GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

NOVEMBER 2016

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

SEGMENTATION OF SKIN CANCER BY USING IMAGE PROCESSING TECHNIQUES

AZHAR KASSEM

MSc. THESIS
DEPARTMENT OF COMPUTER ENGINEERING
PROGRAM OF COMPUTER ENGINEERING

ADVISOR
ASST.PROF. DR. GORKEM SERBES

İSTANBUL, 2016

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

SEGMENTATION OF SKIN CANCER BY USING IMAGE PROCESSING TECHNIQUES

A thesis submitted by Azhar KASSEM in partial fulfillment of the requirements for the degree of **MASTER OF COMPUTER ENGINEERING** is approved by the committee on 15.11.2016 in Department of Computer Engineering, Computer Engineering Program.

Thesis Adviser

Asst. Prof. Dr. Gorkem SERBES

Yıldız Technical University

Approved By the Examining Committee

Asst. Prof. Dr. Gorkem SERBES

Yıldız Technical University _____

Prof. Dr. Nizamettin AYDIN, Member

Yıldız Technical University _____

Asst. Prof. Dr. Cemal OKAN SAKAR, Member

Bahcesehir University _____

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF SYMBOLS

η	Efficiency ratio
σ_B	Between _class variance
σ_G	Global variance
K^*	Optimal threshold
$\oplus$	Dilation
$\ominus$	Erosion
○	Opening
•	Closing

LIST OF ABBREVIATIONS

ADDI   Automatic computer based Diagnoses system for Dermoscopy Images
ANN   Artificial Neural Network
ABCD   Asymmetry, Border, Color, Diameter
BW   Black and White
BCC   Basal Cell Carcinoma
CAD   Computer Aided Diagnosis
CASH   Color, Architecture, Symmetry, Homogeneity
ELM   Extreme Learning Machine
G_T   Ground Truth
GD   Greatest Diameter
GLCM   Gray Level Co occurrence Matrix
GVF   Gradient Vector Flow
KNN   K _Nearest Neighbor
MRF   Markov Random Field
RGB   Red, Green, Blue
SE   Structure Element
SD   Shortest Diameter
SVM   Support Vector Machine
SCC   Squamous Cell Carcinoma
TDS   Total Dermoscopy Score

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

SEGMENTATION OF SKIN CANCER BY USING IMAGE PROCESSING TECHNIQUES

Azhar KASSEM

Department of Computer Engineering

MSc. Thesis

Advisor Asst. Prof. Dr. Görkem SERBES

Skin cancer, if detected early, has a 95-100% successful treatment rate; therefore, its early detection is of vital importance. Skin cancer has diverse types such as melanoma, basal cell and squamous cell carcinoma among which melanoma is the most changeable. Computer aided methods have been developed to assist dermatologists with the diagnosis of melanoma.

A computer based melanoma diagnostic system may provide quantitative and objective analysis of the skin lesions which are instrumental in clinical assessment. Consistent diagnosis capability offered by such a system may also decrease inter-observer variability founded in dermatologist examinations. Additionally this approach automates the analysis, thereby mitigates the amount of repetitive and tedious, hence error-prone, tasks to be done by the physicians.

Many research groups have been developing automatic and semi-automatic computer based strategies to analyze skin lesions and diagnose malignant melanoma. Computer-Aided Diagnostic (CAD) systems need to be very reliable and should demonstrate or achieve high accuracy rates before they can be accepted/used by the clinicians in the daily practice. This thesis study aims at augmenting or expanding existing strategies and develops new techniques to achieve correct, quick and reliable computer based segmentation of malignant melanoma.

Many cascaded stages of operations are required in order to develop a CAD system of melanoma detection; specifically, image acquisition, object segmentation, border detection, feature extraction, and classification. The input to the system is typically skin images obtained using dermoscopy. Dermoscopy images often suffer from artifacts such. as irregular illumination, gel, black frames, ink markings, rulers, air bubbles, yet other naturally occurring structures like blood vessels, hairs, skin lines, and texture can

adversely affect the performance of border detection algorithms. Thus, some preprocessing steps are required to facilitate the segmentation process, such as removal of hairs successfully without lesion distortion.

In this thesis, the first issue tackled is the problem of hair removal. Here we have used two methods, namely the Dull-Razor method for benign melanoma lesions, and morphological bottom-hat filtering for malign melanoma lesions, which yielded inferior results compared the previous method.

The second challenge in this study was to develop an accurate border detection method within segmentation stage. Here we have first used thresholding (Otsu's method) that partitions the input image into foreground and background. Then, the result of this stage, which is a black-white or binary image, is fed into the next step that finalizes the lesion segmentation by tracing the border of lesion. After lesion segmentation, one can obtain the border of lesion, measure its irregularity, measure the area and color of the lesion and so on. These constitute the features that can be used in a machine learning or object classification setting so that the type of the lesion, as malign or benign, can be decided automatically. As our aim was to concentrate on lesion segmentation, the issue of classifier design for melanoma classification, is not dealt with in this study.

We have considered/taken dermatologists border marks as the gold standard (i.e. the true state of the nature) and compared our segmentation results against these results provided in the database, which are based on the decisions of expert dermatologists. Our results show that with our proposed method we can detect melanoma lesions with high accuracy. All the code in this thesis has been developed in MATLAB and the dataset was obtained from (ADDI Project, Automatic computer based Diagnoses system for Dermoscopy Images).

Key words: Basal and Squamous, Melanoma, Dullrazor, Segmentation, ABCD algorithms, Computer-Aided System.

ÖZET

GÖRÜNTÜ İŞLEME TEKNİKLERİ KULLANILARAK CİLT KANSERİ SEGMENTASYONU

Azhar KASSEM

Bilgisayar Mühendisliği Bölümü

Yüksek Lisans Tezi

Tez Danışmanı: Yrd. Doç. Dr. Görkem SERBES

Cilt kanseri erken teşhis edilirse, %95-100 oranında başarılı tedaviye sahiptir, bu nedenle erken teşhisi hayati önem arz eder. Cilt kanseri; Melanom, bazal hücreli ve skuamöz hücreli kanser gibi çeşitli türlerde bulunmaktadır ve bunların içinde en değişkeni Melanom'dur. Bilgisayar destekli yöntemler dermatologlara Melanom teşhisinde yardımcı olmak için geliştirilmektedir.

Bilgisayar tabanlı melanom teşhis sistemi ile deri lezyonlarının nicel ve tarafsız bir şekilde analizini yapmak mümkün olmaktadır. Böyle sistemlerin sağladığı tutarlı teşhis imkânı ile farklı dermatologların teşhisindeki değişkenlik azaltılabilecektir. Ek olarak, bu yaklaşım analizi otomatik hale getirecek ve böylece hekimler tarafından hataya açık tekrar eden sıkıcı işlemleri azaltacaktır.

Melanom teşhisi için otomatik ve yarı otomatik sistemlerin geliştirmesi konusu hakkında çalışan bir çok araştırma grubu bulunmaktadır. Bilgisayar destekli teşhis (BDT) sistemlerinin hekimler tarafından günlük hayatta kullanılabilmesi için çok güvenli ve başarılı hale getirilmeleri gerekmektedir. Bu tezde var olan stratejileri iyileştirerek ve yeni teknikleri de kullanarak hızlı, doğru ve güvenli bilgisayar destekli melonam segmentasyon sistemi geliştirmek amaçlanmıştır.

BDT sistemlerinin geliştirilmesi için bir biri ardına sıralı birçok basamakta işlem yapılmalıdır, örneğin görüntü edinimi, obje segmentasyonu, sınır belirleme, öznitelik çıkarımı ve sınıflandırma gibi. Sisteme girdi olarak dermoskopi kullanılarak elde edilmiş deri görüntüleri verilmektedir. Dermoskopik görüntüler çoğu zaman çeşitli yapay bozulmalara uğramaktadır, örneğin düzensiz ışıklandırma, jeller, siyah

görüntüler, mürekkep lekeleri, işaretleyiciler, hava kabarcıkları gibi. Ayrıca birçok doğal bozulma çeşidi de melanomun bulunduğu alanın sınırlarının tespitini zorlaştırmaktadır, örneğin kan damarları, tüyler, deri çizgileri ve dokudaki değişimler gibi. Bu sebeple, melanom segmentasyon başarısını arttırmak için tüylerin görüntüden temizlenmesi gibi ilgili görüntüyü bozmayan ön-işlemler yapmak gerekmektedir.

Bu tezde, ilk ele alınan problem tüylerin giderilmesidir. Bu amaçla iki yöntem kullanılmıştır, iyi huylu melanom lezyonlarında Dull-Razor yöntemi ve kötü huylu melanom lezyonlarında ilkine göre daha kötü sonuçlar veren morfolojik alt-şapka filtreleme yöntemi kullanılmıştır.

Tezde ele alınan ikinci problem segmentasyon aşamasında lezyonun sınırının belirlenmesidir. Bu aşamada görüntüleri ön-plan ve arka-plan olarak ikiye ayıran eşikleme (Otsu yöntemi) yöntemi kullanılmıştır. Daha sonra sadece siyah ve beyaz piksellerden oluşan bu aşamanın sonuçları lezyon segmentasyonunda sınır belirleme için kullanılmıştır. Lezyon segmentasyonundan sonra, kişi lezyonun sınırını, düzensizliğini, alanını, rengini ve benzeri özelliklerini elde edebilir. Bu değerler lezyonun iyi veya kötü huylu olup olmadığının belirlenmesi için uygulanacak sınıflandırma aşamalarında kullanıla bilinir. Bu çalışmada sadece lezyonun sınırlarının tespiti ile ilgili işlemler yapılmış, melanom sınıflandırma problemine değinilmemiştir.

Dermatoloji uzmanlarınca belirlenen sınırlar altın standart olarak belirlenmiştir ve elde ettiğimiz segmentasyon sonuçları bu altın standart ile karşılaştırılmıştır. Sonuçlar önerilen yöntemin melanom sınırlarını yüksek doğrulukla belirlediğini göstermektedir. Tezde kullanılan tüm kodlar MATLAB ortamında geliştirilmiştir ve kullanılan veri-seti (Dermoskopik Görüntüler için Otomatik Bilgisayar Tabanlı Teşhis Sistemi) adresinden alınmıştır.

Anahtar Kelimeler: ABCD algoritma, Bazal ve Skuamöz, Bilgisayar-Tabanlı Sistem, DullRazor, Melanom, Segmentasyon.

INTRODUCTION

1.1 Literature Review

A skin cancer is of extensive cancers spread among humans, especially in the last two centuries, where recent studies have shown that there is approximately (3.5) million a type of skin cancer posed to workers under the sun because of their working conditions, which makes it imperative for them exposure to the sun for long hours, such as farmers and construction workers, as well as people who are trying to get a tan skin are also susceptible to the disease as. The risk of this disease has encouraged a lot of health organizations and researchers in dermatology to cope with this disease and reduce its spread, through the two routes:

The first aspect: is to educate people not to over-exposure to the sun's rays through posters and television commercials or guidance sessions.

The second aspect: is the expansion of the field of medical research and detection of various types of skin cancer and cooperation with researchers in the field of computer vision and image processing by providing adequate information and images about the disease to encouraging many researchers to work in the field of discover skin cancer and to propose several diagnostic systems, and methods for the detection of skin lesion.

In 1997, number of researchers in the United States from several universities proposed a method for segmentation skin cancer images and other pigmented lesions cancer.

First, minimize a color image into an intensity image and segments the image by using thresholding, then smooth the segmented image by using image edges. Which are also used, to the initial boundary, a closed elastic curve is fitted, and is locally expanded or shrunk to approximate edges in its neighborhood in the area of interest. Then by using the double thresholding to obtain an image area where a lesion boundary occurs, to

centralize the boundary in that area. They used 20 images selected randomly; the results showed that the average error is approximately the same as that has obtained by expert manually segmenting the images [1].

At the same year numbers of researchers from (Cancer Control Research, and British Columbia Cancer, et al). They presented method to remove hairs from an image by using a preprocessing program is called DullRazor. The preprocessing steps to enhance the image in order to achieve satisfactory results in the segmentation step. DullRazor software can be downloaded from [2].

In September 1998, H. Handels, et al. they presented a new method to support computer diagnosis of skin tumors in dermatology. The first step, using co-occurrence matrices to extract features, this features based on Fourier features and fractal features, this feature has described as an optimization problem compared with several approaches like heuristic strategies, greedy and genetic algorithms. The classification rate of the nearest neighbor classifier computed with one-out method is used. Genetic algorithms give the best results. Finally, neural networks with error back-propagation as learning model are trained using the selected feature. Different network topologies are investigated to achieve the classification performance of the neural classifiers. Performance of 97.7% is achieved [3].

According to the method that proposed by M.I. Rajab and M.S. Woolf son, in April 2003, they proposed two approaches to the skin lesion image segmentation problem.

The first one is region-based segmentation method and determines an optimal threshold iteratively by using iso data algorithm. Second method is based on neural network edge detection and a rational Gaussian curve that approximately fits closed elastic curve between the recognized neural network edge patterns. The comparison between the two methods demonstrate that for lesions with a range of different border irregularity properties, the first method provides the best performance over a range of signal to noise ratios and have similar performance when tested on real skin lesions [4].

In June 2004 (Snakes Bulent Erkol1, Randy H. Moss and et al) they published article about detected the border of skin lesions in dermoscopy images using the gradient vector flow (GVF) and a color histogram analysis technique, the number of images that used are 100, that is dividing into two groups, 70 benign and 30 melanoma skin lesion images.

The total errors obtained by the GVF-based method are lower for both melanoma and benign images sets comparison with manually segmented lesions that has been determined by a dermatologist [5].

Another algorithm has been proposed for segmentation of skin cancer images by automatically determines the colors of the lesion, and computing the CIEDE2000 distance in the L*a*b* color space then extracting the energy of some statistical moments of the L* component of the image. Irene Fondón and et al are working in this method in 2007. They are using region growing segmentation, growth parameter depends on the variance on each distance image. The actual grown region is initiated to decide the optimum values of the parameter, the dataset of 20 images are used that achieve excellent results [6]. In the same year, each of (M. Emre Celebi1, and et al) tried to detect border in dermoscopy skin lesion images by using a method of an unsupervised approach depend on an adapted version of the JSEG algorithm by testing a set of 100 dermoscopy images as ground truth that used in this method, the borders have been determined manually from a dermatologist. The results have been compared with the result of another dermatologist and three other automated methods, they achieved the fast and accurate border detection in dermoscopy images [7].

Another research has been presented in Feb 2013, involves a brief outline about segmentation techniques, the most common like thresholding, Model based, Edge detection, Clustering etc., and the comparison between the results of these techniques, they found that Markov Random Field (MRF) is the strongest method of noise cancellation in images, but still thresholding is common and simplest technique for segmentation [8]. In the same year, number of other researchers used ABCD rule Dermoscopy technology to diagnosis the melanoma skin lesion, they did several steps to detect melanoma like:

- Image Acquisition

- Preprocessing

- Segmentation image

- Selection and extraction features

- Classification methods.

The ABCD method includes, Symmetry detection, Border Detection, Color, and Diameter detection to calculate TDV (total dermatoscopy value) then, the value that obtained decides if the cancer is melanoma or benign [9]. Also, In February 2013 another researcher from Bangladesh did the analyzing for different digital images, this analyzing based on unsupervised segmentation techniques, after that, feature extraction techniques were applied on those segmented images. The discussion has based depending on the obtained results [10].

In July 2013 (Nilkamal S. Ramteke1 and Shweta V. Jain, and et al from College of Engineering & Management Nagpur, INDIA) they did enhancement the images by using Wavelet Transformation and decomposition, then calculating diameter of lesion by using histogram analysis, the segmentation techniques is watershed, border detection, decision related with structural nature of lesion, then by using the ABCD rule of cancer diagnostic and calculating TDS(total dermoscopy score) in order to differentiating between the non cancerous(benign), suspicious and cancerous (melanoma) images[11].

In February 2014 from the department of (Computer Science and Biology, Monta Vista High School, Cupertino, CA, USA) many student have been proposed method to melanoma diagnosis after enhanced image processing and segmentation, then features extraction using new and improved techniques. This feature has feed automatically to a multistage classifier such neural network which achieved greater than 93% specificity and greater than 97% sensitivity. They found the lesion images online which it is tested with the trained system have the same sensitivity. Finally, a new approach was discussed by using Dermlite DL1 dermatoscope that can be connected to the iPhone. After taking the lesion image, the physician gets the diagnosis system with a few simple clicks. This new technology could have widespread divergence on melanoma diagnosis, which is achieves higher sensitivity and widespread on melanoma diagnosis and provides an easy to use iPhone app to detect melanoma in early stages without the need for biopsy [12].

In July-2014, simple algorithm has been proposed to detection of skin cancer by finding shape, and infected area. The first stage is a preprocessing to remove the noise by using median filter. The second stage is segmentation by using two methods, thresholding and fuzzy c-means. The third stage is feature extraction by (GLCM) Gray Level Co-Occurrence Matrix and contour signature [13]. At the same time the Education Society

Institute of Technology presented a computer aided method for the detection of Melanoma Skin Cancer by using the tools of Image processing. The input image is the skin lesion image and then by applying sequence of image processing techniques, analyses the image to detect skin cancer. Those tools to checking different parameters of melanoma like Asymmetry, Border, Color, and Diameter, (ABCD) etc... The feature extraction parameter are used to classify the image as benign or melanoma cancer lesion [14].

In 2015 Dr. Poulami Das, Ms. Mayurika Gangopadhyay Assistant Professor, Dept of Computer Science and Engineering, they proposed a method to hair removal form grayscale dermoscopic images is based on mathematical operations, by using a circular mask for removal of non-skin element and then reset the image by normalizing the pixel values. In the case of (RGB) images, based on histogram values, frequency of incident is measured through individual components. Neighborhood operation based on minimum distance between each component is the key strategy. This method preserves the border properties and it is applied on enormous number of image and it produced suitable results [15]. In the same year many researcher from India, they proposed the segmentation techniques based on the mean shift based fuzzy c-means formula which needs less machine time comparing with other techniques and obtaining perfect results for segmentation. The mean shift will closely and quickly realizes cluster centers; the strategy is able to detect regions inside image [16].

By taking advantage of the previous methods carried out by many researchers in the field of image processing techniques to detect skin cancer, in this study it was followed some of the methods used to improve the image, hair removal process is done by using the Dullrazor method as well as the application of another way, a hair removal through use morphological Bothat filter. As for segmentation, by benefiting from the simplest thresholding (Otsu's method), and improving method of tracing boundary of the lesion, calculation some geometric properties like area and perimeter for region of interest in order to extraction and verification the disease.

1.2   Aim of the Thesis

This thesis aims at augmenting or expanding existing strategies and develops new techniques to achieve correct, quick and reliable computer based diagnosing of malignant melanoma, and build a diagnostic system helps dermatologist in determining

and classification of skin cancer, to benign or malignant depending on image processing techniques and new developments in this field. Also, the clinical diagnosis is depending on the dermatologist opinion which may be suffers from many mistakes. This issue highlights the need for receiving other second opinion so this system helps to increasing the accuracy of diagnosis, to decreasing the number of false excision of benign lesions, to speeding the discovery of the disease, in addition to reducing costs imposed on individuals by unnecessary surgeries.

1.3   Hypothesis

In this thesis, the implementation diagnostic system steps, this includes enhancement the input image.

Hair has been removed from the image of the skin through the use of Dullrazor method and bottom hat filter then segmentation of the image, the purpose of segmentation to separate the lesion from the rest of the skin using thresholding (Otsu's method), identify the (ROI) region of interest by tracing the boundary and then determined the area and perimeter of the lesion and find the error percentage between our result and the G_T.
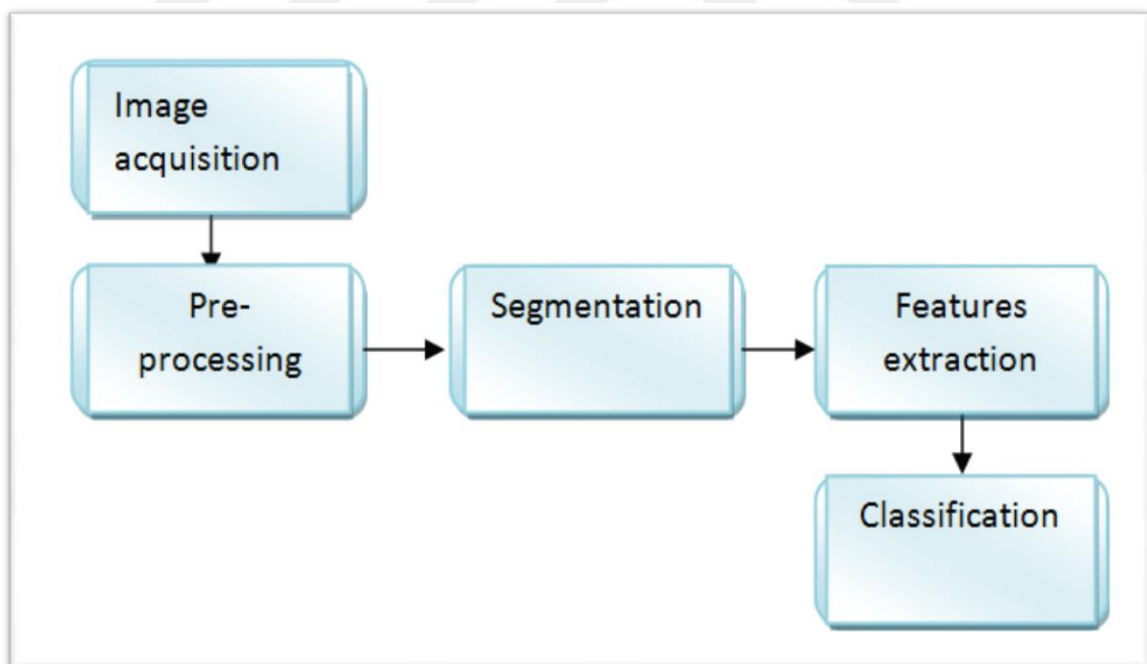
Figure 1.1 Computer aided diagnosis (CAD) system for skin lesions

GENERAL INFORMATION

Skin cancer is the most common cancers in the United States and is most commonly comparison with other cancers, such as colon, breast and prostate cancer. There are approximately 3.5 million type of skin cancer, but there are three types are dominant, which was diagnosed as a 100% state of skin cancer and these species are, basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and malignant melanoma. The latter one is the most dangerous types of skin cancers that lead to death, although the incidence is 5% of the total percentage of diagnosed. May be any one do not know how skin cancer consists, and how it is growth.

Cancer starts in cells in our body. Cells are tiny building blocks which make up the tissues and organs of our bodies. Those cells are divide in a controlled way to produce new cells, and it receives signals from the body telling them when must be divide and grow, or when to stop growing, so when it can't be repaired, it gets a signal to stop working and die. Cancer is developing when the normal cell working wrong and become abnormal. These abnormal cell dividing, produce more and more abnormal cells figure (2.1). These cells can be either benign or melanoma [17]. Benign is not cancerous may grow in the body but not spread to anywhere so it is only causes problems if it puts pressure on nearby organs. Malignant is a cancer can grow into nearby tissue. Cancer cells spread from the first started (the primary site) to other parts of the body. So it can travel through the blood or lymphatic system and when it reaches another part of the body, it begins to grow and form another tumor (secondary cancer or a metastasis). In this chapter we offer a brief background concerning the skin structure, formation of melanoma cancer and also the clinical diagnostic approach taken by dermatologists. This is followed by an abstract of existing CAD systems of malignant

melanoma, including image process and machine learning that based mostly on those systems.
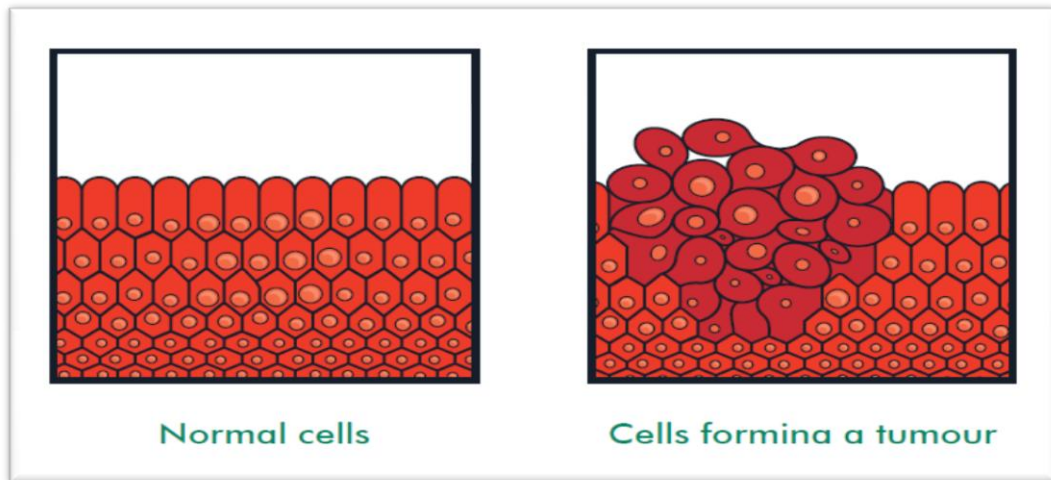
Figure 2.1 Normal and up normal cell [17]

2.1 The Skin Structure

Skin is a largest organ and the first line of defense for a body's that has many purposes, such us:

- Protects the body from infection
- Helps to organize body temperature
- Helps to control fluid loss
- Through the sweat glands. Skin gets rid of waste substances.

The skin is divided into two main layers. The outer layer is (epidermis), and the underneath is the (dermis). Below these, layer of fatty tissue which is deeper. The epidermis contains three types of cells, it is filled by cells known as squamous cells, a rounder cells are founded at the base of the squamous cells known to us a basal cells, in between the second type of cells are other cells called melanocytes, which is responsible of produce the pigment melanin. This pigment gives skin its color. The two main layers are shown in the figure (2.2).
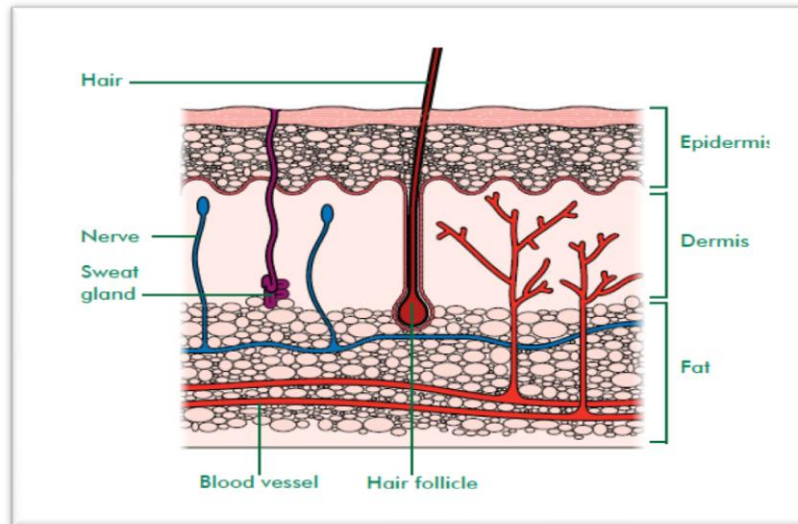
Figure 2.2 Structure of the skin [17]

2.1.1 Types of Skin Cancer

There are three types of skin cancer:

2.1.1.1 Basal Cell Carcinoma (BCC)

A cancer of the basal cells is:

- At the bottom of the epidermis.
- It's very common.
- Very slow-growing
- Almost never spread

But some BCCs are aggressive, and may spread into the deeper layers and sometimes to the bones [17].

2.1.1.1.1 Symptoms of (BCC)

Basal cell cancers may:

- Be smooth and pearly
- Look waxy
- Appear as a firm, red lump
- Bleed sometimes
- Develop a crust or scab
- Begin to heal but never completely heal
- Be itchy
- Look like a flat, red spot that is crusty and scaly

- Develop into a painless ulcer [17]. Figure (2.3) shown images of (BCC).

Figure 2.3 Images of Basal Cell Carcinomas on the Nose [18]

2.1.1.2 Squamous Cell Carcinoma **(SCC)**

A cancer of the keratin cells in the outer layer

- Slow-growing
- If they are left untreated it spread to other parts
- More aggressively [17].

2.1.1.2.1 Symptoms of (SCC)

Squamous Cell Cancers may:

- Look scaly
- Have a hard, crusty scab
- Make the skin raised in the area of the cancer
- Feel tender to touch
- Bleed sometimes [17]. Figure (2.4) shown image for elderly man with badly sun damaged skin has SCC on his face.

Figure 2.4 Squamous Cell Cancer [18].

2.1.1.3 Melanoma

Melanoma is a more dangerous type of carcinomas as it appears in figure (2.5). Its incidence and mortality rates have increased steady worldwide over the past many decades. The seriousness of melanoma rate spread depends largely on thickness and stage of melanoma.

If the abnormally created melanin remained within the epidermis, melanoma is named unchanged. At this stage, it's not dangerous and early detection with immediate surgical excision for the lesion, leading to treatment. However, the first designation of melanoma isn't very simple and there are continuously equivocal cases, the optical and visual properties of melanoma score helps to quickly treated it. When malignant melanocytic penetrate into the dermis leading to changes in skin coloration. This results in advanced melanoma which can be incurable. The characteristics of melanoma are:

- Grow quickly

- Needs to be treated early.

- Rare in people with dark skin [17].

2.1.1.3.1   Symptoms of Melanoma

- Grow into nearby tissue.
- Cancer cells spread to other parts
- Travel through the blood or lymphatic system [17]. Figure (2.5) shown the melanoma image.
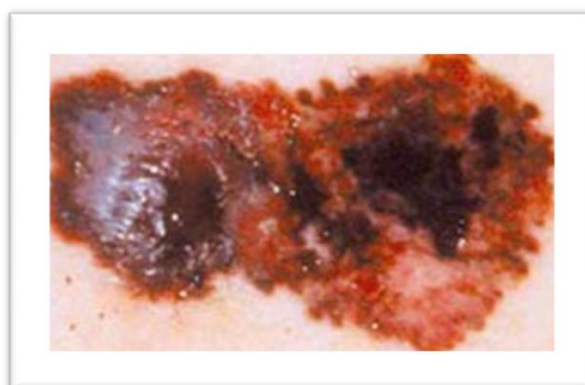
Figure 2.5 Melanoma Images [18]

2.2   How Skin Cancers are diagnosed

Most people, who have suspected of skin cancer, are going to a specialist in their hospital for advice and treatment. A dermatologist is a doctor who specializes in

treating skin diseases, and he is able to learn a lot from a simple examination of the unnatural area of skin. A dermatologist may use an instrument called a dermatoscope, he may be advice to take a tissue sample (biopsy), in this procedure by removing all or part of the affected area, and then send it to the laboratory. A pathologist will be examined the biopsy under a microscope, this is a clinical procedure. After the great development in field of computer science and it was use in a number of medical fields. In addition to the developing in the field of photography and the use of high resolution cameras, which led to build diagnostic systems helps the dermatologist in the diagnosis of skin lesion . Owing to the importance of early and accurate diagnosis of skin cancer, a lot of efforts have made to enhance the clinical diagnosis in the field of dermoscopy technique and digital monitoring; pattern analysis is one of these techniques that contain ABCD rule of dermoscopy, Menzies method, 7-point checklist, and the CASH algorithm. More details about these techniques are presented in the following [19] [20].

2.2.1   7-Point Checklist

In the 7-point checklist, we have three major criteria, include atypical pigment network, atypical vascular pattern, and vascular pattern, each of them have a maximum score of 2. Also we have four minor criteria, include irregular streaks, irregular pigmentation, irregular dots or globules, and regression structures. A minimum total score required is 3 for the lesion to be diagnosed as melanoma [20].

2.2.2   Menzies Method

In this method, two separate groups are defined: the first one consisting of negative features and the second one consisting of positive features. The first group includes axial symmetry of pigmentation with any presence of a singular color. The second group includes blue-white veil, radial streaming, multiple brown dots, peripheral black dots, multiple blue/grey dots, and scar like pigmentation, pseudo pods, multiple colors (5 to 6), and a broadened network. When diagnosing melanomas, the first method must not be missing and at least one must be found from positive feature [20].

2.2.3 CASH Algorithm

The CASH algorithm is a simplified version aimed to be suitable for less experienced testing, which examines

- Color (C)
- Architecture (A)
- Symmetry (S)
- Homogeneity (H).

The accuracy of this method is comparable to other melanoma distinguishing algorithms [20].

2.2.4 ABCD Algorithm

The most commonly used algorithm, the ABCD's method of detection was developed in 1985 to aid in self-diagnosis and monitoring the malignant melanoma [21]. It is based on:

- Asymmetry (A)
- Border (B)
- Color (C)
- Diameter (D) measuring.
- Evolving.

Total Dermoscopy Score

$$TDS = [(A \times 1.3) + (B \times 0.1) + (C \times 0.5) + (D \times 0.5)] \qquad (2.1)$$

Figure (2.6) is illustrated the steps of ABCDE.
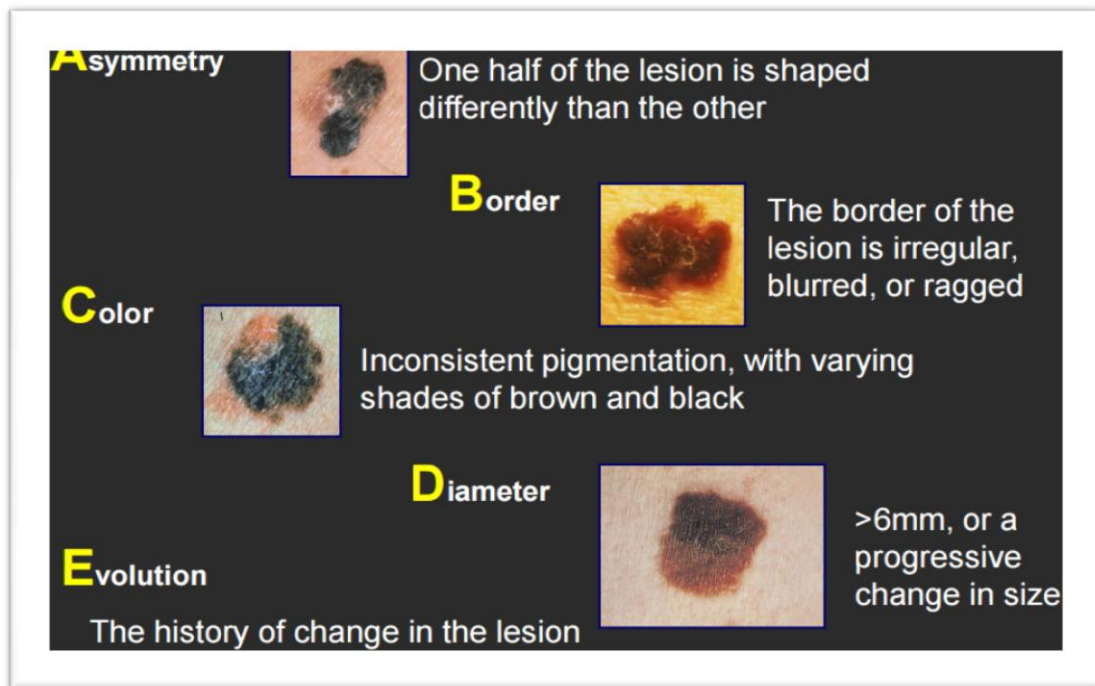
Figure 2.6 ABCDEs of detecting and diagnosing Malignant Melanomas [22]

2.3   Computer Aided Methods of Diagnosing Skin Cancer

The usual clinical practice of melanoma diagnosis is a visual examination by dermatologist. Clinical diagnostic accuracy is a bit substandard [20, 21]. Diagnosis or diagnostics is the process of identifying a skin texture or problem by its signs, symptoms and the result of various diagnostic procedures. The result of these processes is called a diagnosis. The diagnosis system is a system which can be used for analyses any problem by answering some questions which lead to a solution to the problem. The skin cancer detection system is a system to identify and recognize skin cancer symptoms and diagnose it in early stages, according to the above definitions there are some programs have been developed with a combination of the clinical methods that mentioned above to help dermatologists diagnose cancer in skin lesions. The dermoscopy images are an indispensable tool for imaging programs, using different optical methods, multi spectral imaging segmentation, classification, and pattern analysis, those methods are exploited with variables specified by the ABCD algorithm [22]. CAD systems been working with clinicians, dermatologists, to achieve the desired goals, this system can be divided into three steps:

14

- Determination of the tumor area from dermoscopy image
- Extraction of image features
- Building of the classification model and evaluation [20].

The field of dermoscopic image processing is constantly growing and improving. The ultimate goal is to increase accuracy of diagnosing skin cancer in its earliest stages. CAD system generally consist five main components;

Image acquisition, image enhancement, image segmentation (or border detection), feature extraction, feature selection and classification. A summary for these methods were explained in the following sections.

2.3.1   Image Acquisition

It is the primary step. Dermoscopy images are essentially digital photographs/images of enlarged skin lesion, which taken by using typical camera equipped with special lens extension. The lens connected to the dermatoscope acts sort of a magnifier scientific instrument with its own light source that illuminates the skin surface equally. Because of source of illumination integrated into dermatoscopy lens, there happens to be problem with skin reflections, to overcome this problem, liquid is employing as a medium layer between the lens and also the skin. Dermatologist can produce correct documentation that is known the ground truth (G _T) of gathered images, smoothed a path for computer analysis which will later used to classify the input images [14].

2.3.2   Image Enhancement

Image acquisition is the process which done by high resolution cameras and with the possibility to enlarge the image (in the range of 10 times) and special equipment for this purpose and the use of certain materials, such as gel and oil on the patient's skin to reduce the reflection of light on the camera lens. Therefore, some images appear containing various artifacts like ink marking or the presence of a ruler, gel, air bubble, vessel blood and hair. All of these signs is possible to deal with and removed it by using one of the methods of image enhancement, but the biggest challenge is the human body, which it covered by hair and it is different from one person to other, there are dark and thick hair and there are the same skin color and thin. Therefore, some images where disease region, is entirely covered with hair. The hair removal by shaving process, it's

possible to hurt the patient, especially if the illness is the kind of bleeding or squamous. The new developments in the field of image processing have produced a number of methods of hair removal. The accuracy of results is different. In this thesis, hair removal was done by using Dullrazor method which will be explained in the next chapter. Human hair covers the whole body and has a range of different colors, textures, and orientations, skin lesions images always blocking by hair [2]. In dermoscopic images, there are many types of hairs, small and light as well as thick and dark, short or long, due to this, hair can lead to distortion the main information when working with a skin lesion image. Despite the development and fast growth in the field of image processing for dermatological applications, this specific issue had never been completely addressed until the publication of DullRazor by T. Lee (T. Lee, 1997) [2]. Before DullRazor, a few solutions were presented to reform this problem. The simplest way, by removing all images containing hair from the data set, this action causes to reduce the general usage of the algorithm. Simply hair shaving physically is another solution, but shaving can be painful and dangerous process that can cause bleeding, or tissue damage, also this action lead to distortion the features of the lesion that causing errors in detection and diagnosis. Another solution, by go manually through each image and discover each border (hair and lesion). All of these actions still used until DullRazor was published in 1997, that was quite used for hair detection and removal. Dull Razors software is introduced to digitally remove hairs from dermoscopic images. The idea in developing this software was to remove dark thick hairs, but it cannot remove thin hairs, it consist three steps:

- Detecting hair pixels
- Replacing such pixels with non-hair pixels (interpolation)
- Smoothing the final result.

Two disadvantages:

- It can only remove dark hairs from an image.

- It is rather slow.

(Median filtering + Dull Razor)=Remove thin hair

Median filtering for images with thin hairs, and Dull Razors for thick hairs

Using grayscale morphological closing operations in each of the three color band images (Red, Green, and Blue) allows for smoothing of low intensity values, (binary

masks 0° (horizontal), 45°(diagonal), and 90° (vertical). Once the binary hair mask of the image is obtained for each color band interpolation is done across the hairs to replace the hair-pixel with an interpolation for the surrounding non-hair pixel values. An adaptive Gaussian filter is also used to smooth out thin lines left from interpolation and hair removal. The steps of DullRazor technique are shown in the figure (2.7).

Figure 2.7 Steps of hair removal by using dullrazor

From the results it was reveal that three structures element SE, all in different directions, are adequate to detect all of the dark hairs. The structure elements consist of three image masks in three directions of 0° (horizontal), 45° (diagonal), and 90° (vertical), as seen below in figure (2.8) .

Figure 2.8 Structure elements for DullRazor for the generalized closing operation. (a) Horizontal structure element, (b) diagonal structure element, and (c) vertical structure element

Let G_r be the generalized grayscale closing image of the original image red band, O_r. S_0, S_{45}, and S_{90} will be the structure elements presented in figure (2.8), where the variance corresponds to the angle of rotation. This allows Gr to be represented as [18]

$$G_r = |O_r - \max\{O_r \bullet S_o, O_r \bullet S_{45}, O_r \bullet S_{90}\}| \qquad (2.2)$$

Where $\bullet$ is the closing operation. The binary mask of the image (x), is then threshold with a predefined value at the location of (x) by computing [18]

$$M_r(x,y) = \begin{cases} 1, & if \ \ G_r(x,y) > T \\ 0, & if \ \ otherwise \end{cases} \qquad (2.3)$$

This is then repeated for the green and blue band images with the final hair mask for the original image, M given by the union of all three masks.

$$M = M_r \cup M_g \cup M_b \qquad (2.4)$$

Where, Mr, Mg, and Mb are parallel hair masks for corresponding color bands. Once the binary hair mask of the image is obtained for each color band interpolation is done across the hairs to replace the hair-pixel with an interpolation for the surrounding non-hair pixel values. Because of the partial shadow effect (darker pixels around the hair caused by the shadow of the hair on the skin) exact border location of the hairs are difficult to determine, to keep away from this issue, neighbors pixels that chosen for interpolation be 11 pixels away from hair border. A Gaussian filter is also used to smooth out thin lines left from interpolation and hair removal. The second method for hair removal is a morphological bottom hat filter. The principle application of this filter is removing unwanted objects from an image by using a SE. The bottom-hat filter is used for removing dark objects over a light background [34].

Resulting images will be entered to the stage of separating disease from the remainder skin, this process called segmentation.

2.3.3 Segmentation

Segmentation is most important steps in cancer detection. It is the process of partitioning an image into groups of pixels which have same properties and criteria. Segmentation methods can be generally divided into the following categories:

- Histogram thresholding: This method is using to determining one or more histogram values that separate the object from the other parts of image.
- Clustering: This method is using to partitioning of a color (feature) space into regions that have the same properties (identical) by using clustering algorithms.
- Edge-based: This method is using to detect edges between regions by using edge operators.

- Region-based: This method is using to grouping pixels into identical regions by using region merging, region splitting, or both.

- Morphological: This method is using to detecting object's contours from preset seeds by using the watershed.

- Model-based: This method is using to modeling images as random fields in order to determining their parameters by using different procedures to optimization results.

- Active contours: This method is involved (snakes and their variants) is using to detecting object's contours by using curve evolution techniques.

- Soft computing: This method is used for classification of pixels by using soft-computing techniques including neural networks, fuzzy logic, and evaluation computing. Several issues should be taken in account when choosing a segmentation method:

- Scalar vs. vector: Most segmentation methods are designed to deal with scalar images. Whereas numerous vector methods for image segmentation have been developed in the last decade, these methods are hindered by many factors including computational time requirements also choosing a suitable color space.

- Automatic vs. semi-automatic: Many of segmentation methods are automated, while others are required human interaction. Such method like active contour methods is requiring the manual assigning for initial contour, whereas another method like seeded region growing is requiring to specifying initial region seeds.

- Number of parameters: Most segmentation methods need many parameters because its values need to be determined a priori.

In our case, the proper segmentation is difficult to determine, because of the great verities of the lesion shapes, sizes, and colors along with different skin types and textures. In addition, some lesions have irregular boundaries and in some cases there is smooth transition between the lesion and the skin [18]. Our goal is to achieve a practical, simple and easily solution to the problem of border detection in dermoscopy images, and avoiding complications as far as possible. In this thesis we detected the border of the lesion by benefiting from the basic image processing technique of thresholding. The Otsu's method is adopted for determining the threshold in our

proposed method for segmentation stage. Thresholding is the simplest method. The pixels are partitioned depending on their intensity value.

Global thresholding, using an appropriate threshold T:

$$g(i,j) = \begin{cases} 1, for\ f(i,j) \geq T \\ 0, for\ f(i,j) < T \end{cases} \quad\quad\quad (2.5)$$

Where T is the threshold,

g (i, j) =1, for image elements of objects,

And g (i, j) = 0, for the background (or vice versa).

The algorithm of basic thresholding:

Search all the pixels f (i, j) of the image f. An image element g(i, .j) of the segmented image is an object pixel if f(i, j) ≥T, and is a background pixel otherwise.

If objects do not touch each other, and if their gray-levels are clearly distinct from background gray-levels, thresholding is a suitable segmentation method. Such an example is found in figure (2.9) a, original image, and the threshold segmentation result for which is shown in figure (2.9) b, figures (2.9) c and (2.9) d, show segmentation results for different threshold values. Correct threshold selection is lead to successful threshold segmentation; this selection can be determined interactively or it can be the result of some threshold detection method [24].
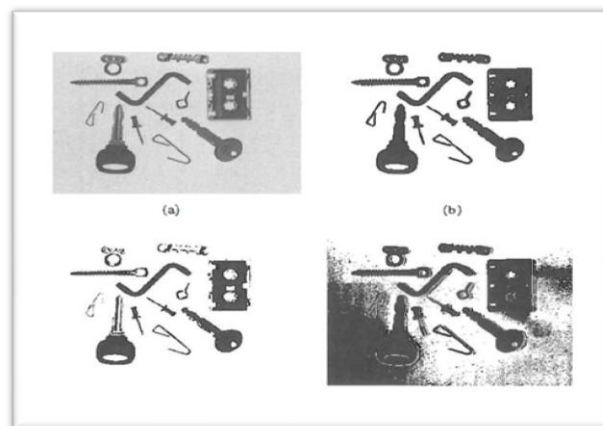
Figure 2.9 Image thresholding, (a) Input image. (b) Thresholding (c) Low threshold. (d) High threshold.[24]

Otsu's method is a part of thresholding techniques aimed in finding the optimal value for the global threshold. It is based on the interclass variance maximization [30]. Well threshold classes have well discriminated intensity values.

Let $M \times N$ indicate to image histogram:

L intensity levels, $[0 \ldots L - 1]$;

n_i #pixels of intensity i:

$$MN = \sum_{i=0}^{L-1} n_i \tag{2.6}$$

Normalized histogram:

$$P_i = \frac{n_i}{MN} \tag{2.7}$$

$$\sum_{i=0}^{L-1} P_i = 1, P_i \geq 0 \tag{2.8}$$

Using K, $0<K<L-1$, as threshold, $T=K$:

Two classes : C_1(pixels in [0, k]) and C_2(pixels in [k+1, L-1])

$$P_1 = P(C_1) = \sum_{i=0}^{K} P_i \tag{2.9}$$

Probability of the class C_1

$$P_2 = P(C_2) = \sum_{i=k+1}^{L-1} P_i = 1 - P_i \tag{2.10}$$

Probability of the class C_2

m_1, mean intensity of the pixels in C_1:

$$m_1 = \sum_{i=0}^{K} i . P(i|C_1) \tag{2.11}$$

$$m_1 = \sum_{i=0}^{K} i \frac{P(C_1|i)P(i)}{P(C_1)} \tag{2.12}$$

$$m_1 = \frac{1}{P_1} \sum_{i=0}^{K} i . P_i \tag{2.13}$$

Where

$$P(C_1|i) = 1, P(i) = P_i \rightarrow P(C_1) = P_1 \tag{2.14}$$

Similarly, m2, mean intensity of the pixels in C2:

$$m_2 = \frac{1}{P_2} \sum_{i=K+1}^{L-1} i . P_i \tag{2.15}$$

Mean global intensity m_G:

$$m_G = \sum_{i=0}^{L-1} i . P_i \tag{2.16}$$

While the mean intensity up to the k level, m:

$$m = \sum_{i=0}^{k} i . P_i \tag{2.17}$$

Hence:

$$P_1 m_1 + P_2 m_2 = m_G \tag{2.18}$$

$$P_1 + P_2 = 1 \tag{2.19}$$

The global variance σ_G^2 :

$$\sigma_G^2 = \sum_{i=0}^{L-1}(i - m_G)^2 . P_i \tag{2.20}$$

The between –class variance, σ_B , can be defined as :

$$\sigma_G^2 = P_1 (m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \tag{2.21}$$

$$= P_1 P_2 (m_1 - m_2)^2 \tag{2.22}$$

$$= \frac{(m_G P_1 - m)^2}{P_1(1 - P_1)} \tag{2.23}$$

The goodness of the choice T=K can be estimated as the ratio η:

$$\eta = \frac{\sigma_B^2}{\sigma_G^2} \tag{2.24}$$

The equation required for the calculation of η, can be obtained from the histogram:

Therefore, for each value of k, η (k) can be computed:

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2} \tag{2.25}$$

Where

$$\sigma_B^2(k) = \frac{(m_G P_1(k) - m(k))^2}{P_1(k)(1 - P_1(k))} \tag{2.26}$$

The optimal threshold value k^* satisfies:

$$\sigma_B^2(k^*) = max_{0<k<L-1}\sigma_B^2(k) \tag{2.27}$$

Figure (2.10) illustrating the result of segmentation stage after applying Otsu's method.
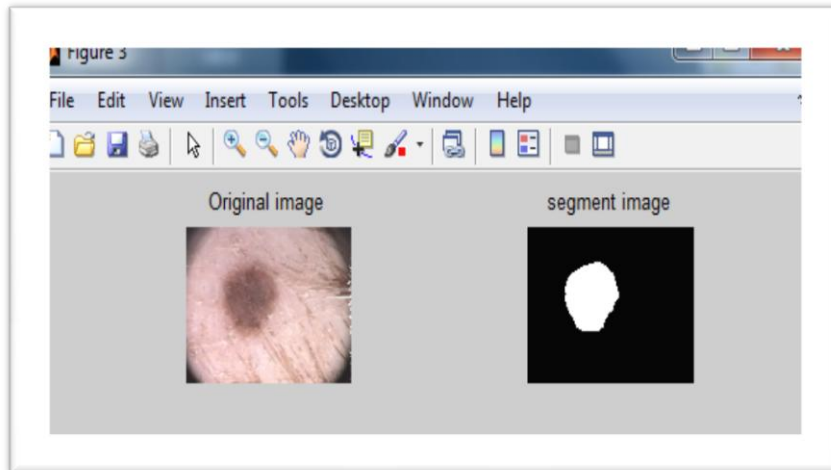
Figure 2.10 Segmentation stage (a) Original image after hair removal (b) Segmented Image after applying Otsu's method

Another technique of segmentation is Morphological image processing which is a set of non-linear operations related to shape features or texture in an image. The mathematical morphology is a tool for extracting components from image that are useful in the representation and formation of region shape, boundaries, skeletons, etc. Morphological operations are suitable for binary images processing because it is rely only on the

relative ordering of pixel values [29], not on their numerical values, and can also be applied to grey scale images.

Structuring Element (SE) is morphological techniques to investigate an image with a small form or template can position at all locations in the image and compared with the matching neighborhood of pixels. Some operations examine whether the element "fits" within the neighborhood, while others test either it "hits" or intersects with the neighborhood [33]. The figure (2.11) illustrating the inquiry of an image with (SE).
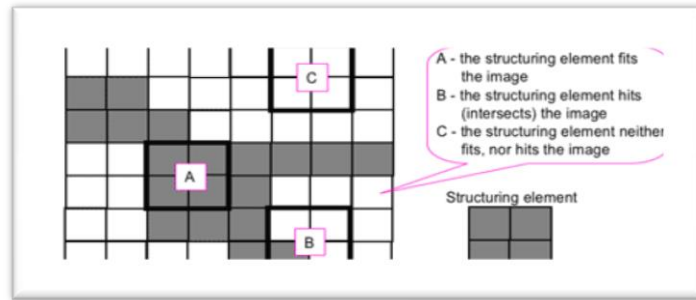
Figure 2.11 Investigation of an image with (SE). [33].

The fundamental of morphological operation is (erosion & dilation) as shown in figure (2.12). The erosion is shrinking the components of an image, while dilation expanding them. The erosion of the binary image A by the structuring element B is defined by:

$$A \ominus B = \{z \in E | B_z \subseteq A\} \tag{2.28}$$

Where B_z is the translation of B by the vector z

The dilation of A by the structuring element B is defined by:

$$A \oplus B = \{z \in E | (B^s)_z \cap A \neq \emptyset\} \tag{2.29}$$

Where B^s denotes the symmetric of B, that is,

$$B^s = \{x \in E | - x \in B\} \tag{2.30}$$
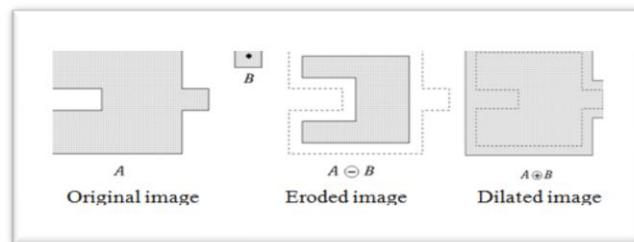
Figure 2.12 Morphological erosion and dilation [31].

Opening operator is the dilation of the erosion of a set A by a structuring element B [35]

23

$$A \circ B = (A \ominus B) \oplus B \tag{2.31}$$

Where $\ominus$ and $\oplus$ denote erosion and dilation,

The opening operation satisfies the following properties [36]:

1. Opening of A by B is a subset of A.
2. If C is a subset of D, then C $\circ$ B is a subset of D $\circ$ B.
3. (A $\circ$ B) $\circ$ B = A $\circ$ B.

Closing operator is a dilation followed by erosion. The closing of an image A by the structuring element B is denoted by $A \bullet B$ and is defined as

$$A \bullet B = (A \oplus B) \ominus B \tag{2.32}$$

If an image A is unchanged by closing with B, it's known as closed with respect to B [36].

The closing operation satisfies the following properties:

1. A is a subset of closing of A by B.
2. If C is a subset of D, then C $\bullet$ B is a subset of D $\bullet$ B.
3. (A $\bullet$ B) $\bullet$ B = A $\bullet$ B.

Opening and closing with an identical structuring component is employed to eliminate specific image details smaller than the structuring element but the general form of the objects is not distorted. Closing connects objects that are close to one another, fills up tiny holes, and smoothes the object by filling up narrow gulfs. Meanings of near, small and narrow are related onto dimensions and also the form of structure element [24]. Figure (2.13) is shown the effect of opening and closing operators.
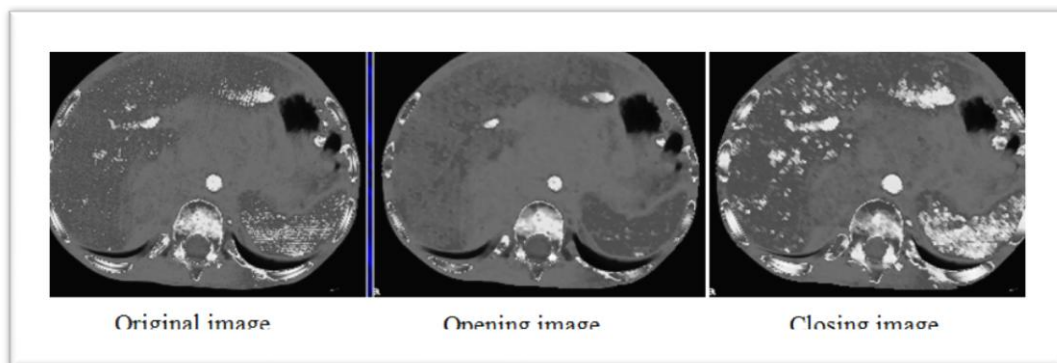
Figure 2.13 Opening and closing operators [31].

Combining openings and closing with image subtraction results in top-hat and bottom-hat transformations [34]. The top-hat transformation of a gray-scale image (f) is defines as (f) minus its opening:

$$T_{hat}(f) = f - (f \circ b) \tag{2.33}$$

The bottom-hat of a gray-scale image (f) is subtracted closing of f from f:

$$B_{hat}(f) = f - (f \bullet b) \tag{2.34}$$

The principle application of those filters as shown in figure (2.14), to removing unwanted objects from an image by using an SE .the top-hat filter is used to remove light objects over a dark background and correcting the effects of non-uniform illumination while, the bottom-hat filter is used for dark objects over a light background [34].
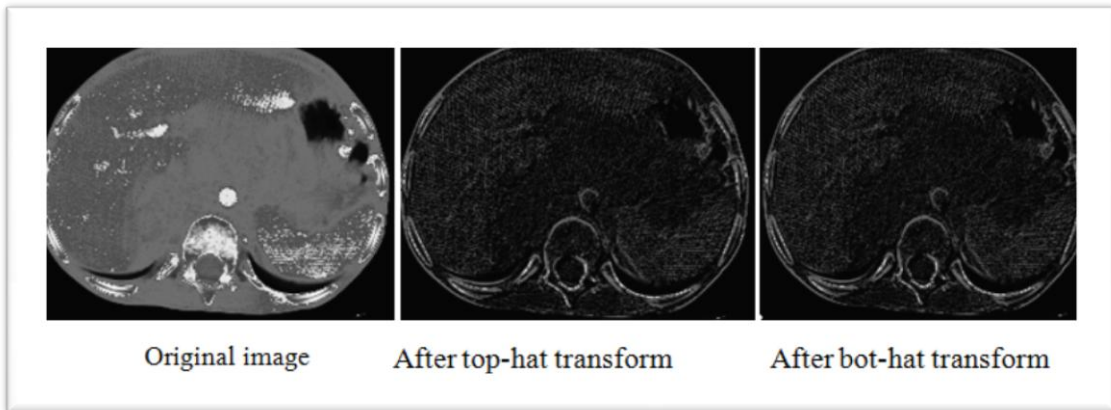
Figure 2.14 the effect of top-hat and bot-hat [31].

2.3.4   Feature Extraction

The main features of the melanoma skin lesion are its geometric feature. So, we aimed to extract the geometric features of segmented skin lesion. The standard geometry features (Area, Perimeter, Greatest Diameter, Circularity and Irregularity Index). From the results of segmenting images which are the black and white (BW) and after separating the region containing only skin lesion, this region is analyzed to extract its geometrical features. The different features extracted are as follows:

Area (A): Number of pixels of the lesion.

Perimeter (P): Number of edge pixels.

Greatest Diameter (GD) which is represented the Major Axis Length: Is length of the line that connecting the two farthest boundary points and passing through region lesion centroid

$$(x_c , y_c) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n} \right) \qquad (2.35)$$

Where n is the number of pixels inside the lesion, (x_i, y_i) is the coordinates of the ith lesion pixel.

Shortest Diameter (SD) which is represented Minor Axis Length: Is length of the line that connecting the two nearest boundary points and passing through lesion region centroid. While Circularity Index (CRC): It is responsible about the shape uniformity [33].

$$CRC = \frac{4\pi A}{P^2} \qquad (2.36)$$

Irregularity index (IrA):

$$IrA = \frac{P}{A} \qquad (2.37)$$

Irregularity index (IrB):

$$IrB = \frac{P}{GD} \qquad (2.38)$$

After feature extraction and selection, it will be fed to classifier, which we have numerous classifiers that will be explain within the next section.

2.3.5 Classification

Classification stage: Is the stage which used results that have been obtained from the segmentation stage and feature extraction in order to reach the diagnosis of the input image. There are two different types of classification of dermoscopic images: The first type is based on the rating that the disease is either malignant or benign (This means the division of skin cancer into two categories) and the resulting data will be either zero or one, and it is supported vector machines. The second type tries to model P (y|x); this not only a class label for an information item, however conjointly a chance of sophistication membership. The members of the second type are Logistic regression, k -nearest neighbors, artificial neural networks and decision trees. Although they vary significantly in building associate and approximation to p (y|x) from data, therefore we can recognize the classification methods as:

26

- k-Nearest Neighbor Algorithm

- Decision Trees

- Logistic Regression

- Artificial Neural Network

- Support Vector Machines

- Extreme Learning Machine

Figure (2.15) is shown the different types of classification methods.
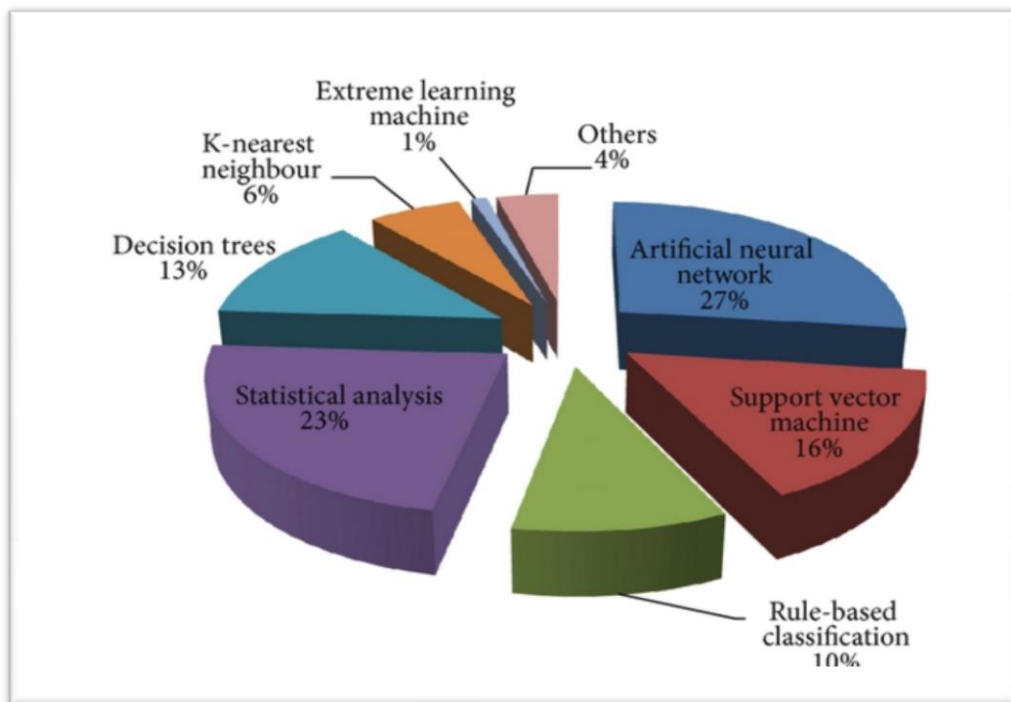
Figure 2.15 Classification Methods [23]

2.3.5.1 K-Nearest Neighbor Algorithm (K-NN)

The k-nearest neighbor classifier [26, 27] is a nonparametric method of pattern recognition. For a lesion belonging to the test set which is called (query vector), it is found that the vectors are the closest to the test set in the training set. The unclassified sample is then assigned to the class represented by the closest neighbors. The most critical requirement of this classifier is having a training set involving enough examples of each class of pigmented lesions to sufficiently represent the full range of measurements that can be expected from each class. Good feature selection and weight definition could improve the performance of this classifier [28]. K-NN algorithm permits extraction and visualization of the "most similar" cases to those at hand. These

cases resemble the medical result and give dermatologists a chance to directly compare between unknown lesions with other known skin lesions. This case can provide an advantage in areas where black-box models are inadequate. K-NN classifier is failure in irrelevant features case. This classifier can be used for the evaluation of feature subset selection; because it allows incorporating/eliminating features easily and it has low computational cost [28].Figure (2.16) is shown the simple presentation of NN algorithm.
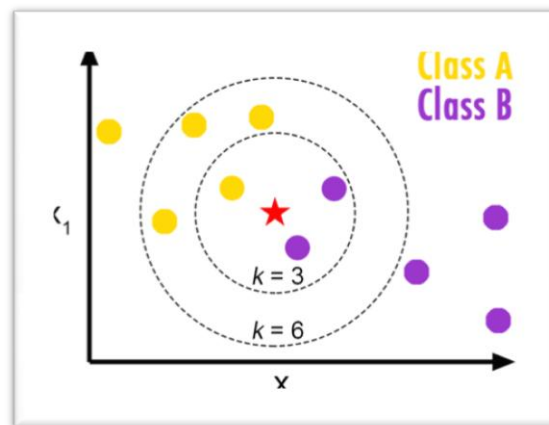
Figure 2.16 Basic examples about Nearest Neighbor Algorithm [23]

2.3.5.2 Decision Trees

The decision tree belongs to the supervised machine learning techniques. It is most common use due to it is simplicity in constructing, efficient and simple representation, lead to easy understood by humans. This algorithm repeatedly divides the data set according to a criterion that maximizes the separation of the data by identifying a variable and a threshold in the range of this variable and divides the data set into two groups. The good selection to the variable and threshold minimizes the difference measures in the resulting groups. The information gain is unimportant standard, this means that at each split, maximizes the decrease in entropy, and due to this split could estimate the ratio p (y|x) of y class elements over all elements of the leaf node that contains data item [23].

2.3.5.3 Logistic Regression

Logistic regression is an algorithm that based on separating hyper plane between two groups of data sets, using the logistic function to locate distance from the hyper plane as a probability of class membership. Despite the model is linear in parameters and can only calculate linear decision boundaries, but it is a widely used predictive model in medical applications. The advantage of this method its ease of use comparable with other algorithms, allowing translation of results as probabilities and the ability of varies selection. From comparative study between the logistic regression performing and artificial neural networks, showing that they are on the same level and support vector machines, and both have capable of implementing nonlinear separating surfaces [23].

2.3.5.4 Artificial Neural Network (ANN)

Artificial neural network is one of the great vital parts of soft computing. It consists of several tiny processing units (the artificial neurons) are highly interconnected. Information fed to the ANN is modeled after the human brain. The ANN is an iterative process requires many availability of training set, the system distinguished capacity to obtain idea from complex data and take out patterns and determine orientation that are too difficult to be determined by humans or any other computer skills. Nowadays the ANNs are used in a lot of research on dermoscopic image analysis [23]. The general working mechanism for artificial neural network present in figure (2.17).
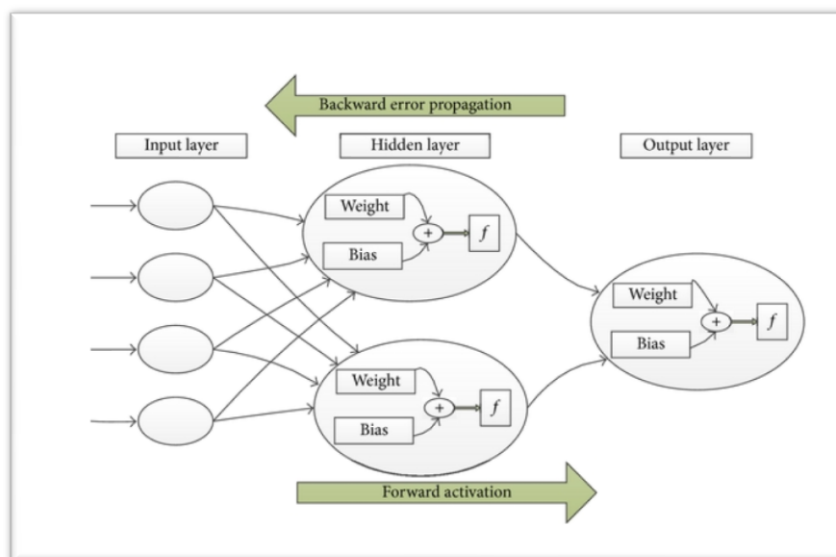
Figure 2.17 Working Mechanism of Artificial Neural Network. [23].

2.3.5.5 Support Vector Machines (SVMs)

Support vector machines (SVMs) are a machine learning paradigm depending on the theory of statistical learning. The medical literature reported about the performances of the machine learning algorithms. Algorithmically, SVM is solving an affected quadratic optimization problem by building optimal separating boundaries between data sets. For comparative with the basic training algorithm that can only build linear separators, different element functions (i.e., linear, radial basis function, polynomial, and sigmoid) can be used involving varying degrees of nonlinearity and flexibility in the model [23]. Figure (2.18) is shown the principle of support vector machine.
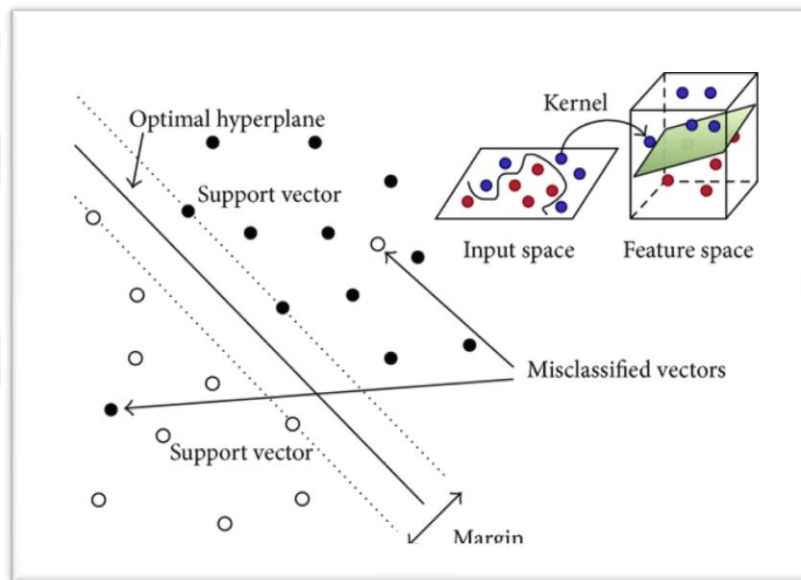
Figure 2.18 Principle of Support Vector Machine [23]

2.3.5.6 Extreme Learning Machine (ELM)

The ELM is the feed forward network, similar to the other networks, consists of three layers and the difference is that the hidden elements can be independent from the target functions and training data, this reason leads to provide better generalization performance and it can impact much faster as compared to the other traditional algorithms. The important features that can be acquired from using extreme learning machine, simple math, simple tuning free three step algorithm, the learning speed is quite fast. This learning algorithm looks much simpler than many other learning algorithms like SVM and NN [29].

Material and Methods

3.1   Material

Skin cancer is a widespread disease in the last two centuries, possible affect any age without exception, and the various colors of skin (light and dark) with that little spread among children. So, dermatologists made valuable efforts to diagnose this disease in the early stage, the interested on this disease has encouraged a number of researchers in the field of computer vision and image processing to build systems which are helping dermatologist in the diagnosis of disease. The dermatologist is dealing with a patient in a medical clinic, while the researchers in the field of image processing are dealing with images in order to diagnose disease (data set), which was identified by the skin experts to serve as a reference for researchers to make sure of the results accuracy they got from diagnostic systems, provide a reliable second opinion to dermatologists is the aim of such systems. Through mathematical and statistical analysis for digital images, can these systems realizing the features and characteristics cannot be identified with the naked eye. CAD system is using to diminishing the mistakes that could be found in dermatologist examinations and reduces the amount of repetitive and boring tasks that be done by clinicians. Several groups around the world are trying to assist the dermatologist to evaluate of pigmented skin lesions by developing effective systems and devices using for that purpose. One of these groups is the Hospital Pedro Hispano, Matosinhos, Portugal. It has provided PH², which is a dermoscopic image database that is specified to the Dermatology Service. The PH² dataset has been worked on intense way to get used in benchmarking and research purposes, the main purpose is to facilitate comparative research on both segmentation and classification algorithms related with dermoscopic images. Images were obtained under the same conditions through Tubingen Mole Analyzer system using a magnification of 20 xs. They are 8-bit RGB color images with a resolution of 768x560 pixels.

The specified dataset include 200 dermoscopic images of melanocytic lesions in total, containing 80 common nevi, 80 atypical nevi, and 40 melanomas. The PH² it has medical explanations to all images called medical segmentation of the lesion, clinical and histological diagnosis and the examination of multiple dermoscopic features (colors; pigment network; dots/globules; streaks; regression areas; blue-whitish veil). The description of features is shown in the figure (3.1).

Figure 3.1 Legends of Samples used in dataset

The examination of these dataset was done under supervision of an expert dermatologist for each feature. The PH² Database package (dataset and medical explanations) and the PH² Browser can be downloaded from the two links below after participant form has been filling:

PH² Dataset - download link

PH² Browser - download link

The appreciation for each parameter was performed by aiding expert dermatologist, taking in account the following parameters:

Criterion	PH² Segmentation
Clinical Diagnosis	0 - Common Nevus
	1 - Atypical Nevus
	2 - Melanoma
Lesion Segmentation	Available as a binary mask (with the same size of the original image).
Color Segmentation	Available as a binary mask (with the same size of the original image) (If available).
Asymmetry	0 - Fully Simmetry
	1 — Asymmetry in One Axis
	2 - Fully Asimmetry
Pigment Network	AT - Atypical
	T - Typical
Dots/Globules	A - Absent
	AT - Atypical
	T - Typical
Streaks	A - Absent
	P - Present
Regression Areas	A - Absent
	P - Present
Blue Whitish Veil	A - Absent
	P - Present
Colors	1 - White
	2 - Red
	3 - Light-Brown
	4 - Dark-Brown
	5 - Blue-Gray
	6 - Black

Figure 3.2 Criteria of Parameter in PH^2 dataset

This research study used two types of clinical diagnostic images (80 Benign and 25 Melanoma) and made comparison between our segmentation results and original dataset images, all results were obtained by using MATLAB version R2014a. Therefore, from the above clarification we can address our problem in the following steps:

- Hair removal.
- What is the best technique to extract the region of interest (ROI) and border detection from a given dermoscopy image?
- How can we determine the accuracy of the results obtained and their conformity with the ground truth?

33

3.2 Methods

The proposed methods have the following stages:

Stage 1: Dull razor algorithm for hair removal and bottom hat filter.

Stage 2: Mean shift filtering algorithm, which produces a filtered image.

Stage 3: Segmentation by using Otsu's thresholding

Stage 4: Features selection such (area, perimeter, and border irregularity).

3.2.1 Stage 1: Hair Removal

Steps of the stage1 (hair removal by using Dullrazor method, number of images under test are 80 images) can be arranged in the figure (3.3).

Step 1: Loading an image and finding size

Step 2: Masks for closing operations

     mask_0 = [0 1 1 1 1 1 1 1 1 1 1 1 0];

     mask_45 = [0 0 0 0 0 0 0 0; 0 1 0 0 0 0 0 0; 0 0 1 0 0 0 0 0; 0 0 0 1 0 0 0 0 0; 0 0 0 0 1 0 0 0 0; 0 0 0 0 0 1 0 0 0; 0 0 0 0 0 0 1 0 0; 0 0 0 0 0 0 1 0; 0 0 0 0 0 0 0 0];

     mask_90 = mask_0';

Step 3: Separating image into RGB layers (addition of median filter for better hair masking)

Step 4: Closing mask on red image & green &blue

Step 5: Max of all masks

Step 6: Generalized grayscale closing image red & green &blue

Step 7: Union of hair masks for three color bands

Step 8: replace remove hair

Step 9: Follow interpolation from dull razor

Step10: Generate Zero image with three shell RGB of size (m x n)

Figure 3.3 Steps to remove hair by using dullrazor method

Also implementation of morphological bottom hat filter for hair removal (the number of images under test are 25 images), the following figure illustrates steps of removed hair by using bottom hat filter.

Step 1: load image

Step 2: chose the suitable structur element

Step 3: applid the morophological bottom _hat filter on the orginal image

Step 4: converted the image from rgb2gray

Step 5: replaced the (ROI) in the orginal image with the (RIO)in the image which under bothat filter

Step 6: hair removal by appliding closing operation with structure element in direction 0°

Step 7: hair removal by appliding closing operation with structure element in direction 180°

Figure 3.4 Steps to remove hair by using bottom hat filter

3.2.2 Stage 2: Mean Filtering

The mean filter produces a blur image, the image resulting in less of high-frequency image detail and brightness differences across boundaries.

3.2.3 Stage 3: Otsu's Thresholding

Otsu's thresholding is the best method for estimate optimal thresholding because it automatically performed clustering by separating the image into foreground and background in order to tracing boundaries of lesion. Steps of detecting border of lesion are shown in figure (3.5)
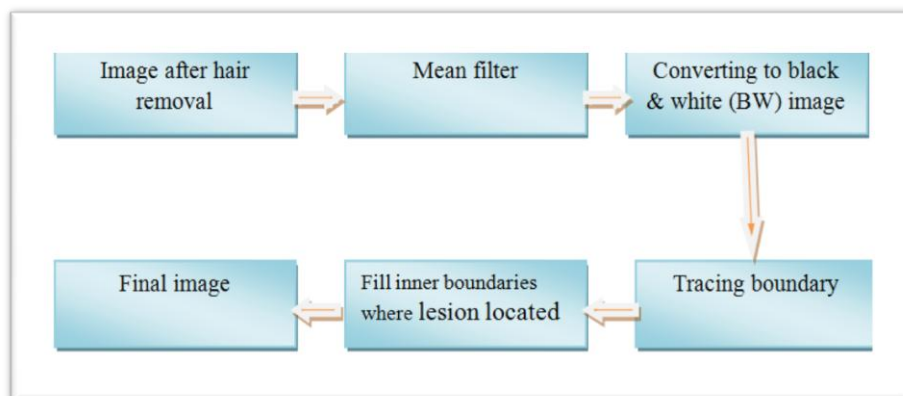
Figure 3.5 Steps of segmentation

3.2.4 Stage 4: Feature Selection

In this stage, the main features that can be selected from segmented image are their Geometric Features. Hence, adopted some standard geometry features like (Area,

Perimeter, Irregularity Index). Calculating area and perimeter for both our segmented and G_T images, and calculate the difference between areas; also calculate the variation in border irregularities. Steps of selection geometrical features can be arranged in the figure (3.6) and difference between areas in the figure (3.7).

Step 1: regionprops to estimate (Area, Perimeter, Centroid)
Step 2: compute Our _segment image perimeter
Step 3: compute G _T image perimeter
Step 3: compute the area of Our _segment image
Step 4: compute the area of the G_T image
Step 5: find the different area by subtract Our_segment image from G_T image
Step 6: calculate the border irregularity of our _segment image
Step 7: calculate the border irregularity of G _T image
Step 8: find the percentage different between the border irregularities

Figure 3.6: Steps of Feature Selection

G_T image   Our segmented image   Diff=|G_T−Our image
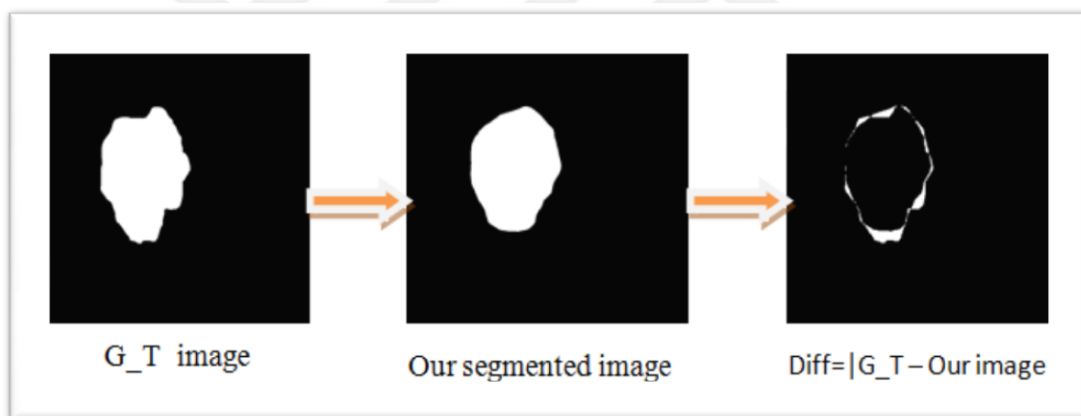
Figure 3.7 Subtract Our_image from G_T image

RESULTS AND DISCUSSION

Results that were obtained from implementation the four stages in chapter three will be explained in this chapter with discussion.

4.1   Stage1: Results of Hair removal

By using Dullrazor method for hair removal, results are shown in the following figures:

Figure 4.1Implementation of Dullrazor method

Figure 4.2 Benign images before hair removal
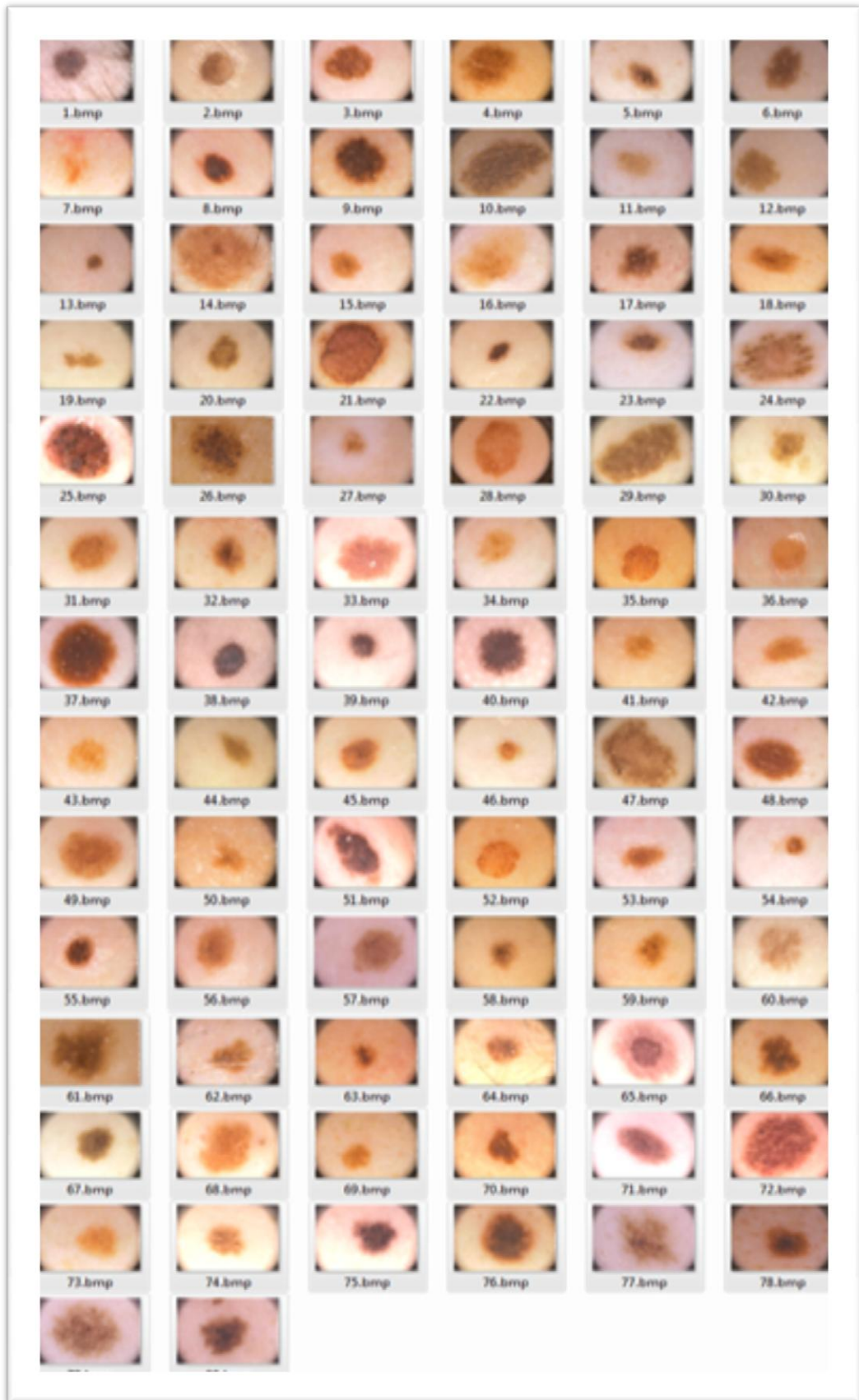
Figure 4.3 Benign images after hair removal by using Dullrazor method

Results of hair removal by using bottom hat filter.

Figure 4.4 Melanoma images before hair removal
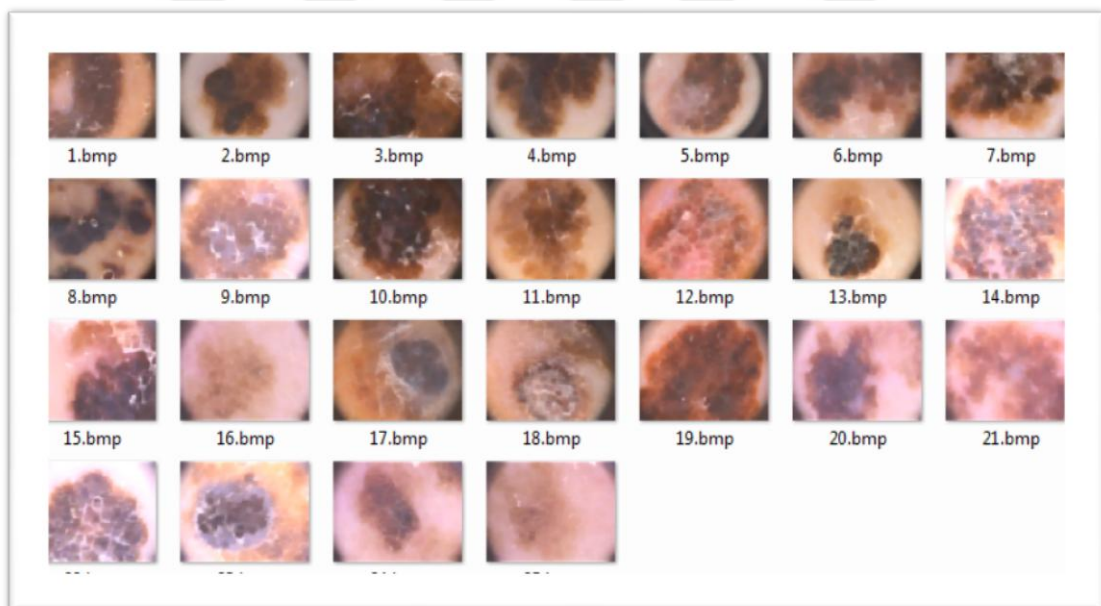
Figure 4.5 Melanoma images after hair removal by using bottom hat filter

4.2 Stage 2 & 3: Results of Segmentation

Results of mean filter and segmentation are shown in the figure (4.6).
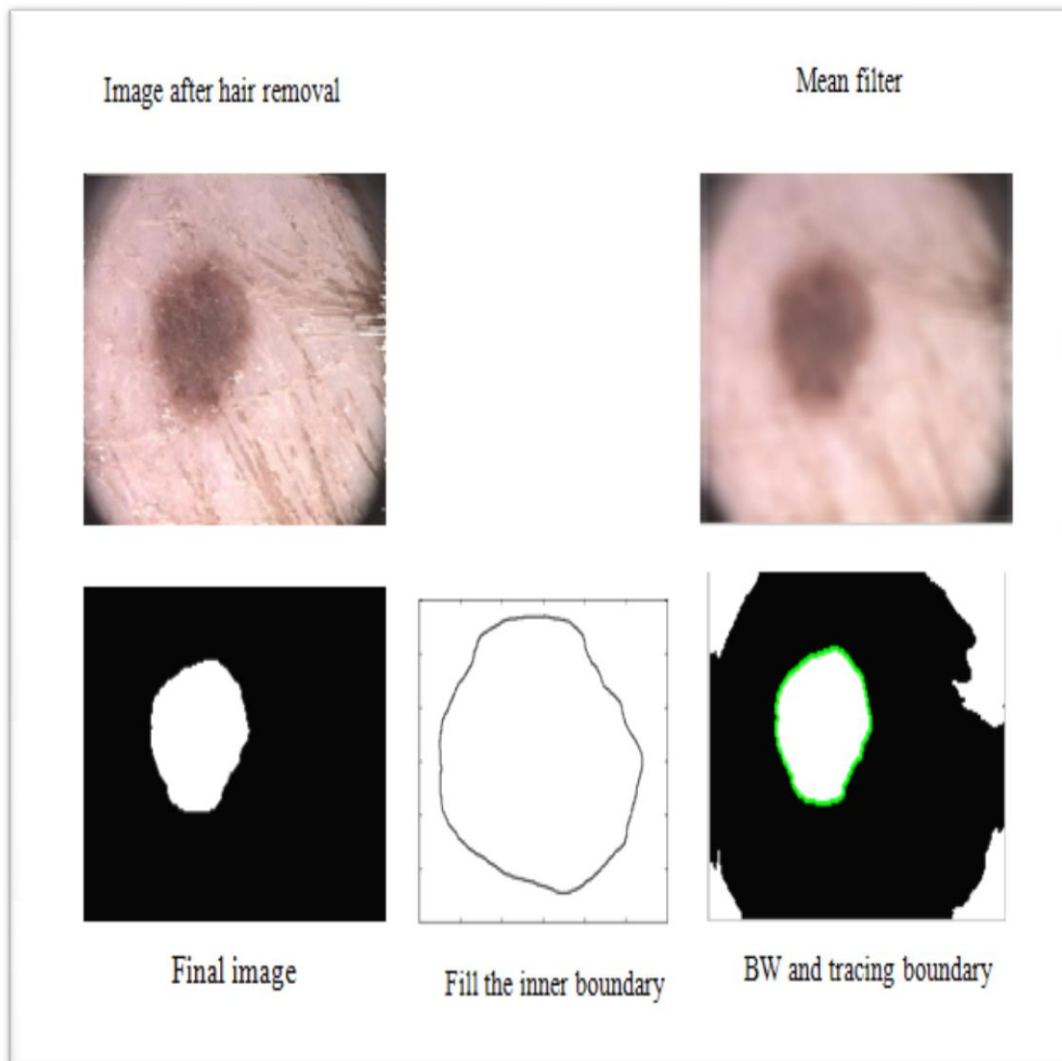
Figure 4.6 Results of Segmentation image by using Otsu's method

4.3 Stage 4: Results of Feature Selection

Features selection and find the different between our segmented and G_T images and find the variation percentage between the border irregularities. The following tables demonstrate the results of stages 3&4.

Table 4.1 Results of our segmentation and the G_T segmentation and the subtraction ratio for benign
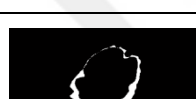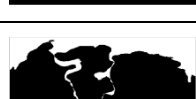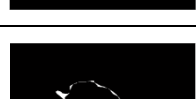
Image No	Our segmentation bw1	Ground truth bw2	Image subtraction diff=(abs(logical (bw1)-bw2))	Diff%
1.				1.24%
2.				3.62%
3.				1.20%
4.				6.56%
5.				2.25%
6.				3.28%
7.				35.00%
8.				1.14%
9.				2.11%
10.				7.67%
11.				0.83%

Table 4.1 (cont'd)

12.				3.35%
13.				0.83%
14.				20.19%
15.				2.69%
16.				29.28%
17.				4.51%
18.				6.52%
19.				1.47%
20.				2.66%
21.				4.49%
22.				1.37%
23.				3.80%
24.				8.60%
25.				4.83%

Table 4.1 (cont'd)

26.				7.62%
27.				0.58%
28.				4.79%
29.				9.55%
30.				6.21%
31.				4.67%
32.				5.05%
33.				7.08%
34.				24.51%
35.				6.13%
36.				46.32%
37.				7.31%
38.				1.03%
39.				0.64%

Table 4.1 (cont'd)

40.				2.11%
41.				5.62%
42.				3.34%
43.				2.29%
44.				3.21%
45.				2.44%
46.				2.21%
47.				12.04%
48.				4.64%
49.				6.07%
50.				49.87%
51.				5.02%
52.				9.20%

Table 4.1 (cont'd)
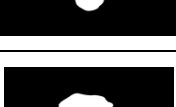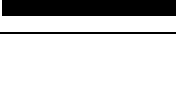
53.				3.16%
54.				0.44%
55.				1.21%
56.				6.87%
57.				3.71%
58.				6.91%
59.				6.29%
60.				5.20%
61.				14.02%
62.				7.77%
63.				4.26%
64.				3.23%
65.				26.85%

Table 4.1 (cont'd)

66.				5.88%
67.				2.62%
68.				9.70%
69.				1.96%
70.				1.68%
71.				4.66%
72.				11.00%
73.				10.00%
74.				3.57%
75.				5.61%
76.				5.26%
77.				11.00%
78.				4.74%

Table 4.1 (cont'd)

79.				11.50%
80.				5.69%
			Average	=7.31%

Table 4.2 Results of our segmentation and the G_T segmentation and the subtraction ratio for melanoma
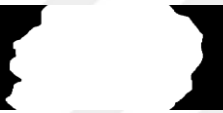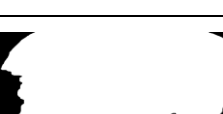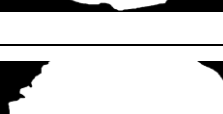
Image No	Our segmentation bw1	Ground truth bw2	Image subtraction diff=(abs(logical (bw1)-bw2))	Diff%
1				16.91%
2				7.30%
3				21.48%
4				13.81%
5				55.86%
6				18.94%

48
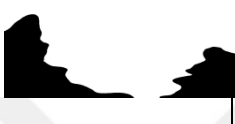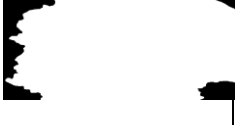
Table 4.2 (cont'd)

7				18.50%
8				60.50%
9				28.54%
10				14.44%
11				22.32%
12				58.51%
13				15.22%
14				21.33%
15				24.86%
16				23.70%
17				76.08%

Table 4.2 (cont'd)
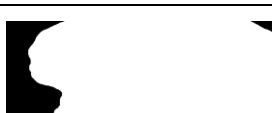
18				33.86%
19				9.62%
20				34.83%
21				39.71%
22				5.65%
23				72.86%
24				58.12%
25				16.83%
			Average	=30.79%

Table 4.3 the ratio of areas difference and the ratio of irregularity difference between our segmentation and G_T for benign

Image No	Area2 of G.T (a2)	Area1 of OUR_SKIN (a1)	Difference | a2 - a1 |	Ratio= Difference/a2	BIR1= area1/perimeter1	BIR2= area2/perimeter2	DFB=(BIR2-BIR1)/BIR2
1.	30542	7242	23300	0.7629	21.9745	41.1248	0.4657

Table 4.3 (cont'd)

2.	5948	26400	20452	3.4385	35.8135	19.1268	0.8724
3.	12825	6141	6684	0.5212	20.0934	26.0965	0.2300
4.	2060	11013	8953	4.3461	25.6302	10.9080	1.3497
5.	37335	1511	35824	0.9595	10.2751	45.4421	0.7739
6.	6054	38501	32447	5.3596	32.1306	19.2143	0.6722
7.	26147	4285	21862	0.8361	16.7762	38.4848	0.5641
8.	10984	6966	4018	0.3658	16.8190	24.5317	0.3144
9.	11205	8025	3180	0.2838	21.8499	26.2077	0.1663
10.	4242	6927	2685	0.6330	21.1222	13.9314	0.5162
11.	7435	4989	2446	0.3290	17.5815	20.8076	0.1550
12.	8570	5967	2603	0.3037	17.9463	21.0785	0.1486
13.	31670	6814	24856	0.7848	20.3061	43.1247	0.5291
14.	3856	30707	26851	6.9634	43.9193	14.7115	1.9854
15.	9846	2949	6897	0.7005	13.7385	24.4870	0.4389
16.	31198	12040	19158	0.6141	26.4240	44.1105	0.4010
17.	26050	25602	448	0.0172	35.0645	41.7201	0.1595
18.	24125	23998	127	0.0053	38.6362	37.6398	0.0265
19.	3857	19134	15277	3.9609	31.9668	14.8032	1.1594
20.	21095	4046	17049	0.8082	15.6194	34.6282	0.5489
21.	33166	5963	27203	0.8202	9.0918	42.4097	0.7856
22.	7454	26886	19432	2.6069	36.4869	22.6985	0.6075

Table 4.3 (cont'd)

23.	10233	11601	1368	0.1337	27.9012	21.6740	0.2873
24.	15056	6152	8904	0.5914	18.8509	31.4347	0.4003
25.	9796	11995	2199	0.2245	28.2423	22.2815	0.2675
26.	19960	6472	13488	0.6758	20.2744	32.8336	0.3825
27.	8012	15297	7285	0.9093	30.6105	20.6007	0.4859
28.	10507	23162	12655	1.2044	16.3765	26.0707	0.3718
29.	7186	6781	405	0.0564	16.7892	22.5454	0.2553
30.	31851	37539	5688	0.1786	30.2849	45.0364	0.3275
31.	7758	27036	19278	2.4849	42.0001	22.1013	0.9003
32.	5011	7797	2786	0.5560	22.5223	17.9051	0.2579
33.	11031	5054	5977	0.5418	18.4871	26.5613	0.3040
34.	15651	13859	1792	0.1145	29.3640	30.0064	0.0214
35.	6414	15142	8728	1.3608	30.8048	19.5052	0.5793
36.	7917	2756	5161	0.6519	13.0469	20.4692	0.3626
37.	9122	5720	3402	0.3729	18.8944	22.3529	0.1547
38.	6872	7752	880	0.1281	22.5795	18.3125	0.2330
39.	9324	4761	4563	0.4894	16.8971	23.7362	0.2881
40.	3855	7895	4040	1.0480	23.0257	16.2232	0.4193
41.	35248	2405	32843	0.9318	12.9354	45.4405	0.7153
42.	17629	27341	9712	0.5509	35.9044	32.8531	0.0929
43.	21494	14585	6909	0.3214	31.5009	37.2328	0.1539

Table 4.3 (cont'd)

44.	17953	17651	302	0.0168	33.2017	28.9425	0.1472
45.	6705	4259	2446	0.3648	16.2278	17.3864	0.0666
46.	17109	38121	21012	1.2281	24.4644	25.6194	0.0451
47.	10723	13903	3180	0.2966	26.8412	26.7366	0.0039
48.	7017	5196	1821	0.2595	12.5277	20.7153	0.3952
49.	2810	4937	2127	0.7569	18.3080	13.9161	0.3156
50.	6202	2556	3646	0.5879	13.4509	19.9371	0.3253
51.	13424	5451	7973	0.5939	19.0523	26.0139	0.2676
52.	16398	8913	7485	0.4565	23.7912	30.7770	0.2270
53.	7845	13949	6104	0.7781	29.7074	19.9783	0.4870
54.	8681	3299	5382	0.6200	14.3912	22.8308	0.3697
55.	5526	4547	979	0.1772	16.0822	17.8172	0.0974
56.	13814	9721	4093	0.2963	23.5031	26.9660	0.1284
57.	30454	10672	19782	0.6496	23.7523	38.8545	0.3887
58.	10330	21256	10926	1.0577	32.0777	23.7438	0.3510
59.	5877	5348	529	0.0900	16.2564	17.9471	0.0942
60.	7979	3072	4907	0.6150	13.5950	22.2032	0.3877
61.	33030	5887	27143	0.8218	19.7935	45.0634	0.5608
62.	14143	15428	1285	0.0909	28.3060	26.8666	0.0536
63.	10125	10275	150	0.0148	24.4449	23.3539	0.0467
64.	22673	8424	14249	0.6285	23.4415	35.2676	0.3353

Table 4.3 (cont'd)

65.	6632	16300	9668	1.4578	30.6513	20.0022	0.5324
66.	11493	5340	6153	0.5354	18.9588	22.7346	0.1661
67.	8009	32553	24544	3.0646	27.2349	20.5983	0.3222
68.	14755	6940	7815	0.5297	19.4069	29.4569	0.3412
69.	35472	11701	23771	0.6701	27.3962	44.9968	0.3912
70.	12054	28256	16202	1.3441	37.0116	26.1887	0.4133
71.	9164	5491	3673	0.4008	17.9156	21.9834	0.1850
72.	13755	6801	6954	0.5056	20.3889	26.8487	0.2406
73.	20990	10055	10935	0.5210	25.2517	34.8721	0.2759
74.	22912	17684	5228	0.2282	33.5532	29.6586	0.1313
75.	12364	15708	3344	0.2705	27.6257	25.6058	0.0789
76.	30043	9242	20801	0.6924	24.5384	36.1023	0.3203
77.	9658	22448	12790	1.3243	36.0190	18.1897	0.9802
78.	16075	6805	9270	0.5767	21.3014	28.1432	0.2431
79.	6924	12319	5395	0.7792	27.6856	19.0352	0.4544
80.	16508	16460	48	0.0029	32.1502	29.8209	0.0781

Table 4.4 the ratio of areas difference and the ratio of irregularity difference between our segmentation and G_T for melanoma

Image No	Area2 of G.T (a2)	Area1 of OUR_SKIN (a1)	Difference | a2 - a1|	Percentage= Difference/a2	BIR1= area1/perimeter1	BIR2= area2/perimeter2	DFB=(BIR2 -BIR1)/BIR2
1.	42124	31888	10236	0.2430	30.0236	50.4089	0.4044

Table 4.4 (cont'd)

2.	48927	47006	1921	0.0393	35.0179	46.0689	0.2399
3.	39489	24885	14604	0.3698	30.2388	44.3710	0.3185
4.	47680	9487	38193	0.8010	13.2195	51.7886	0.7447
5.	24683	14624	10059	0.4075	23.6482	33.6282	0.2968
6.	57560	44396	13164	0.2287	27.5691	57.0304	0.5166
7.	42712	26458	16254	0.3805	32.4404	48.8318	0.3357
8.	35749	28941	6808	0.1904	26.0868	44.0337	0.4076
9.	61707	11841	49866	0.8081	25.3763	65.7454	0.6140
10.	37192	25137	12055	0.3241	16.1733	36.5387	0.5574
11.	52673	52476	197	0.0037	45.1257	55.3713	0.1850
12.	31730	27060	4670	0.1472	37.7520	39.5836	0.0463
13.	47871	25124	22747	0.4752	31.8752	47.8047	0.3332
14.	58534	32940	25594	0.4373	27.2991	60.8468	0.5513
15.	42992	40263	2729	0.0635	45.3414	51.3094	0.1163
16.	64455	16686	47769	0.7411	30.1767	65.8773	0.5419
17.	55556	17471	38085	0.6855	30.0558	59.3516	0.4936
18.	37665	31509	6156	0.1634	31.6088	45.3875	0.3036
19.	51903	42510	9393	0.1810	31.6714	55.0377	0.4246
20.	45207	42293	2914	0.0645	37.3907	47.2667	0.2089
21.	29947	6724	23223	0.7755	10.2043	37.1579	0.7254
22.	47744	39100	8644	0.1810	27.4157	45.0733	0.3918
23.	46210	40331	5879	0.1272	33.4655	48.2512	0.3064
24.	61204	23849	37355	0.6103	26.8931	65.6190	0.5902

Table 4.4 (cont'd)

25.	56604	40123	16481	0.2912	38.4691	56.8305	0.3231

4.4 Discussion

The process of obtaining the results was passed through four stages, the first stage is hair removal, which was very successful and the hair was removed from all images without any distortion or corruptions the images. The challenge that we faced in the segmentation stage to separate the lesion from remainder skin, many of images were almost compatible with the ground truth with a few differences, especially for benign images. After testing 80 images which been diagnostic as benign images according to the G_T, we found that the identical ratio was estimated at 93% for all images under test but after excluding the worst images, it became 95%. The average ratio of border irregularity is 37.81% as shown in the table (4.5). For melanoma images, there were wide differences. After testing 25 images which been diagnostic as melanoma images according to the G_T. The identical ratio is estimated at 70% for all images under test but after excluding the worst images, it became at 80%, and the average ratio of border irregularity is 39.90%, this difference due to that dermatologist in some images (example number 23 & 24 in melanoma images) marked all regions around lesion that may be infected not only the lesion under interest, this leads that the final results are not completely identical as shown in the following images:
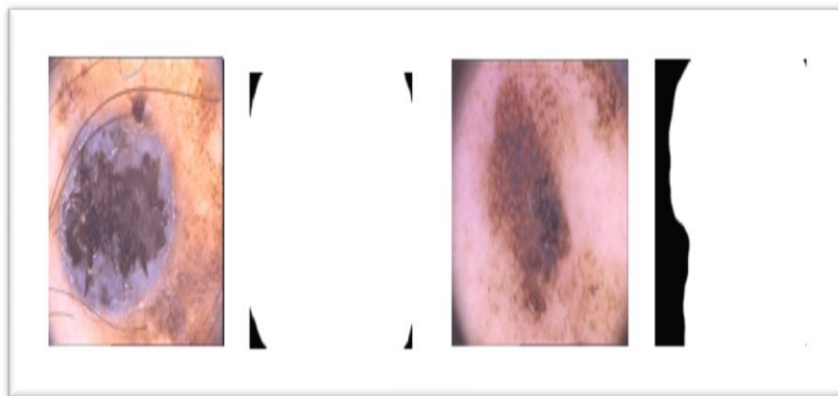
Figure 4.7 Difference between the original images boundaries and boundaries assigned by dermatologist

Table 4.5 Comparison results for benign and melanoma images

Total # of images which have been tested	The percentage difference (±0.5%)	The percentage of identical (±0.5%)
80 for benign	7%	93%
73 for benign	5%	95%
25 for melanoma	30%	70%
19 for melanoma	20%	80%

CONCLUSION AND FUTURE WORK

5.1   Conclusion

Skin cancer if detected early, has a 95-100% successful treatment rate, therefore early detection is vital. Computer aided methods have been developed to assist dermatologists. The goal of any imaging methodology used in dermatology is to detected melanoma and diagnosis in early stages, because by dependence on it, the treatments are affected. Investigations shows, that in early diagnosis more than 90% can be treated while in late diagnosis less than 50%. The diagnosis and successful treatments often complete with stable monitoring for suspicious skin lesions. Skin cancer founds in diverse types such as Melanoma, Basal cell and Squamous cell carcinoma among which Melanoma is the most changeable. The Computer vision can play important role in Medical Image Diagnosis and has been proved by many existing systems this with highly capable. Despite the great advances in the field of imaging technology, but it is the biggest challenge remains the clinical diagnosis of skin cancer, especially for the physician, which deals with the disease since its initial stages and diagnostic is the first point of contact with patients, in order to deportation to dermatology community. As a result, the diagnosis skin cancer operation has become occupies great importance in recent years, and it became the focus of attention of specialists and diagnostic regimes. The aim of this research is to assist physicians in the diagnosis of skin lesion through quantitative analysis of the lesion. Diagnostic systems have played an important role in the registration and follow-up the patient's condition and the progression of the disease and ways to treat it. It is still work in progress in order to get effective results and it needs the concerted efforts of doctors and accurate work to provide assistance to researchers in this field.

With the aim of contributing to the development and improvement of some methods in the field of diagnosis, and developing new technologies to be accurate, fast and reliable

helped in the diagnostic process; this thesis presented a segmentation technique that can help in the development of a diagnostic system for skin cancer. CAD system involves several steps including image acquisition, segmentation or border detection, feature extraction and selection, and classification. In this thesis we have discussed a segmentation of skin cancer by using image processing techniques. The data set used in this thesis has been downloaded from (ADDI Project) and all results were obtained by using MATLAB version R2014a. After following diagnoses system steps, the first step is image enhancement; dermatology images often contain artifacts such as irregular illumination, dermoscopic gel, hair pixels, oil bubbles, skin lines, and ruler markings. As a result, there is a need for robust methods to remove artifacts and detect lesion borders in dermoscopy images. Those artifacts may be occluding some of the information about the lesion such as its boundary and texture. Problems of oil or air bubbles possible to overcome them by using a new dermoscopic and polarized light, but removing hairs remains a serious challenge the needing for an effective and automatic method preserves all the lesion features to overcome this problem.

There are a few methods described in the literature review that deal with the problem of hair removal. This study mainly is focused on hair removal by using two methods (Dullrazor and bottom hat filter), those methods are using to remove unwanted hair from the images. Human's hair covers the entire body and has a range of different colors, textures, and orientations. Skin lesions images always have a possibility of being obscuring by hair. Due to this, hair can cause major information corruption when working with a skin lesion image. Despite rapid growth in the field of image processing for dermatological applications, this particular issue had never been fully addressed until been publication of DullRazor in 1997. DullRazor done by performing the following steps:

1. Locate the dark hair by a generalized grayscale morphological closing operation.
2. Replace the hair pixels by non-hair pixel, interpolation.
3. Smoothness of the resulting image.

When the binary hair mask of the image is obtained for each color band, interpolation is done across hairs to replace the hair-pixel with an interpolation for the surrounding non-hair pixel values. Because of the partial shadow effect (darker pixels around the hair caused by the shadow of the hair on the skin) exact border location of the hairs are

difficult to determine, surrounding pixels that using for interpolation chosen 11 in length away from hair border. A mean filter is also used to smooth out thin lines left from interpolation and hair removal. We can realized that steps of the stage1 (hair removal by using Dullrazor method on 80 images) are working effectively enough to produce satisfactory results for hair removal. Also implementation of morphological bottom hat filter for hair removal by utilizing (25) images, take advantage of the morphological operators characteristics (closing operator in two orientation 90°,180°) we get the hair removal successfully.

Segmentation is the second step and it is one of the important steps in cancer detection. It concerned with dividing an image into meaningful regions. Segmentation methods can be classified into the;

Histogram thresholding, clustering, edge-based, region-based, morphological, model-based, active contours (snakes and their variants), soft computing, scalar vs. vector processing, automatic vs. semi-automatic, number of parameters. In our case, segmentation is based on thresholding. Threshold is one of the widely methods used for image segmentation. Otsu's method thresholding using to automatically performed clustering by reducing a gray level image to a binary image, and selecting optimal threshold value (T). Then all gray level values are beyond this T value will be classified as black (0), and values above T value as white, then isolated the lesion from the rest skin and traced lesion boundaries.

In CAD system, feature extraction process is added to determine the accuracy of the results, these features is obtained from (ROI) which was determined in the previous process (segmentation), this region must be similar to the region that was determined by the expert in the ground truth. In order to measure border features, geometrical measurements was applied, such as area, perimeter, irregularity index, etc. to analyses the general shape of the lesion. There are several methods that can be used in clinical diagnosis, the common one is ABCD rule for dermoscopy and pattern analysis. A denoted to Asymmetry, B denoted to Border, C denoted to Color and D to Diameter.

It can be concluded from the results that the system can be used by patients and dermatologist to diagnose the skin cancer. This tool is more useful for region that far away from city or rural areas where experts in the medical field or dermatologist may not be available. Due to the development in the field of image acquisition and provides the appropriate tools, a patient become able to making initial diagnosis using this system, it can serve the purpose of automatic diagnostics of the skin cancer.

5.2 Future Work

Throughout the thesis we tried to offer a detailed explanation about the basic components for diagnosis system of skin cancer, and we discussed a number of directions to achieve progress and development of this system to serve all people (dermatologists and patients), to bring it up to a level that can be relied upon diagnosis. The CAD system should be experimenting on thousands Dermatology lesions, and different skin colors, as well as development in the field of image acquisition by taking images in different lighting conditions. All this requires the active participation and regulated by the various medical institutions.

The CAD system should be tailored to suit the needs of: expert users including experienced dermatologists, who need system to add second opinion to its clinical diagnosis and non-expert users including general practitioners, who needing system to get better diagnostic accuracy and gaining more of trust.

The ultimate target of this system is people/patients who should able to do a preliminary analysis to their skin at home. In order to accommodate certain requirements we should rise to the extra technology like using ultra sound to extract information about the depth and surface of the lesion. Also patients should helping doctors by giving accurate information, over a period of time, taking series of images in order to provide more accurate analyzing about given lesion. Further improvements within this field of research including the results of border detection, other metrics and factors can be incorporated in the performance index. The technique of feature extraction can also be further advanced by applying supervised pattern analyzing techniques to detect accurate differential structures such as radial streaming, globules, dots, etc. from implementation another segmentation technique and trying to reach the classification stage in order to diagnosis the cancer into benign or malignant, we become able to making comparison between results after classification.

REFERENCES

[1]   Xu, L., et al., (1999). "Segmentation of Skin Cancer Images", Image and Vision Computing, 17(1): 65-74.

[2]   Lee, T., et al., (1997). "Dullrazor®: A Software Approach to Hair Removal from Images" Computers in biology and medicine, 27(6):533-543.

[3]   Handels, H., et al., (1999). "Feature Selection for Optimized Skin Tumor Recognition Using Genetic Algorithms", Artificial Intelligence in Medicine, 16(3): 283-297.

[4]   Rajab, M I., Woolfson, M S., and Morgan, S P., (2014). "Application of Region-Based Segmentation and Neural Network Edge Detection to Skin Lesions", Computerized Medical Imaging and Graphics, 28(1): 61-68.

[5]   Erkol, B., et al., (2005). "Automatic Lesion Boundary Detection in Dermoscopy Images Using Gradient Vector Flow Snakes" , Skin Research and Technology, 11(1): 17-26.

[6]   Fondón, I., Serrano, C. and Acha, B., (2007). "Segmentation of Skin Cancer Images Based on Multistep Region Growing", MVA2007 IAPR. Machine Vision Applications Conference, 18-16 May 2007, Tokyo, JAPAN.

[7]   Celebi, E., M., et al., (2007). "Unsupervised Border Detection in Dermoscopy Images", Skin Research and Technology, 13(4): 454-462.

[8]   Yogamangalam, R. and Karthikeyan, B., (2013). "Segmentation Techniques Comparison in Image Processing", International Journal of Engineering and Technology , 5(1): 307-313.

[9]   Oommachen, N., et al., (2013). "Melanoma Skin Cancer Detection Based on Skin Lesions Characterization", IOSR Journal of Engineering (IOSRJEN), 3(2): 52-59.

[10]   Bhuiyan, M.A.H., Azad, I. and Uddin, M.K., (2013). "Image Processing for Skin Cancer Features Extraction", International Journal of Scientific and Engineering Research, 4(2): 1-6.

[11]   Ramteke, N.S. and Jain, S.V., (2013). "ABCD Rule Based Automatic Computer-Aided Skin Cancer Detection Using MATLAB®", International Journal of Computer Technology and Applications, 4(4): 691-697.

[12]   Cheerla, N. and Frazier, D., (2014). "Automatic Melanoma Detection Using Multi-Stage Neural Networks", International Journal of Innovative Research in Science, Engineering and Technology, 3(2): 9164- 9183.

[13]   Sonali, J. and Kamat, K., (2014). "Segmentation Based Detection of Skin Cancer", 13th IRF International Conference, 20th July-2014, Pune, India.

[14]   Sanjay J., Mehran K. and Vaishali G., (2015). " Skin Cancer Detection Using Digital Image Processing ", International Journal of Scientific Engineering and Research, 3 (6):138-140.

[15]   Das, P. and Gangopadhyay, M.M., (2015). "Removal of Hair Particles from Skin Disease Images Using Pixel Based Approach", International Journal of Computer Science and Information Technologies,6(5):4154-4158.

[16]   Gajbar, A.M. and Deshpande, A.S., (2015). "Detection and Analysis of Skin Cancer in Skin Lesions", International Journal of Advanced Research in Computer and Communication Engineering, 4(2):1173-1178.

[17]   Bath-Hextall F., et al., (2007). Interventions for Preventing Non-Melanoma Skin Cancers in High-Risk Groups, UK.

[18]   Madison K., (2013). Automated Curved Hair Detection and Removal in Skin Images to Support Automated Melanoma Detection, MSc Thesis, Case Western Reserve University, Department of Electrical Engineering. Cleveland, Ohio.

[19]   Iyatomi, H., (2010). "Computer-Based Diagnosis of Pigmented Skin Lesions" INTECH Integrated Technologies, 33(2):183-199.

[20]   Garnavi, R., (2011). Computer-Aided Diagnosis of Melanoma, PhD thesis, The University of Melbourne, Department of Electrical and Electronic Engineering,. Australia.

[21]   Pehamberger, H., Steiner, A. and Wolff, K., (1987). "Pattern Analysis of Pigmented Skin Lesions", Journal of the American Academy of Dermatology, 17(4):571-583.

[22]   Rigel, D.S. and Carucci, J.A., (2000). "Malignant Melanoma: Prevention Early Detection and Treatment", Cancer Journal for Clinicians, 50(4): 215-236.

[23]   Masood, A., and Ali Al-Jumaily, A., (2013). "Computer Aided Diagnostic Support System for Skin Cancer", International Journal of Biomedical Imaging, 3(11): 1-20.

[24]   Sonka, M., Hlavac, V. and Boyle, R., (2014). Image Processing Analysis and Machine Vision. United States of America.

[25]   http://www.cs.uu.nl/docs/vakken/ibv/reader/chapter10.pdf . 30 March 2016.

[26]   Cover, T., and Hart, P., (1967). "Nearest Neighbor Pattern Classification", IEEE Transactions on Information Theory, 13(1): 21-27.

[27]   Dasarathy, B,V., (1991). Nearest Neighbor ({NN}) Norms:{NN} Pattern Classification Techniques, California.

[28]   Burroni, M., et al., (2004). "Melanoma Computer-Aided Diagnosis Reliability and Feasibility Study", Clinical cancer research, 10(6): 1881-1886.

[29]   Huang, G.B., Zhu, Q.Y. and Siew, C.K., (2006). "Extreme Learning Machine: Theory and Applications", Neurocomputing, 70(1): 489-501.

[30] Sultana, A., et al., (2012). "Preliminary Work on Dermatoscopic Lesion Segmentation", 20th European Signal Processing Conference (EUSIPCO), 27-31 Aug. 2012 , IEEE.

[31] http://www.inf.u-szeged.hu/ssip/1996/morpho/morphology.html , 12 June 2016.

[32] Qidwai, U. and Chen, C.H., (2009). Digital Image Processing: an Algorithmic Approach with MATLAB., CRC press. London.

[33]   Efford, N., (2000). Digital Image Processing: a Practical Introduction Using Java. Boston.

[34] Soille, P., (2004). Erosion and Dilation In Morphological Image Analysis , Ispra (Va), Italy.

[35] https://en.wikipedia.org/wiki/Closing_(morphology) , 18 March 2016.

[36] Tiwari, S., et al., (2013). "Morphological Algorithms for Image Enhancement", International Journal of Mathematical Archive (IJMA), 4(10): 2229-5046.

.

CURRICULUM VITAE

PERSONAL INFORMATION

Name Surname   : AZHAR KASSEM FLAYEH

Date of birth and place   :26/8/1968 Baghdad

Foreign Languages   :English

E-mail   :flowers201075@yahoo.com

                               azharkassem@yahoo.com

EDUCATION

Degree	Department	University	Date of Graduation
Graduate	Computer Engineering	Yidiz Technical University	2016
Higher Diploma	Information Technology	Iraqi Commission for Computers &Informatics	2011
Undergraduate	Physics	Baghdad University	1990
High School	Science	Almuroj Secondary School	1986

PUBLISHMENTS

Conference Papers

1. Flayeh, A., (2016). "Detection of Skin Cancer by Using Image Processing Techniques", ICAS 2016" "International Conference on Advances Science, 31 August-2 September 2016, Istanbul, Turkey.