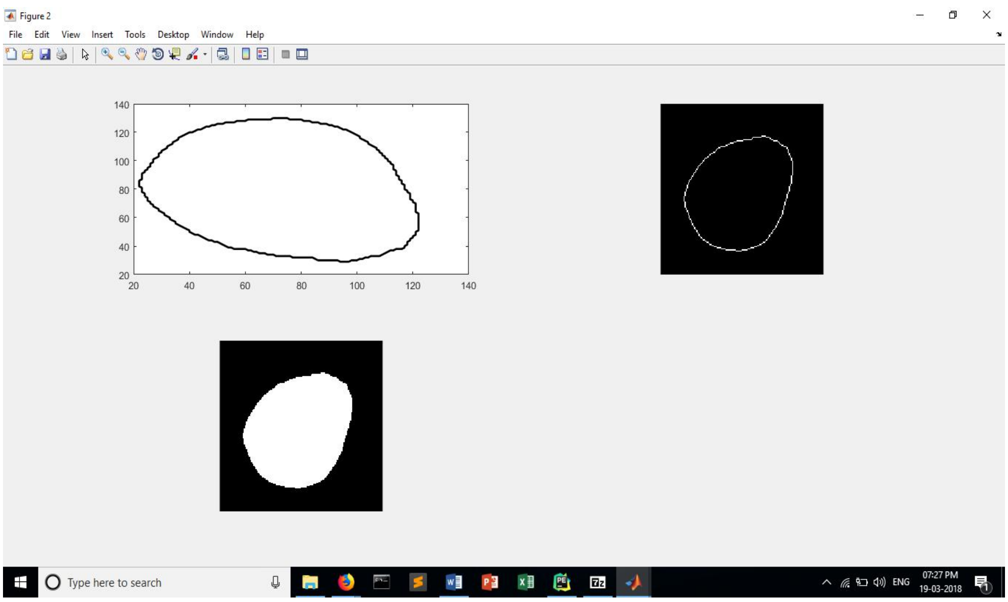
Thus the skin lesion examined is **<u>non-cancerous</u>**

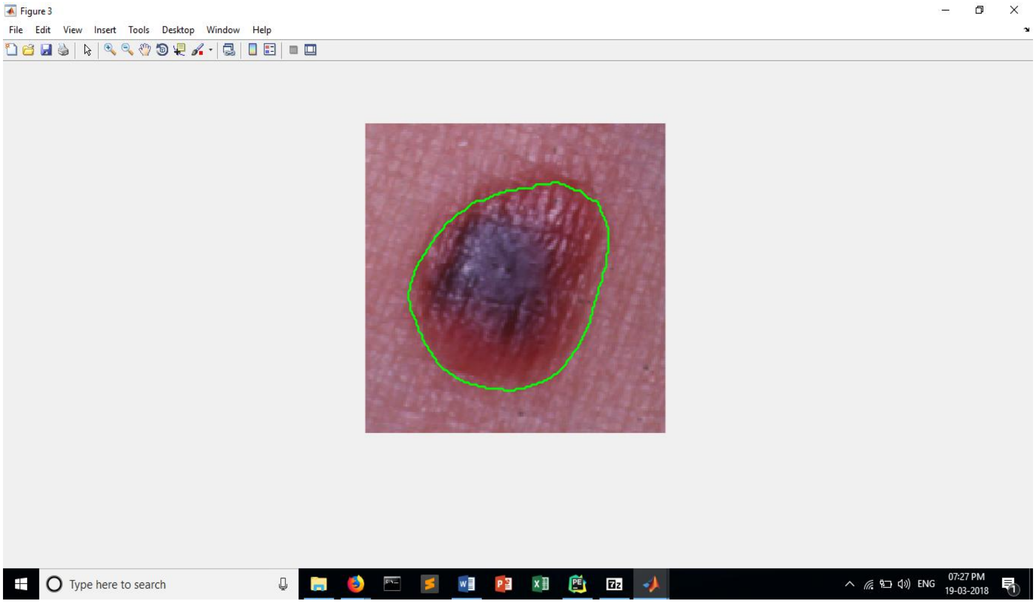
2.)MALIGNANT IMAGES:

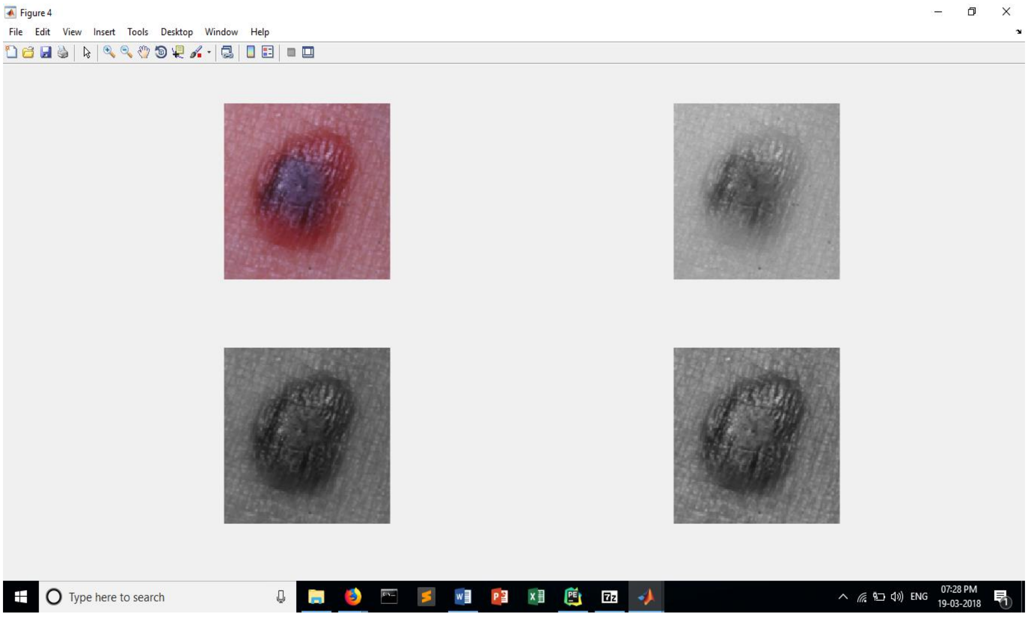
a.)Input image:

Outputs:









ANALYSIS:

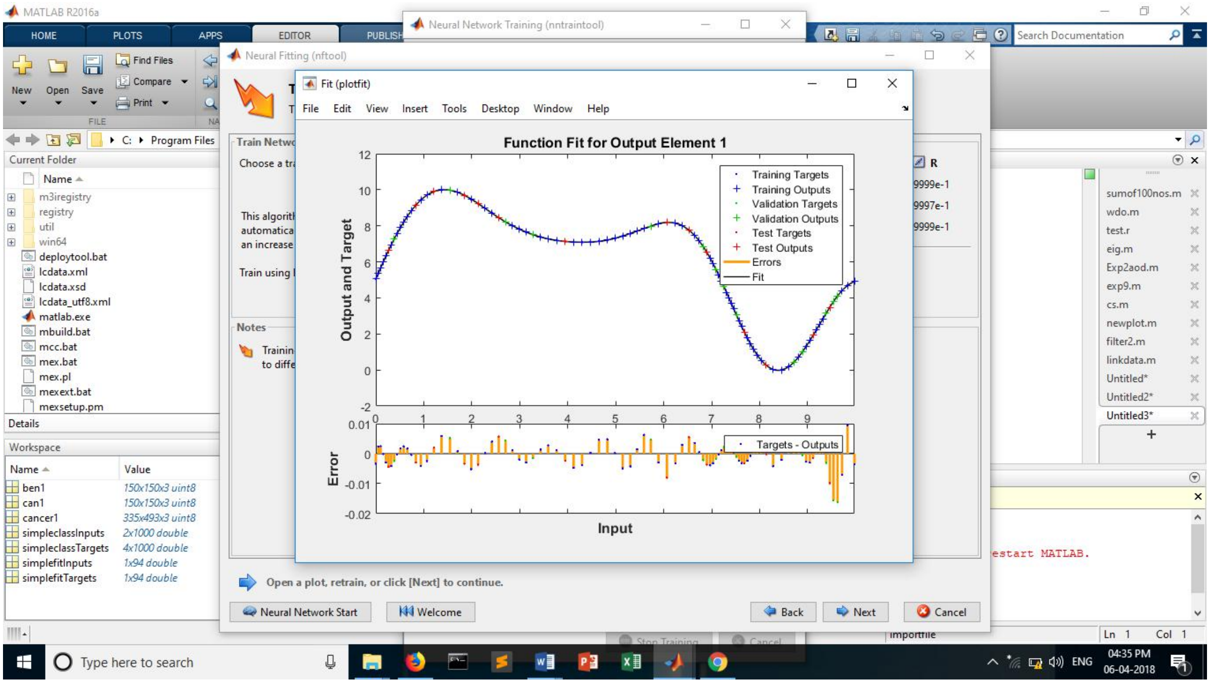
- Asymmetry. **Is** asymmetric
- Border irregularity. The edges **are** ragged, notched, or blurred.

- Color. The color (pigmentation) is **non- uniform**. Shades of tan, brown, and black **are** Present. **It has** a mottled appearance.

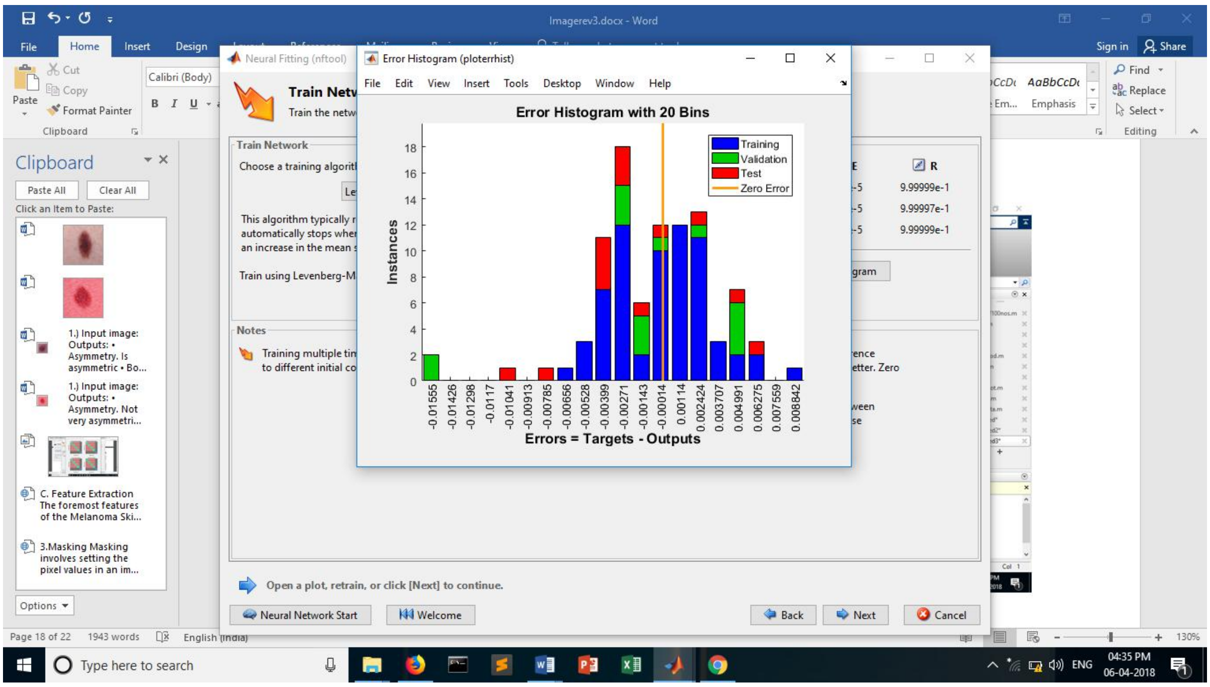
- Diameter. The size of the mole is **greater** than 1/4 inch (6 mm), about the size of a pencil eraser..

Thus the skin lesion examined is **<u>Cancerous</u>**

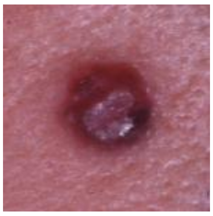
Plotfit



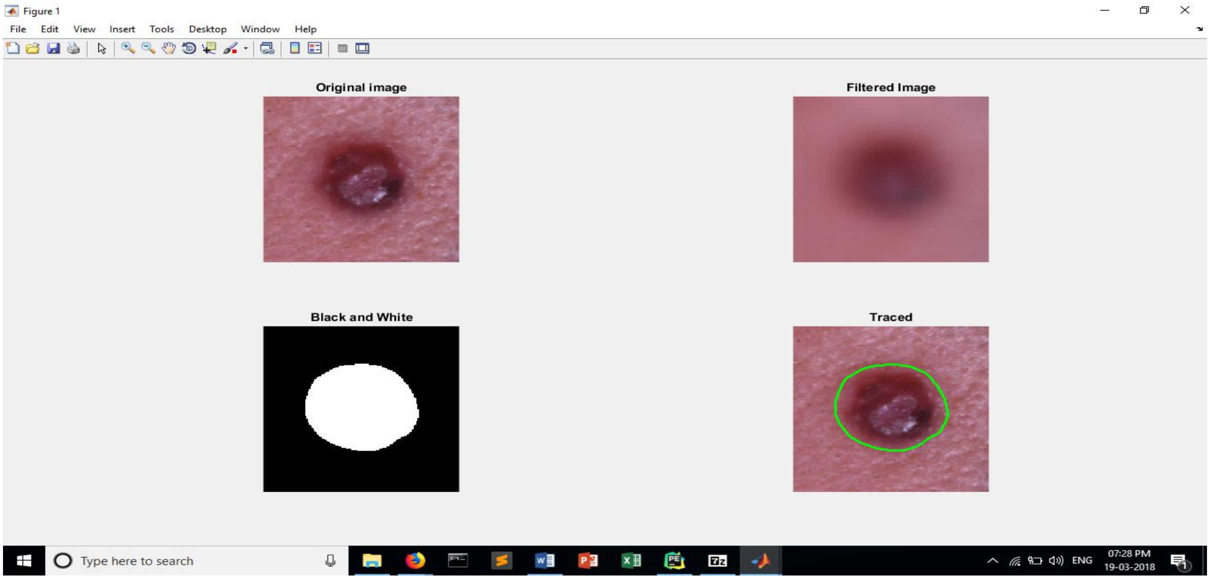
Plot error histogram

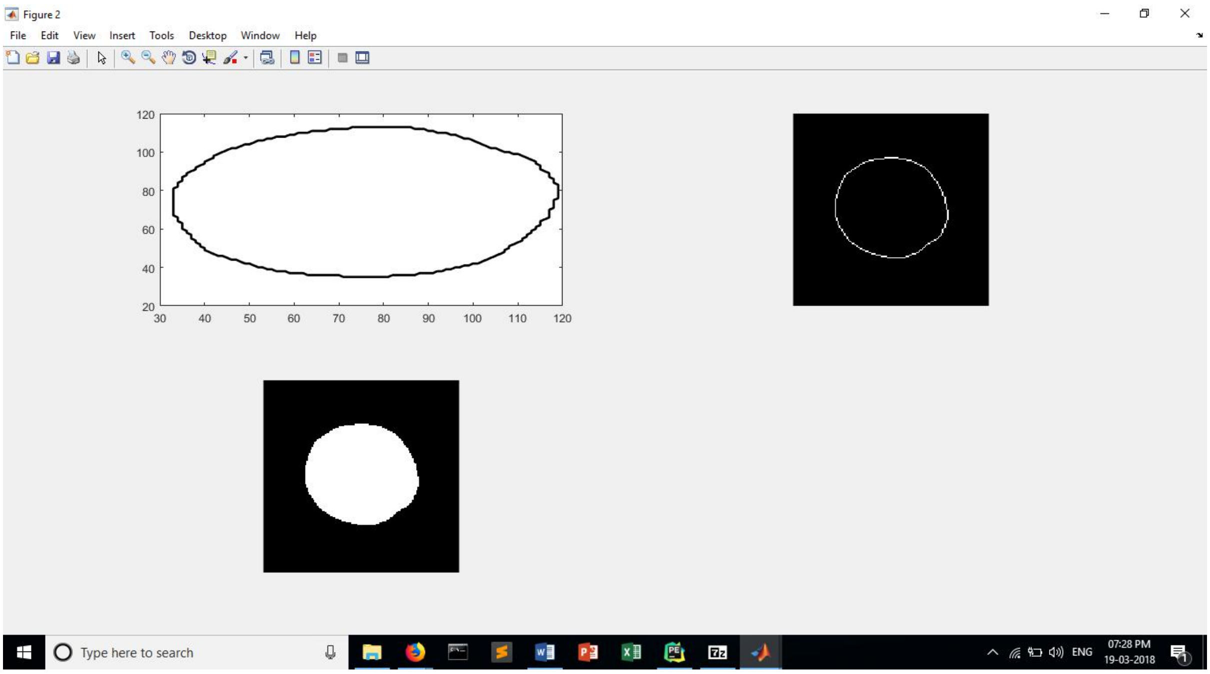


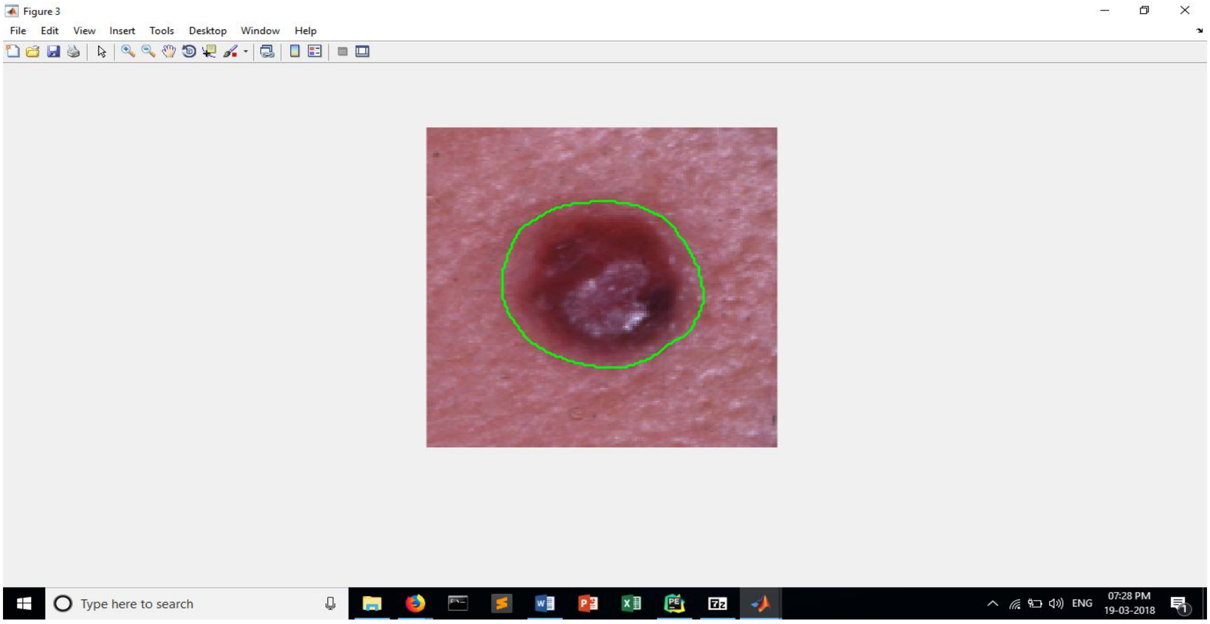
b.)Input image:

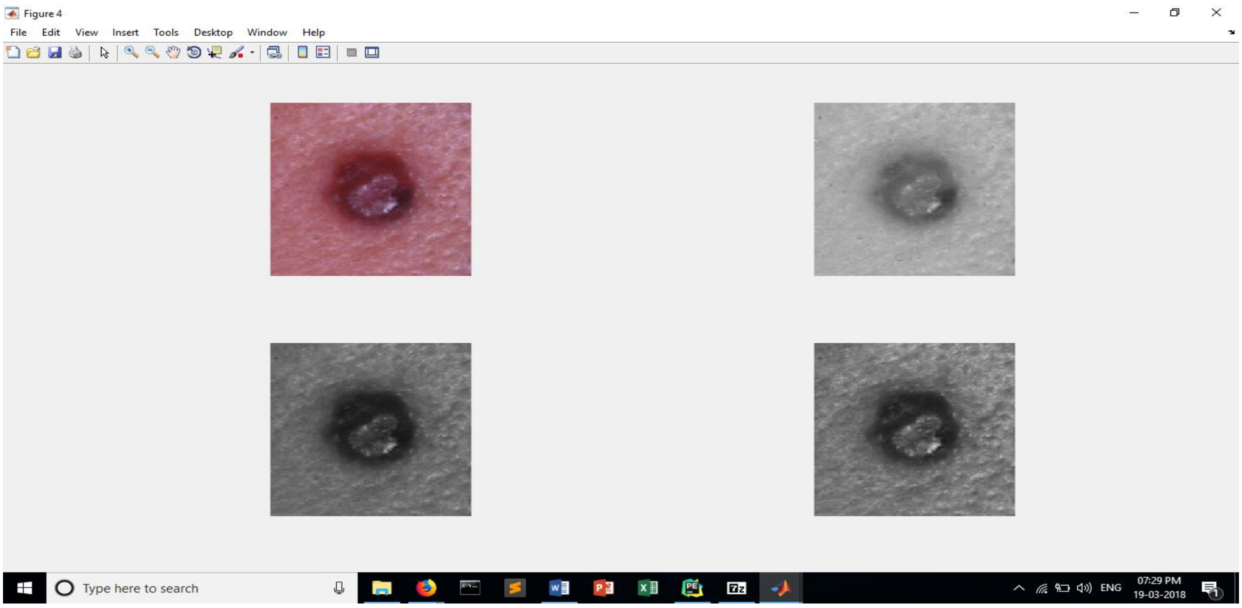


Outputs:









ANALYSIS:

- Asymmetry. **Is** asymmetric
- Border irregularity. The edges **are** ragged, notched, or blurred.

- Color. The color (pigmentation) is **non- uniform**. Shades of tan, brown, and black **are** Present. **It has** a mottled appearance.

- Diameter. The size of the mole is **greater** than 1/4 inch (6 mm), about the size of a pencil eraser..

Thus the skin lesion examined is **Cancerous**

APPENDICES

CODE:

```
clear all
close all
clc
%k parameter can be changed to adjust intensity of image
ei=25;
st=35;
%k=10
k=ei*st;
I = imread('benign1.bmp');
%h=filter matrx
h = ones(ei,st) / k;
I1 = imfilter(I,h,'symmetric');
figure
subplot(2,2,1),imshow(I), title('Original image');
subplot(2,2,2), imshow(I1), title('Filtered Image');
IG=rgb2gray(I1);
%Converting to BW
I11 = imadjust(IG,stretchlim(IG),[]);
level = graythresh(I11);
BWJ = im2bw(I11,level);
dim = size(BWJ)
IN=ones(dim(1),dim(2));
BW=xor(BWJ,IN);  %inverting
subplot(2,2,3), imshow(BW), title('Black and White');
%Finding of initial point
row = round(dim(1)/2);
col = min(find(BW(row,:)))
%Tracing
boundary = bwtraceboundary(BW,[row, col],'W');
subplot(2,2,4),imshow(I), title('Traced');
hold on;
%Display traced boundary
plot(boundary(:,2),boundary(:,1),'g','LineWidth',2);
hold off
% figure
% plot(boundary(:,2),boundary(:,1),'black','LineWidth',2);

nn=size(boundary);
KM=zeros(dim(1),dim(2));
ii=0;
%Create new matrix with boundary points. there fore we can get rid off
%other distortions outside boundaries
while ii<nn(1)
ii=ii+1;
KM(boundary(ii,1),boundary(ii,2))=1;
end
figure
subplot(2,2,1),plot(boundary(:,2),boundary(:,1),'black','LineWidth',2);
subplot(2,2,2),imshow(KM)
%Fill inner boundaries where lesion is located
KM2 = imfill(KM,'holes');
subplot(2,2,3),imshow(KM2)
KM1=xor(KM2,IN);
```

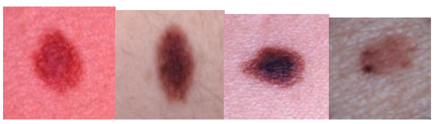
```
% subplot(2,2,4),imshow(KM1)
%Geometrical center
IVx=[1:dim(2)];
IVy=[1:dim(1)];
IMx=ones(dim(1),1)*IVx;
IMy=ones(dim(2),1)*IVy;
IMy = imrotate(IMy,-90);
Koordx=IMx.*KM2;
Koordy=IMy.*KM2;
xmean=mean(Koordx,2);
yc=round(sum(xmean.*IMy(:,1))/sum(xmean));
ymean=mean(Koordy);
xc=round(sum(ymean.*IVx)/sum(ymean));
figure
imshow(I)
hold on
plot(boundary(:,2),boundary(:,1),'green','LineWidth',2);
hold on
plot(xc,1:dim(1),'red','LineWidth',2);
plot(1:dim(2),yc,'red','LineWidth',2);
hold off
% ID=im2double(I);
ID1(:,:,1)=im2double(I(:,:,1));
ID1(:,:,2)=im2double(I(:,:,2));
ID1(:,:,3)=im2double(I(:,:,3));
figure
subplot(2,2,1), imshow(ID1);
subplot(2,2,2), imshow(ID1(:,:,1));
hold on
plot(xc,1:dim(1),'red','LineWidth',2);
plot(1:dim(2),yc,'red','LineWidth',2);
hold off
subplot(2,2,3), imshow(ID1(:,:,2));
subplot(2,2,4), imshow(ID1(:,:,3));

%---------------------------------
```

DATASET USED:

We used the following images:

The benign cancer (non-cancerous)images:



The malignant cancer (Cancerous) images:

The data set was taken from:

http://www.advancedsourcecode.com/melanomaprot.zip

The images we took from this reference have been taken from a verified internet site

CONCLUSION

We all know that skin cancer has multiplied to such an extent that it's very important to detect the disease at its initial stages. In order to solve this issue I came up with the method of image segmentation to detect early sign of skin cancer due to raised concentration in certain parts of the skin. I have used a mat lab code to detect the same and prove its efficiency. For formulating the code I have used the k-means algorithm.

It can be concluded from the network results that the suggested system can be capably used by patients and physicians to diagnose the skin cancer more exactly. This tool is useful for the rural areas where the experts in the diagnosis field may not be applicable. Since the tool is made more feasible and robust for images acquired in any conditions, it can deliver the purpose of automatic diagnostics of the Melanoma Skin Cancer. In future, we could develop a computer algorithm for skin cancer diagnosis using Support Vector Machine, which is also an emerging technology nowadays.

REFERENCES

1. Pehamberger H, Binder M, Steiner A, Wolff K. In vivo epi luminescence microscopy: improvement of early diagnosis of melanoma. J Invest Dermatol,100:356S–62S, 1993.

2. A. Bono, S. Tomatis, and C. Bartoli, The ABCD system of melanoma detection: A spectrophotometric analysis of the asymmetry, border, color, and dimension, "Cancer", vol. 85, no. 1, pp. 72–77, January 1999.

3. Bafounta ML, Beauchet A, Aegerter P, Saiag P. Is dermoscopy (epi luminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests. Arch Dermatol,137:13,43–50. 2001.