Detection of Melanoma in Dermoscopy Images

Using Local Binary Patterns

Author

Sidra Naeem

NUST201362458MCEME35213F

Supervisor

Dr. Farhan Riaz

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY ISLAMABAD

OCTOBER, 2015

Skin Cancer

Author

Sidra Naeem

NUST201362458MCEME35213F

A thesis submitted in partial fulfillment of the requirements for the degree of MS Computer Engineering

Thesis Supervisor

Dr. Farhan Riaz

Signature:_____

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY ISLAMABAD

OCTOBER, 2015

Declaration

I hereby ratify that I have developed this thesis titled as " Detection of Melanoma in Dermoscopy Images Using Local Binary Patterns " wholly on the basis of my personal efforts under the earnest toil and sincere guidance of my supervisor Dr. Farhan Riaz. All the sources used in this thesis have been cited and contents of the work have not been plagiarized. Any section of the presented work has not been submitted for degree of qualification to any other university.

Signature of Student

Sidra Naeem

NUST201362458MCEME35213F

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical and spelling mistakes. Thesis is according to the format given by the university.

Signature of Student

Sidra Naeem

Signature of Supervisor

Dr. Farhan Riaz

Copyright Statement

Acknowledgments

"In the name of Allah, the most Merciful, the most Compassionate"

First and above all, I praise Allah, the almighty for providing me this opportunity and granting me the capability to proceed successfully.
I would like to express my deepest gratitude to my supervisor Dr.Farhan Riaz for his excellent guidance, motivation, enthusiasm, and immense knowledge. His invaluable help of constructive comments and suggestions throughout the thesis work have contributed to the success of this research.
I acknowledge the contributions of Dr.Ali Hassan as marvelous. I am deeply grateful to him for the weekly meetings that helped me sort out the technical details of my work.
I would also like to thank the rest of my thesis committee members: Dr.Saad Rehman and Dr.Usman Akram for their encouragement, help and insightful comments.
One person who has been very supportive to me is Dr.Sajid Siraj. I would like to thank him, who as a good colleague, was always willing to help and give his best suggestions.
In the end, my deepest gratitude goes to my beloved parents. My parents have been the single greatest investor in my career. I would like to share the profound gratitude from my deep heart to my beloved parents, my sisters and brother for their love and continuous support. My sincere thanks and prayers go with my nephew, Muhammad Ibrahim, who brings such joy to our lives.
I warmly thanks my dearest friend Miss Rida Nisar who made available her support in a number of ways.

Abstract

This thesis is motivated by the potential gains that can be achieved by the use of computer assisted decision systems (CAD) for diagnosis of melanoma in the skin using dermoscopy. A CAD system provides quantitative and objective evaluation of the skin lesion versus the subjective clinical assessment. It automates the skin lesion analysis, and reduces the amount of repetitive and tedious tasks to be done by physicians. This research is mainly focused on the computer vision perspective to design a CAD system which will facilitate the physicians. A complete pattern recognition system that includes three vital stages to conform the analysis of skin lesions by the clinicians: segmentation, feature extraction and classification. The data-set contains images and annotations provided by physicians.
Segmentation is an imperative preprocessing step for CAD system of skin lesions. Segmentation is performed using active contours with creasness features. Feature extraction of segmented skin lesions is a pivotal step for implementing accurate decision support systems. Physicians are interested in examining a specific clinically significant region in a lesion. Such a region is expected to have more information in the form of texture that can be relevant for detection. In case of detection of melanoma various local features for example pigment network and streaks usually occur in peripheral region of the lesion. This led to the extraction of peripheral part for feature extraction instead of whole lesion processing. We propose novel techniques for feature extraction on peripheral part of the lesion using joint histogram of multiresolution Local Binary Pattern along with the contrast of the patterns. Classification results obtained from the proposed feature matrix were compared with some other texture descriptors, showing the superiority of our proposed descriptor.

Contents

List of Figures

List of Tables

1 Introduction

Abnormal growth of cells is the primary cause of cancer also known as malignancy. According to World Health Organization 2015 (WHO) death toll reached 14.1 million due to cancer world wide. There are various types of cancer for example skin, breast, colon and lung etc. The mortality trends and statistics shows that skin cancer is the fourth most deadly form of cancer worldwide as shown in Figure 1.1. Usually, skin cancer is divided into two main types: melanoma and non-melanoma. Melanoma is among the most deadliest type of skin cancer, rapidly increasing in the world. The American Cancer Society estimates that more than 135,000 new cases of melanoma are diagnosed in the US. In 2015, an estimated 73,870 new cases of melanoma are diagnosed, with about 42,670 in males and 31,200 in women [1].

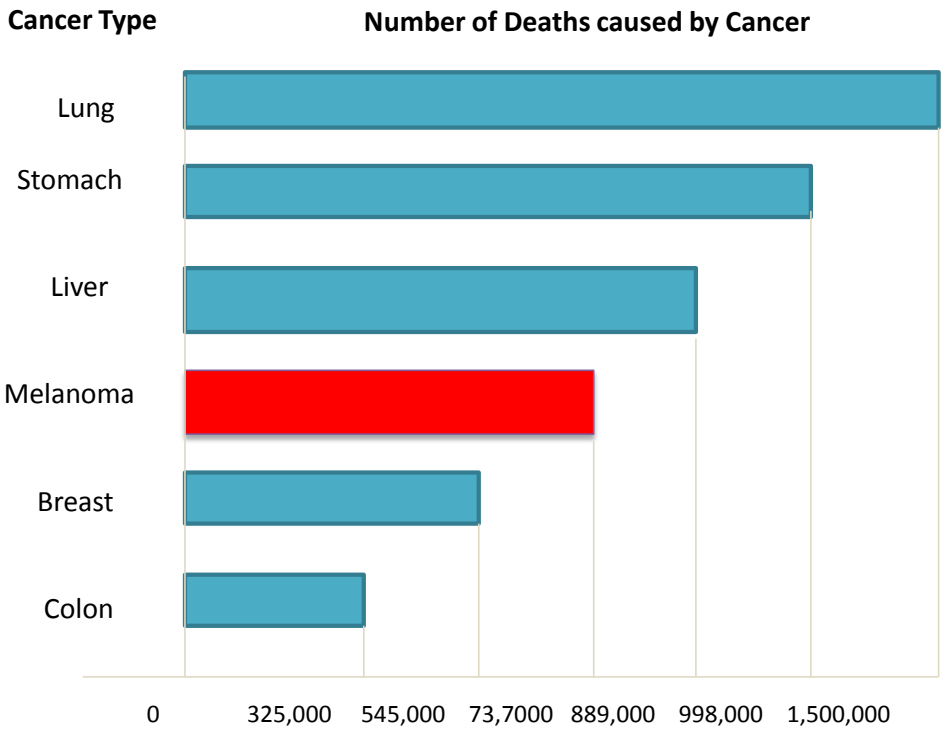

Figure 1.1: Cancer mortality rate (Reprinted from [1])

Vital warning symptoms and signs of melanoma includes mole that is varying in its size, shape, or color or the appearance of a new growth of the mole on the skin. Along with Melanoma, Non-melanoma skin cancers (NMSC) consists of basal[1] and squamous cell carcinomas [2]. NMSC are also very difficult to diagnose at its early stage because of similarity in appearance with normal skin [15]. The estimated occurrence of NMSC cases in 2015 is about 3.5 million cases. NMSC is highly curable if properly diagnosed at its initial stage. For an early detection of skin cancer the best way is to recognize new or changing skin growths, mainly those that look different from other moles.

1.1 Types of Skin Cancer

Like all body tissues our skin comprises cells:- basal cells, squamous cells and melanocytes [16]. Basal, squamous and melanocytes cancers are named after the skin cell in which the cancer develops. The two skin cancers basal and squamous are combined together as Non Melanoma Skin Cancers (NMSC). Merkel cell tumors and dermato fibrosarcoma protruberans are other kind of skin cancer.

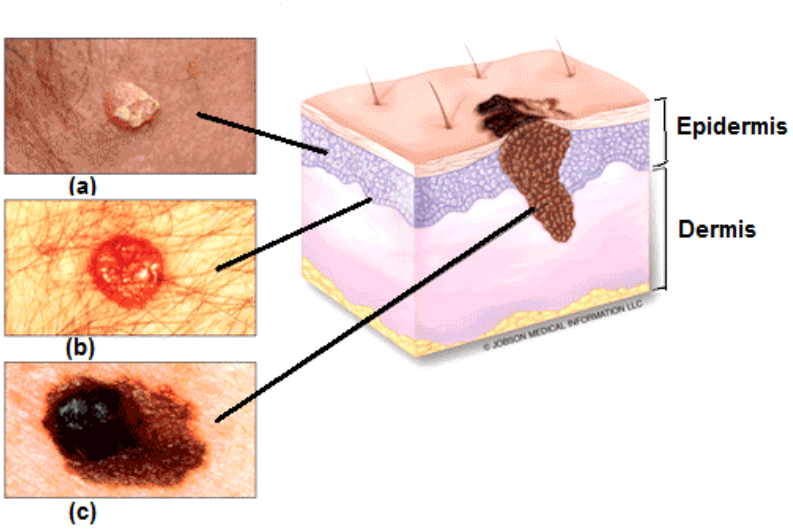

Figure 1.2: Types of Skin Cancer (a) Squamous Cell Carcinoma, (b) Basal Cell Carcinoma and (c) Melanoma (Reprinted from [2])

[1]Basals cell carcinoma (BCC) are the lesions that arise in the skin's basal layer, the deepest layer of the epidermis.

[2]Squamous cell carcinoma (SCC) are the lesion that arise in the squamous cells, which is the upper layer of the epidermis

1.1.1 Melanoma Skin Cancer

American Cancer Society reported that about 73,870 cases of melanoma will be diagnosed in 2015. Melanoma is the less common type of skin cancer, accounting for less than 5% of all other type of skin cancer. However, it is by far the most aggressive kind of skin cancer because it is more possible to metastasize than other skin tumors. This characteristic makes melanoma the deadliest form of skin cancer [1]. The rate of melanoma has increased significantly in the last decade. Consequently, it has received attention both from the public health field, with medical prevention campaigns, and from the cancer research field. Melanoma starts from the cells known as melanocytes which controls the pigment in our skin. Melanoma commonly look like a mole at early stage. Mostly melanoma are in black or brown color but some times they contain variation in color. Main cause of the melanoma is Ultra-violet (UV) exposure [1]. It has been found that the early detection of melanoma in the patients can significantly increase the survival rates for the patients [17].

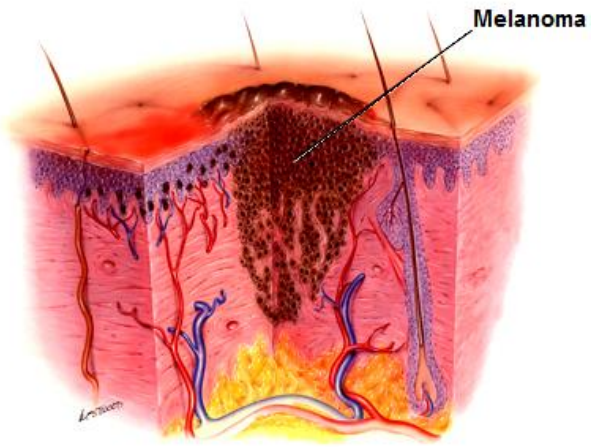

Figure 1.3: Melanoma cells extending from the surface of the skin (Reprinted from [2])

Based on clinical significance there are four main types of Malignant Melanoma (MM) as discussed below [18]:

- Superficial spreading melanoma (SSM)

SSM is a form of melanoma in which the malignant cells tend to stay within the tissue of origin for a prolonged period. At first, SSM grows horizontally on the skin surface. This is known as the radial growth phase. The lesion presents as a slowly, enlarging flat area of discoloured skin.

- Nodular melanoma (NM)

In NM, malignant melanoma cells proliferate downwards through the skin. This is known as vertical growth. The lesion presents as a nodule (lump) that has been rapidly with the passage of time. A NM can penetrate deeply within the skin within a few months of its first appearance.

- Lentigo melanoma (LM)

LM is an early form of melanoma in which the malignant cells are confined to the tissue of origin, the epidermis. LM is similar to the SSM, and normally looks like a flat lesion and in brown or dark brown color.

- Acral lentiginous melanoma

Acral lentiginous melanoma is a form of melanoma characterized by its site of origin: palm, sole, or beneath the nail. It is more common on feet than on hands.

1.1.2 Non-Melanoma Skin Cancer

NMSC comprises basal and squamous cell carcinomas. They are named after the types of skin cells from which these cancers develop. BCC develops from the basal cells. These cells are present in the deepest layer of the epidermis and around the hair follicle. While SCC begins form the cells named as keratinocytes, which are present in the epidermis. Non Melanoma skin cancer is considered to be a mixture of both these types of cancers. These are not very deadly but treatment for removal is painful [19].

1.2 Diagnosis of Melanoma cancer

For the effective melanoma diagnosis, dermatologist will carefully examine growths, moles, and dry patches. If the mole appears melanoma to the dermatologist then a sample of skin tissue will be taken from the suspicious area and sent to a lab to be clearly examined under a microscope. This procedure is known as skin biopsy. There are several ways for skin biopsy. The choice of the procedure for biopsy depends upon the size of suspicious area. But this strategy can prove problematic if there are multiple suspected moles present in the body. The decision to biopsy all of the suspected moles is a challenging tasks for dermatologist [19]. So dermatologists will look for the new method such as dermoscopy, which will allow him for better evaluation of the lesion. Dermoscopy can also give a dermatologist more confidence that whether a biopsy is inevitable or can be avoided.

1.2.1 Dermoscopy

Dermoscopy is a non invasive procedure used for *in vivo* examination of skin lesions. This procedure allows evaluation of specific, minute pigment patterns, enhancing differential diagnosis. The physician apply gel on the skin and inspects it with a magnification instrument (dermatoscope, stereomicroscope, or a digital imaging system), which amplifies the lesion 6–100×, magnification depends upon the instrument used [2]. This magnification allows the recognition of several surface and subsurface structures, that are invisible to naked eye [20]. The features of a brown spot become more prominent and the pattern of pigment can be seen clearly with the help of dermoscopy. Dermoscopy helps to determine the depth of the pigment, which is one of the most crucial information required for the evaluation of skin cells.

1.3 Challenges in diagnosis of Melanoma cancer

Early detection of melanoma is the ultimate goal for physicians because, it eventually increase the survival rate [17]. One of the main factor affecting survival for the people with melanoma is thickness of the tumor when it is diagnosed. Thickness indicate the staging process for melanoma. The most important warning sign for melanoma is a spot that is changing in its size, shape, and color. However some melanoma do not meet these criteria due to similarity in appearance with the benign lesions [21]. Differentiating melanoma in its early stages from other benign skin lesions remains a difficult task even for experienced dermatologists. Several screening and scoring methods (section 4.2) have been widely used to improve diagnostic accuracy, but efficient methods to extend the diagnostic capabilities are still lacking.

1.4 Need for CAD System in Dermatology

The survival rate of patients suffering from melanoma increases at its early detection. To measure, manage, diagnose and treat cancer at its early stage is a trivial task for the physicians because of no clear symptoms [19]. So medical imaging has turned into an important part in the early diagnosis of cancer. Medical imaging is the first step for preventing the cancer spread by making its early detection and it is possible to cure the cancer altogether. Computer aided diagnosis (CAD) systems are being used by clinicians for detection and characterization of the diseases. CAD systems can be used to assist physicians by providing second opinion and may also help the physician to use the varying technological development in the field of dermoscopy.

1.5 Computer Vision based CAD System

Designing a CAD system for the skin cancer detection, require three components: data record, Computer Vision (CV) and Human Computer Interaction (HCI). The data record system is a twofold mechanism; firstly it contains the detailed information of the patients for the treatment and secondly it consists of queries to retrieve/store the data effectively. Once images are acquired and stored, the next step is to process them. The CV module has to process and analyze the various patterns present in the images to make an efficient diagnosis. Finally, HCI module provides user interface to the designed CAD system. To perform appropriate actions HCI module can use some voice commands for interface handling, or recognize some kind of gestures.

Therefore, these three modules build up a CAD system based on CV approach to provide improved diagnostic capabilities by providing second opinion to the physicians. A CAD system is supposed to be more robust in terms of time complexity. The computerized analysis of dermoscopic images can be useful to assess and detect different features from which dermatologists derive their conclusion.

1.6 Objective of this Thesis

From the requirements presented by the CAD systems for dermoscopic images, we focus our research on the CV dimension to detect melanoma. In past several CAD systems have been presented to assist dermatologist. Dermoscopy is an procedure used by the physicians. The thesis was written with the purpose of identifying the aspects related to the development of CAD system, from the physiologic nature underlying the formation of skin cancer, to the techniques required and explored to implement a system for melanoma detection. Various CAD systems adopt different pattern recognition approaches i.e. by the extraction of features (color, texture and shape) from each dermoscopy image. Melanoma can be distinguished form benign lesions on the basis of various visual clues. The survival rate of patients suffering from melanoma increases at its early detection. But melanoma at its initial stages have features similar to the benign lesions. To handle this issue there is a need for an automated system that differentiate melanoma from benign at its initial stage. In this research, two different experimental setups have been presented to explain the potential of the proposed automated systems. The experimental setups contains distinct visual features for the diagnosis of melanoma. For our experiments data acquisition was done from one of the major Portugal hospital discussed in chapter 5.

1.7 Contributions

Main objective of this thesis are as follows:

(1) Visual Descriptors: We proposed an automated system based on visual descriptors for melanoma detection from dermoscopy images. Visual descriptors contains asymmetry index for shape feature and standard Local Binary Pattern for extracting differential structures from the images. The automated system contains segmentation, feature extraction followed by classification.

(2) Extraction of Peripheral part from skin lesion: Instead of processing of the whole lesion, we extract peripheral part of the skin lesion based on Computer Vision techniques. This peripheral part contains significant information in the form of local dermoscopic patterns.

(3) Joint Histogram of Multiresolution Local Binary Pattern (LBP) and Local Contrast: The detailed analysis for peripheral region of the lesion is achieved using multiresolution Local Binary Pattern. This methodology integrates multiresolution LBP patterns with the strength of the pattern i.e Contrast. The proposed novel methodology gives better classification results in comparison with the other descriptors presented in this thesis.

1.8 Overview of the Thesis

Chapter 1 explains the motivation of an automated diagnosis system for melanoma detection. It also encompasses mortality statistics, symptoms, causes and types of skin cancer. Chapter 2 focus on some basic skin imaging techniques used for screening malignant melanoma. Chapter 3 describes the conventional pattern recognition structure centered on CV for the design of automated CAD system for melanoma. Chapter 4 describes the background of computer vision techniques for diagnosis of melanoma based on pattern analysis. Chapter 5 includes description of dataset. Proposed methodology and contribution are represented in Chapter 6. Finally, Chapter 7 concludes the thesis with relevant areas to be further investigated.

2 Skin Imaging Techniques

Dermatology is a technique in which majority of skin examination can be done by the visual inspection. Dermatology is mostly *non invasive* [1] procedure [22]. The diagnosis made by the physicians is very subjective to the color, distribution of the lesion based on the distribution and color of a lesion. There exists several limitations in the diagnosis of skin diseases for example depth of the lesion, size and various internal features of superficial lesions [19]. Such limitations result in need for an objective *in vivo* observation and evaluation of the skin. Dermoscopy is a clinical procedure, widely used by the dermatologists to carry out an *in vivo* observation of skin lesions. The critical analysis for the characteristics of skin is presented below.

2.1 Introduction to Skin Lesions

The skin is made up of an epidermis, dermis and subcutaneous fat layers. Several skin diseases affects a specific layer of the skin. For the extraction of skin features it is very helpful to use multi layer skin model. The most widely used is the four layer skin model which contains epidermis, dermis and subcutaneous as shown in Figure 2.1 [23].

2.1.1 Epidermis

The epidermis layer as shown in Figure 2.1 consists of connected tissues. It has melanin producing cells, the melanocytes. Melanocytes basically absorbs most of the light.

2.1.2 Dermis

Dermis layer consists of further two sub layers: papillary and reticular dermis. Dermis consists of collagen fibers, sensors, blood vessels and nerve ends. While collagen fibers are thinner and they act as backscattering layer in papillary dermis.

[1] *non invasive* is a procedure not involving the making of a relatively large incision in the body.

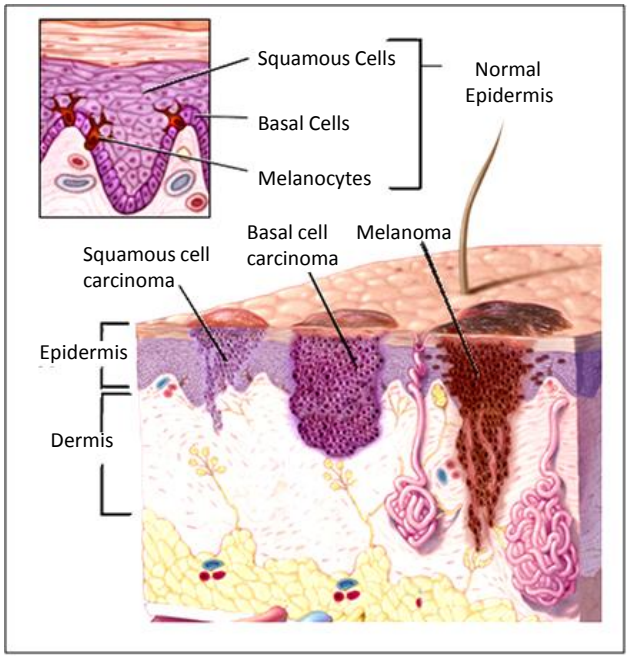

Figure 2.1: Skin layers and cancer generation (Reprinted from [2])

2.1.3 Subcutaneous fat

This layer consists of adipose tissue that are separated by connective tissue consisting of blood vessels, lymphatics and nerves.

2.2 Skin Imaging Techniques

Skin imaging includes various modifications of electromagnetic wave imaging such as optical, infrared, nuclear magnetic resonance, multispectral imaging, acoustical wave imaging and mechanical wave imaging. Different illumination method called epiluminence microscopy (ELM, or dermoscopy) can be used to get the image from the various layers of the skin. The light is directed straight in to the skin layers and then reflected back through skin providing detailed information about the layers of the skin [24]. Another appealing solution of obtaining detailed information from skin is multi spectral photography, which uses narrow frequency bands of illumination light.

Some of the skin imaging techniques are described in next section.

2.2.1 Working Principle of Dermatoscopy

Dermatoscopy widely known as ELM or Dermoscopy is a diagnostic procedure that is used mostly in dermatology for the identification and diagnosis of skin lesions [25]. This procedure allows the detailed identification of the skin structures that are not visible by the naked eye [20].

Principle

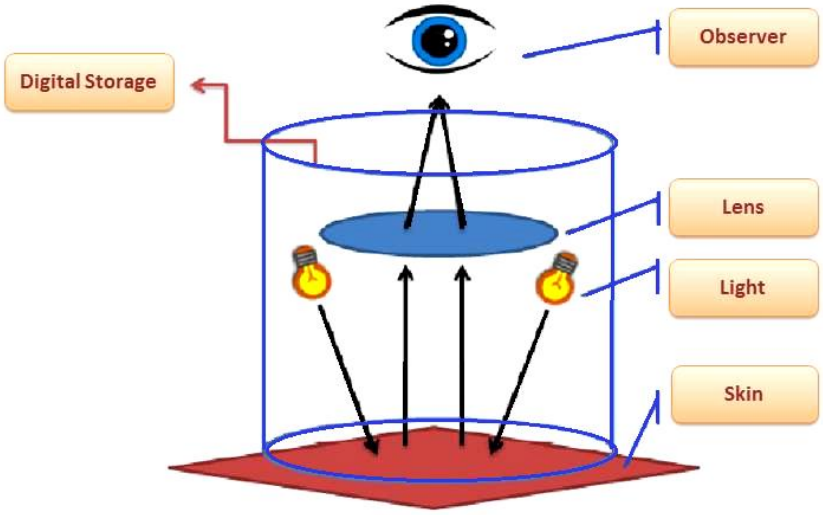

Figure 2.2: Dermatoscopy Principle (Reprinted from [3])

The process of dermatoscopy is achieved by applying an oil immersion on the skin. Lens of a microscope is positioned directly to the surface of the skin. The lighting can magnify the skin by revealing the pigmented structure along with various shades of color. It allows detailed examination of the epidermis layer and provides the ability to assess structures as deep as in the reticular dermis, and the ability to record images. ELM devices are being used by physicians to get better visual inspection of skin. ELM is an inexpensive and useful tool to aid dermatologist by providing correct assessment of malignant lesions. However, the diagnostic is qualitative and potentially subjective, so it is dependent on the dermatologists experience for a correct assessment [26].

2.2.2 Multispectral imaging

ELM of imaging procedure make use of surface reflectance dominant illumination techniques to allow the visualization of subsurface structures and colors. The structures and colors of these subsurface along with their location and pattern have been

shown to improve a clinician's ability to diagnose early melanoma. Multispectral imaging is an emerging technique that captures image data at particular frequencies across the electromagnetic spectrum. The wavelengths may be separated by the use of filters that are sensitive to specific wavelengths. This can allow extraction of detailed information the naked eye fails to capture [27].

2.2.3 Optical Spectroscopy

Optical spectroscopy also termed as Reflectance spectroscopy is the study of light that has been reflected or scattered from a liquid, solid, or gas. As photons enters the surface, some of the photons are reflected from the surface, and some are absorbed. The photons that are reflected from surface are called scattered photons. To measure scattered photons spectrometer is used. The reflection and scattering property of photons depends on biochemical composition and the wavelength of light. Normally melanoma cells contain various optical characteristics from those of normal cells [28].

2.2.4 Other Image Acquisition Techniques

In literature several new technologies exist for the diagnosis of melanoma for example Computed tomography(CT) have also been used [29] in order to detect melanoma and track both progress of the disease and response to treatment. Positron emission tomography(PET) [30] has also been proven to be a extremely sensitive to light and suitable diagnostic procedure for various kinds of melanoma.

2.3 Conclusion

From the above discussion, it would be fair enough to conclude that there are several skin imaging technologies which are used for skin cancer diagnosis. Dermoscopy is a diagnostic procedure, that is widely used by the dermatologists to diagnose the skin pathologies. This diagnostic procedure allows the effective recognition of structures that are not visible to the naked eye. The images obtained from the dermoscopy allows the enhancement of different dermoscopic features and differential structures that can be analysed by some computer vision techniques.

It is vital to mention that the images acquired from the dermoscopy contain visual features: texture and differential structures. For computer vision part of CAD system of dermoscopy, texture features and differential structures should be given a higher priority for the diagnosis of melanoma cancer. Following this concept, the scope of this work is to use differential structures and texture features, which would

not only be able to generate optimal results but also a competent method to fulfill the requirement of more general computer vision module for CAD system.

3 Pattern Recognition

Computer-aided diagnosis (CAD) systems are increasingly being used as an aid by clinicians for detection and characterization of the diseases. The fact that computers have the capability of storing and processing large amount of data, perform complex calculations with high reproducibility makes them useful for implementing decision support systems. Designing a CAD system related to pattern recognition techniques is an important challenge in dermoscopy. Conventionally, a physician observes the affected lesion to have a detailed information of the tissue. Shape, color and texture are the visual features associated with the skin examination. These features provide physicians some important information about the skin lesion. In view of computer vision, this manual diagnostic procedure of the skin lesion involves pattern recognition (PR) system. Our automated systems for dermoscopy involves the full pipeline of a PR system as shown in Figure 3.1.

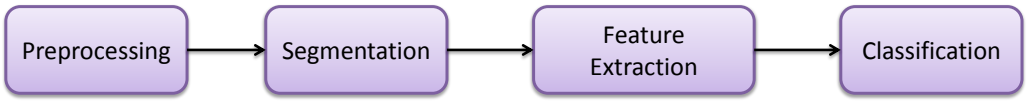

Figure 3.1: Steps of Pattern Recognition System

The first step after acquiring the images from ELM, is the preprocessing to reduce the artifacts for accurate detection of melanoma. Two types of artifacts: hair and reflections caused by the dermoscopic gel. These artifacts must be removed prior to the next step of the CAD system. Preprocessing step is followed by image segmentation technique. A better performance of segmentation means that clinically, relevant information is filtered and redundant information is discarded. Accuracy of the image segmentation effects the subsequent steps thus making segmentation one of the most crucial problem to tackle for automatic classification of skin lesions. After image segmentation, different type of visual features (color, shape and texture) can be quantified to make a feature vector. This feature vector is designed to ensure that a decision boundary is drawn between data samples of different classes. Finally, class labels are assigned to each subject on the basis of some common characteristics of feature vector. This is called classification.

3.1 Preprocessing

Preprocessing is a crucial step for an efficient CAD system of melanoma detection. Its purpose is to prepare the image for the segmentation procedure, by eliminating undesired artifacts and enhancing the image contrast of the lesion to facilitate the border detection step, while retaining its most important features.

3.1.1 Removal of Artifacts

Dermoscopic images often contain some artifacts for example smooth contrast, reflection due to dermoscopic gel, hair and air bubbles that complicate the border detection [31].

One possible way to address the removal of these artifacts is image smoothing using filters for example median filters and Gaussian filters etc. By choosing the adequate filter parameters, these can reduce the effects of artifacts without losing image information. The result of segmentation is poor due to insufficient contrast and sometimes due to uneven transitions between lesion and skin. For contrast enhancement, histogram stretching, histogram equalization, homomorphic filtering, and high pass filtering can be used.

The preprocessing block executes two main tasks: the detection and removal of the artifacts. Hair and reflection artifacts occur in dermoscopic images caused by the dermoscopic gel. Detection and removal of these artifacts are described in [32].

3.2 Segmentation

Segmentation serves the purpose of delineating the region of interest (ROI), which in a CAD system for skin lesions is the lesion's area. This is an essential step for a reliable functioning of the whole system. Accurate lesion segmentation is a challenging task. A correct segmentation allows different types of features to be extracted from the lesion region, which in turn leads to a more accurate classification. However, for the development of an automated diagnostic system for skin lesions, it is important to develop automatic segmentation techniques.

Segmentation of the images is divided into two types [4].

- **Segmentation by clustering**: This technique groups together data items according to some criterion to make a cluster.
- **Segmentation by fitting**: In this technique the model is associated with the grouped data items.

3.2.1 Segmentation by Clustering

Clustering is the process to identify how different types of data are related and creating new segments based on specific criteria. Clustering finds the relationship between data points so they can be segmented. Clustering criteria is specific to the application. In clustering, pixels may grouped together if they have the similar texture and color. There are two main techniques for clustering.

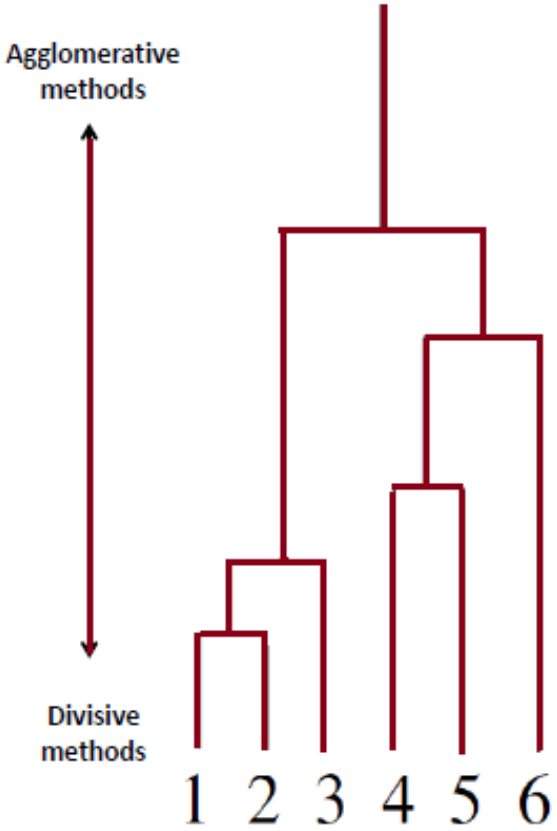

Figure 3.2: Dendogram corresponds to different segmentation techniques [4]

- **Divisive clustering:** This is a "top down" technique. All data items starts from one cluster, and splits iteratively as one moves down the hierarchy.
- **Agglomerative clustering:** This is a "bottom up" technique. All data items starts in its own cluster, and merge pairwise as one moves up the hierarchy.

There are various challenges associated with clustering techniques. One of the challenges is to decide the number of the cluster at the start of the algorithm. Second challenge is what is inter cluster distance? A good clustering technique should have

high similarity between data items within one cluster (known as intra-class similarity) and low similarity between data items from other clusters (known as inter-class similarity) as shown in Figure 3.3. In literature various techniques available for segmentation by clustering. K means is the most widely used example which is discussed below:

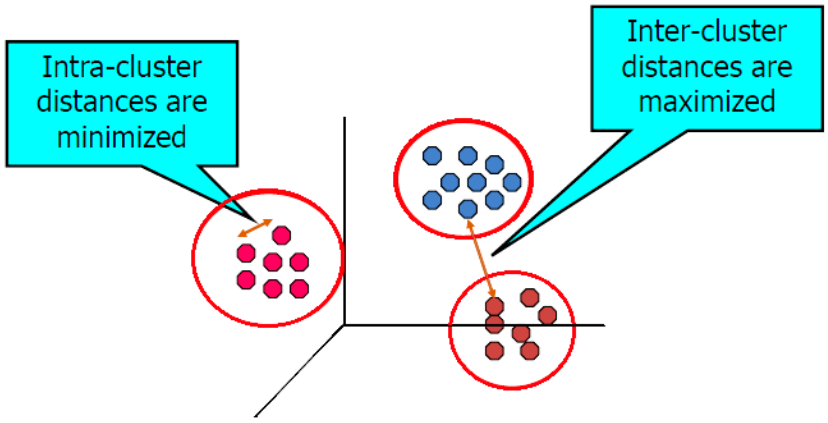

Figure 3.3: Intra-cluster and Inter-cluster distances of clusters (Reprinted from [5])

3.2.1.1 Segmentation by K means clustering

K-Means [33] is a least squares partitioning method that divides data items into K clusters. It partitions the data points into k clusters which is fixed initially. k centers should be chosen appropriately locations because different locations may result in different clustering at the end. Lets say the i^{th} cluster center be c_i.The data items need to be clustered are $\{x_1,x_2,x_3,...,x_m\}$. A starting point is selected randomly by choosing cluster centers and then find the Euclidean distance of each point from each cluster's center. Assign points to the cluster it is nearest to.

$$\sum_{k=1}^{m}\sum_{l=1}^{i}||x_k - ci||^2 \tag{3.1}$$

3.3 Feature Extraction

As referred in section 1.3, malignant melanoma are difficult to differentiate from other pigmented skin lesions, especially in its early stages [17]. However, it is crucial that a CAD system for melanoma detection is capable of doing this differentiation with at least the same accuracy of a dermatologist if it is to be accepted in a clinical setting. It is not intended to replace the professional opinion, but it would constitute an important tool to improve biopsy decision making. Similarly in contrast to the various visual diagnosis procedure for skin cancer, the CAD systems search for features in the lesion regions and combine them to characterize the lesion. The lesions will therefore be characterized by a feature vector, having n dimensional vector, containing the measures of the selected features that correspond to the region of interest in the image. Feature extraction can be divided into three major groups: texture, shape and color features. The features studied for each group are described in the following sections.

3.3.1 Texture Features

Texture feature extraction from skin lesion is motivated by its diagnostic value in the differentiation of malignant and benign tumors. Texture feature extraction is an efficient method to find the structure, orientation and roughness of the regions in an image. Texture extraction methods include statistical, structural, model and filtering-based methods [34].

3.3.1.1 Statistical methods

Statistical methods deals with statistical characteristics as well as spatial distribution of the gray level values in an image. They are used to measure the spatial distribution of gray level values by finding local features at each point in the image and then from the distributions of local features to derive a set of statistics.

First order statistics

First order statistics operates on individual pixels of an image. Since in a gray scale image every pixel corresponds to an unsigned integer. If k represents total bits used to indicate pixel values, a specific pixel value can range from 0 to 2^k -1 different values. The spatial distribution of a gray scale image can be analyzed by finding the probability distribution of pixel intensities. The probability distribution at different pixel intensities can be calculated with the help of histogram. The appearance of

the histogram indicates the kind of texture present in an image. Several statistical measures which describe the histograms are statistical moments, such as first order moment is mean and second order moment is standard deviation, third order moment is skewness and so on [35].

Second order statistics

These methods are based on pixel values along with pixel's neighborhood values. Gray Level Co-occurrence Matrices (GLCM) and Local Binary Patterns (LBP) are the examples of mainly used second order statistics.

Originally proposed by Haralick [36], GLCM first find the co-occurence matrix of an image and then calculate the texture features from this matrix.

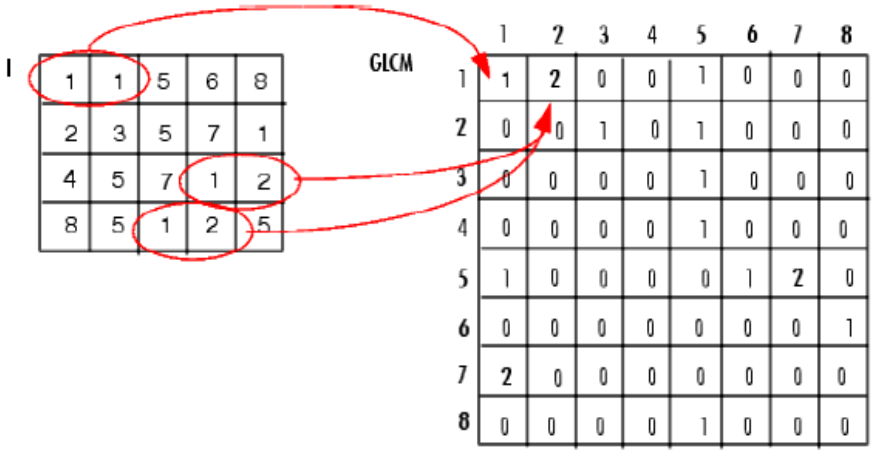

Figure 3.4: Gray Level Co-occurrence matrix [6]

LBP originally presented by Ojala et al. [37] is gray-scale and invariant texture descriptor that creates LBP codes at every pixel in the image by thresholding all the neighborhood pixesl with the value of central pixel and then concatenate the result in the form of a pattern as shown in Figure 3.5. It unifies the structural and statistical information of the texture using a histogram of the LBP codes.

3.3.1.2 Model based methods

These are used in the texture synthesis field. These methods are based on image construction model that can be helpful to illustrate image texture as well as used to synthesize an image. In these models intensity value of each pixel in an image

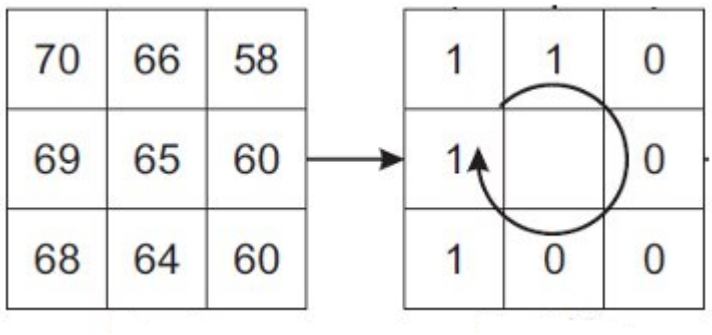

Figure 3.5: Local Binary Patterns [7]

depends on the intensities of its neighboring pixels values. There are basic three model based techniques: Markov Random Fields [38], fractals [39] and Multiresolution Autoregressive features [40].

3.3.1.3 Structural methods

Textures are made up of some basic primitives termed as 'texels'. Structural based methods based on the analysis of 'texels' which usually appear in near regular repetitive spatial arrangements of an image. These are achieved with the help of some placement rules and so a texture is identifiable using texels and different placement rules. Consequently, key components of the resulting models are structure and spatial characterization of texture elements. Shapes and edges are the examples of texture elements.

3.3.1.4 Signal processing methods

Statistical approach have been widely used for feature extraction, but the main issue with statistical approach is that image information related to frequency domain is lost. According to the research, human brain performs frequency analysis of an image [41]. The basic visual cortex of the human beings can be recognized as a mixture of frequency selective and orientation sensitive filters. Signal processing techniques involves filtering of the images in frequency, spatial or in spatial-frequency domain [42].

Spatial domain filtering:

This technique is the most widely used technique to take the information of texture characteristics from an image. Spatial domain filtering is the low level signal processing technique which is used mostly for noise removal, sharpening of an image,

detection of edges in an image etc. This technique involves the convolution of an image with a pre-defined mask using Equation 3.2.

$$f(t) * h(t) = \int_{-\infty}^{+\infty} f(\tau)\, h(t-\tau) \tag{3.2}$$

Here $f(t)\, and\, h(t)$ represents the function and mask for convolution respectively.

Edge in an image is considered to be the most vital and basic texture feature which is the boundary between different regions having various gray level features [43]. An edge occurs where the intensity distribution finds a discontinuity. Prewitt mask is shown in Figure 3.6.

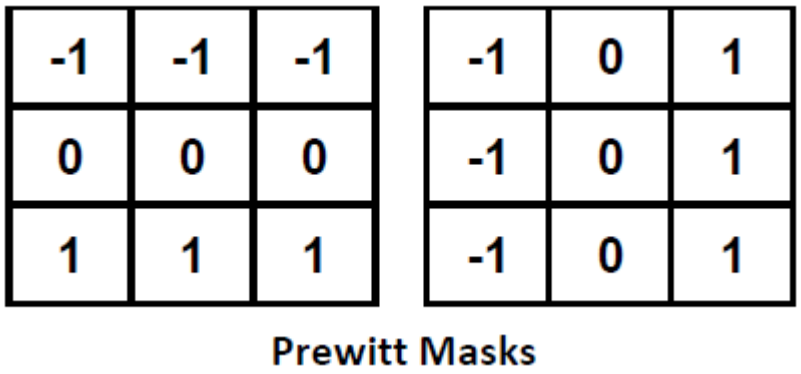

Figure 3.6: Prewitt filter used for Edge detection (Reprinted from [8])

Frequency domain filtering:

Convolution in time domain is computationally complex. In order to overcome this constraint, filtering in frequency domain is preferred. Since the masks in the spatial domain have their specific counterparts in frequency domain which can be used for frequency domain filtering. Spatial domain filtering is equivalent to multiplication of the image with the mask in frequency domain, so most of the computational time is reduced in frequency domain. Along with less computational extensive, filtering in frequency domain presents an alternative representation of the images which is very helpful for rotation invariant feature extraction. Images in frequency domain consist of magnitude and phase. Phase component deals with the rotation of an image while magnitude is kept constant, therefore the shift property [44] of Fourier Transform allows research on rotation invariant descriptors while working in the frequency domain.

Spatial-frequency domain filtering:

Frequency domain filtering is computationally less complex but comes with a constraint that it ignores all the spatial information of an image. This constraint is resolved by the use of wavelets transform. The wavelet transform is a method to handle the constraints of the fourier transform using Equation 3.3.

$$H_{\psi}(a,b) = \frac{1}{\sqrt{|\alpha|}} \int_{-\infty}^{+\infty} \Psi(\frac{x-b}{a})\, h(x)\, dx \tag{3.3}$$

Decomposed signal is $h(x)$, ψ indicate the mother wavelet, scaled by the factor of a and shifted by the factor of b to get different self similar wavelets indicating different filters in the frequency domain, and obtaining a decomposed signal $H_{\psi}(a,b)$ in result.

3.3.2 Shape Features

Shape features are relevant for the diagnosis of melanoma. In contrat to benign lesions, melanoma pertains asymmetric shape, irregular borders, and disordered architecture. Many shape parameters are of simple computation, which makes them valuable if they prove to have good differentiation potential.

3.3.2.1 Simple Shape Features

The most widely used shape features are the lesion area, compactness, rectangularity and length of its major and minor axis. The compactness ($C = \frac{4\pi A}{P^2}$) of the lesion is the ratio between the area of the lesion and the area of a circle with the perimeter (P) same as that of the lesion. The lengths of the major and minor axis are the maximum and minimum diameters of the lesion respectively. Lesion rectangularity is defined as the "ratio between the area of the skin lesion and the area of that smallest rectangle which is able to contain the whole lesion" [45], [46].

3.3.2.2 Shape Asymmetry

For melanoma detection, the analysis of shape asymmetry is a very significant factor, as the local appearance of certain structures or colors is most indicative for a melanoma. According to dermatologists, melanoma usually grow in an asymmetric way while benign tumors are symmetrical [47]. A melanoma is asymmetric, if after dividing the melanoma over its principal axes, the two halves do not match.

3.3.2.3 Fourier Descriptors

The Fourier Descriptors also provide shape related information that is invariant to translation, scaling and rotation of the images. Fourier descriptor corresponds to the coefficients of the Discrete Fourier Transform obtained from a shape signature [48]. Fourier descriptor is applied on the boundary points of the lesion.

$$z(n) = x(n) + iy(n) \tag{3.4}$$

i indicate imaginary numbers, and n represents number of boundary points.

3.4 Classification

The last step in the CAD system is the classification, which is responsible of classifying the test samples. On the basis of previously selected features, the system needs to find the class label to which the lesion belongs to. This is the final goal of a CAD system for melanoma detection. Classification methods are broadly divided into two groups; Supervised and Unsupervised learning.

3.4.1 Supervised Learning

In supervised learning, training data which is the input vector comprises their corresponding labels. Classifiers based on supervised learning method examines the training data samples and conclude an objective function from that training data to make a decision boundary between data samples of different classes. When the testing stage came the designed classifier is used to assign the class to the test value. There are variety of supervised learning classifiers found in literature, adapted according to the need of application. In this thesis we have used Support Vector Machines and K-Nearest Neighbor classifiers. These two classifiers are used to classify features, obtained with novel feature extraction methodologies.

3.4.1.1 Support Vector Machines (SVM)

SVM was presented by Vapnik in 1979 [49], is a supervised learning algorithm. It is capable of building optimal separating boundaries between input data sets by solving a constrained quadratic optimization problem. The basic training algorithm for

SVMs is only capable of constructing linear separators; however, there are numerous kernel functions available for example linear, sigmoid, radial basis function, and polynomial functions to include varying degrees of nonlinearity and flexibility in the model.

With the help of nonlinear mapping function $\rho(.)$, SVMs transforms the input feature space x into higher dimension such that $y = \rho(x)$. In this case there is a seperate decision boundary occur between different classes with the help of hyperplane. For each of the n different patterns from input data, lets assume that z = ± 1 indicating that whether pattern belongs to class $\omega 1$ or $\omega 2$. A linear discriminant in higher dimensional y space is

$$g(y) = a^T y \tag{3.5}$$

here the weight vector and transformed pattern vectors are in higher dimensional space. The separating hyperplane between different classes ensures

$$zg(y) \geq 1 \tag{3.6}$$

SVM main objective is to find the hyperplane having the maximum margin for separation; assumption for good generalization of classifier is that the distance between margins is larger Figure 3.7. The distance from a hyperplane to a transformed feature space y is $\frac{g(y)}{||a||}$ and a positive margin b exists.

$$\frac{zg(y)}{||a||} \geq b \tag{3.7}$$

The objective is to calculate the weight vector a that maximizes b. Support vectors are present to the margins of the hyperplane. Equation 3.7. The support vectors are the training data samples, very trivial to classify since they are very helpful for classification task.

3.4.1.2 K- Nearest Neighbor (KNN)

The k-Nearest Neighbors (kNN) is one of the simple supervised learning method for classification used to evaluate data. It computes the distance function which is the

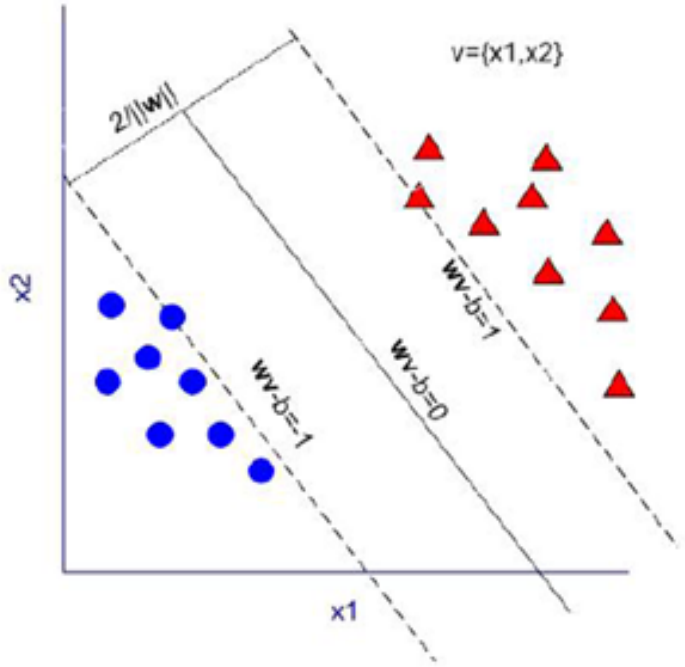

Figure 3.7: SVM hyperplanes for linearly separable data [9]

difference between the features belonging to the image in the test set and all the images in the training set.

The simplest form of KNN is when we consider only one neighbor $k = 1$ for the classification, the image to be tested is classified with the same label as the image in the training set which has the lower value of distance [50]. Let x_i be the input samples, the test case is assigned to the nearest neighbor where y_i represents the new data under consideration as shown in Equation 3.8:

$$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{3.8}$$

In addition to the classification based on one neighbor $k = 1$, it is possible to perform a classification using a higher number of neighbors. In this scenario, the method is similar to the one already described above, but instead of being classified by only one image from the training set, the image from the test set is classified based on the k closer images. Thus, if the positive classe is in the majority, the test image will be classified as 1 (melanoma). Otherwise, it will be classified as 0 which means that it belongs to the negative class (benign lesion).

There are different distances that can be applied to compare features [51]. The Euclidean and Bhattacharyya are two distances commonly used in order to classify dermoscopic images using kNN classifier.

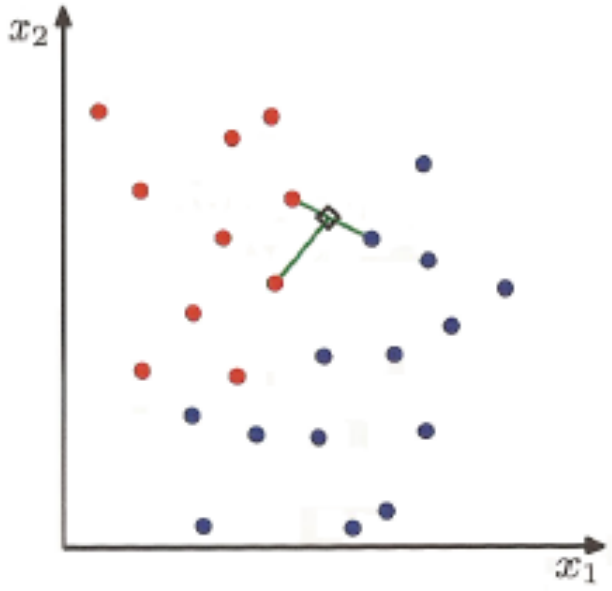

Figure 3.8: Representation of data points being classified by k-Nearest Neighbors classifier. The parameter k equals a value of 3 and the point takes the label of the class that is in majority which is the red one (Reprinted from [10])

There are two main challenges when applying KNN Classifier. The first is that we have to store the entire training set in memory in classification stage. Second challenge is: What is the value of k? Solution is that the best value of k depends upon the data; usually, larger value of k reduce the effect of noise on the classification, but at the same time decision boundary between classes become less distinct. Use k that gives lowest average error over the N training examples.

3.4.2 UnSupervised Learning

In this technique, data samples without class labels are used for classification. Though, unsupervised learning also encompasses various other methods that try to explain various key features of the data points. Various examples of unsupervised learning method include K means clustering described in subsubsection 3.2.1.1, self organizing maps, mixture models [52].

3.5 Assessment of Classifier

Defining a decision boundary on the trained data samples is the main task of the classifier. After classification of the data, there should be a criterion to decide whether the classifier is appropriate or not. To measure the performance of the classifier, model selection and assessment parameters are used that are described below.

3.5.1 Model selection of the classifier

Selection of a suitable model for training of the classifier is the main problem in real pattern recognition systems. Model selection deals with the distribution of data samples for training and testing phases. Hold out and random subsampling are the methods used to select model for appropriate classification.

Hold out methods

Data set is divided into two parts: training and a testing part. The training part is used to train the classifier and boundaries are drawn; and the testing part is used to assess the performance of the classifier. The results from hold out method might be misleading, because the training part does not completely define the characteristics of the data samples.

Random subsampling

Random subsampling is also method for model selection which gives better assessment of the true performance of the classier. With this method, data set is divided into fixed n partitions, n-1 partitions are used for the training, and the last set is used for the testing of the classifier. Using n fold cross validation, n classification models are built, to test on each n partition, and the classification results should be the average of over all n test sets. Leave-one-out (LOO) method is the limit scenario, in which only one data sample is used for testing of the classifier.

3.5.2 Assessment metrics

To measure the performance of the classifier in an objective way various assessment metrics are used. These metrics indicate the behavior of a classifier, that how accurately it is classifying the data. For unbalance data, assessment metrics are very helpful to analyze the performance of the classifier. These metrics includes numerical measures, confusion matrix and receiver operator characteristic curve.

3.5.2.1 Numerical Analysis

Numerical analysis describes the performance of the classier. Numerical measures include various metrics for evaluation of the performance. It usually comprise the following:

- True Positive (TP): The result which is truly positive when compared to actual class is called true positive.
- True Negatives (TN): The result which is truly negative when compared to actual class is called true negative.
- False Positive (FP): The result which is not positive actually but the classifier classifies it positive.
- False Negative (FN): The result which is not negative actually but the classifier classifies it negative.

These four quantities are used to derive specificity, sensitivity and accuracy for performance measure. Table 3.1 contains various performance evaluation measures.

Performance evaluation metrics		
Specificity	$\frac{TN}{TN+FP}$	Percentage of correctly classified negative samples
Sensitivity	$\frac{TP}{TP+FN}$	Percentage of correctly classified positive samples
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	Percentage of correctly classified samples

Table 3.1: Classifier performance assessment metrics

3.5.2.2 Confusion matrix

Confusion matrix is a specific table layout that allows visualization of the performance of a classifier. It is used when there is uneven number of samples of various classes. Columns indicates the samples in the predictive class label while rows represents the samples of the actual class. Confusion matrix enables an intuitive understanding of the system performance.

Actual Class	**Predicted class**		
		Positive	**Negative**
	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

Table 3.2: Confusion matrix

3.5.2.3 Receiver Operating Characteristics curves

When the training data is unbalanced, accuracy is not an appropriate measure to assess the performance of the classifier because it is strongly biased to vote the majority class labels. To avoid this problem in unbalanced data sets, the Receiver operating characteristics (ROC) curve is often used, as it exemplifies the performance of a classifier without considering the class distributions. The ROC curve is a graphical plot that visualize the performance of a binary classifier system as its threshold is varied. In an ROC curve, TP is plotted as a function of 1-TN. Values of sensitivity are calculated by varying the various cut off points of TN Figure 3.9.

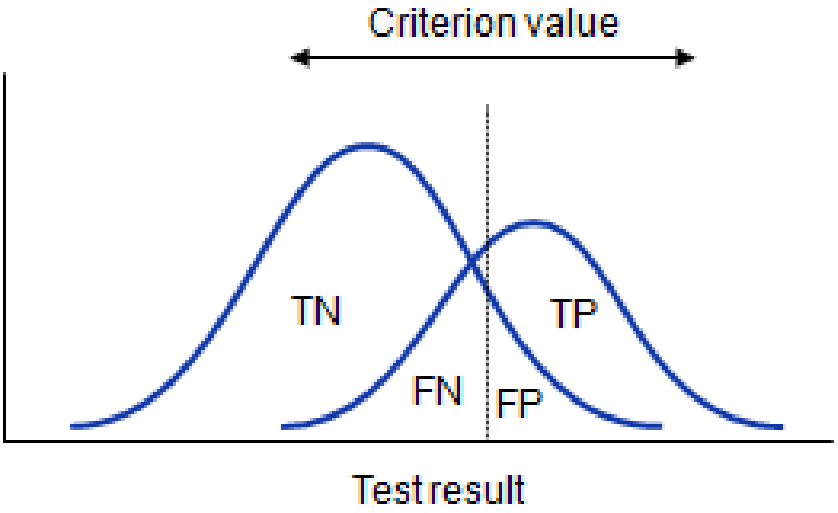

Figure 3.9: ROC curve (Reprinted from [11])

4 Computer Vision Techniques in Dermoscopy

In literature several CAD systems have been proposed for melanoma detection in which some systems try to mimic the performance of dermatologists by extracting several features using pattern recognition approach. In this chapter we will describe various state-of-the-art techniques in computer vision (CV) for melanoma detection, following by pattern analysis of melanoma. These patterns are local and global patterns which are discussed in the coming sections.

4.1 Clinical Signs of Malignant Melanoma

Melanoma is highly curable, with surgical excision of the primary cancer, if detected in its earliest stages [17]. The most important warning sign for melanoma is a spot that is changing in size, shape, and color [53]. Asymmetry of the mole is a visual sign for melanoma detection. Clinical significance of asymmetry according to dermatologists is that; melanoma grow in an asymmetric way while benign tumors are symmetrical. Along with asymmetry of the mole, border of the mole is also a parameter for the diagnosis of melanoma. Border of the most early melanoma are irregularly shaped. Physicians while analyzing melanoma also check the variation in colors as melanoma mostly comprises of multiple colors. Melanoma have also various differential structures in terms of pigment network and streaks.

In view of the above clinical sign and symptoms asymmetry, border, color, and differential structures help physicians to diagnose melanoma at early stage. These four parameters builds up an ABCD system. ABCD system [54] is the tool for detecting melanoma. These parameters are used for characterizing between malignant and benign skin lesions. In literature various scoring methods are available that assist in diagnosing of melanoma as described in section 4.2.

4.2 Scoring Algorithms

The main problem in the diagnostic of malignant melanoma is that it is highly dependent on subjective judgment. Several scoring systems such as ABCD rule, 7-

point checklist, Menzies method and Three point checklist help to improve diagnostic performance of the general physicians.

ABCD

ABCD [54] rule is a semi-quantitative analysis of four parameters as the most crucial for melanoma diagnosis: (A)symmetry is computed by comparing the similarity between the two half portion of the lesion; (B)order which defines the irregularity of lesion; (C)olor indicate the variation in colors and (D)ifferential Structures detect the presence of various structures (pigment networks, structureless areas, dots, globules, streaks) in the lesion. These parameters weighs differently, and are summed up to obtain a final quantitative result that relates to the probability of a lesion being a malignant melanoma. There is a formula to score and differentiate between benign or melanoma. The formula for ABCD rule is:

$$Score = (A_{score} * 1.3) + (B_{score} * 0.1) + (C_{score} * 0.5) + (D_{score} * 0.5) \tag{4.1}$$

The interpretation of the total score says if the score is < 4.75, the lesion is a benign; Score between 4.75 and 5.45 is considered a suspicious lesion; and the score > 5.45 illustrate that the lesion is melanoma. ABCDE is also a modified ABCD that adds E for Elevation or Evolution for melanoma diagnosis.

Menzies Method

The Menzies method was originally developed to differentiate invasive melanoma from other pigmented lesions. This method [55] search for negative features, namely symmetry, presence of a one color (usually dark brown and blue); while positive features look for the presence of blue-white veil, multiple brown dots, peripheral black dots and various colors. At least one positive feature must be found to classify a lesion as melanoma.

7 point checklist

One of the scoring methods known as 7 point checklist [56] which considers seven dermatologic parameters for melanoma. These parameters are divided into three major and four minor criteria. The three major criteria comprise of atypical pigment network, vascular pattern and blue-whitish veil and four minor criteria include irregular streaks, pigmentation, dots/globules and regression structures. A minimum of 3 points is required to classify lesion as melanoma.

Three point checklist

The three-point checklist [57] consists of three dermoscopic patterns which are used for detection of melanoma: 1) asymmetry of the lesion; 2) atypical network 3) blue white structures. One score point is assigned to each criterion. Score value of 2 is considered as lesion, so biopsy of the lesion is sufficient.

CASH

CASH algoirthm [58] is also a scoring method. CASH is an acronym for Color, Architectural disorder, Symmetry and Homogeneity of dermoscopic structures. This method is based upon evaluating a pigmented lesion by the presence of few versus many colors, architectural order versus disorder, symmetry of shape and pattern versus asymmetry and homogeneity versus heterogeneity of dermoscopic structures. Score of 8 form all the dermoscopic features is in CASH algorithm is considered as melanoma.

Even though these scoring systems have improved the overall diagnostic agreement of pigmented skin lesions, the main reason for which they were designed was not fully attained, because concordance between observers and experts is low [59].

4.3 Pattern Analysis for Melanoma

Pattern Analysis, is deemed as the most effective technique for the corrected diagnosis of melanoma, and it is deemed superior to the other algorithms as discussed in section 4.2. In 1987 Pattern Analysis, was proposed by Pehamberger [60] for the diagnosis of melanoma. This section explains the most important dermoscopic patterns of the lesions. Now a days, pattern recognition is widely used method for melanoma detection[61].

Pattern Analysis in dermoscopic images identify certain features, that are global and local. Global features provides a quick classification of the lesion, and they usually present in the whole lesion. Local features provides individual or grouped characteristics that usually present in some part of the lesion. Some of the findings concluded from this observation of local and global patterns were discussed in [62]; the multicomponent is the global feature was the most prognostic for melanoma detection, whereas the homogeneous, cobblestone and globular were most prognostic for benign lesions. While atypical pigmented network, regression structures and irregular streaks are the local features closely related to melanoma.

4.3.1 Local Pattern Analysis

Local patterns are pigment network, streaks and vascular structures [62]. Local patterns can be in irregular & regular or atypical & typical types . Figure 4.1 shows

example of local patterns. In this thesis we analyse two local patterns (pigment netwrok and streaks) for melanoma detection. So their clinical significance and automated systems for their detection is discussed below:

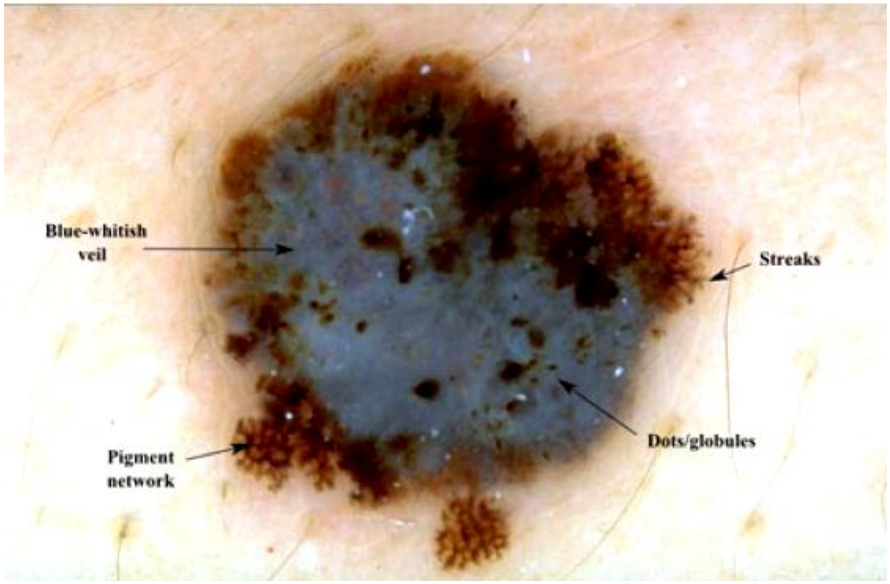

Figure 4.1: Examples of local patterns [12]

4.3.1.1 Pigment Network Detection and Analysis

Pigment network is the most common pattern in melanoma detection [62]. It contains "pigmented network lines and hypo-pigmented holes".

Clinical Significance

In melanoma, pigment network is lightly pigmented. Thin and Light brown network lines fade progressively at the peripheral part of the lesion. Holes are in regular and narrow in structure. In melanoma, pigment network usually ends abruptly at the peripheral part of the lesion [12].

Types

A pigment network has two types Typical and Atypical . Typical pigment network is "light to dark brown network in color with smaller, equally spaced network holes". Network lines are usually thin, distributed more or less regularly all over the lesion and normally thinning out at the periphery. Atypical pigment network is: "a black, brown or gray network with irregular holes and thick lines".

Automated detection of pigment network

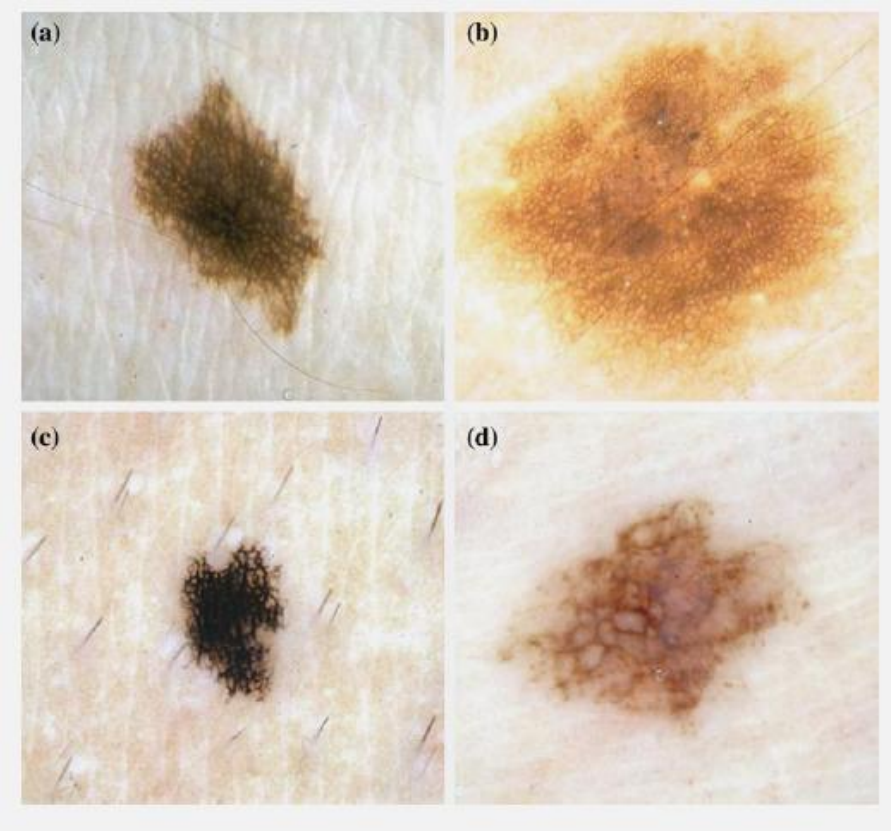

Figure 4.2: Lesion example containing pigment network. (a and b) are typical, whereas (c and d) are atypical pigment network [12]

Anantha et al. [63] describes the detection of melanoma with the help of local feature i.e. pigment network. The proposed automated system use two statistical texture identification methods; one is Neighboring Grey-Level Dependence Matrix (NGLDM), and the other is the Lattice Aperture Waveform Set (LAWS).

The automated system presented by Anantha et al. diagnose melanoma on the basis of pigment network. So Betta et al. [64] further presented a system for classification of pigment network as atypical/typical. The proposed automated system was based on the already discussed methodology in [65] of pigment network detection. The authors use two different analysis: structural and spectral analysis. The first analysis of the images search for some basic structures such as edges or lines. In the second analysis the author used fourier transform of the image to calculate various texture patterns.

Di Leo et al. [66] enhance the research presented by Betta et al. [64] to find atypical pigmented network. Atypical pigmented network is a feature closely related to melanoma [64]. Detection of pigment network followed by classification using color and geometric features.

Shrestha et al. [67] presented classification of pigment network using texture measurements that could discriminate melanoma commonly referred as Atypical Pigment Network (APN) from benign lesions. The proposed methodology use Gray Level Co-Occurrence matrix (GLCM) from the luminance part of an RGB color image.

Sadeghi et al. [68] proposed classification of dermoscopic feature pigment network. The proposed system consists of pigment network detection using Laplacian of Gaussian (LOG) filter to detect sharp changes of intensity. Lines and holes were detected and various structural and geometric features were calculated.

Wighton et al. [69] presented a technique for pigment network detection using supervised learning. The authors first converted the image to CIE L*a*b* and every color channel was filter with various Gaussian and Laplacian of Gaussian (LOG) filters at multiple scales, so that feature matrix can be obtained for each pixel. Finally, the posterior probabilities were modeled as multivariate Gaussian distributions.

4.3.1.2 Streaks Detection and Analysis

Streaks are in "brownish color with black linear structures of variable thickness" that are present in benign and malignant lesions. Streaks are usually placed at the peripheral part of the lesion . For melanoma, streaks may be in irregular or regular [70].

Clinical Significance

Streaks is a terminology used interchangeably with radial streaming. Streaks are local dermoscopy features, however they can correlate with a global pattern of skin lesions called a starburst pattern if symmetrically arranged over the entire lesion.

Automated detection of Streaks

Betta et al. [65] calculated streaks as in "finger like irregularities having brown color" at the lesion boundary. For the detection of finger like irregularities, gray scale conversion is required and then binary images were computed. To find contours blob finding algorithm is used. Then, an irregularity ratio was evaluated. Finally, the presence of streaks was calculated if image contains irregular boundaries and brown pigmentation.

Sadeghi et al. [71] proposed a four step methodology for streaks detection:- preprocessing, contour detection, feature extraction and classification as presence or absense of streaks. In the preprocessing step, images were segmented. Luminance

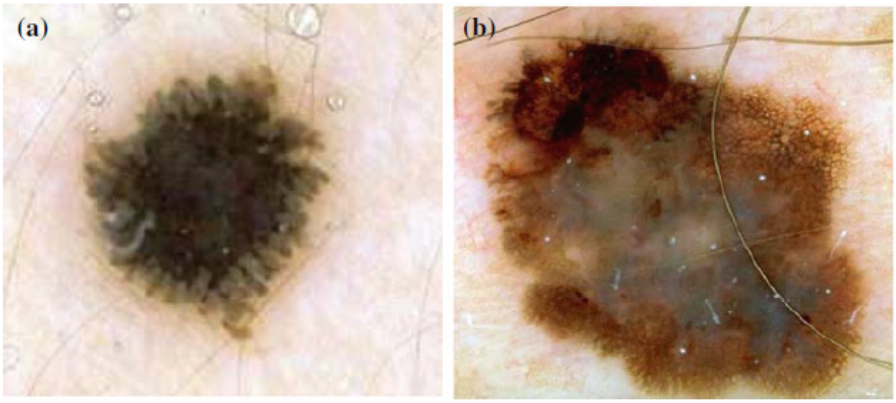

Figure 4.3: Lesion example containing (a) regular streaks, (b) irregular streaks [12]

component (L*) was used for the analysis. Streaks can be considered as linear structures with a Gaussian cross section profile near the boundary.

Sadeghi et al. [72] extend their already published methodology [71]. In this version, they presented an technique that classify a lesion into three class problem as : absence, regular and irregular streaks. The technique focus on valid streak lines from the candidate streak lines to reduce false positive streaks. The proposed novel geometric features are helpful to identify the presence or absence of streak lines.

In this thesis, for melanoma detection we focus on two main local patterns i.e. streaks and pigment network.

4.4 Discussion

Pattern analysis is widely used for effective diagnosis of melanoma. In fact, it is considered more effective as compared to the other scoring algorithms (as discussed above) for diagnostic efficiency by the physicians. Patterna are local or global to detect melanoma or benign lesion.

To design a CAD system using pattern recognition approach, image segmentation is the vital step as it is helpful in analyzing specific features from melanoma. In this research along with manual annotated masks, we use automatic segmented masks. From the clinical findings it is evident that differential structures for example pigment network and streaks are more intuitive for melanoma detection. Skin lesion contains peripheral part which has high texture containing pigment network and streaks. Since both of these features are local features, melanoma can be distinguished from benign lesion on the basis of local features. In light of these facts, we focus on local patterns located at the peripheral part of the lesion and extract some visual features for melanoma detection.

5 Materials

Collection of the data and its manual annotation is the most important steps for the designing of CAD systems. The key stages of the pattern recognition system depends on the accurate data. So it is essential to gather the data along with its manual segmented masks from the expert physicians. These manual segmented masks identify clinically appropriate areas and their final classification about diagnosis in the images. In this thesis, the images were acquired from one of the Hospitals of Portugal using dermoscopy as an imaging modality.

5.1 PH^2 DATASET

The PH^2 Dataset was built up through a joint research collaboration between the Universidade do Porto, T´ecnico Lisboa, and the Dermatology service of Hospital Pedro Hispano in Matosinhos, Portugal. The dermoscopic images were obtained under the same conditions through Tuebinger Mole Analyzer system using a magnification of 20. They are 8 bit RGB images with a resolution of 768x560 pixels. Every image contains following parameters:

- Manual annotation of the skin lesion
- Clinical and histological diagnosis
- Dermoscopic features

5.1.1 Analysing Manual Annotation

Dermatologists performed the manual segmentation and annotation of the images using a customized annotation tool for dermoscopic images, called DerMAT [73]. In this database, the manual annotation of each image is available as a binary mask, in which pixels having intensity value 1 represents the segmented lesion, while pixels having intensity value 0 represents to the background. This binary mask has the same size of the original image and, hence, it can be used to extract the boundary points of the lesion.

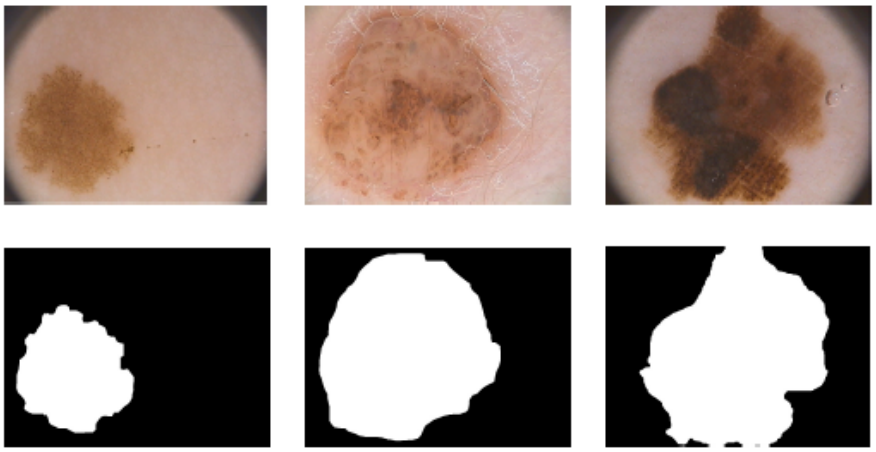

Figure 5.1: Manual segmentation of three different melanocytic lesions

5.1.2 Clinical diagnosis

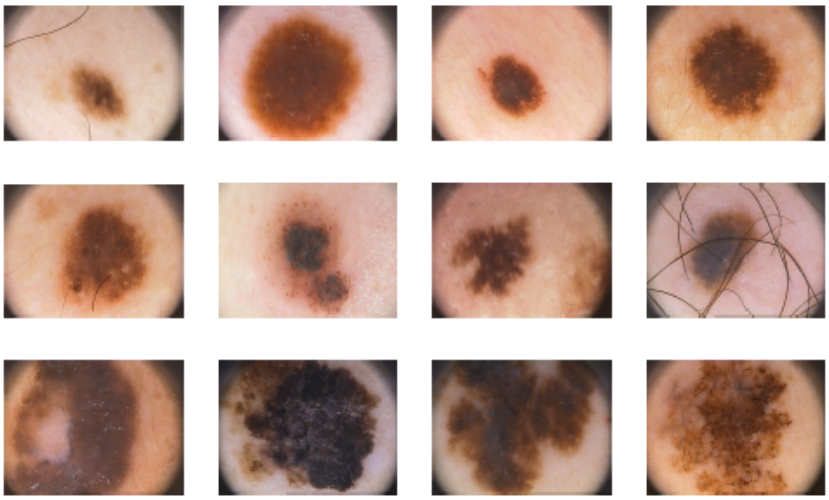

Figure 5.2: An illustrative collection of images from PH2 database, including common nevus (1st row), atypical nevus (2nd row) and melanoma (3rd row).

This image dataset contains total 200 dermoscopic images, containing 80 common nevus [1], 80 atypical nevus, and 40 melanoma. All images are either from the skin type II or III, according to the Fitzpatrick skin type classification scale [74]. Therefore, the skin colors represented in the PH^2 database may vary from white to cream white. The melanocytic lesions can be divided in two main classes depending upon their nature: benign lesions (which include common and atypical nevus) and malignant lesions (or melanoma). Hence, each image of the dataset is classified into

[1] Nevus are basically benign lesion.

common nevus, atypical nevus, or melanoma.

Class I was defined as common nevi.

Class II was defined as atypical nevi .

Class III was defined as melanoma.

Clinical findings indicate that Class I images are benign lesions, Class II images are also considered benign lesions. Class III shows melanoma lesions.

PH^2 Dataset		
Class I	Class II	Class III
Common Nevus	Atypical Nevus	Melanoma
80	80	40

Table 5.1: Configuration of PH^2 dataset containing three classes

For the detection of melanoma we consider two classes (Melanoma and Nevi). For detection we have two class labels (0 is for nevi and 1 is for melanoma). Table 5.2 shows composition of PH^2 dataset containing two classes used for detection of melanoma.

PH^2 Dataset	
Class I	Class II
Nevi	Melanoma
160	40

Table 5.2: Configuration of PH^2 dataset containing two classes

5.1.3 Dermoscopic features

The set of features that are present in the PH^2 dataset corresponds to those features that the dematologists of Hospital Pedro Hispano consider more relevant for melanoma detection. These features have to be evaluated in the most widely used diagnosis procedures, such as the ABCD rule, 7 point checklist and the Menzies method. Various features and their evaluation process are described below.

1. Asymmetry

Asymmetry is one of the important feature for diagnosing melanoma. Asymmetry has the maximum weight in ABCD rule. Asymmetry is evaluated according to the lesion contour, colors, and shape. There are three labels for the asymmetry evaluation: 0 indicated fully symmetric lesion; 1 is for asymmetric lesion with respect to one axis; and 2 for asymmetric lesion with respect to two axes.

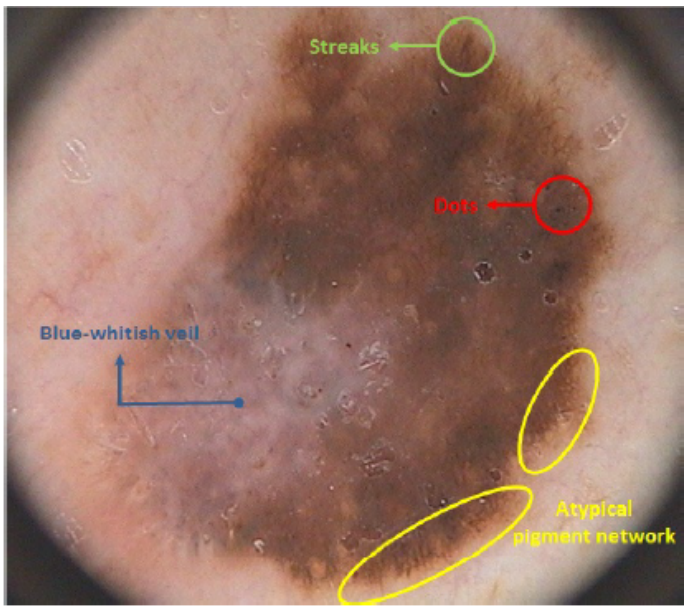

Figure 5.3: Dermoscopic features identification (Reprinted from [13])

2. Pigment Network

Pigment network is a grid like network having pigmented lines (brown or black) and hypopigmented holes [2]. This structure has a crucial role in the distinction between melanoma and non melanoma lesions. The pigment network structure was visually examined by the dermatologist, and classified as typical or atypical.

3. Color

Six different colors are considered for melanoma detection. These colors contains white, red, light brown, dark brown, blue gray, and black [2].

4. Dots & Globules

As demonstrated in Figure 5.3, this feature is in spherical or oval, varying in size and colors. The existence of this feature is also particularly useful for melanoma detection [2]. This features were visually examined by dermatologists, and categorically classified as presence or absence . Presense of dots & globules in a lesion are additional classified as regular or irregular depending upon their structural distribution in the lesion.

5. Blue-whitish veil

This feature can be considered as a confluent, opaque, having irregular blue pigmentation. Its existence indicate that the lesion is melanoma [2]. This feature is labeled as presence or absence, in the image. Blue whitish veil in melanoma is illustrated in Figure 5.3.

6. Regression areas

This feature is in white, scar-like depigmentation often combined with pepper like regions [2]. This feature is further classified in two categories i.e. presence and absence in the skin lesion.

7. Streaks

Streaks are like finger projections of the pigment network from the peripheral part of the lesion. Presence of streaks is a sign of melanoma [2]. Streaks are further classified as presence or absence. Figure 5.3 illustrates the presence of a streak in a skin lesion.

5.2 Conclusion

This section presents a dataset of dermoscopic images, PH^2, acquired at Pedro Hispano Hospital. This database includes medical annotation of all the images namely medical segmentation of the lesion, clinical diagnosis and dermoscopic criteria (asymmetry, colors and the presence of typical and atypical differential structures).

6 Proposed Methodology

A complete pattern recognition system as illustriated in Figure 3.1 comprises three main parts; image segmentation, feature extraction and classification. For automatic image segmentation we have used level set creasness functions and manual annoatation is provided by the physicians. In this research, we have proposed two main different techniques for feature extraction. The first feature extraction method deals with the visual descriptors i.e. Shape and Texture descriptors; the second module extract features form dermoscopic images using joint histogram of multiresolution LBP and local contrast. In this research we also proposed a methodology for lesion processing that is instead of processing whole lesion, we focus only on the peripheral part in view of its clinical significance. At the end, classification results are compiled using different classifiers and a comparison is drawn.

6.1 Automated System Based on Visual Descriptors

CAD systems adopt various pattern recognition approaches i.e. by the extraction of various features (color, texture and shape) from each dermoscopy image followed by the training of a classifier. Clinical findings have shown that an early detection of melanoma can be done by an inspection of visual characteristics of some specific regions (lesions) of the skin. This research quantifies the effects of using visual characteristics from the lesions on the identification of melanoma. Effectively, a pattern recognition system is designed that includes three vital stages that conform to the analysis of skin lesions by the clinicians: segmentation, feature extraction and classification.

The proposed automated system involves three different steps:- 1) segmentation using level sets method, 2) feature extraction from two attributes in the ABCD rule i.e. asymmetry index and differential structures features 3) classification using Support Vector Machine (SVM) and K-Nearest Neighbor (KNN).

6.1.1 Segmentation

Segmentation is an vital step in the automatic classification of skin lesions because accuracy of this block effects the successive steps. But, segmentation is difficult

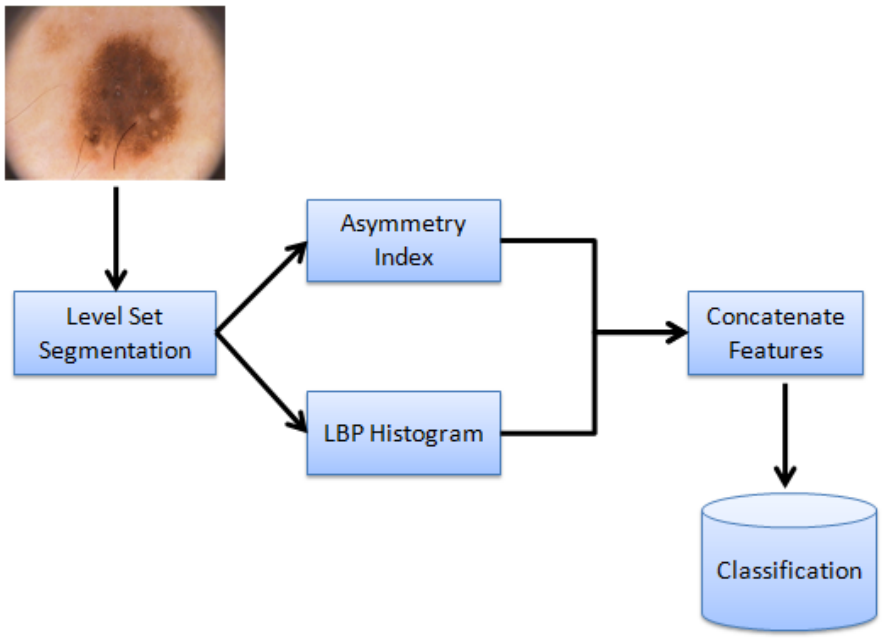

Figure 6.1: Proposed automated system for melanoma detection based on visual descriptors

step because of the extensive range of lesion shapes, colors, sizes and various skin tones. Additionally, the smooth transition of between the skin and the lesion can make correct segmentation of the lesion a challenging task to handle. To address these challenges, various segmentation techniques have been presented which can be grouped as threshold, edge and region based method. When the contrast between lesion and the skin is high, thresholding based methods illustrate very good results. Edge based methods fail when the skin and the lesion are separated by smooth boundaries. In this paper, we implement level sets curve evolution for the segmentation of dermoscopy images given its significance in the literature [75].

6.1.1.1 Automatic Segmentation using level sets curve evolution function

Active contours are dynamic fronts which move within the images towards object boundaries. In the level set formulation, the dynamic fronts $\mathbb{C}$ are usually characterized using zero level set $C(t)$.

$C(t) = \{(x, y)|\phi(t, x, y) = 0\}$ of an underlying function $\phi(t, x, y)$.The generic form

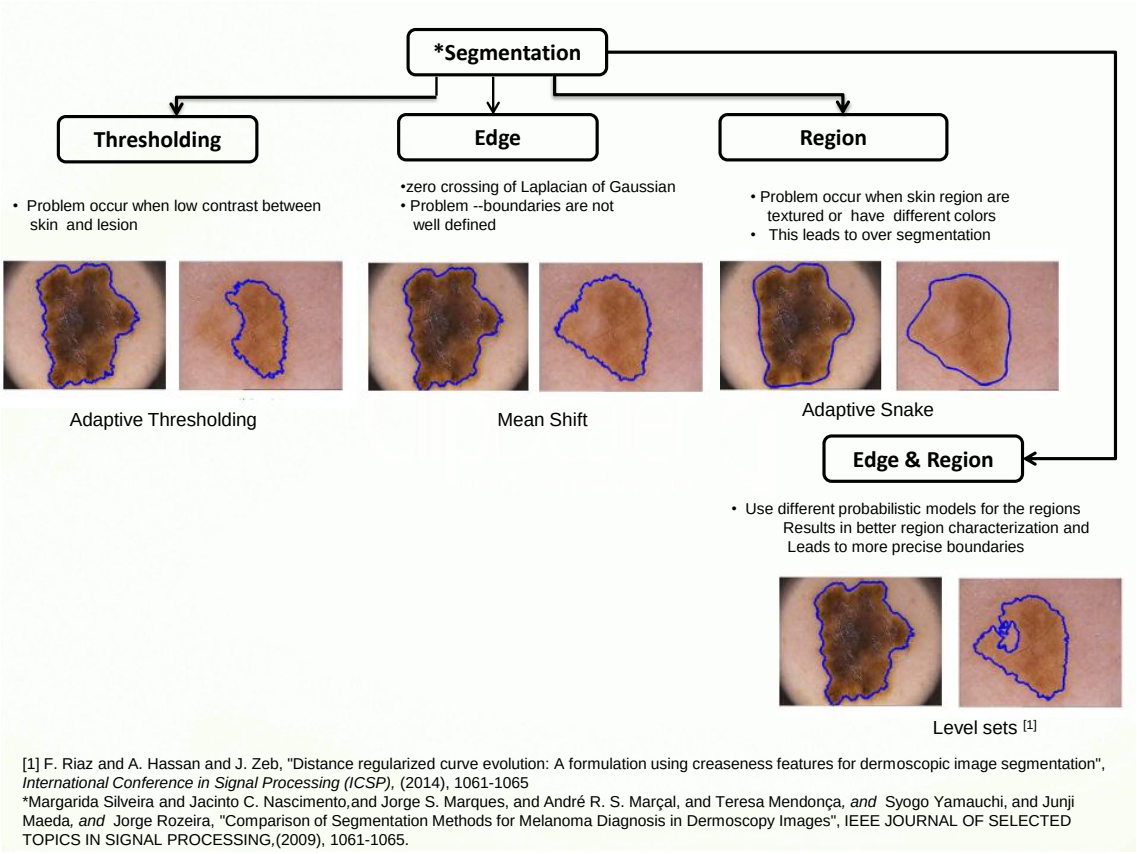

Figure 6.2: Segmentation of Dermoscopy Images ([14])

of a level set function ϕ is shown in Equation 6.1 [76]:

$$\frac{\partial \phi}{\partial t} = F|\nabla \phi| \tag{6.1}$$

The function F is known as the speed function.

Traditional level sets has a limitation that is the active contour can have sharp or/and flat shapes etc. A solution to this problem is, before evolution initialize as signed distance function and then during evolution periodically reinitialize as a signed distance. The process of re-initialization is quite difficult. Another technique, termed as variational level sets for curve evolution was used that actually takes the simple finite difference scheme and it avoids the re-initialization phase. The curve evolution can be achieved by:

$$\varepsilon(\phi) = \mu P(\phi) + \varepsilon_e(\phi) \tag{6.2}$$

$P(\phi)$ is defined as internal energy which is a signed distance function [14] and $\varepsilon_e(\phi)$ term is known as external energy function. In this research we are interested in devising an external energy function which depends on the image data, while keeping the internal energy function same as that in traditional level set formulation. For the calculation of external energy terms in the level sets framework, creasness features is obtained using the multilocal level set extrinsic curvature with improvement by structure tensor features (MLSEC-ST) originally proposed by Lopez et al. [77]. This operator has the ability to enhance ridges and valleys in an image. We can find the magnitude of creasness from the images using Equation 6.3:

$$\mathcal{F}(m,n) = 1 - e^{-(\lambda_1(m,n)-\lambda_2(m,n))^2} \tag{6.3}$$

where,$\lambda_1 > \lambda_2 \geqslant 0$ are the eigen values of structure tensor function $S(m,n)$, while n represents rows and m represents columns of gray scale image. Empirical findings indicate that creaseness features improve lesion boundaries while concealing the local image texture.

6.1.2 Feature Extraction

Segmentation is followed by feature extraction. This is the key point of the classification and has to be adequate to facilitate a good system detection rate. Feature extraction is broadly divided into three main categories (color, texture and shape) as discussed in section 3.3. Concerning the image features for dermoscopic images, physicians use several visual cues, such as differential structures, number of colors, symmetry, and border transitions. Inspired from the ABCD rule, we have extracted two different types of features: texture features (for differential structure detection) and shape features (for calculating lesion asymmetry) via computer vision techniques. We intend to mimic the human classification of dermoscopy images using shape and differential structures from the images.

6.1.2.1 Asymmetry Index

Typically for objects that are characterized by a random contour, the orientation of the objects can manipulate with the metric of asymmetry which is highly undesirable in our scenario. This is because, irrespective of the viewing angle of the observer (physician), the measure of asymmetry has to be the same.

Clinical Significance

Analysis of asymmetry is a very important fact, in the diagnosis of melanoma, since the local appearance of certain structures or colors is most indicative for a melanoma. According to dermatologists, melanoma grow in an anarchic way (asymmetric tumor) while benign tumors are symmetrical.

Asymmetry Index Calculation

Accordingly, we have devised an asymmetry metric that measures the principal axis of the shape by performing the principal component analysis (PCA) of the segmented lesion. The lesion data is projected onto the transformed subspace which reorientates the lesion according to its principal axis. This is transformed by the measurement of asymmetry of the lesion. Asymmetry is evaluated by translating the lesion to the center of the image, followed by finding the difference between one of the sides of the axis and the mirrored image of the other side using the XOR operator as shown in Figure 6.3. Asymmetry index is measured using Equation 6.4

$$AsymmetryIndex = \frac{A_D}{A} \times 100 \tag{6.4}$$

Here A_D defines the difference between two sides of the lesion while A represents the complete area of the lesion for the binary case.

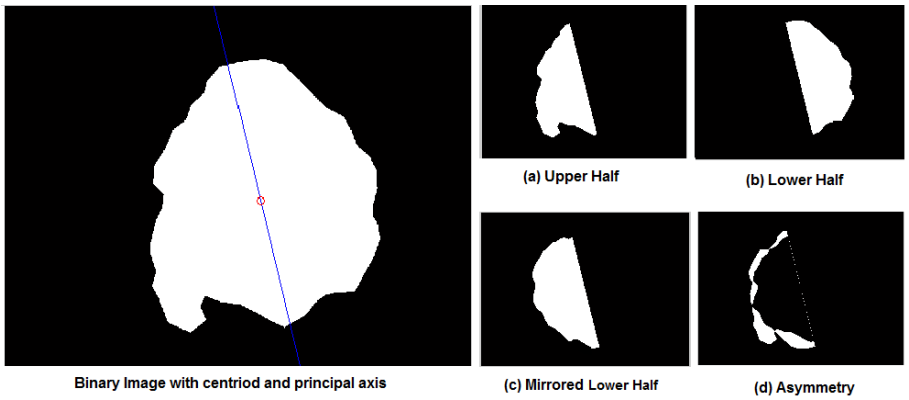

Figure 6.3: Calculating asymmetry index: (a) upper half, (b) lower half, (c) mirrored of upper half (d) asymmetry.

6.1.2.2 Local Binary Pattern (LBP)

Pattern analysis and the identification of "differential structures" according to the ABCD rule represents the basis for dermoscopic diagnosis. Differential structures

deal with the presence of dermoscopic structures such as pigment network, regression structures, dots, globules and so on as described in subsection 5.1.3.

Motivation of differential structures

The differential structures in dermoscopy images can be analyzed using the local binary pattern (LBP) texture descriptor. It is important to note that the analysis of differential structures is the most common among ABCD clinical rule providing adequate visual features for dermoscopic image analysis.

LBP Methodology

Originally presented by Ojala et al. [78] LBP is gray scale and invariant texture descriptor that creates LBP codes at every pixel in the image by thresholding all the neighborhood pixel with the value of central pixel and then concatenate the result in the form of a pattern as shown in Figure 6.4. For basic LBP the thresholding function can be obtained using Equation 6.5:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^P, s(x) = \{ \begin{matrix} 1 & x \leq 0 \\ 0 & x > 0 \end{matrix} \tag{6.5}$$

where g_c and g_p shows gray level values of center pixel and also its neighbor respectively, while p is the neighbor's index. P represents the total number of the neighbors in a circular set surrounding a pixel at a radius of R from g_c.

After generating LBP codes of an image, LBP histogram is created to count the occurrences of the LBP codes in an image. The generated LBP histogram is further used as a feature.

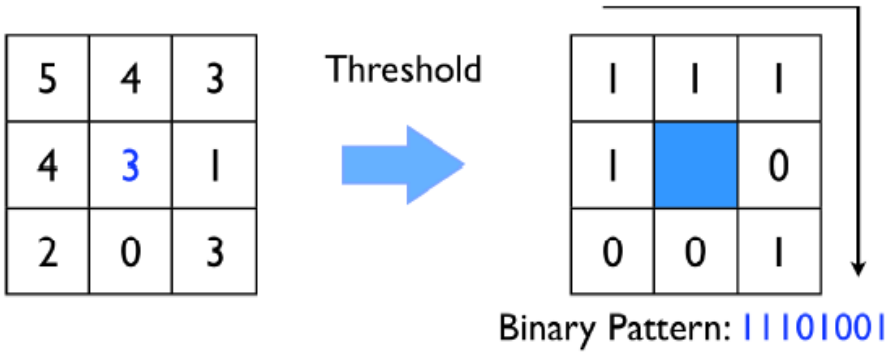

Figure 6.4: Local Binary Pattern

The LBP pattern used for detection in this research is uniform patterns that is basically an extension to the standard LBP. Motivation behind using this uniform

patterns is that some binary patterns occur normally in texture images than others. A LBP pattern is uniform if its uniformity measure is at most 2. To quantify the uniformity of the LBP $U(LBP_{P,R})$ is a parameter defined in Equation 6.6:

$$U(LBP_{P,R}) = \sum_{p=1}^{P-1} |sign(g_p - g_c) - s(g_{p-1} - g_c)| \tag{6.6}$$

The motivation for using uniform LBPs is their ability to detect the significant intrinsic characteristics of textures like spots, line edges, edges and corners as shown in Figure 6.5.

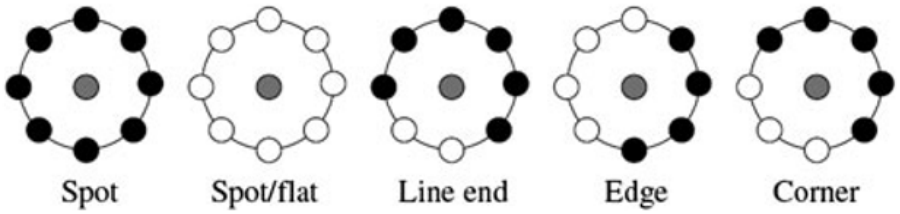

Figure 6.5: Local Binary Patterns Structures [7]

6.1.3 Experiments

Our objective is to analyze the performance of our automated system for detection of melanoma. We will empirically demonstrate the experimental setup used to evaluate significance of visual descriptors for melanoma detection. For the segmentation of dermoscopic images we have used active contours with creasness function followed by feature extraction using two different features: the rotational invariant asymmetry and standard LBP histograms. This selection is based on the clinical significance of asymmetry and texture in the detection of melanoma according to the ABCD rule. The performance of the proposed automated system is tested on manual and automatic segmented images. We performed an experimental setup that is Concatenation of "A" (Asymmetry) and "D" (Differential structures) features from ABCD rule. The objective of this experiment is to asses the detection results from a combination of features and validate the clinical significance of the ABCD rule. Effectively, we concatenate the rotational invariant asymmetry with LBP histograms for feature extraction, followed by the detection of melanoma.

6.2 Joint Histogram of LBP with Local Contrast

Next dimension in this research is the variant of LBP used for feature extraction. LBP is a gray-scale invariant texture operator, that neglects the strength of the pattern (usually known as contrast of the LBP patterns). Pattern information is independent of the gray scale where as contrast relies completely on the gray scale or its variation. Patterns and contrast are the two measures of the texture that can be combined in a very useful way.

In this section we propose a novel texture descriptor. Figure 6.6 shows block diagram of the proposed system. Segmentation using level set is described in subsubsection 6.1.1.1. Now for the feature extraction we will study the combination of LBP pattern along with the local contrast.

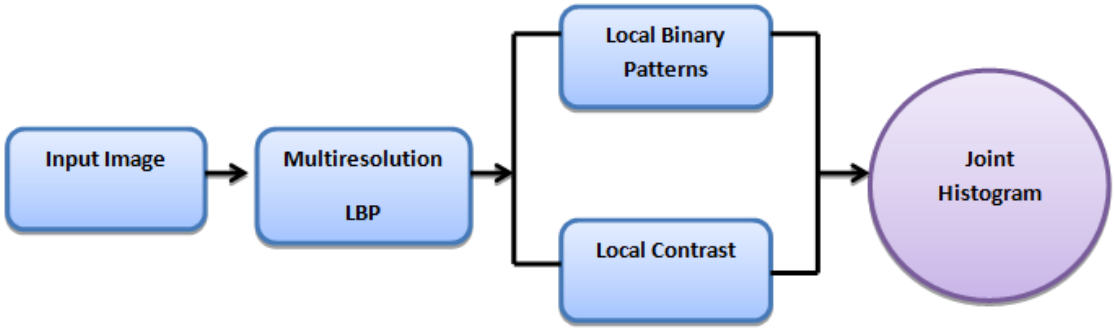

Figure 6.6: Proposed system for melanoma detection using joint histogram of LBP and local contrast

6.2.1 Feature Extraction using Joint Histogram of LBP pattern with local contrast

An enhancement in the standard LBP texture can be obtained if the contrast of the LBP patterns is also taken into account. This is mainly because; image texture is described by the local contrast and underlying pattern at a particular spatial location. More specifically, the observers typically examine the images with different levels of visual perception (detail) depending on the richness of texture content (contrast). Thus, an enhancement in the description of texture content can be obtained if the strength of LBPs is also incorporated in the texture descriptor. To take the strength of LBPs into account, we have used various scales of LBPs for feature extraction. The multi scale LBPs can be obtained by varying the values of R to obtain the patterns at various radii from the center pixel. Let $r_1, r_2, ...r_k$ be various radius at which the LBPs are calculated. At each pixel position, the absolute difference between the center pixel and its neighbors can be given as:

$$C^k(g_c) = \sum_{p=o}^{P-1} |(g_p^k - g_c)| \tag{6.7}$$

where $C^k(g_c)$ represents the contrast of an underlying LBP patterns at the resolution r_k. The LBP having the highest strength consists of the most relevant pattern and can be obtained as follows:

$$ULBP_{gc} = arg\, max_{ULBP(P,r_k)}\{C^k(gc)\} \tag{6.8}$$

Therefore, the LBP at g_c is represented by the pattern, which exhibits the maximum strength i.e., $C^k(g_c)$ when analyzed at various resolutions at the pixel g_c . The LBP codes of an image along with the contrast of the patterns are used to create a joint histogram. Joint histogram of LBP patterns and contrast is created visiting each (x, y) location once. At each x, y location, increment histogram bin (b1, b2), where b1 is the 1D histogram bin for the value at x; y in LBP pattern and b2 is the 1D histogram bin for the value at (x; y) in contrast. Result of joint histogram is (b1xb2) two dimensional array. This generated joint histogram is further used as a feature.

6.2.2 Experiments

We perform an experimental setup that attributes the superior performance of joint histogram of LBP patterns with the contrast of the pattern in order to capture the micro structures in the images. These micro-structures are effectively representative of the differential structures in the images. Local contrast in multi scales of LBP approximates the human observers (clinicians) who try to visualize the image texture at various levels of attention. In our experiment, for the extraction of multi resolution LBPs we have used the radii of 1.0, 1.5 and 2.0 and at each resolution, $N = 8$ neighbors were considered. The performance of the proposed automated system is tested on manual and automatic segmented images. The experimental setup validates that joint histogram of LBP pattern and local contrast has a significant role in the detection of melanoma.

6.3 Melanoma Detection based on Peripheral region

Automated system for melanoma detection systems can be helpful in improving our ability to screen the dermoscopy by providing a second opinion to the physician

using pattern recognition approaches. One of the most fundamental challenges in computer vision is the segmentation or extraction of region of interest (ROI). Need for extracting ROI is motivated by the procedure of manual diagnosis where the physician is interested in examining a specific clinically relevant region in a lesion. This region is expected to contain maximum information that can be helpful for diagnosis of melanoma. In this section we have proposed a technique for region extraction which contains clinically more relevant information in the form of texture for melanoma detection i.e. peripheral area, followed by feature extraction and classification. In contrast to the above mentioned proposed automated system section 6.1 the major difference occur in this experiment is, instead of using whole lesion only a subset of lesion is extracted relevant to contain more pattern in terms of pigment network, streak etc. Therefore, we analyse the significance of lesion peripheral part.

6.3.1 Skin Lesion Periphery

In almost all the clinical dermoscopy methods, dermatologists look for the presence of specific visual features to diagnose melanoma. Then, these features are analyzed for irregularities and malignancy. The most important diagnostic features of melanoma are; pigment network, streaks, hypopigmentation, dots, globules discussed in section 4.3. The presence of these dermoscopic features in image varies in size, color, texture, and location on the lesion image. While examination physicians focus on some specific parts of the lesion for example if border of the lesion is irregular and asymmetric then the lesion is usually malignant tumor. Skin lesion contains two main regions 1) center and 2) peripheral part.

Clinical findings have concluded that various diagnostic feature of lesions contains significant information in the form of shape, color and texture at the peripheral region of the lesion. In case of pigment network, light brown network lines are thin and fade gradually at the periphery. In melanoma, the pigment network usually ends abruptly at the periphery and has irregular holes, thickened and darkened network lines, and tree like branching at the periphery [12]. So by analyzing the peripheral part of the lesion, detection of pigment network can be done. Similarly streaks are the local dermoscopic feature that are typically placed at the periphery of a lesion in the form of brownish-black linear structures [70]. Streaks are usually placed at the peripheral part of the lesion and are not always connected to the lines of the pigment network. In case of malignant melanoma streaks can be irregular, and they are not evenly distributed. Therefore pigment network and streaks have clinical association with the peripheral region of the lesion. The presence of pigment network and streaks indicate the presence of high texture in the images, which requires a fine and detailed analysis of the images. Multi-resolution analysis is a significant aspect for the fine and detailed analysis of the peripheral region. This is because the peripheral part has different spatial characteristics which means these should be analyzed at multi scales.

Figure 6.7 shows the two regions extracted from the skin lesion image. Peripheral region contains rich texture, while center region contains low texture information. Therefore, if we calculate peripheral part of the lesion, then it would help us to detect melanoma on the basis of some local features.

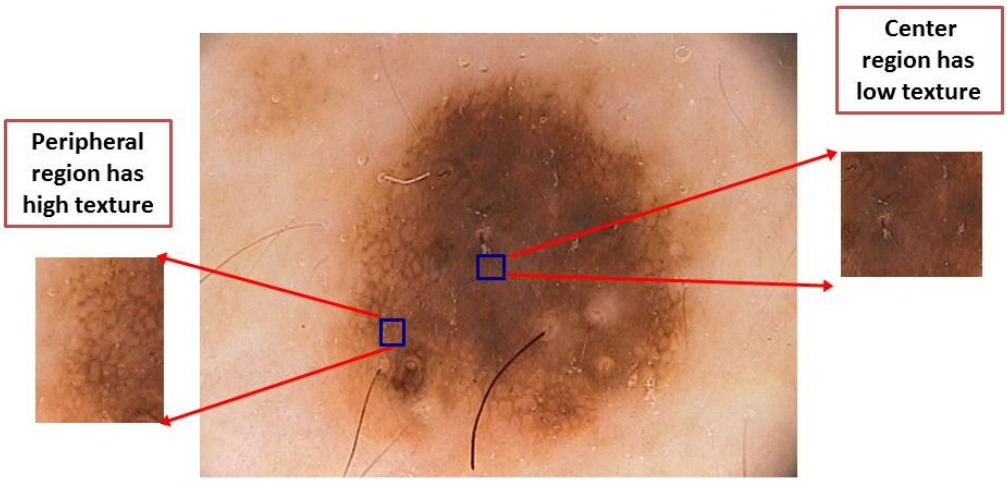

Figure 6.7: Peripheral and Center Regions of skin lesion

From literature we didn't find any method that detect melanoma only on the basis of peripheral part of the lesion. In this research we focus our study on the analysis of peripheral part of the lesion. The first step in the designing of CAD system is the lesion border detection. After determining the border of the lesion, we divide skin lesion into two parts i.e. peripheral and central. Peripheral part of the lesion is evaluated by translating the lesion to the center of the image followed by extracting donut that is similar in shape to the original lesion. This donut is some area of the whole lesion obtained by moving inward some pixels from the border coordinates. The peripheral part of the lesion is obtained by subtracting the donut from the whole lesion. Finally, the central part is obtained by removing the peripheral part from the whole lesion.

6.3.2 Feature Extraction

Next step after region extraction from the lesion is feature extraction. For feature extraction we use joint histogram of multiresolution LBP pattern and local contrast. This is described detailed in subsection 6.2.1.

6.3.3 Experiments

In our above two experiments, we analyse whole lesion and proposed an automated systems for melanoma detection. In this experiment, instead of focusing on the whole skin lesion, examination of the peripheral part is more sufficient for melanoma detection. We have used Matlab routine (resample Matlab function) to make a donut from the lesion. We vary the size of donut from 5 pixels and onwards to observe the change occur in peripheral region. Considering small pixels (for example 5 pixels) from the border coordinates gives very thin layer of the peripheral part which means we have very small amount of texture information left in peripheral region. Similarly if we go maximum pixels inwards from the border coordinates then we get the peripheral part which contains almost whole lesion. So both of these extreme cases (considering small pixels or maximum pixels from the border) are not useful for us for melanoma detection. In our research we focus on that region which contains high texture in the lesion. Hence we consider the intermediate state between selection of pixels going inwards form the border. Feature extraction comprises joint histogram of LBP patterns and local contrast of the patterns.

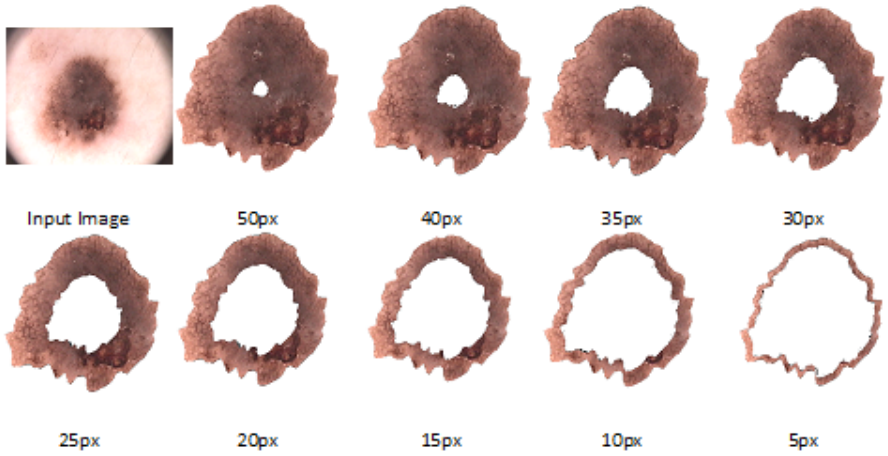

Figure 6.8: Varying the size of peripheral region in term of pixels (px) from the border coordinates

6.4 Classification Results

Our objective of the proposed automated system is to diagnose melanoma. Next step after feature extraction form the proposed pattern recognition systems is classification. For classification, we have used support vector machine (SVM) [79] and K-Nearest Neighbor (KNN) [50]. We have used the SVM linear kernel and value

of K=1 in our implementation. For this Weka data mining tool was used in our experimental results. All the experimental results were obtained using 10-fold cross validation.

Table 5.2 shows that we have an imbalance data, so we used average overall accuracy (Acc), sensitivity (SE) and specificity (SP) as appropriate performance measures to assess the performance of the classifier, as discussed in subsection 3.5.2.

6.4.1 Assessment of differential structures and shape features for melanoma detection

The objective of this classification involves the combined usage of texture and shape features for the detection of melanoma. Experiments show that when both LBP features and asymmetry index (AI) are used for feature extraction, good detection rates are obtained. The observation is consistent for both manual annotations and automatic segmentation of the images.

Method	Classifier	Manual Segmentation			Automatic Segmentation		
		Acc %	SE %	SP %	Acc %	SE %	SP %
LBP + AI	SVM	89.9	89.5	75.5	86.5	89.8	69.9
	KNN	83.5	88.4	60	86	90.7	65.7

Table 6.1: SVM and KNN classification results for the detection of melanoma when asymmetry index (AI) and Local Binary Pattern (LBP) are concatenated to form a descriptor for manual and automatic segmented images.

6.4.2 Assessment of Joint Histogram of Multiresolution LBP pattern with local contrast

We attribute the superior performance of joint histogram of LBP patterns with the contrast of the pattern (JHLBP_C) in order to capture the micro-structures in the images which are effectively representative of the differential structures in the images. Local contrast in multi scales of LBP approximates the human observers (clinicians) who try to visualize the image texture at various levels of attention, giving superior classification results in the identification of melanoma. For performing this comparison, we select state of the art techniques i.e. Standard LBP, Homogeneous Texture (HT) and Autocorrelation Homogenous Texture (AHT) . The test bed used for all these methods is the same as that used for our proposed method. Our experiments demonstrate that the HT shows worst performance amongst state-of-the-art methods. Standard LBP shows relatively low classification rates. We suspect that this is because of the lack of multiresolution analysis in standard LBPs. Our other

novel descriptor i.e. joint histogram of multiresolution LBP and local contrast shows better performance as compared to the other descriptors considered in this research.

Classifiers	Methods	Manual Segmentation			Automatic Segmentation		
		Acc %	SE %	SP %	Acc %	SE %	SP %
SVM	HT	81.01	64.1	86.2	79	79.1	63.0
	AHT	82.9	79.0	78.4	80.9	81.5	64.5
	LBP	86	88.3	71.4	82.5	86.8	64.7
	JHLBP_C	**90.5**	**91.2**	**86.2**	**89.5**	**91.6**	**78.7**
KNN	HT	79.5	69	71	66.5	78.1	40.9
	AHT	81.6	70.8	60.2	69.4	79.9	55.8
	LBP	82.5	85.8	62.5	79.5	87.4	63.7
	JHLBP_C	**86**	**88.3**	**71.4**	**82.5**	**84.15**	**64.7**

Table 6.2: Comparison of SVM and KNN classification results obtained using proposed features i.e. Joint Histogram of Multi resolution LBP and local contrast (JHLBP_C) and state-of-the-art feature extraction methods.

6.4.3 Assessment of Peripheral region

The goal of this experiment is to analyze the significance of peripheral part of the lesion. In our experiment, we used 20 pixels from the border coordinates to get peripheral region. The selection of the pixels from the border for peripheral region was done empirically by evaluating the performance of peripheral layer of over a range of possible pixel values and then selecting the appropriate ones for our dataset. Our experiments demonstrate that the joint histogram of multiresolution LBP and contrast on peripheral region (JHLBP_C_P) shows better performance.

JHLBP_C_P	Classifiers	Manual Segmentation			Automatic Segmentation		
		Acc %	SE %	SP %	Acc %	SE %	SP %
	SVM	**92.5**	**92.3**	**93.1**	**91.5**	**91.3**	**92.5**
	KNN	88	90.4	75	87.5	89.9	74.19

Table 6.3: SVM and KNN Classification results of melanoma detection using Joint Histogram of Multiresolution LBP pattern and Local Contrast on Peripheral Region (JHLBP_C_P).

Classifier	Segmentation	5px			10px			15px			20px			25px			30px		
		Acc%	SE%	SP%	Acc%	SE%	SP%	Acc%	SE%	SP%	Acc%	SE%	SP%	Acc%	SE%	SP%	Acc%	SE%	SP%
SVM	Manual	87.7	84	79	90.3	87	64	90.9	89	54.9	**92.5**	**92.3**	**93.1**	91.2	91	64	90.5	89	74
	Automatic	85.5	79	68.5	85.1	85	90.1	90.5	90	90.7	**91.5**	**91.3**	**92.5**	89	90.4	86	88.7	82	68
KNN	Manual	76	68	74.1	81.5	85	69.3	86.4	88	89.5	**88**	**90.4**	**75**	85.4	86.5	78	83.1	80	65
	Automatic	80.9	81	75.6	83.1	84	81.9	85.3	85	76.5	**87.5**	**89.9**	**74.19**	87.7	79.5	74	85.9	80	70

Table 6.4: SVM and KNN Classification results of melanoma detection using Joint Histogram of Multiresolution LBP pattern and Local Contrast on Peripheral Region (JHLBP_C_P).

7 Conclusions and Future Work

In this chapter, we will conclude the thesis objectives that are used for designing an automated system for melanoma detection. The objectives of the thesis lead us to present novel methods. In this section, we will discuss the objective and the results that were achieved through our methodology. Finally, we will conclude the thesis with some future area to be investigated.

7.1 Discussion on objectives

Computer Vision and pattern recognition techniques are widely used in medical imaging to automate diagnosis of different diseases and treatment. Therefore it aids the physicians on a broader scale for effective diagnosis. We discussed various challenges associated with early detection of melanoma and the design of computer vision modalities to handle such challenges. We deduced that visual descriptor is the primary content which is most crucial step in correct and timely detection of melanoma. Our main objective is to design an automated system that can detect whether the lesion is benign or melanoma.

7.2 Discussion on results

In this research, an automated system for melanoma detection using multiresolution LBP patterns and local contrast of the pattern on peripheral part of the lesion has been proposed. The full pipeline of a pattern recognition system including segmentation, feature extraction and classification has been used. Before feature extraction, physicians are interested in examining a specific clinically significant region in a lesion. Such a region is expected to have more information in the form of texture that can be relevant for detection. In case of melanoma detection, various local features, for instance, pigment network and streaks are helpful. The presence of these local features usually occur in peripheral region of the lesion. So peripheral region contains rich texture content in the form of pigment network and streaks, while center region contains low texture information. Evaluation of peripheral region for melanoma detection yields good classification results as compared to the evaluation of the whole

lesion. Also instead of processing the whole lesion we focus our research on peripheral part of the lesion for better analysis of the features. Secondly, Segmentation is an imperative preprocessing step for CAD system of skin lesions. For the segmentation we have used active contours with creasness function followed by feature extraction using multiresolution LBP pattern along with its contrast (strength). Effectively this enhances the effectiveness of standard LBP. Multiresolution LBP and contrast are optimally combined using joint histogram. This selection is based on the clinical significance of differential structures in dermoscopic images. The evaluation performance of the presented automated system is tested on manual and automatic segmented images using overall accuracy, sensitivity and specificity. The best performance of the system, SE=92.3% and SP=93.1% was achieved by manual annotated masks on peripheral region. While the performance of the system SE=91.3% and SP=92.5% was achieved by automatic segmented masks on peripheral region. In view of the classification results, minor degradation was observed when automatic segmentation methods are used, instead of manual ones. According to these results, we conclude that joint histogram of LBP pattern and contrast features on peripheral part has s significant role in the detection of melanoma. Our results shows the significance of peripheral region for melanoma detection. In addition, theses results indicate the superiority of the proposed feature matrix in comparison with the other state-of-the-art feature extraction methods.

7.3 Future Work

Although good results are obtained, there is still scope for improvement in each step of the pattern recognition pipeline for melanoma detection. In dermoscopy images, there are some situations where there is a smooth transition between the skin and the lesion. Active contours fail to detect true boundaries of the lesions in such cases. The need to devise more adequate image features for use by active contours may improve the segmentation results. Since the dermoscopy images suffer from reduced color spaces, therefore more adequate color descriptors can be constructed. This can be done by adapting the color spaces to the specific scenario of dermoscopy potentially giving better results for the identification of melanoma. In the future, it would be interesting to enrich this variant of LBP with the best set of shape and color features. Furthermore, it would also be interesting to analyze the behavior of this variant of LBP descriptors when applied to larger databases.

Nomenclature

Acc Accuracy

AI Asymmetry Index

CAD Computer Aided Diagnosis

CV Computer Vision

GLCM Gray Level Co-occurance matrix

HCI Human Computer Interaction

HSV Hue, Saturation, Value

KNN K Nearest Neighbor

LBP Local Binary Pattern

LOG Laplacian of Gaussian

MM Malignant Melanoma

NMSC Non Melanoma Skin Cancer

PR Pattern Recognition

ROC Receiver operating characteristics

ROI Region of Interest

SE Sensitivity

SP Specificity

SVM Support Vector Machines

TN True Negative

TP True Positive

WHO World Health Organization

Bibliography

[1] Cancer facts and figures. American cancer society, atlanta, ga, usa, 2015, tech. rep. 2015.

[2] Dermoscopy Tutorial. http://www.dermoscopy.org/atlas/base.html.

[3] Ashfaq A Marghoob, Josep Malvehy, Ralph P Braun, and Alfred W. Kopf. *An Atlas of Dermoscopy.* CRC Press, 2004.

[4] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach.* Prentice Hall,, 2003.

[5] Cluster analysis. https://en.wikipedia.org/wiki/clusteranalysis.

[6] GLCM. http://matlab.izmiran.ru/help/toolbox/images/enhanc15.html., Mathworks.

[7] A. Sousa. Analysis of color and texture features of vital stained magnification endoscopy images for computer assisted diagnosis of precancerous and cancer lesions. *Phd Thesis, Faculdade de Engenharia da Universidade do Porto.*, 2008.

[8] Prewitt Masks. https://en.wikipedia.org/wiki/prewitt-operator.

[9] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern classification. *ohn Wiley and sons Inc*, 2000.

[10] KNN. https://en.wikipedia.org/wiki/k-nearest-neighbors-algorithm.

[11] ROC. https://en.wikipedia.org/wiki/receiver-operating-characteristic.

[12] Jacob Scharcanski and M. Emre Celebi. *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, 2014.

[13] T. Mendonca, P.M. Ferreira, J.S. Marques, A.R.S. Marcal, and J. Rozeira. Ph2 - a dermoscopic image database for research and benchmarking. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 5437–5440, July 2013.

[14] F. Riaz, A. Hassan, and J. Zeb. Distance regularized curve evolution: A formulation using creaseness features for dermoscopic image segmentation. *International Conference in Signal Processing (ICSP),*, 12:1061–1065, 2014.

[15] Non Melanoma Skin Cancer. http://www.cancer.net/cancer-types/skin-cancer-non-melanoma.

[16] R Rox Anderson and John A Parrish. The optics of human skin. *Journal of Investigative Dermatology*, 77:13–19, 1981.

[17] I. Maglogiannis and C. N. Doukas. Overview of advanced computer vision systems for skin lesions characterization. *IEEE transactions on information technology in biomedicine*, 13, 2009.

[18] N. H. Cox, T. C. Aitchison, J. M. Sirel, and R. M. MacKie. Comparison between lentigo maligna melanoma and other histogenetic types of malignant melanoma of the head and neck. scottish melanoma group. *JBritish ournal of Cancerv*, 73, 1996.

[19] American Academy of Dermatology. https://www.aad.org/dermatology-a-to-z/diseases-and-treatments/m-p/melanoma/diagnosis-treatment.

[20] C. Barata, M. Ruela, M. Francisco, T. Mendoncža, and J. S. Marques. Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal*, 2013.

[21] Vaidyanathan S, Mansour P, Hughes PL, Selmi F, Singh G, Pulya K, and Soni BM. Challenges in diagnosis and treatment of a cervical spinal cord injury patient with melanoma, adenocarcinoma, and hepatic and osteolytic metastases: need to implement strategies for prevention and early detection of cancer in spinal cord injury patients. *Case Reports in Oncological Medicine,Hindawi*, 2012.

[22] P. Carli, V. De Giorgi, E. Crocetti, F. Mannone, D. Massi, A. Chiarugi, and B. Giannotti. Improvement of malignant/benign ratio in excised melanocytic lesions in the dermoscopy era : A retrospective study. *The British Journal of Dermatology*, pages 135–150, 2004.

[23] R Rox Anderson and John A Parrish. The optics of human skin. In *Journal of Investigative Dermatology*, volume 77, pages 13–19, 1981.

[24] T. L. Hwang, W. R. Lee, S. C. Hua, and J.Y. Fang. Cisplatin encapsulated in phosphatidylethanolamine liposome enhances the in vitro cytotoxicity and in vivo intratumor drug accumulation against melanomas. *J. Dermatol. Sci.,*, 46:11–20, 2007.

[25] Argenziano G, Soyer HP, De Giorgi V, Piccolo D, Carli P, and Delfino M. Dermoscopy: A tutorial. *Milan: EDRA Medical Publishing&NewMedia*, 2002.

[26] H. Lorentzen, K. Weismann, C. S. Petersen, F. G. Larsen, L. Secher, and V. Skodt. Clinical and dermatoscopic diagnosis of malignant melanoma. *Acta Derm Venereol*, 79:301–304, 1999.

[27] M. Elbaum, A. W. Kopf, and H. S. Rabinovitz et al. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study. *Journal of the American Academy of Dermatology*, 44:207–218, 2001.

[28] Joshua B Fishkin, Olivier Coquoz, Eric R. Anderson, Matthew Brenner, and Bruce J. Tromberg. Frequency-domain photon mi-

gration measurements of normal and malignant tissue optical properties in a human subject. *Applied Optics*, 36:10–20, 1997.

[29] J. Solomon, S. Mavinkurve, D. Cox, and R. M. Summers. Computer assisted detection of subcutaneous melanomas. *Acad. Radiol*, 11:678–85, 2004.

[30] C. Pleiss, J. H. Risse, H. J. Biersack, and H. Bender. Role of fdg pet in the assessment of survival prognosis in melanoma. *Cancer Biother. Radiopharm*, 22:740–747, 2007.

[31] M. Elgamal. Automatic skin cancer images classification. *International Journal of Advanced Computer Sciences and Applications*, 4, 2013.

[32] Barata.C, Marques.J, and S. Rozeira.J. A system for the detection of pigment network in dermoscopy images using directional filters. *IEEE Trans. Biomed. Eng.*, 59:2744–2754, 2012.

[33] J. T. Tou and R. C. Gonzalez. Pattern recognition principles. In *Massachusetts: Addison-Wesley*, 1974.

[34] X. Yuan, Z. Yang, G. Zouridakis, and N. Mullani. Svm-based texture classification and application to early melanoma detection. *Proc. 28th IEEE EMBS Annu. Int. Conf., New York,*, pages 4775–4778, 2006.

[35] Akira Hayashi Nobuo Suematsu, Yoshihiro Ishida, and Toshihiko Kanbara. Region-based image retrieval using wavelet transform. *Digital Image Processing, Springer-Verlag Berlin, Germany,*, 2005.

[36] R. M. Haralick, K. Shanmugan, and I. Dinstein. Texture features for image classification. *IEEE transactions on Systems, Man and Cybernatics*, 36, 1973.

[37] T. Ojala, M. Pietikainen, and D. Harwood. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29, 1996.

[38] R. C. Dubes and A. K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 16, 1989.

[39] A. A. Pentland. Fractal based description of natural scenes. *IEEE transactions on Pattern Recognition and Machine Intelligence*, 6, 1984.

[40] J. Mao and A. K. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25, 1992.

[41] F. W. Campbell and J. G. Robson. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, 1968.

[42] J. J. Liu M. H. Bharati and J. F. Macgregor. Image texture analysis: methods and comparisons. *J. Chemometrics Intell*, 2004.

[43] R. E. Woods R. C. Gonzalez and S. L. Eddins. Digital image processing. *Prentice Hall*, 2008.

[44] M. A. Kutay C. Candan and H. M. Ozaktas. The discrete fractional fourier transform. *IEEE transactions on Signal Processing*, 48, 2000.

[45] M. E. Celebi, H. A. Kingravi, and B. Uddin et al. A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics*, 31:362–373, 2007.

[46] F.. Chawla Ecral, A..Stoecker, Lee W. V., Moss H.-C., and R. H. Neural network diagnosis of malignant melanoma from color images. *IEEE Trans on Biom Eng*, 41:837–845, 1994.

[47] V. T. Y. Ng, B. Y. M. Fung, and T. K. Lee. Determining the asymmetry of skin lesion with fuzzy borders. *Comput. Biol. Med*, 35:103–120, 2005.

[48] D. Zhang and G. Lu. A comparative study on shape retrieval using fourier descriptors with different shape signatures. *Journal of Visual Communication and Image Representation*, 14:41–60, 2003.

[49] V. Vapnik. Estimation of dependences based on empirical data. *Springer Verlag, Newyork*, 1979.

[50] J. C. Russ. The image processing handbook. In *6th ed. New York USA CRC Press*, 2011.

[51] M. Sadeghi, T. Lee, H. Lui, D. McLean, and M. S. Atkins. Global pattern analysis and classification of dermoscopic images using textons. *SPIE Medical Imaging Conference*, 8314, 2012.

[52] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern classification. *John Wiley and sons Inc*, 2000.

[53] Scope A, Dusza SW, and Halpern AC et al. The ugly duckling sign agreement between observers. In *Arch Dermatol. 2008*, volume 144, pages 58–64, 2008.

[54] A. R. W. Stolz and A. B. Cognetta. Abcd rule of dermatoscopy:a new practical method for early recognition of malignant melanoma. *Eur. J.Dermatol*, 4:521–527, 1994.

[55] K. C. a. M. S. Menzies and C. Ingvar. Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. *Archives Dermatol*, 132:1178–1182, 1996.

[56] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions.comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Arch. Dermatol*, 134:1563–1570, 1998.

[57] H.P. Soyer, G. Argenziano, I. Zalaudek, R. Corona, F. Sera, R. Talamini, F. Barbato, A. Baroni, L. Cicale, and et al. A. Di Stefani. Three-point checklist of dermoscopy. *Journal of Dermatology*, 208:27–31, 2000.

[58] Henning JS, Dusza SW, Wang SQ, and et al. The cash (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *J Am Acad Dermatol*, pages 45–56, 2007.

[59] I. Stanganelli, M. Burroni, S. Rafanelli, and L. Bucchi. Intraobserver agreement in interpretation of digital epiluminescence microscopy ii. *Journal of the American Academy of Dermatology*, 33:584–589, 1995.

[60] Pehamberger, Steiner H., and K. A., Wolff. In vivo epiluminescence microscopy of pigmented skin lesions. *Pattern analysis of pigmented skin lesions. J. Am. Acad. Dermatol*, 17:571–583, 1987.

[61] Rezze, G.G., De Sa, B.C.S., and R.I. Neves. Dermoscopy: the pattern analysis. *Anais Brasileiros de Dermatologia*, 81:261–268, 2006.

[62] Argenziano, Soyer G., Chimenti H.P., Talamini S., Corona R., Sera R., Binder F., Kopf M., and A.W. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. *J. Am. Acad. Dermatol*, 48:680–693, 2003.

[63] Anantha, Moss M., R.H.. Stoecker, and W.V. Detection of pigment network in dermatoscopy images using texture analysis. *Comput. Med. Imaging Graph*, 28:225–234, 2004.

[64] Betta. G, Di Leo. G, Fabbrocini. G, Paolillo. A, and Sommella. P. Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. *IEEE International Workshop on Medical Measurement and Applications, MeMeA, art.*, 1644462:63–67, 2006.

[65] Betta. G, Di Leo. G, Fabbrocini.G, Paolillo. A, and Scalvenzi. M. Automated application of the 7-point checklist diagnosis method for skin lesions: estimation of chromatic and shape parameters. *Conference Record IEEE Instrumentation and Measurement Technology Conference*, 3:1818–1822, 2005.

[66] Di Leo.G, Paolillo. A., Sommella.P, and Liguori.C. An improved procedure for the automatic detection of dermoscopic structures in digital elm images of skin lesions. *In Proceedings of IEEE Conference Virtual Environments Human-Computer Interfaces and Measurement Systems*, pages 190–194, 2008.

[67] B. Shrestha, J.Bishop, K. Kam, X. Chen, R.H. Moss, W.V. Stoecker, S. Umbaugh, R.J. Stanley, M.E. Celebi, A.A. Marghoob, G. Argenziano, and H.P.Soyer. Detection of atypical texture features in early malignant melanoma. *Skin Res. Technol*, 16:60–65, 2010.

[68] Sadeghi. M, Razmara. M, .P Wighton, Lee. T.K., and Atkins. M.S. Modeling the dermoscopic structure pigment network using a clinically inspired feature set. *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics. LNCS,Springer, New York*, 6326:467–474, 2010.

[69] Wighton.P, Lee. T.K, Lui. H, McLean. D.I, and Atkins. M.S. Generalizing common tasks in automated skin lesion diagnosis. *IEEE Trans. Information Technology in Biomed*, 15:622–629, 2011.

[70] Braun. R.P, Rabinovitz. H.S, Oliviero. M, Kopf. A.W, and Saurat. J.H. Dermoscopy of pigmented skin lesions. *J. Am. Acad. Dermatol*, 52:109–121, 2005.

[71] Sadeghi.M, Lee. T.K, McLean. D, Lui. H, and Atkins. M.S. Oriented pattern analysis for streak detection in dermoscopy images. medical image computing and computer-assisted intervention. *MICCAI In International Conference on Medical Image Computing and Computer Assisted Intervention*, 15:298–306, 2012.

[72] Sadeghi. M, Lee. T, Lui. H, McLean. D, and Atkins. S. Detection and analysis of irregular streaks in dermoscopic images of skin lesions. *IEEE Trans. Med. Imaging (in press)*, 2013.

[73] P. M. Ferreira, T. Mendonc, a.J. Rozeira, and P. Rocha. An annotation tool for dermoscopic image segmentation. In *in Proceedings of the 1st International Workshop on Visual Interfaces for Ground Truth Collection in Computer Vision Applications. ACM,*, pages 5:1–5:6, 2012.

[74] S. Sachdeva. Fitzpatrick skin typing: Applications in dermatology. *Indian J Dermatol Venereol Leprol*, 75:93–96, 2009.

[75] M. Silveira, J. C. Nascimento, J. S. Marques, A. R. S.Marcal, T. Mendonca, S. Yamauchi, J. maeda, and J. Rozeira. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE Journal of Selected Topics in Signal Processing*, 3:35–45, 2009.

[76] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79, 1988.

[77] J. S. A. M. Lopez, F. Lumbreras, and J. J. Villanueva. Evaluation of methods for ridge and valley detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21:327–335, 1999.

[78] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, no. 7:971–987, 2002.

[79] I. Guyon, J. Weston, and S. Barnhilland V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46,no.1-3:389–422, 2002.