Segmentation and Lesion Detection in Dermoscopic Images

A thesis submitted as partial fulfillment of the requirement of Doctor of Philosophy (Ph.D.)

by

Khalid Ahmad A Eltayef

Computer Science Department

Brunel University London

November 2017

Declaration

I, Khalid Eltayef, hereby declare that the work presented in this thesis is completely my own and has not been submitted for any other degree qualification. Parts of the work has been previously published in conference or journal papers, this has been mentioned in the thesis. Where I have consulted the work of others, this is always clearly stated.

List of Publications

The following conferences publications have been produced as a direct or indirect result of the research that is discussed in this thesis:

1. Eltayef, Khalid, Yongmin Li, and Xiaohui Liu. "Detection of pigment networks in dermoscopy images." Journal of Physics: Conference Series. Vol. 787. No. 1. IOP Publishing, 2017.

2. Eltayef, Khalid, Yongmin Li, and Xiaohui Liu. "Detection of melanoma skin cancer in dermoscopy images." Journal of Physics: Conference Series. Vol. 787. No. 1. IOP Publishing, 2017.

3. Eltayef, Khalid, Yongmin Li, and Xiaohui Liu. "Lesion segmentation in dermoscopy images using particle swarm optimization and markov random field." International Symposium on Computer-Based Medical Systems - (IEEE CBMS 2017).

4. Eltayef, Khalid, Yongmin Li, Bashir I. Dodo, and Xiaohui Liu. "Skin Cancer Detection in Dermoscopy Images Using Sub-Region Features." Intelligent Data Analysis (IDA 2017).

This thesis is dedicated with endless love and gratitude to: my wife Sara who has been always supportive, encouraging despite worked hard to see that I succeed in life, and my beautiful daughters Nadeen and Maryia who with their love and jokes brought happiness and cheerfulness to my life.

Abstract

Malignant melanoma is one of the most fatal forms of skin cancer. It has also become increasingly common, especially among white-skinned people exposed to the sun. Early detection of melanoma is essential to raise survival rates, since its detection at an early stage can be helpful and curable. Working out the dermoscopic clinical features (pigment network and lesion borders) of melanoma is a vital step for dermatologists, who require an accurate method of reaching the correct clinical diagnosis, and ensure the right area receives the correct treatment. These structures are considered one of the main keys that refer to melanoma or non-melanoma disease. However, determining these clinical features can be a time-consuming, subjective (even for trained clinicians) and challenging task for several reasons: lesions vary considerably in size and colour, low contrast between an affected area and the surrounding healthy skin, especially in early stages, and the presence of several elements such as hair, reflections, oils and air bubbles on almost all images.

This thesis aims to provide an accurate, robust and reliable automated dermoscopy image analysis technique, to facilitate the early detection of malignant melanoma disease. In particular, four innovative methods are proposed for region segmentation and classification, including two for pigmented region segmentation, one for pigment network detection, and one for lesion classification. In terms of boundary delineation, four pre-processing operations, including Gabor filter, image sharpening, Sobel filter and image inpainting methods are integrated in the segmentation approach to delete unwanted objects (noise), and enhance the appearance of the lesion boundaries in the image. The lesion border segmentation is performed using two alternative approaches. The Fuzzy C-means and the Markov Random Field approaches detect the lesion boundary by repeating the labeling of pixels in all clusters, as a first method. Whereas, the Particle Swarm Optimization with the Markov Random Field method achieves greater accuracy for the same aim by combining them in the second method to perform a local search and reassign all image pixels to its cluster properly. With respect to the pigment network detection, the aforementioned pre-processing method is applied, in order to remove most of the hair while keeping the image information and increase the visibility of the pigment network structures. Therefore, a Gabor filter with connected component analysis are used to detect the

pigment network lines, before several features are extracted and fed to the Artificial Neural Network as a classifier algorithm. In the lesion classification approach, the K-means is applied to the segmented lesion to separate it into homogeneous clusters, where important features are extracted; then, an Artificial Neural Network with Radial Basis Functions is trained by representative features to classify the given lesion as melanoma or not. The strong experimental results of the lesion border segmentation methods including Fuzzy C-means with Markov Random Field and the combination between the Particle Swarm Optimization and Markov Random Field, achieved an average accuracy of 94.00% , 94.74% respectively. Whereas, the lesion classification stage by using extracted features form pigment network structures and segmented lesions achieved an average accuracy of 90.1% , 95.97% respectively. The results for the entire experiment were obtained using a public database PH2 comprising 200 images. The results were then compared with existing methods in the literature, which have demonstrated that our proposed approach is accurate, robust, and efficient in the segmentation of the lesion boundary, in addition to its classification.

Acknowledgements

First of all I would like to express my sincere gratitude to my supervisor, Dr. Yongmin Li (Brunel University) for his significant advices, mentor-ship, unconditional support and encouragement throughout my PhD research. Our discussions always gave me the confidence of being on the right path. I also owe special thanks to Professor Xiaohui Liu (Brunel University), my second supervisor, for his suggestions and encouraging comments throughout this research.

Special thanks to my wife Sara and my beautiful princesses Nadeen and Mariya who have been hugely supportive throughout the four years it has taken to achieve this work. Also I thank all my family and all my relatives in Libya for their continued support.

Finally, I cannot forget to thank all my colleagues for their support including Faris Alwzinani, Mahir Arzoky, Bashir Dodo, Mashael Al-luhaybi, Liela Yousefi, Nadia Hussain, Weibo Liu, Najeeb Gambo, Khalipha Abubaker, and all PhD student. As well as thank all the lectures and stuff in Computer Science Department. Also I would like to thank Aljabal Algharby University for giving me the opportunity to study in the UK.

Contents

List of Figures

List of Tables

List of Abbreviations

AC Accuracy Rate.

ANN Artificial Neural Network.

Bg Background.

BoF Bag of Features.

CAD Computer Aided Diagnosis

DSC Dice Similarity Coefficient.

FCM Fuzzy C-means.

Fg Foreground.

FN False Negative Rate.

FP False Positive Rate.

MRF Random Markov Filed.

PSO Particle Swarm Optimization.

SE Sensitivity Rate.

SP Specificity Rate.

SVM Support Vector Machine.

TN True Negative Rate.

TP True Positive Rate.

Chapter 1

Introduction

1.1 Introduction

Skin cancer can be classified into melanoma and non-melanoma. Melanoma is a malignancy of the cells that give the skin its colour (melanocytes) and can invade nearby tissues. Moreover, it can spread throughout the human body and may cause death. Conversely, non-melanoma seldom spreads to other parts of the body. Malignant melanoma is considered one of the most fatal forms of human skin cancers, which led to a raised mortality rate. In the last few decades, the incidence of melanoma has increased significantly, especially among white-skinned people who are exposed to the sun. For instance, in North America, melanoma became the fifth common cancer among males and the sixth common cancer among females, while the fourth most common in Australia [41] [151] [137] [19] [143]. Moreover, according to reports from the western world, melanoma is considered the seventh most common malignancy in women and the sixth most common in men [148]. Nevertheless, it is also the most treatable type of skin cancer if detected or diagnosed at an early stage [126]. With an early diagnosis, melanoma can often be cured with a simple excision; thereby reducing the mortality rate.

Dermoscopy is one of the major tools used in the diagnosis of melanoma. It is widely used by dermatologists due to its value in detecting melanoma in its early stages. It provides better visualisation of several pigmented structures, such as streaks, dots, pigment networks and blue-white areas, which are invisible to the naked eye [159] [142] [44] [5]. One or two features

alone are not sufficient to identify the given lesion as melanoma. Thus, by using dermoscopy images, dermatologists become more confident in distinguishing the types of lesions.

In melanoma detection, dermatologists adopt the ABCD rule to analyse four parameters (Asymmetry, Border, Colors and Diameter), to diagnose melanoma at an early stage. In addition, 7-point checklist criteria also used for the same purpose [41] [77] [125] [15].

In addition to the presence of hair and reflection in images, many melanoma borders are often invisible or fuzzy, which makes visual identification very difficult for skin cancer experts. Moreover, the interpretation of the images is time consuming and subjective, even for trained dermatologists. Computer Aided Diagnosis (CAD) became essential to eliminate all these issues and assist the specialists and physicians to interpret images clearly, and reach the correct decision for their diagnosis. In addition, such systems reduce the time required for diagnosis and increase the accuracy of the final results. CAD systems have been proposed by many research groups to identify various structures in medical images. Depending on medical knowledge, CAD systems try to mimic the performance of the dermatologists for determining the given lesion as normal or abnormal skin [24] [19] [105] [134]. Integrating dermoscopy techniques with CAD systems has become a crucial research field in recent years, since it assists the physicians in obtaining meaningful information from images. Subsequently, these can be used to identify melanoma correctly.

Although the use of several methods in melanoma disease diagnosis has been seen to evolve swiftly in recent years, the evaluation of the region structures, such as region borders, pigment network, dots/globules, and streaks, can be complicated and require more time and attention. It has been considered that an accurate and efficient detection of these structures can aid in early melanoma detection. Several segmentation approaches, based on the early detection of melanoma fail to extract the whole and complete structures, due to the quality of the images (noise, low contrast or intensity inhomogeneity). In addition, many methods are source-dependent; in other words, they cannot perform well on images from different sources.

The pigment network and the region border are considered among the key dermoscopic structures that refer to melanoma or non-melanoma disease. For instance, the presence of pigment network structures on the pigmented regions is an unmistakable sign of the lesion's

origin, because its appearance is a reference to the existence of melanin in deep skin layers, which is regarded as a hallmark of melanocytic nevi. In addition to the pigment network, the asymmetry, border irregularity, colour and diameter structures can be obtained from the region shape, which play a key role in melanoma diagnosis. Therefore, effective segmentation of these significant structures can improve the productivity of dermatologists by reducing the time required in the diagnosis process.

This study compiles different fields, such as image pre-processing, segmentation, feature extraction, and classification, all of which make important contributions to the extraction of the skin region structures and its classification. The focus of this thesis is to address issues in the segmentation of the pigment network structures and skin border regions, and distinguish skin lesion types, which include:

- Removal of reflection artifacts and hairs from images with no impact on pigment network lines.
- Detection and segmentation of pigment network structures from dermoscopy images with no restrictions on the sizes and shapes of it.
- Accurate segmentation of skin region boundaries over dermoscopy images without affecting the region border.
- classification of skin lesion based on several extracted features.

The proposed approach could be used to support a non-invasive diagnosis technique in modern dermatology for early detection of melanoma diseases and treatment evaluation.

1.2 Dermoscopy Image Analysis and Early Detection of Melanoma Disease

It is important to highlight the main purpose of this research and gain some knowledge about dermoscopy image analysis. Increasing the number of images to be examined and assisting the expert dermatologists to obtain a more accurate diagnosis of melanoma, in addition to

reducing examination time are the key motivation for the necessity of developing automated image analysis tools. Doctors are compelled to perform the image analysis manually (visual interpretation), which is extremely time consuming and tedious. Frequently, the outcomes of dermoscopy image analysis are subjective and vulnerable to human error, which means the accuracy of diagnosis depends on the experience of the dermatologists [85]. Therefore, the need for reliable, robust and accurate computerised techniques is fundamental and necessary. This can decrease the taken time for image examination and increase the accuracy of the final diagnosis.

Early detection of melanoma disease has been shown to increase the probability of patients survival. Meanwhile, its mortality rate has also risen significantly, in cases where it is not detected early. Malignant melanoma is always curable if identified and treated early. Conversely, late diagnosis makes the disease difficult to cure and may lead to death, since it is capable of advancing and spreading to other parts of the body. Melanoma is considered the deadliest form of skin cancer. According to the report produced by the American Cancer Society (ACS), about 10,130 fatalities and 76,380 new cases of melanomas are estimated to be diagnosed in 2016, with incidences rising every year [107].

The principal challenge facing modern dermatology presently is to provide an effective automated analysis tool, capable of performing early diagnosis of melanoma disease and reducing the required time for the diagnostic process while maintaining a high standard of accurate diagnosis.

1.3 Dermoscopic Image Structures Extraction

As explained previously, the morphology of the dermoscopic structures such as pigment networks and region borders are the key indicator of malignant melanoma diseases. Skin cancer experts spend a great deal of time and energy for the purpose of evaluating these significant structures. However, the image processing techniques, such as segmentation and classification, can allow an automated analysis of dermoscopy images; consequently, only complicated images are reviewed by dermatologists. Several segmentation and classification methods have been

developed, including:

- Location and segmentation of pigment network structures.
- Location and segmentation of region boundary structures.
- Segmentation of other structure such as streaks and dots/globules.
- Classification of expected lesions as melanoma or not.

This work focuses on the segmentation of the pigment network structures with the delineation of the region borders, and the classification of the given lesions. In recent years, many algorithms have been published for the segmentation and the classification of these structures. These algorithms segment the region of interest based on the pixel intensity, colour and texture, whereas classify the lesions depends on the shape and size. It is important to highlight that most of these methods fail to provide full segmentation of the dermoscopic structures due to the presence of imaging artifacts, such as hair, lightening reflection, low contrast between the region of interest and the image background (healthy skin) and intensity inhomogeneity. Furthermore, no accurate classification method works well for any image sources.

1.3.1 Challenges in Dermoscopic Image Structures Detection and Skin Lesion Classification

Various segmentation methods of pigment network and region border structures have been implemented for the purpose of helping the expert dermatologists for obtaining an accurate diagnosis of melanoma. These two structures are vital features in identifying melanoma disease. For instance, the lesion border structure provides crucial information for accurate diagnosis of melanoma. Many clinical features, such as asymmetry, border irregularity and shape size, are calculated from the lesion border. In addition, the extraction of other significant features such as globules, pigment networks and blue-white areas, critically depends on the accuracy of border detection. However, an accurate segmentation of these dermoscopic structures is a complex task for many reasons including the low contrast between the region of interest and the background (smooth transition between the lesion and the surrounding

healthy skin), the existence of noise in the image such as reflection artifacts, skin lines and air bubbles, and the intensity inhomogeneity. In addition, the variation of skin lesion shapes, sizes and colors. Moreover, the presence of hair causes a significant degradation in the performance of the automated pigment network segmentation techniques, as well as lesion border structures detection. In addition, extracting the desired features from images to provide an accurate diagnosis of melanoma can be difficult due to the same issues described earlier, since these artifacts often cause to occlude the targeted feature. Figure 1.1 shows many different challenges of the pigment network and lesion border segmentation. The challenges mentioned are summarised below:

- Poor contrast between the lesion and the surrounding healthy skin.
- Invisible or fuzzy lesion borders.
- Presence of several noises artifacts such as air bubbles, gel, skin lines, and lightening reflection in images.
- The high degree of visual similarity between melanoma and non-melanoma lesions.
- Variegated coloring inside the lesion area.
- The overlapping between the hair and the pigment network lines.
- variation of lesion color, shapes, and sizes.
- Obtain images from multiple sources.

1.4 Thesis Aims

This work aims to provide accurate, robust, reliable and automated dermoscopic image analysis technique, to allow for early detection of malignant melanoma disease

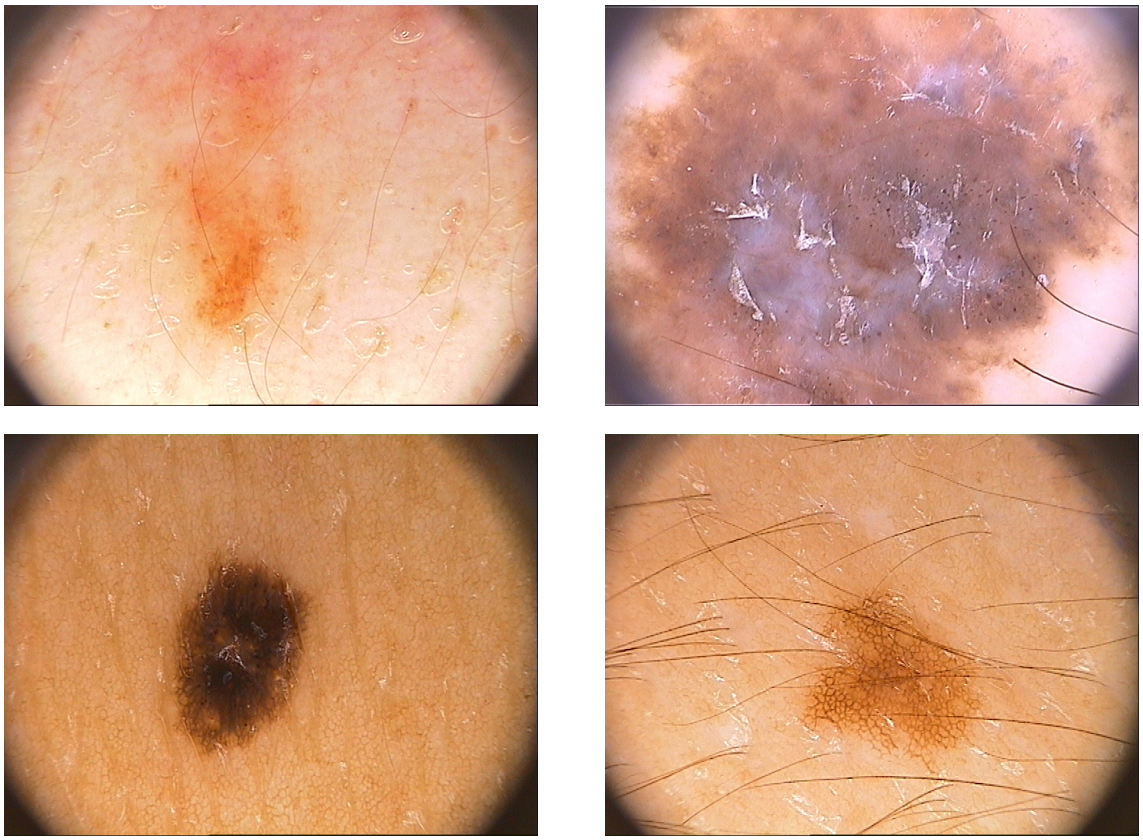


Figure 1.1: Illustration of different challenges of lesion segmentation with pigment network detection: low contrast and invisible lesion boundary (top left), variegated coloring (top right), presence of skin lines (bottom left) and existence of hair with air bubbles (bottom right). (Source: [101]).

1.5 Thesis Objectives

In particular, the specific objectives of the work described in this thesis are as follows:

- Develop a new method to reduce the noise from dermoscopic images without corrupting image information, and produce images more suitable for further processing.
- Implement a new algorithm to extract the actual lesion boundaries in dermoscopic images.
- Develop an automatic algorithm for the detection of pigment network structures.
- Identify skin lesion types based on the presence or absence of the pigment network structures.
- Develop a fully automatic approach for melanoma detection in dermoscopy images based on a suitable number of features with an ideal process of extracting them.

1.6 Thesis Motivations

The specific motivations can be illustrated as follows:

- Melanoma is one of the most rapidly increasing cancers globally.
- It is the most dangerous form of human skin cancer.
- The mortality rate caused by melanoma has been increased significantly in the last decades.

1.7 Contributions

Dermoscopic image analysis has been a research topic that influences a large number of medical image analysis research groups worldwide. With the deployment of more sophisticated non-invasive imaging tools (dermoscopy), many segmentation and classification algorithms have been implemented to extract the structures of the pigmented regions, in order to obtain

the amount of key information for the lesion, to assist in the classification process. These techniques are used frequently to help dermatologists perform efficient and reliable skin cancer disease assessment.

Dermoscopic structures segmentation techniques were applied on dermoscopic images to extract several clinical features such as lesion boundaries, pigment networks, dots/globules, and streaks. In attempting to apply these methods to real life, dermoscopic images, noise, intensity inhomogeneity, algorithmic complexity, and many other imaging artifacts all became key obstacles to be overcome in automated dermoscopy structures segmentation.

To address the above issues, the existing techniques require more robust image processing operations, as well as the incorporation of some prior knowledge about different dermoscopy structures. The focus was placed on noise estimation operations that were used to improve the appearance of significant features in the image. A careful selection and combination of processing techniques were conducted to allow complete extraction of the different dermoscopy structures. This research incorporates some of the previous work in image segmentation and classification, to guide the development of novel segmentation and classification methods designed to provide and accurately extract the features from images, to identify the lesion types correctly.

The work presented in this thesis provides the fundamental knowledge for segmenting the significant structures of images, such as pigment network structures and regions borders, in addition determining the lesions type, in order to allow complete use of the dermoscopy image analysis tools to day-to-day work in modern dermatology.

The main contributions of this thesis are outlined below:

- Segmentation of lesions using FCM and MRF: The skin lesion boundary plays a vital role in early detection of melanoma disease, since the ABCD-E rule is dependent on its accurate segmentation. However, the primary drawback of the low and/or medium level methods, such as threshold, edge-based, region-based, K-means and Fuzzy C-means, is that they deal with pixels individually and only by their intensity values. Therefore, they lack the capacity to model the overall appearance of a local neighbourhood region. Also, the overlapping between the hair and the lesion boundaries could cause an in-

correct detection of the lesion. In order to address this issue, the Gabor filter, image sharpening, Sobel filter and image inpainting methods were implemented respectively for the purpose of removing imaging noise. The Fuzzy C-means (FCM) was applied to images to estimate the initial parameters for the Markov Random Field (MRF) method. The primary objective of image segmentation using the MRF method is to minimise the energy function or maximise the probability of pixel allocation to a cluster by using Maximum a Posteriori (MAP). The pre-processed images gave the algorithm an ability to delineate the lesion boundary properly.

- Segmentation of lesions using PSO and MRF: As aforementioned, the lesion borders are considered one of the main key dermoscopic structures used in identifying malignant melanoma disease at its early stage. Consequently, the actual lesions have to be determined and extracted accurately. However, accurate segmentation of lesions is a challenge as mentioned earlier (low contrast edges and intensity inhomogeneity). Therefore, to handle this issue and segment the real lesion, we propose a combined method using the Particle Swarm Optimization (PSO) and the MRF. The PSO algorithm is able to generate a solution, and share it with the MRF approach to perform a local search in each iteration. This gives the algorithm robustness since it provides the best solution in each iteration and used the neighbors of every pixel to assign it again. Finding the appropriate class of each pixel in the image was performed by integrating the PSO with the MRF method.

- Pigment network detection: The main aim of detecting the pigment network structure in dermoscopy images is to identify skin lesions type (melanoma or non-melanoma), since it plays a vital role for early melanoma detection. However, imaging noise such as reflection artifacts and hairs cause significant degradation to the performance of automated segmentation of it. Therefore, careful pre-processing experiments were conducted, to define unique processing operation capable of decreasing the noise without corrupting image information, as seen with many mathematical image processing operators. Consequently, the directional Gabor filter, image sharpening, Sobel filter and image inpainting

methods were carried out respectively, in order to detect and remove the hair from images. Applying Sobel filter on sharpened images was a new method, and it was able to detect most of the hair correctly. Thus, enhancing the appearance of the pigment network structure. Removing the majority of the hair from images and keeping the image information allows a complete segmentation of the target structures. In terms of pigment network detection, the directional Gabor filter was also applied with different parameters from those used to detect the hairs. Finally, a connected component analysis with adaptive threshold method was used to extract the large mesh of pigment network lines.

- Sub-region feature extraction for lesion classification: To improve the quality of the existing diagnostic systems, a comprehensive approach was developed, including the whole process of image enhancing, segmentation of lesions, feature extraction and lesion classification. Feature extraction seeks to train the classifier algorithm to classify the lesions as melanoma or benign, as features are typically responsible for pattern recognition. Too many irrelevant features complicate the classifier and require more computational time, which simultaneously reduces the classification accuracy. To address this issue, a robust technique for extracting several numbers of features is performed by applying the K-means method to the segmented image (lesion), in order to separate each homogeneous set of pixels in one group (cluster). Therefore, the required features were extracted at the sub-region (cluster) level. The features were then fed to an Artificial Neural Network for final melanoma classification. The idea of separating image pixels into homogeneous groups and extracting desired features was succeeded in describing images properly. Our comprehensive method is capable of achieving high results in distinguishing given lesions without the need for user interaction. We believe that using the proposed approach could reduce the time and the work required for conventional supervised dermoscopy structures segmentation and classification.

1.8 Thesis Outline

This section outlines the structure of this thesis, which comprises seven chapters, including conclusions. Every chapter is an independent section of research in its own right. All chapters follow the same format, which includes the introduction of the primary problem, demonstration of the techniques used and the experimentation results. As with all traditional researches, collectively the chapters provide an evolution of ideas, following a story regarding the use of the CAD systems in modern dermatology. Figure 1.2 illustrates the structure of the main chapters.

Chapter 2 provides general background information relevant to each chapter. It includes the current imaging tools employed to capture the photographs of skin cancer and a brief introduction to the dermoscopic image analysis. In addition, the chapter presents some previous work, including skin cancer image segmentation and classification.

Chapters 3 and 4 will introduce the segmentation of the skin lesion area. Chapter 3 provides the lesion border segmentation using the Fuzzy C-means (FCM) and The Markov Random Field (MRF) approaches for the purpose of extracting the actual lesion borders. While Chapter 4 improves the lesion boundary extraction by combining The Particle Swarm Optimization approach with the MRF method.

Chapter 5 will introduce the detection of the pigment network structure using directional Gabor filter, image sharpening, Sobel filter and image inpainting method to remove most of the hair, and enhance the lines of the pigment network, which is needed to be segmented.

Chapter 6 will introduce a comprehensive approach consists of image enhancement, segmentation, feature extraction and classification, for the purpose of malignant melanoma identification. With respect to the feature extraction, we provide a different method for extracting desired features from homogeneous clusters, by applying K-means approach on segmented lesions.

Chapter 7 provides a conclusion of the work presented in this thesis, research limitations, and highlights future research directions currently under investigation.

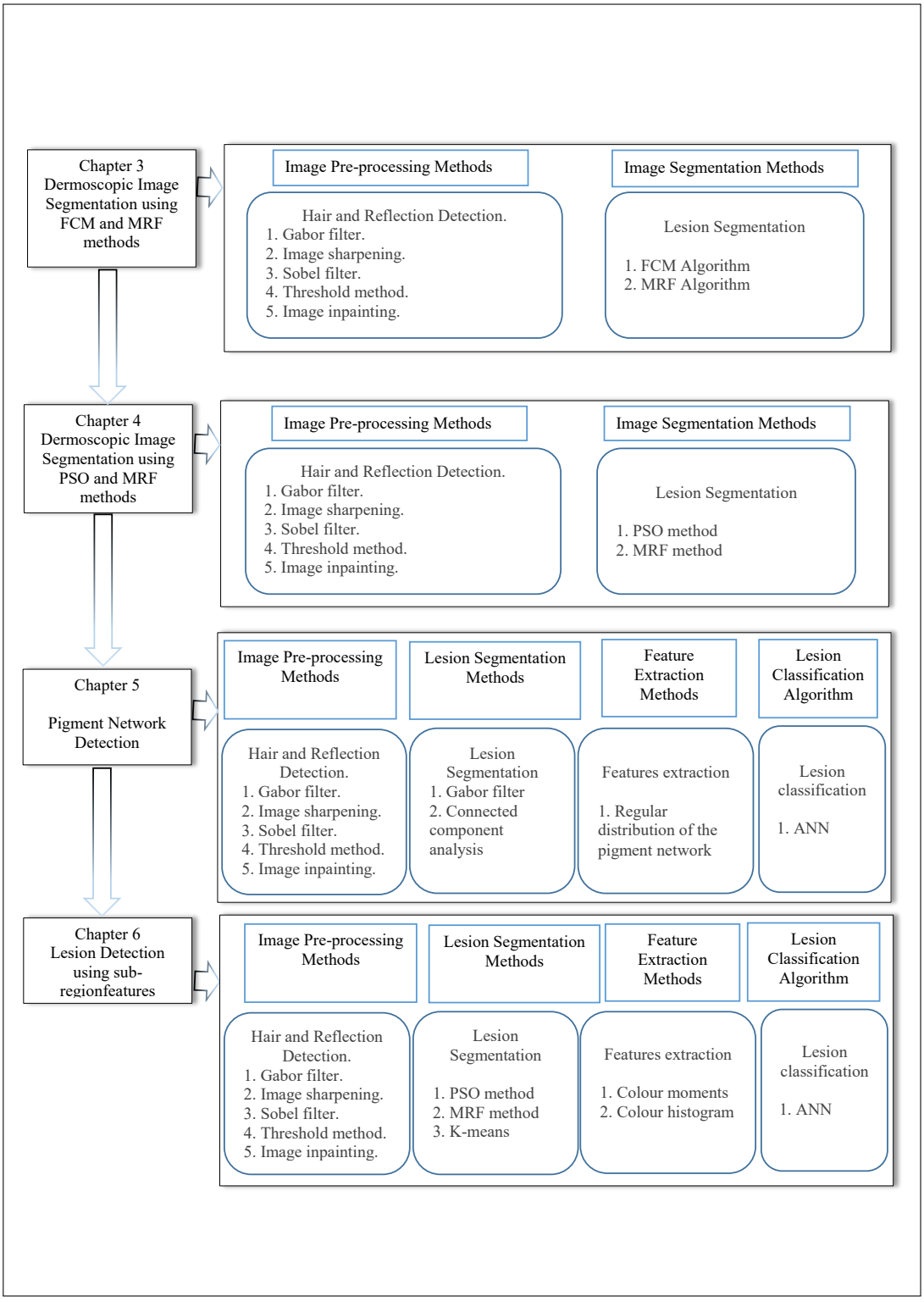


Figure 1.2: A diagram showing the structure of the technical chapters.

Chapter 2

Background of Dermoscopy Images and Literature Review

2.1 Introduction

Skin cancer is considered one of the most serious public health problems since it increases significantly the mortality rate, especially if it is not detected early. With late diagnosis, melanoma typically does not respond to treatment and can result in death. Thus, successful treatment of melanoma is through performing a simple surgical procedure, is wholly reliant on the early detection of the affected lesion area.

Since the early diagnosis of malignant melanoma significantly raises the survival rate of the patients, non-invasive imaging tools, such as dermoscopy, have emerged and developed in recent years. These tools assist the screening process and boost diagnostic accuracy. Dermoscopy allows the examination of skin lesions at a higher magnification, which in turn leads to provide more details of the morphological structures. This tool is capable to produce a better visualization of several clinical features from the pigmented skin lesion, which are invisible to the naked eye examination.

Several medical diagnosis methods are adopted, such as the 3-point checklist, the ABCDE rule, 7-point checklist, Menzies method, and pattern analysis, in order to guide dermatologists

in diagnosing melanoma skin cancer. Generally, the obtained clinical features must be assessed by expert doctors using several dermoscopic criteria, such asymmetry, border, colours, differential structures for the purpose of producing the final clinical diagnosis.

Several computer aided diagnosis (CAD) systems have been implemented and proposed by many research groups to reduce the time required for diagnosis, assist the clinical evaluation of dermoscopic structures, and increase the accuracy of the results. Depending on medical knowledge, CAD systems try to mimic the performance of skin cancer experts for distinguishing the pigmented skin lesion type. Overall, A CAD system has usually four stages, namely: image pre-processing, region segmentation, feature extraction, and lesion classification. The combination of CAD systems and dermoscopy techniques is capable to help and support the doctors for obtaining an accurate and correct diagnosis of the disease.

This chapter is structured as follows: Section 2.2 provides an overview of dermoscopy. The dataset description is presented in Section 2.3. Section 2.4 provides an overview of an image segmentation and previous works. Previous works related to lesion segmentation are presented in Section 2.5. Finally, Section 2.6 provides an overview of skin lesion classification methods developed previously.

2.2 Dermoscopy

Dermoscopy is a non-invasive diagnostic technique used to check a variety of patterns and clinical features in affected skin lesions that are invisible to the naked eye. These features are very useful for the characterization of skin lesions type. Dermoscopy has several other names such as Epiluminescence Light Microscopy (ELM), dermatoscopy or skin surface microscopy. This technique is widely used by dermatologists due to its value in providing meaningful clinical features from pigmented skin lesions. In particular, the technique consists of placing water, oil, alcohol or gel on the surface of the affected area. Thus, the lesion is inspected using a dermatoscope, a camera or a digital imaging system. Depending on the instrument used, the magnication of a given lesion ranges from 6x to 40x and even up to 100x [70]. With the magnication, several clinical structures become visible and can be used as an input in several

Table 2.1: Different datasets used by researchers. SE-sensitivity and SP-specificity.

Classification Algorithm	Reference	SE	SP	Images/Melanomas
ANN	[123]	93.3%	93.8%	588/200
	[78]	85.9%	86.0%	1258/198
	[104]	67.5%	80.5%	180/72
	[125]	78.4%	95.7%	98/51
SVM	[45]	92.0%	93.3%	564/88
	[120]	72.4%	72.4%	358/134
AdaBoost	[103]	92.0%	70.0%	152/42
	[41]	90.0%	77.0%	655/511

procedures that are used by physicians to distinguish between different types of skin lesions (melanocytic and non-melanocytic), as well as in the detection of melanomas. Examples of dermoscopic criteria and more information about dermoscopic clinical structures are illustrated in Appendix A.

2.3 Dataset Description

Most of the CAD systems are trained and tested using dermoscopic databases that are obtained at one or more hospitals. Each researcher tends to use their own dataset, which differ in size, number of melanomas, number of non- melanoma and acquisition setups. Table 2.1 illustrates a number of previous works that used various datasets with a different number of images. For instance, the authors in the work presented in [123] used their archive images, which include about 4200 cases. While the work presented in [78] used data provided from Keio University and Vienna University. Moreover, authors in [104] used a database of over 180 images, which was validated by a survey of dermatologists in the CHUT (Centre hospitalier Universitaire de Tlemcen, Algeria). On this basis, we observed that most of these datasets are not publicly available. This led to the inability to make a direct comparison between approaches, which is required to understand the value of each method

It is worth mentioning that many research groups use a large commercial dataset, which usually comes with medical information, such as the diagnosis and the evaluation of lesions.

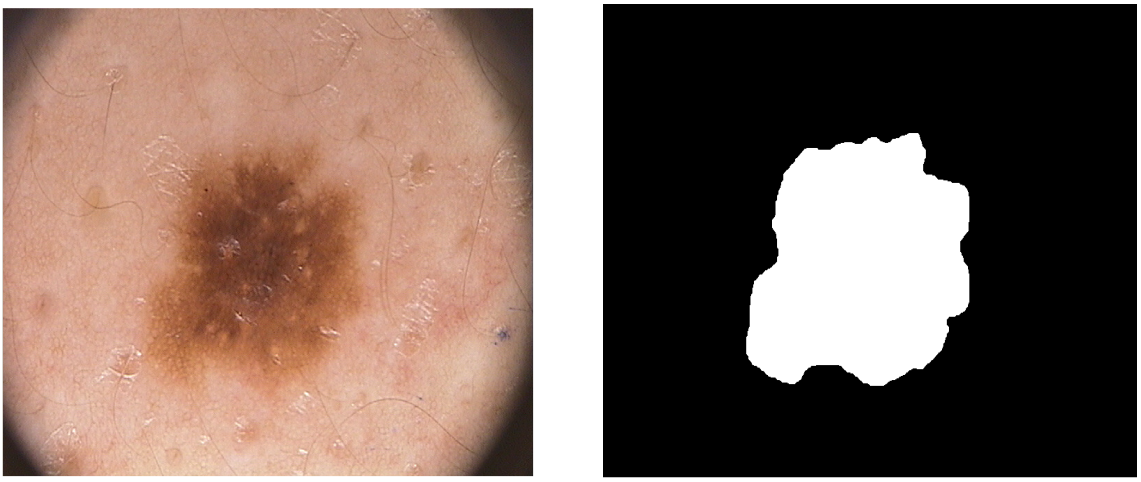


Figure 2.1: Manual segmentation of pigmented skin lesion. Left: original dermoscopy image. Right: manual segmentation mask(ground truth). (Source:[101])

The use of these datasets decreases the variability between systems, since they are trained using approximately the same set of images. However, commercial databases can be expensive and very difcult to acquire.

To the best of our knowledge, the first publicly available dataset of dermoscopic images is PH2 dataset. it was released by the dermatology service of Hospital Pedro Hispano, Portugal [101]. The dermoscopic images were acquired under the same situation through Tuebinger Mole Analyzer system using a magnification of 20x. The images are 8-bit RGB colour component with a resolution of 768 x 560 pixels. The dataset consists of 200 dermoscopic images, including 80 common nevi, 80 atypical nevi, and 40 malignant melanomas. The database also contains the manual segmentation of each image, which is represented as a binary mask. The pixels with an intensity value of 1 correspond to the segmented lesion (foreground or ROI), whereas pixels with a value of 0 correspond to the healthy skin (background). The ground truth images are essential and significant, for the evaluation of the segmentation phase of CAD systems. An example of a dermoscopic image with its ground truth mask (manual segmentation) can be seen in Figure 2.1 .

The customised annotation tool (known as DerMAT) was used by dermatologists to perform the manual segmentation and annotation of the images [61] [101]. Figure 2.2 illustrates an example of the manual segmentation of two lesions using the DerMAT software.

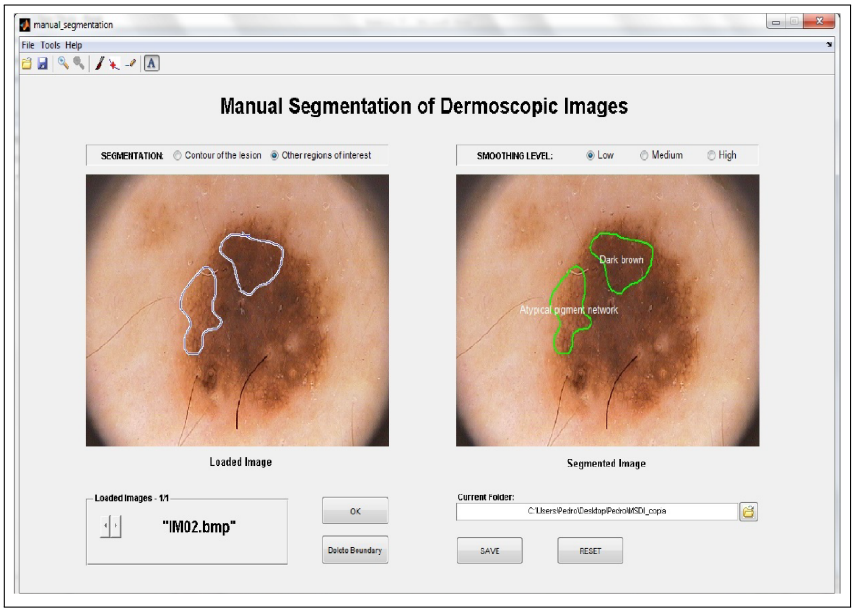


Figure 2.2: Example of DerMAT interface for the delineation and labeling of multiple regions of interest in the image. (Source:[101])

Furthermore, the dataset includes the evaluation of all images, which is significant for the evaluation while performing the classification step of CAD systems. The evaluation of each image performed by skin cancer experts followed the parameters below:

- Manual segmentation of the pigmented skin lesion, known as ground truth.
- Clinical diagnosis.
- Dermoscopic criteria (asymmetry, colors, pigment network, dots, globules, streaks, regression areas and bluewhitish veil).

Recently, the International Skin Imaging Collaboration (ISIC) dermoscopic archive contains over than 13,000 dermoscopic images, which were collected from leading clinical centers internationally, and acquired from a difference devices [1]. Based on our knowledge, no published works containing the whole images are available so far; therefore and due to the amount of available information in the PH2 dataset, such as providing the ground truth, lesion type and the information of presence or absence of relevant medical criteria, we have been motivated to use it as a source to test and evaluate our approaches. Moreover, different researchers

can acquire it and compare their results with those obtained using different methods. This also encouraged us to use it since several works have been published already based on this dataset; thus it also helped us perform the comparison. It is desired that PH2 can be used for a fair comparison between different systems, which is why it is adopted in the work described in this thesis.

2.4 An Overview of an Image Segmentation and Previous Works

Before reviewing previous works in dermoscopic images, clarification is required of the meaning and significance of image segmentation process. This can be referred to as the partitioning of an image into disjoint regions (groups) that are homogeneous with respect to a chosen property such as luminance, texture, and color [44]. Segmented images capable to improve image analysis by providing unobserved information (useful features) from original images. In case of melanoma detection using dermoscopy images, the purpose of image segmentation is to extract the affected region (lesion area) from the background (normal healthy skin). Image lesion segmentation is a crucial stage as it allows the identification of various clinical features, locally and globally. Moreover, the boundary of the segmented lesion provides useful information for use in lesion analysis. Efficient lesion segmentation leads to desired feature extraction, correct lesion classification and therefore accurate diagnosis. Consequently, the final diagnostic results of melanoma disease depend largely on the performance of the lesion segmentation stage [107].

According to the work presented by Celebi et al[44], the algorithms of image segmentation techniques can be classified into the following categories:

- Histogram thresholding. These methods include the determination of one or more histogram threshold values that extract the objects (ROI) from the background [38] [39] [96].
- Clustering. These methods involve the partitioning of a colour space into homogeneous groups using unsupervised clustering approaches such as k-means and fuzzy c-means methods [93] [100][42].

- Edge-based. These methods include the detection of edges between various regions using edge operators (Sobel and Canny edge detector).
- Region-based. These methods involve the clustering or grouping of image pixels into homogeneous regions using region merging, region splitting, or both [149] [161].
- Morphological. These methods comprise the detection of object contours from predefined seeds using the watershed transform.
- Model-based. These methods involve the modeling of images as random fields whose parameters are determined using various optimization procedures.
- Active contours (snakes and their variants). These methods involve the detection of object contours using curve evolution techniques.
- Soft computing. These methods involve the classification of pixels using soft-computing techniques including neural networks, fuzzy logic, and evolutionary computation.

The segmentation of dermoscopic structures and the distinction between lesions type in dermoscopic images are considered one of the challenging issues, due to the containing of several artifacts such as uneven illumination or lightening reflection, hair, and presence of noise. Furthermore, the low contrast between the lesion and its surrounding skin, and the variegated colouring inside the affected lesion. Therefore, several methods attempted to apply image processing techniques for different purposes, including image enhancement, pigmented lesion segmentation, and lesion classification. In this thesis, we categorised the previous works as skin lesion segmentation and skin lesion classification as presented in the following sections.

2.5 Previous Works Related to Lesion Segmentation

In dermoscopy image analysis, the morphology of the lesion boundaries (shape and size) indicator for assessing the presence of melanoma disease. Therefore, numerous methods have been proposed to perform the segmentation of the affected lesions over skin images, in order to produce a new approach capable to support the dermatologists for the final clinical diagnosis.

Hence, we review several different methods applied to skin cancer images as a way to extract the affected lesion areas, using fully-automatic and semi-automatic algorithms.

A fully automated segmentation method is based on a threshold for dermoscopic images, as described in Kruk et al[88]. The method uses histogram-based thresholding for all three RGB colour components.

Celebi et al [43] presented a fusion based method to segment the accurate lesion boarders in dermoscopic images. In their work, they applied several thresholding methods for a group of images using a fusion of four algorithms, which includes Otsu's algorithm [115], Kapur's algorithm [83], Huang's algorithm [73], and Kittler's algorithm [87].

Also, Humayun et al [74] proposed a multi threshold method, which divides the image histogram iteratively into multiple classes by selecting the threshold values for each class, based on intensity level using Otsu's method.

In another method of lesion border detection, Abbas et al [8] presented an unsupervised approach for multiple lesion segmentation using a modified region-based active contour (RACs) [92]scheme. Iterative histogram thresholding was performed on luminance image to initialise the level set automatically. Then a localised region-based active-contour model was applied to the segmented lesions. Their method was able to segment the lesion border properly whether the boundary lesion was clear, and had a good contrast with the normal skin (background). However, it failed to detect the lesions that transition smoothly with the surrounding healthy skin.

Abbas et al[7] proposed a new technique for lesion boarder delineation. Their method began with a pre-processing step where several artifacts are removed using homomorphic filtering and weighted median filtering. Then, the least-squares method (LSM)[155] was performed to obtain edge points. Therefore they utilised the dynamic programming(DP) [92] [78] technique to locate the optimal boundary of the lesion.

Zhou et al [161] introduced a new mean shift [48] based gradient vector flow (GVF) [150] algorithm that drives the internal/external energies towards the correct direction. they combined a mean field term within the standard GVF objective function to perform the segmentation process of skin lesions in dermoscopy images.

Yukse et al. [156] proposed a method to extract the boundary of pigmented skin lesion from healthy skin.The authors applied type-2 fuzzy logic technique for the purpose of determining an automatic threshold value, which was used to classify the image pixels as foreground (lesion area) and background (healthy skin).

Celebi et al. [59] presented a method to segment skin lesion in dermoscopy images through statistical region merging method. The method is a technique developed to segment images based on region growth and merging approaches.

An automatic iterative stochastic region-merging method was implemented by Wong et al [149] to segment the affected skin lesions from macroscopic images. In terms of the algorithm process and as an initial step, every single pixel in the image was assigned to a unique region, then based on a region merging likelihood function, these regions were merged with other regions in a stochastic manner. Therefore, the regions formed during the initial phase were merged using the same stochastic region merging process to obtain the final image segmentation of the skin lesions. This process continues refining the segmentation results until a stopping criterion is met.

Based on clustering algorithms, which are grouping a set of homogeneous pixels in one group, Lee and Chan [93] proposed a method based on fuzzy c-mean approach (FCM) by using type-2 fuzzy set algorithm (Zadeh [157]), for the purpose of obtaining an optimum threshold value, and delineating the cancerous boundaries from the skin images correctly. They also utilised the 3D colour constancy algorithm to minimise the shadows and affects of skin tone variations in images during the pre-processing phase.

Another approach based on fuzzy c-mean clustering and density-based clustering (DBSCAN) was applied to the segmentation of lesions from surrounding healthy skin on mobile platforms Mendi et al [100].

The K-means method was also adopted by Castillejos et al. [42] as a clustering algorithm. The authors presented a novel approach to segment and detect the border of the skin lesion based on the wavelet transform for K-Means, Fuzzy C-Means and Cluster Preselection Fuzzy C-Means methods.

Zhou et al. [160] proposed a new type of dynamic energy for the segmentation of pigmented

skin lesions in dermoscopy images. The authors combined the classical gradient vector flow (GVF) model with the mean shift method to improve the segmentation accuracy.

Threshold method was used by several researchers such as Alocon et al[12] and Pirnog et al [118] for the purpose of extracting the lesion area from surrounding health skin, using dermoscopic images. In the work presented by Pirnog et al [118], the RGB image was converted into HSV representation, the image histogram of the saturation (S) image component was computed. Therefore, the threshold value was determined and used to segment the image into two parts: foreground, which was denoted by the lesion area, and background, which was indicated the normal skin.

A new mean shift approach based on FCM method is proposed by Zhou et al [162] to extract the pigmented skin lesions from normal healthy skin. It is more effective than the FCM method and less computational time than the mean shift method.

Abbas et al [9] proposed a novel perceptually oriented approach for melanoma border detection by combining region and edge-based segmentation techniques. In their method, the RGB images were transformed to CIE L*a*b* color space, then lesion contrast was enhanced by adjusting and mapping the intensity values of the lesion pixels in the specified range using the three channels of CIE L*a*b*. Therefore, a hill-climbing method was adopted to detect the region-of-interest (ROI) and an adaptive thresholding was applied to determine the optimal lesion border.

Another novel unified approach was proposed by Abbas et al [10] for automatic lesion detection. A pre-processing step was performed by normalised smoothing filter (NSF) to reduce the background noise. Then, the mixture models technique was utilised to initially segment the lesion area roughly. Afterwards, local entropy thresholding was performed to extract the lesion candidate pixels. Finally, the authors applied the morphological reconstruction algorithms to refine the lesion segmentation process.

A new approach for lesion segmentation is proposed by Pennisi et at [117]. Closing operation is used to remove the hair and few outlier pixels. Subsequently, two segmentation processes are implemented in parallel; thereby providing two different images. One image is constructed by detecting the skin region, while the other is created by applying edge detected

through Delaunay Triangulation. Finally, the authors combine these two images to extract the final lesion area.

Bi et al [33] proposed a new automatic melanoma detection method for dermoscopic images via multi-scale lesion-biased representation and joint reverse classification. They attempted to represent the skin lesions using multiple of closely-related histograms derived from different rotations and scales of the image.

After removing the hair and several reflection artifacts from images, the region-growing method was implemented, for the purpose of extracting and segmenting the affected lesion area Jaworek et al [81].

A novel automatic segmentation approach was proposed by Garnavi et al. [65]. The authors used colour space analysis and clustering based on histogram thresholding to find out the optimal colour channel. Therefore, the skin lesion segmentation was performed using well-known Otsu's thresholding method.

Fuzzy c-means (FCM) thresholding technique was used to classify image pixels into two categories for the purpose of obtaining an accurate segmentation of the pigmented lesion [139].

An automatic adaptive threshold (AT) is used by Silveira et al [134]. With respect to a segmentation step, the pixels whose intensity value were greater than the threshold were classified as a region of interest (lesion), while the others pixels were classified as background (healthy skin).

Hwang et al [75] used Gabor filters approach to extract several texture features form dermoscopy images and applied the g-means clustering method [72]. The g-means approach automatically determined the number of clusters and separated the given image into regions, which are homogeneous in texture.

Dermoscopy image segmentation based on texture distinctiveness (TD) is proposed by Jeffrey et al [69]. Their method attempted to capture the dissimilarity between different texture distributions based on TD metric.

Barata et al [27] proposed two systems for melanoma detection in dermoscopy images. The texture and colour features were used in accordance with local and global features. Histogram computation, peak detection and threshold estimation were implemented to obtain an adaptive

threshold and segment the pigmented skin lesion.

A new automated method based on the JSEG algorithm [54] was developed by Celebi et al [45], and used for lesion border detection. The method was started by removing the noise from images as a pre-processing step. Therefore, the lesion boundary was delineated based on the calculation of three approaches including j-images, region-growing, and region merging. For the purpose of refining the final segmentation results, the post-processing stage including removal of the unwanted objects that belong to the background, elimination of the isolated regions, and combining of the remaining regions was performed.

An adaptive filter inspired by Swarm Intelligence (SI) optimization algorithms was developed by Nowak et al[111], for the purpose of detecting pigment network structures. At the beginning of filtration process, the filters (agents) are applied randomly to sections of the image, where each adapts its output based on its neighbours. Agents share information with other agents located in immediate vicinity. This is a new approach to the problem of dermatoscopic structure detection, and the authors claim it is highly flexible, as it can be applied to images without the need of previous pre-processing step. However, the method inherited high computation complexity of the optimisation problems, which renders it very difficult to develop and fine tune. For instance, processing one image can take up to 5 minutes depending on the agent's count, iterations and image size.

Recently, convolutional neural networks have been adopted for the purpose of increasing the segmentation accuracy of medical images. Qi et al [119] applied a Fully Convolutional Neural Network (FCNN) with a pre-trained VGG 16-layer net [95], by replacing all fully connected layers by convolution layers, with randomly initialized weights for skin lesion segmentation. Their pre-trained network eliminated the need for a larger training data. In addition, stochastic gradient descent was used for fine tuning the network.

Furthermore, based on deep learning approaches, a new method for accurate segmentation of skin lesions was proposed by Jafari et al [80]. After the input image was pre-processed by removing unwanted objects, the extraction of color and texture descriptors from images were used to feed the deep Convolutional Neural Network (CNN) and outputting a label for each image pixel. In line with this, the mask referred to as the region of interest (affected lesion)

was extracted from its surrounding skin. As a final step, the mask was refined by applying post processing operations.

Another approach based on the deep convolutional neural network for skin lesion detection was proposed by Yuan e al[154]. Several effective training strategies were implemented to solve the limited data problem during training a deep network. The segmentation performance of the lesion boundaries was improved by the use of techniques such as adam optimization [84], batch normalization [76], and Jacquard index [79] based loss function.

A novel computational method was presented by Oliveira et al [113] for extracting skin lesion features from images based on asymmetry, border, colour and texture analysis, for the purpose of determining skin lesions type. Their approach was based on an anisotropic diffusion filter, an active contour model without edges and a support vector machine to reduce the noise presented in images, segment the lesion borders and perform lesion classification stage respectively.

2.6 Previous Methods Related to Skin Lesion Classification

In recent years, there has been an increasing interest in early detection of skin cancer using CAD systems. Most of these diagnosis techniques are based on the ABCD rule and the 7-point checklist criterion [49] [116] [18] [23] [46] [15]. This section introduces several different methods used to differentiate between melanoma and non-melanoma lesions.

Colour distribution in the RGB colour components (mean RGB distance, maximum distance, and variance) was used by Seidenari et al[132] to distinguish between three types of lesions including melanomas, atypical nevi, and clearly benign nevi.

Relative colour histogram analysis technique was also adopted by Stanley et al [141] to evaluate pigmented skin lesion discrimination based on colour feature calculations in different regions of the skin lesion in dermoscopy images. The purpose of this was to distinguish between melanomas and benign or atypical nevi .

Sheha et al [133] presented an automated method for melanoma diagnosis using a set of dermoscopy images. The authors extracted several features based on gray level Co-occurrence

matrix (GLCM), and they used multilayer perceptron classifier (MLP) to classify between Melanocytic Nevi and Malignant melanoma.

The system proposed by Iyatomi et al[78] is an Internet-based melanoma screening system that uses different features to distinguish between lesions types. A total of 428 image-related objective features with reference to the ABCD rule and differential structures of the lesions are calculated and used as input to the neural network classifier.

Situ et al [135] presented a new method for melanoma detection using a bag-of-features (BoF) approach. They represented each image as a set of several patches sampled from a 16 $\times$ 16 regular grid placed on the pigmented skin lesion. In order to describe each patch, they used wavelets and Gabor-like [131] filters, leading to a total of 23 features. Therefore, two different classifiers (naive Bayes and SVM) were applied and compared.

Sadeghi et al.[126] aim to detect the holes of the network. Three different stages were implemented to detect the holes. First, a Laplacian of Gaussian (LoG) filter is used in order to find meshes or cyclic structures. Second, an 8-connected components analysis is applied to transform the filtered image to a set of graphs. Finally, a search is conducted for loops or cyclic graphs. Then, the distance between the holes is measured and based on this distance a new graph is created. This graph is used to detect the pigment network. The same research group presented a new method for classifying between Absent, Typical and Atypical [127]. They detected the net structure and extracted structural, geometric, chromatic and texture features, and used these features to train a boosting classifier.

Two different techniques structural and spectral are proposed by Betta et al.[31] They started to combine these techniques to perform the detection of pigment network. Firstly, they used structural techniques to search for simple shape structures, like lines or dots. Secondly, the spectral technique is based on a Fourier analysis of the gray-level image. A sequence of Fast Fourier Transform, high-pass filtering, Inverse Fast Fourier Transform and thresholding were applied. Eventually, the mask from each technique is combined together to provide a network image. This study has been updated and proposed by Di Leo et al[55]. They classified the areas which constitute the network image as either atypical or typical, the existence of typical pigment network or the absence of network are both included in the same class. Chromatic

and spatial features such as mean and standard deviation related to the obtained structures were extracted and used to train a decision tree algorithm.

Grana et al[71] Proposed an algorithm to undertake the detection of the pigment network. They used Gaussian derivative kernel to detect the net edges and Fisher linear technique to provide the optimal thresholds. Morphological masks are used to complete the line linking process. However, the experiments which have been performed are not focused on the task of distinguishing between pigment network and no pigment network.

Barata et al[25] proposed a new approach to extract pigment networks from dermoscopy images using a bank of direction filters and many morphological operations. Two distinctive properties: region pigment network intensity and geometry were used, and several features were extracted. Next, an Adaboost algorithm was used to classify the given region as either normal or abnormal.

Based on colour discrimination, Barata et al [26] proposed a method to evaluate the importance of color in the keypoint detection. They applied Harris Laplace detector method [106] and its colour extensions to compare the performance of gray scale with that of colour sampling methods. In every image, square patches were extracted from each interest point, and feature vectors were created using two different types of features, SIFT feature [97] and color-SEFT feature [144] as a patch descriptor. The authors applied the Bag-of-Features (BoF) model [136] to equalise the number of extracted patches from all images. Thus, the decision rule was computed using the k-Nearest Neighbor (kNN) classifier to identify melanoma lesions, and quite promising results were obtained. However, using the square patches to extract features from images requires more computational time and this could be expensive when the patch size is large. Moreover, the authors used only 176 dermoscopy images from a total of 200 images.

Furthermore, a Bag-of-Features (BoF) model for the classification of melanoma in dermoscopy images was implemented by Barata et al [24]. The authors used two different types of local descriptors: colour and texture, and their performance was evaluated separately, and then compared to assess their ability to describe the different dermoscopic features. The same research group presented a new method to classify skin cancer images as either melanoma or

non-melanoma [22]. Colour features and texture features were used based on colour histograms in three different colour (HSV, L*a*b*, and Opponent) and gradient related histogram. All used features were extracted globally and locally from each image. Therefore, the authors investigated the best way to combine the features by applying two strategies (early and late fusion). A Random Forests classifier yielded the best results.

Celebi et al [46] proposed a machine learning method for automated quantification of clinically-significant colours in dermoscopy images. The K-means clustering approach was used to cluster each image with an optimal K value, which was estimated separately using five commonly used cluster validity criteria.

An automatic framework for detection of melanoma from dysplastic nevi was proposed by Rastgoo et al [121]. They combined several extracted features such as colour, shape, size and texture features with well-known texture features such as local binary pattern, grey-level co-occurrence matrix, a histogram of gradients and the Gabor filter. Support Vector Machines (SVM), gradient boosting and random forest methods were used to evaluate the performance of their work.

A new methodology for color identification in dermoscopy images was introduced by Barata et al [23]. The authors used the Gaussian mixtures model to learn a statistical model for five colors (black, dark brown, light brown, blue-gray and white). Therefore, the learned mixtures were used to assess the colors of a larger set of images.

Alfed et al [14] proposed a new method for melanoma diagnosis. They used a bank of direction filters to segment pigment networks from images, then extracted a few features from the segmented image and used Artificial Neural Network (ANN) as a classifier. The same group introduced a new method for improving a bag-of-words approach by combining color histogram features and first order moments with the Histogram of Oriented Gradients (HOG) [13]. Three classifiers methods were used in their work K-Nearest Neighbors (KNN), SVM and AdaBoost, where the SVM achieved the best results.

The detection of melanoma lesions in dermoscopy images was addressed by Marques et al [99]. The authors used two types of features to represent the image including color and texture features. The colour features consist of an information obtained from a number of colors

(L*a*b) and its distribution in the pigmented skin lesion. While the texture features contain information about the differential dermoscopic structures such as dots, globules, streaks, and pigment network which are appeared in the lesion. The multiple region approaches was applied to obtained two boarder regions close to the boundary (inner region and boarder region). Therefore, the colour and texture histograms were computed for the two regions separately, and their outcomes were concentrated to obtain an extended feature vector to represent the image. A classifier was trained and detected melanoma lesions with high sensitivity. However, not all dataset images were used to perform the experiments; indeed, the authors selected only 163 cases from a total of 200 images.

Riaz et al [122] also used the color distribution in the skin lesions and the texture as fundamental visual characteristics to represent the image. In terms of texture features, the scale adaptive patterns were extracted at each pixel in the image using the strength of the Local Binary Pattern (LBP) [112], followed by constructing a histogram. The standard histogram of the HSV colour space was applied, in order to acquire a colour feature vector. To be able to represent the image, the extracted features (color and texture) were combined to form a feature vector for an image. The differentiation between melanoma and non-melanoma lesions was performed by applying the support vector machines as a classifier algorithm.

Four colour constancy algorithms including GrayWorld [35], max-RGB [91], Shades of Gray [62], and General Gray World [145] were investigated by Barata et al [21], based on the assumption of correcting the colour variation using the colour constancy algorithm [68] to identify melanoma lesions using multisource skin images. The normalisation of the RGB color component was performed using the four mentioned algorithms for the purpose of estimating the color of the illuminant. The authors used SIFT features [97] algorithm to extract several local features from each patch in the image. Therefore, five different BoF systems were trained and tested, including one that refers to the non-normalised image (original) and four indicating normalised images, using the k-means approach to create visual words, and the SVM method to classify the histogram of visual words. The obtained results from the Shades of Gray algorithm were higher than those from the non-normalised images. However, The manual segmentation (ground truth images) performed by the dermatologist was used to

prevent an incorrect segmentation of the pigmented region, and increase the accuracy of lesion classification.

A novel computational method was presented by Oliveira et al [113] for extracting skin lesion features from images based on asymmetry, border, colour and texture analysis, for the purpose of determining skin lesions type. Their approach was based on an anisotropic diffusion filter [28], an active contour model without edges [47] and a support vector machine [37] to reduce the noise presented in images, segment the lesion borders and perform lesion classification stage respectively.

Another work presented by Yu et al [153] adopted residual learning techniques to train deep neural networks. Thus, they designed a Fully Convolutional Residual Network (FCRN) comprising more than 50 layers for both the segmentation and the classification phases to obtain an accurate skin lesion recognition. They combined their proposed FCRN (used for segmentation) and other very deep residual networks (used for classification) to form a two-stage framework. This framework enabled the classification network to extract more specific and representative features based on segmented results instead of the whole images. The support vector machine (SVM) and the Softmax classifiers were used to obtain two predictions of the lesions and their average was performed to acquire the final results.

Yang et al [152] also presented a new approach to perform the segmentation and the classification of the given skin lesions, by using a multi-task deep neural network. Their network consisted of three components for the robust analysis of the skin dataset including segmentation of the lesion boundaries, categorization of dermoscopic data into melanoma or non-melanoma, and detection of features to categorize the lesions into melanoma or non-melanoma.

To improve the quality of existing diagnostic systems, provide meaningful information such as actual lesions boundaries, and assist clinicians who are not fully experienced in dermoscopy images observation, and make a right decision for final disease diagnosis in less time, we propose four main methods for increasing the accuracy of lesion segmentation and classification. The process related for each method is presented in one chapter of the next four chapters.

Chapter 3

Segmentation of Pigmented Regions in Dermoscopy Images

3.1 Introduction

Correct segmentation of the lesion area in dermoscopy images is fundamental and essential for increasing the effectiveness of the subsequent stages, such as feature extraction and lesion classification, because it strongly affects their results. Segmented lesions allow dermatologists to perform large population vision screening exams for the early detection of melanoma diseases and treatment evaluation. An accurate segmentation of the pigmented regions plays a vital role in identifying melanoma disease properly. It derives border structure information, such as asymmetry, diameter and border irregularity of the lesion area, which are essential in melanoma diagnosis. Moreover, important features such as globules, blue-white areas and atypical pigment network, can be extracted only when high accuracy detection of the lesion border is achieved [60]. Consequently, lesion segmentation phase is the most critical step in identifying melanoma at an early stage. However, lesion border detection in dermoscopy images is a challenging task, since the presence of noise such as air bubbles, reflections, skin lines and hairs in the images cause huge errors in the segmentation. In addition to the noise, the intensity inhomogeneity and the low contrast of the lesion borders cause a significant

degradation to the performance of automated skin segmentation techniques. The intensity inhomogeneity of dermoscopy images is attributed generally to the acquisition of the image under different situations of illumination. Thus, a robust and reliable automated method for skin region detection is essential in computer-aided diagnosis.

This chapter presents a new automatic method for removing the noise and detecting the affected lesions on dermoscopic images. In the first method, the image is pre-processed for the purpose of removing an unwanted object (noise) and enhancing its quality. Moreover, it facilitates the segmentation process and enhances the accuracy of the final results. The method is based on applying four approaches to remove most of the noise from images. A sample threshold is used to detect the lighting reflection, and the Gabor filters using a bank of 64 directional filters with image sharpening and Sobel filter are implemented to detect and extract the hair. Image inpainting is curried out to fill in the unknown regions by replacing their pixels. In terms of the second method and for the purpose of delineating the affected lesion, the Fuzzy C-Means (FCM) is implemented, in order to provide the mean and the variance of each cluster, which can be used as initial parameters for the Markov Random Field (MRF) approach. The MRF classify the image pixels as either skin lesion (foreground) or healthy skin (background) based on the minimum cost of the energy function. The process repeated labeling the pixels to the clusters based on the highest probability (minimize the energy function), until stopping criteria are met. As a final step, a morphological operation is performed, for the purpose of filling in the holes inside the segmented region. Upon comparison, the proposed method provided good performance in achieving automatic image segmentation over dermoscopy images. This chapter is structured as follows: Section 3.2 provides an overview of the FCM and MRF methods. The proposed method in details is presented in Section 3.3. Section 3.4 illustrates the results and discussions. Finally, a summary of the chapter is presented in Section 3.5.

3.2 Overview of the FCM and MRF Methods

3.2.1 The FCM Method

The basic idea of the FCM method is to find the centre of each cluster iteratively by adjusting their position, and evaluating an objective function [42]. This process is more flexible, since partial membership can be introduced to other clusters. The method optimises the process by minimising the cost function.

$$F = \sum_{J=1}^{N} \sum_{i=1}^{c} u_{ij}^{m} \parallel xj - vi \parallel^{2}$$

where N number of pixels, c number of clusters, u_{ij} refers to the degree of membership of pixel xj in the ith cluster, vi represent the cluster centre, and m indicates the fuzzifier. The membership function and cluster centre are iteratively updated by:

$$uij = \frac{1}{\sum_{k-1}^{c} (\frac{\|xj - vi\|}{\|xj - vk\|})^{2/(m-1)}}$$

$$vi = \frac{\sum_{j=1}^{N} u_{ij}^{m} xj}{\sum_{j=1}^{N} u_{ij}^{m}}$$

3.2.2 The MRF Method

The MRF method is a statistical model, but can be used for segmentation methods, it was introduced in image analysis by Geman and Geman [66]. MRF theory provides a tool for modelling a vision problem within the Bayes framework using spatial continuity. The image pixels are indexed by a rectangular patch S and each image pixel s is characterised by the grey level y_s from the set $y = y_s : s \in S$. The labelling process consists of accurately labelling each image pixel $s \in S$ with a class label representing the pattern class in the image. A label set is defined as$\Lambda = 1, 2, ..C$ where C is the number of classes. A labelling is indicated by $x = x_s : x_s \in \Lambda, s \in S$ where $x_s = l$ denotes that the class label l is assigned to the pixel s. The goal is to discover the labelling $\hat{x}$ of the image, which is the estimation of the true but

unknown labelling x^*. According to the MAP estimate [66], we have

$$\hat{x} = argmax P(x|y)$$

According to Bayes' theorem we have

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

where $P(x)$ is the prior density of the labelling x and $P(y|x)$ is the conditional probability density of the image y. The prior probability of the image $P(y)$ is independent of the labeling x; therefore, it is rewritten as follows:

$$\hat{x} = argmax\{P(y|x).P(x)\}$$

The Gaussian distribution is used with the assumption of the existence of Gaussian noise in images. Therefore, the possibility of pixel s with the assumption of belonging to class x_s is equal to y_s and can be calculated as follows:

$$P(y_s|x_s) = \frac{1}{\sqrt{2\pi\sigma_{xs}^2}} \exp\{-\frac{(y_s - \mu_{xs})^2}{2\sigma_{xs}^2}\}$$

Based on the conditional independence assumption of y, the conditional density $P(y|x)$ takes the form of

$$P(y|x) = \prod_{s \in S} P(y_s|x_s)$$

Therefore, $P(y|x)$ can be written as follows.

$$P(y|x) = \prod_{s \in S} [\frac{1}{\sqrt{2\pi}} \exp(-\frac{(y_s - \mu_{xs})^2}{2\sigma_{xs}^2} - \log(\sigma_{xs}))]$$

This can be written as

$$P(y|x) = \frac{1}{(2\pi)^{\frac{s}{2}}} \exp[-\sum_{s \in S} (\frac{(y_s - \mu_{xs})^2}{2\sigma_{xs}^2} + \log(\sigma_{xs}))] \tag{3.2.1}$$

where μ_{xs} and σ_{xs} indicate the mean and variance of the class x_s respectively.

The previous model is based on 2D MRF and assumes that the adjacent pixels have the same class label. The Hammersley-Clifford theorem [30] establishes a relation between the MRF and Gibbs distribution. According to this theorem, the previous model $P(x)$ is given by a Gibbs distribution with respect to the neighbourhood system N, and it takes the form of

$$P(x) = \frac{1}{Z} \exp[-\sum_{c \in C} V_c(x)] \tag{3.2.2}$$

where Z is the normalisation constant or partition function, $V_c(x)$ is the potential function for clique c and C is the set of all cliques in the image. According to the equations (3.2.1) and (3.2.2), we can rewrite the posterior probability as:

$$P(x|y) =\propto exp[-U(x)]$$

where the energy function $U(x)$ has the form

$$U(x) = [\sum_{s \in S} \frac{(y_s - \mu_{xs})^2}{2\sigma_{xs}^2} + \sum_{s \in S} \log(\sigma_{xs}) + \sum_{c \in C} V_c(x)] \tag{3.2.3}$$

3.3 Proposed Method

The proposed approach is divided into two steps, image pre-processing and segmentation. The pre-processing stage includes reflection artifact, hairs detection and removal. The Directional Gabor filters, image sharpening, Sobel filter with image inpainting methods are implemented to extract most of the hair, and a threshold algorithm is performed to detect reflection artifact from images. While the segmentation step includes applying the FCM method on dermoscopy images in order to get two benefits: estimate the initial parameters and segment the lesion area, then implementing the MRF method based on the previous parameters. The MRF iterate assigning each pixel to a cluster based on minimising the energy function, until there are no more changes to the label. Therefore, the segmented image is refined by applying a morphological operation. The scheme of lesion detection is illustrated in Figure 3.1, and full

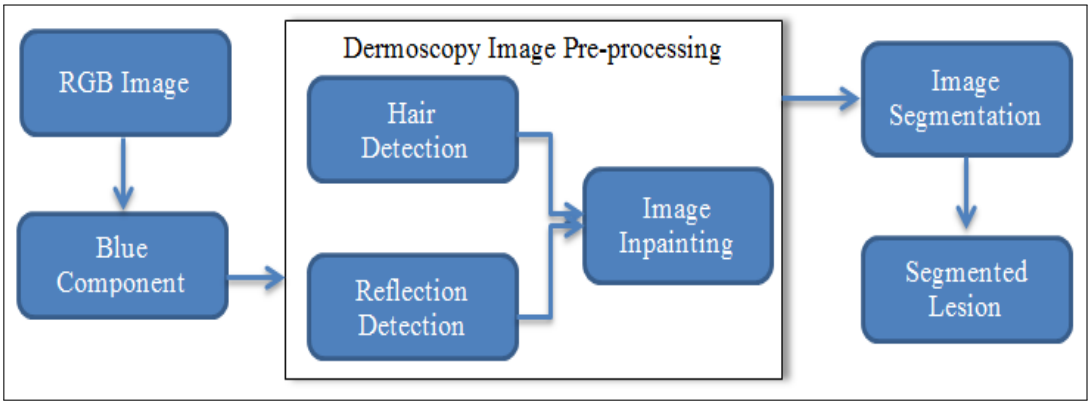


Figure 3.1: Diagram of melanoma detection system.

details are provided in the following section.

3.3.1 Dermoscopic Image Pre-processing

The purpose of image pre-processing is to improve the image data by removing unwanted distortions from the image or enhancing image features, which can be used as input to other image processing techniques. Pre-processing is an important step in medical image analysis, in order to avoid the negative effect of noise in extracting the object. It also facilitates the segmentation and the classification processes. Typically, dermoscopic images do not have the expected quality to perform the diagnostic analysis. Thus, the step of image pre-processing is very important and necessary, in order to reduce the number of artifacts and noise from images. Dark hair, skin lines, oil, air bubbles and lightening reflection caused by placing gel or oil before capturing the image are present in almost every image (see Figure 3.2), which in turn affects the segmentation step; thereby leading to the wrong diagnosis of the disease. Our process of image enhancement involves two key operations: hairs, reflection artifact detection and removal. According to [134], the blue channel of the colour images is selected, as it has been proven experimentally to provide the best discrimination in most dermoscopy images.

3.3.1.1 Reflection Detection

Reflection artifacts and air bubbles appear like noise in dermoscopy images due to the placement of oil or gel on the lesion area before capturing the image (see Figure 3.2). A simple

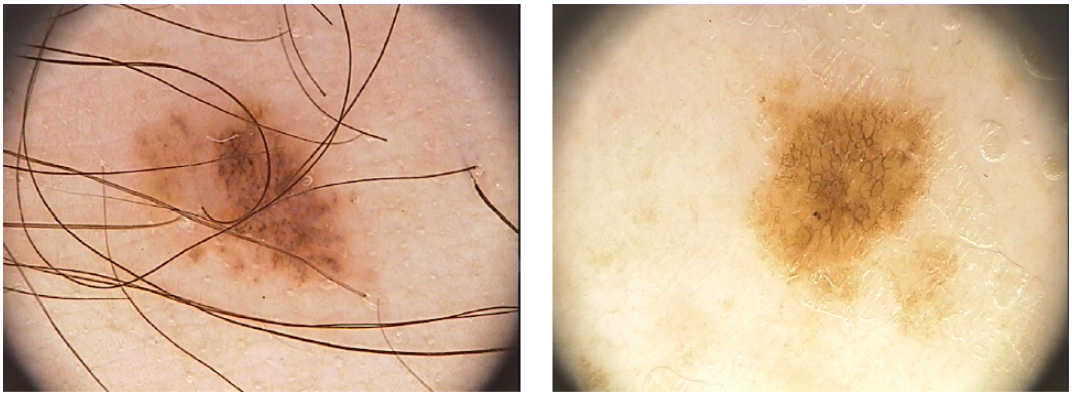


Figure 3.2: Examples of two dermoscopic images show different types of noise. Left: presence of hairs. Right: presence of gel and reflection artifacts. (Source:[101])

threshold algorithm is applied to detect this type of noise. Every pixel (x, y) can be detected and classified as a reflection artifact if its intensity value is higher than threshold T_{R1} and if its intensity value minus the average intensity $I_{avg}(x, y)$ of its surrounding neighbourhood is higher than threshold T_{R2}, *i.e.*

$$\{I(x,y) > T_{R1}\} and \{(I(x,y) - I_{avq}(x,y)) > T_{R2}\}.$$

where I is the image, $I_{avg}(x, y)$ is the average intensity value in a local neighbourhood of the selected pixel, which is computed using a local mean filter with dimensions 11x11 and T_{R1}=0.7, T_{R2}=0.098. An example of lightning reflection detection is illustrated in Figure 3.3.

3.3.1.2 Hair Detection

Most of the dermoscopic images contain hairs, which could subsequently affect the outer borders of the lesion area during the segmentation step, since their shapes are largely similar. Thus, this leads to the wrong detection of the lesion or make the borders invisible. Based on our knowledge, median filter, adaptive threshold and morphological operations such as Top Hat filter (Opening and Closing image) were widely used for this purpose [41] [125] [65] [139].

In our approach and with the purpose of enhancing the image quality, the Directional Gabor filters are implemented to extract hair artifacts from dermoscopic images. However, the parameters used in the Gaussian filters at each stage are different [25]. A bank of 64

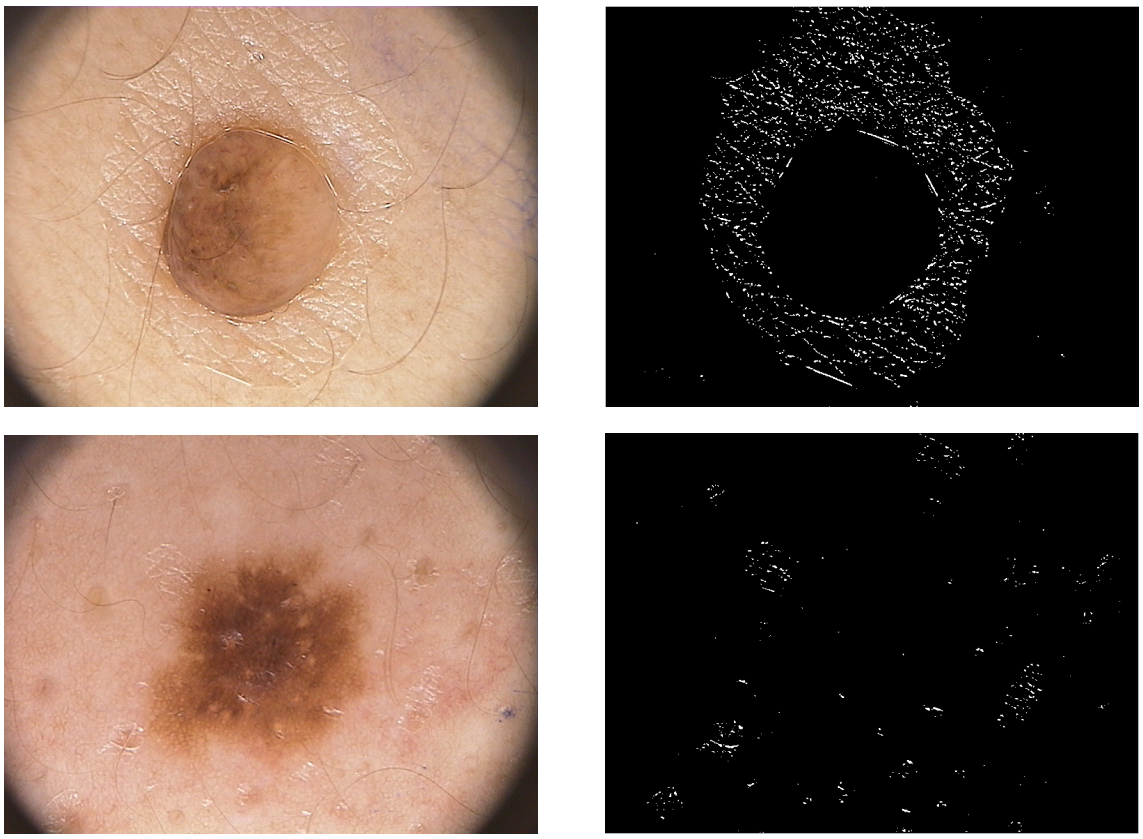


Figure 3.3: Two examples illustrate the reflection artifact detection. original image (Source:[101])(first column) and reflection detection masks (second column).

directional filters has been used to perform hair detection. Let

$$h\theta_i(x,y) = G_1(x,y) - G_2(x,y). \tag{3.3.1}$$

be the impulse of a directional filter with angle θ_i where $\theta_i \in [0,\pi]$, i=$0\cdots$ 64 and G_k is a Gaussian filter:

$$G_K(x,y) = C_K \exp\{-\frac{x'^2}{2\sigma_{x_k}^2} - \frac{y'^2}{2\sigma_{y_k}^2}\}, k = 1,2. \tag{3.3.2}$$

The difference between two consecutive filters θ_i and θ_{i+1} is constant and equal to $\frac{\pi}{N}$. Therefore, $N+1$ filter are used. In (3.3.2), C_K is a normalization constant and the values of (x', y') are dependent on or related with (x, y) by a rotation of amplitude θ_i.

$$x' = x\cos\theta_i + y\sin\theta_i.$$

$$y' = y\cos\theta_i - x\sin\theta_i.$$

Parameters value have been obtained experimentally and set as: σ_{x1} =20, σ_{y1}=6, σ_{x2}=20 and σ_{y2}=0.5. The size of the mask filter is 41 x 41. The difference of Gaussians in Equation (3.3.1) is used because it ensures greater enhancement of directional structures while removing the background effect. The image I is filtered by each directional filter. The output of the ith directional filter is given by the following convolution:

$$I_i(x,y) = h\theta_i(x,y) * I(x,y).$$

After filtering the image by each directional filter, the combination of the output of $N+1$ directional filters are estimated, and the maximum output at each pixel (x, y) is selected according to the following equation:

$$J(x,y) = max_{i\in\{1,2,..N+1\}}(I_i(x,y)).$$

where J is the final image obtained from Gabor filter, this image contains both: hairs and

pigment network lines, since their shapes are very similar.

Adjusting the threshold is tricky. Therefore, we used a different approach than that used by Barata et al[25], in order to be able to extract vast majority of the hair from images without extraction the lines of pigment network. Image sharpening is applied to the outcome image from Gabor filter, to increase the contrast between different colours, and sharpen the transition from black to white, where all edges become sharp. Then, we compute the gradient vector at each point in the sharpened image using Sobel masks. A pair of 3×3 convolution kernels are used to respond maximally to edges running vertically (Gx) and horizontally (Gy) relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels were applied separately to the sharpened image, in order to produce separate measurements of the gradient component in each orientation (Gx) and Gy). Then we combined them together to find the absolute magnitude of the gradient at each point using the following equation.

$$|G| = \sqrt{(Gx^2 + Gy^2)}$$

Sobel filter is applied to the sharpened image with a threshold of 0.0145. This process ignores all edges that are weaker than the threshold value, and detects the remaining edges as hair. Thus, most of the hair is detected and classified as a noise. Figure 3.4 illustrates the hair detection results by applying the proposed approach algorithm.

It is possible that the algorithm performs well in terms of hair detection and achieves reasonable results. However, it failed to identify small hairs in few cases, as illustrated in Figure 3.5. This is because the hair is quite thin and difficult to detect. In addition, in many cases, the hair tends to be grey in colour; thereby rendering it more challenging to identify, especially if present in low contrast images.

The outcomes of the previous steps are two binary masks for each image: one indicates the reflection artifacts mask, and the other refers to the hair. The OR gate is used to match these two masks, in order to obtain one mask for each image. Therefore, the morphological operation such as dilation is implemented with a disk of radius r, which was empirically chosen to be 3. The reason of that was to allow a simple dilation of hair lines, as well as guaranteeing

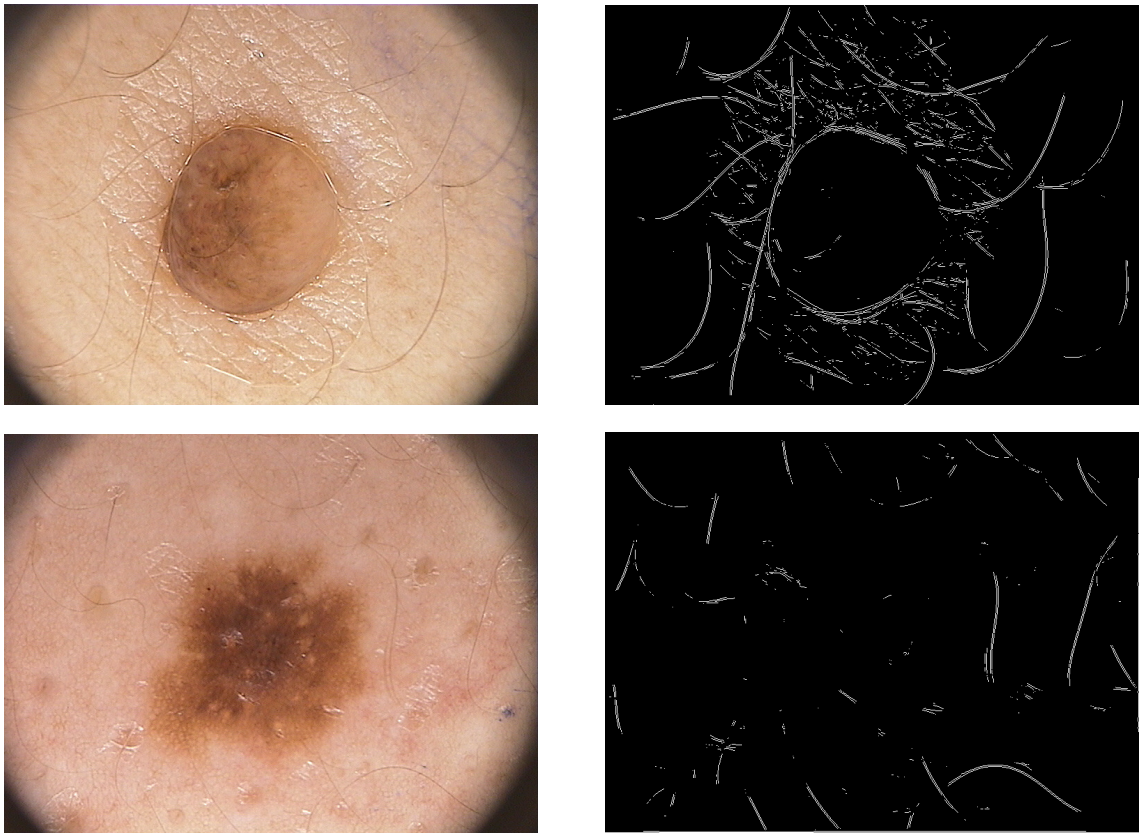


Figure 3.4: Two examples show the hair detection stage. original image (Source:[101])(first column) and hair detection masks (second column).

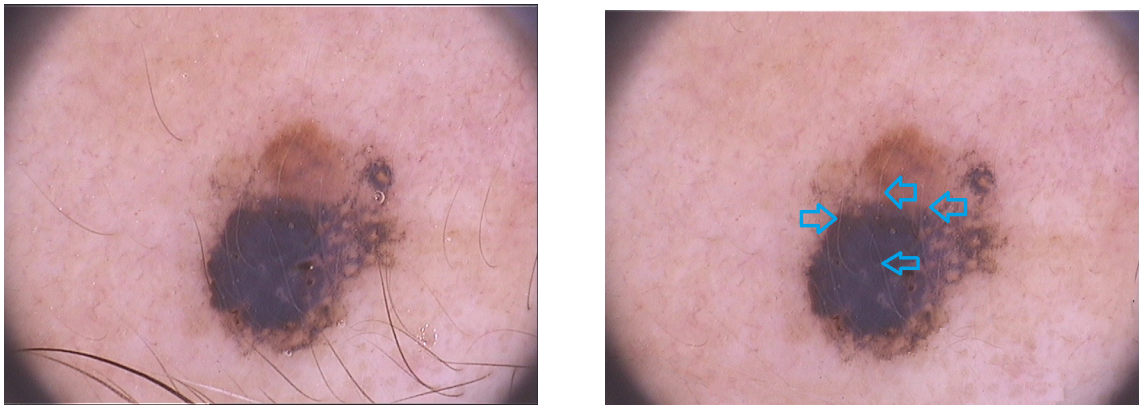


Figure 3.5: Example refers to the fail hair detection. Left: original image (Source:[101]). Right: presence of thin hair in cleaned (pre-processed) image.

the majority of the hair is removed from images. This step will increase the accuracy of hair removed and increase the performance of the image inpainting stage in the next step.

3.3.1.3 Image Inpainting

Image inpainting is a technique that provides a means for reconstruction of damaged area from the image, or fills its missing regions. Thus, it works by removing undesirable objects from images by using their surrounding information. [29]. Most of the inpainting techniques works as:

- Select the regions that will be inpainted (usually manually).
- Propagate the colour information inward from region borders (known regions are used to fill in the unknown area.)

Based on the concept of image inpainting technique, we could use it as an approach to perform hair removal in dermoscopy images. After the reflection artifacts and hair are detected and dilated, their binary masks are multiplied and matched with the original images. This step leads to the appearance of gaps, which indicate the unknown regions in the images. As can be seen in Figure 3.6, the unknown regions are brighter than their surroundings. Therefore, these regions could be filled by propagating the information from surrounding neighbourhood (known regions). In each image, the brighter pixels (unknown regions) are replaced by a new value obtained from its neighbours. Patch priorities (data term and confidence term) are computed in the borders of the unknown regions based on their neighbours. The patch with the highest priority is filled out with data extracted from the source region, and then the patch priorities are updated. This process continues until no more gaps exist [52]. Patch priority $P(p)$ is defined as the product of two terms:

$$P(p) = C(p)D(p)$$

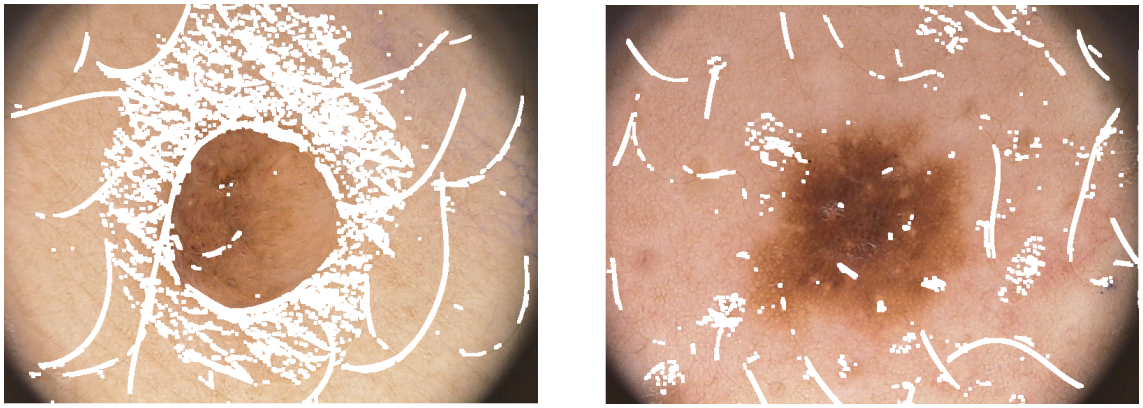


Figure 3.6: Two examples indicate the appearance of the gaps in the images. The pixels value of the hair, lightening reflection and gel were replaced with 255 based on their masks.

where $C(p)$ call the confidence term and $D(p)$ the data term, and they are defined as follows:

$$C(p) = \frac{\Sigma_{q \in \Psi p \bigcap \bar{\Omega}} C(q)}{|\Psi_p|} \quad , \quad D(p) = \frac{|\nabla I_{\frac{\perp}{p}}.n_p|}{\alpha}$$

where $|\Psi_p|$ is the area of Ψ_p , α is a normalisation factor ($e.g., \alpha = 255$ for a grey level image), $\delta\Omega$ is a target region border, and n_p is a unit vector orthogonal to the front $\delta\Omega$ in p point (perpendicular). The priority of every border patch is computed with distinct patches for each pixel on the boundary of the target lesion. More details of the inpainting method are provided in Appendix B.

Three examples are illustrated in Figure.3.7. It can observed that the inpainting method can remove the hair by replacing its pixel value with the new colour value obtained from its neighbours. The final images obtained are clean and will be used in the subsequent step.

3.3.2 Skin Lesion Segmentation

The segmentation stage is one of the most important and challenging steps in image processing. It must be fast and accurate, because the subsequent steps such as feature extraction and classification are largely dependent on its performance. As mentioned previously, segmentation process for dermoscopy images is extremely difficult, due to the existence of several factors such as the low contrast between the lesion and the healthy skin, variance of colours

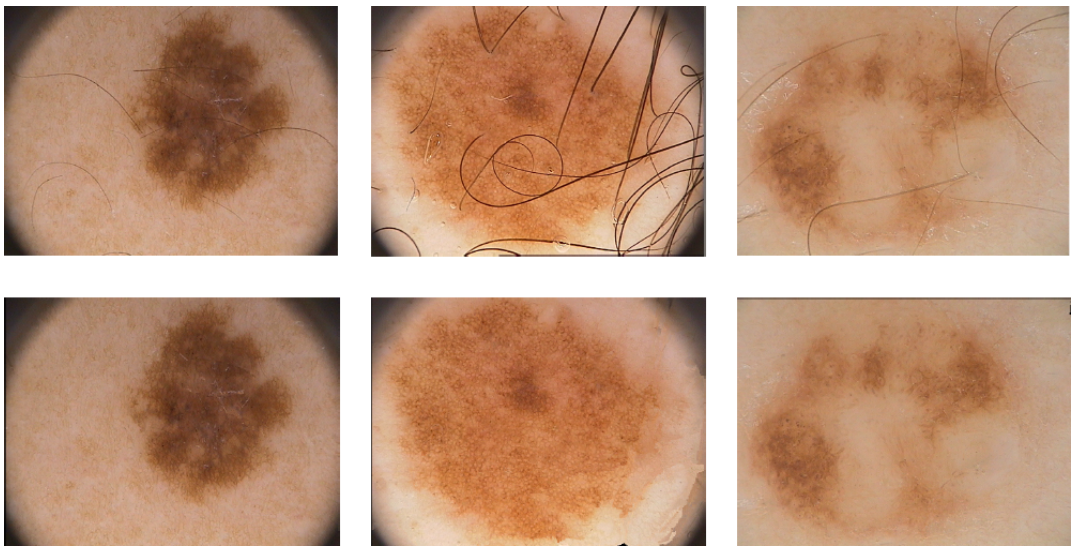


Figure 3.7: Examples of applying inpainting process: original dermoscopic images (Source:[101]) (first row) and inpainted image (second row)

inside the pigmented region and other artifacts. However, it is very helpful to dermatologists because information can be acquired, such as asymmetry, border irregularity, colour and diameter, which play an essential role in melanoma diagnosis. Therefore, by applying a suitable segmentation method to delineate the whole lesion areas from images will be of great benefit in the diagnostic process. Accordingly, the FCM and the MRF were incorporated to perform the final segmentation of all images.

The FCM method is used to initiate the segmentation process. The pixels of an input image is divided into two clusters: cancerous pixels as the foreground (ROI) and normal skin pixels as the background. The major drawback with the FCM method is that it address pixels individually only by their intensity values; thus, it lacks the capacity to model the overall appearance of a local neighbourhood region. The MRF method is implemented to refine the previous segmentation of the image to address this issue. The main goal of the image segmentation using the MRF method is to minimise the energy function or maximise the probability of pixel allocation to a cluster by using Maximum A Post Priority (MAP) equation (3.2.2). The iterated conditional modes (ICM) method is performed to minimise the energy function. In our work, we assume that one pixel has 8-neighbours. Then, the clique

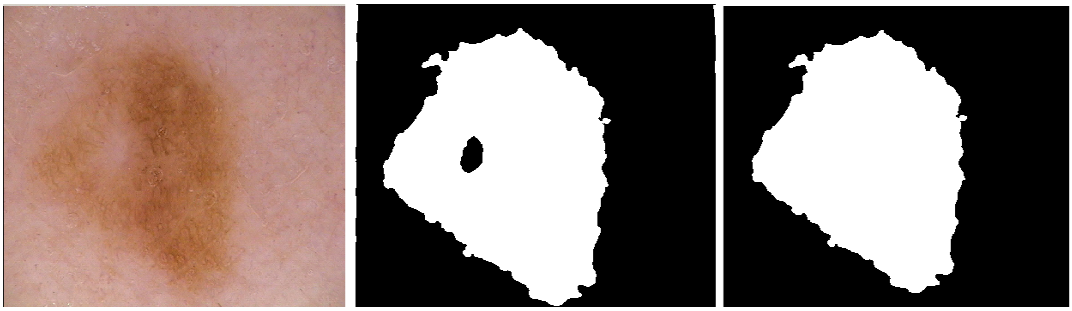


Figure 3.8: Examples of fill in the holes inside the segmented lesion in dermoscopy images: pre-processed image (first column), holes inside the segmented lesion (second column) and final segmented lesion (third column).

potential is defined on pairs of neighbouring pixels:

$$V_c(xi, xj) = (1 - i_{xi,xj})$$

where $i_{xi,xj} = 0$ if $xi \neq xj$ and 1 if $xi = xj$.

We obtained the mean and the variance, which can be used as initial parameters for the MRF method to refine the image segmentation. The MRF method is then iterated as follows. Each pixel is assigned to a cluster based on the highest probability $P(y|x)$ using equation (3.2.1). A calculation is then made of the previous probability using equation (3.2.2). All pixels will be assigned again to different classes by obtaining the minimum cost, equation (3.2.3). This process continues until there is no further change between clusters.

Typically, many parts belonging to the region of interest (lesion area) are classified as background (healthy skin), as indicated in Figure 3.8, which appear as holes inside the segmented lesion. This, subsequently, led to under-segmentation. Therefore, to address this issue and obtain a clear segmentation of the complete regions, a morphological operation such as filling, which is a Matlab function, was used to fill in these small holes [138]. This process allowed the missing pixels of the detected lesion area to be filled in. Moreover, the segmented lesion appeared as one connected region, as presented in 3.8.

3.4 Results and Discussions

The proposed approach was tested on the PH2 dataset [101] which provides 200 RGB dermoscopic images. The images are divided into benign lesions (80 common mole and 80 atypical nevi) and malignant lesions (40 melanoma images). Manually segmented images were also available and used as ground truth. Four different criteria are used to evaluate the performance of our proposed method; namely, sensitivity (SE), specificity (SP), accuracy (AC) and dice similarity coefficient (DSC) [90]. These measurements criteria are used widely in the literature to measure the performance of skin lesion segmentation and they are defined as:

$$SE = \frac{TP}{(TP + FN)}$$

$$SP = \frac{TN}{(TN + FP)}$$

$$AC = \frac{(TP + FN)}{(TN + TP + FN + FP)}$$

where TP, TN, FP and FN indicate to true positive, true negative, false positive and false negative respectively. TP represents the number of pixels which are part of the lesion that classified correctly by both the system and the expert dermatologists. TN represents the number of pixels which are part of the background skin that classified correctly by both the system and the expert dermatologists. FP represents the number of pixels which are classified as a part of the lesion by the system but labelled as a part of the background by the expert. Finally, FN represents the number of pixels which are classified as a part of the background by the system but labelled as a part of the lesion by the dermatologists.

With respect to the dice similarity coefficient (DSC), it is defined as a measure of overlap between the true and the estimated classes, and can be calculated by the following formulation:

$$DSC = 2\frac{(PPV_i * TP)}{(PPV_i + TP)}$$

where PPV_i indicates the *Positive Predictive Value* and can be calculated using the following

equation:

$$PPV_i = \frac{TP}{(TP + FP)}$$

Segmentation of skin cancer was conducted using the FCM and the proposed method. Therefore, a comparison of these methods was undertaken with the lesions obtained by an expert dermatologist, in order to evaluate the performance of the proposed method. Figure 3.9 illustrates the results of the manual segmentation by expert dermatologists together with the results of two methods. In general, the proposed approach has the best performance in terms of the accuracy. For instance, in the second row, the FCM method did not detect the whole lesion area, which leads to misclassification. In addition, the edges of the test images are not close to the real boundaries of the lesions, which means part of the lesion was classified as background. Images in the third row show the results obtained by our method. We can observe that the obtained edges are closer to the boundaries of the skin lesions and their shape is similar to that of the ground truth (first row) compared with the previous method.

Although the quantitative comparison between various methods is difficult because different datasets and criteria have been used. In addition to the visual observations, we were able to perform the quantitative evaluation by comparing the performance of our method against selection of six well-known segmentation methods; namely, FCM [32], J-image segmentation (JSEG) [54], Statistical Region Merging (SRM) [110], Otsu [114], Level Set [51] and the Automatic Skin Lesion Method (ASLM) [117], which have been already considered for dermoscopy image segmentation, as they have used the same dataset and evaluation criteria. The experimental results of the total dermoscopy images are presented in Table 3.1. The performance results of J-image segmentation (JSEG) [54], Statistical Region Merging (SRM) [110], Otsu [114], Level Set [51] and an Automatic Skin Lesion Method (ASLM) [117] were generated from their original manuscripts, while the performance results of FCM [32] was obtained by our implementation. The performance of the dierent methods was generated using all 200 dermoscopic images. Our proposed method has the highest average sensitivity = 0.9320, specificity = 0.9800, accuracy = 0.9400 and with an average dice similarity coefficient of = 0.9105, it out-

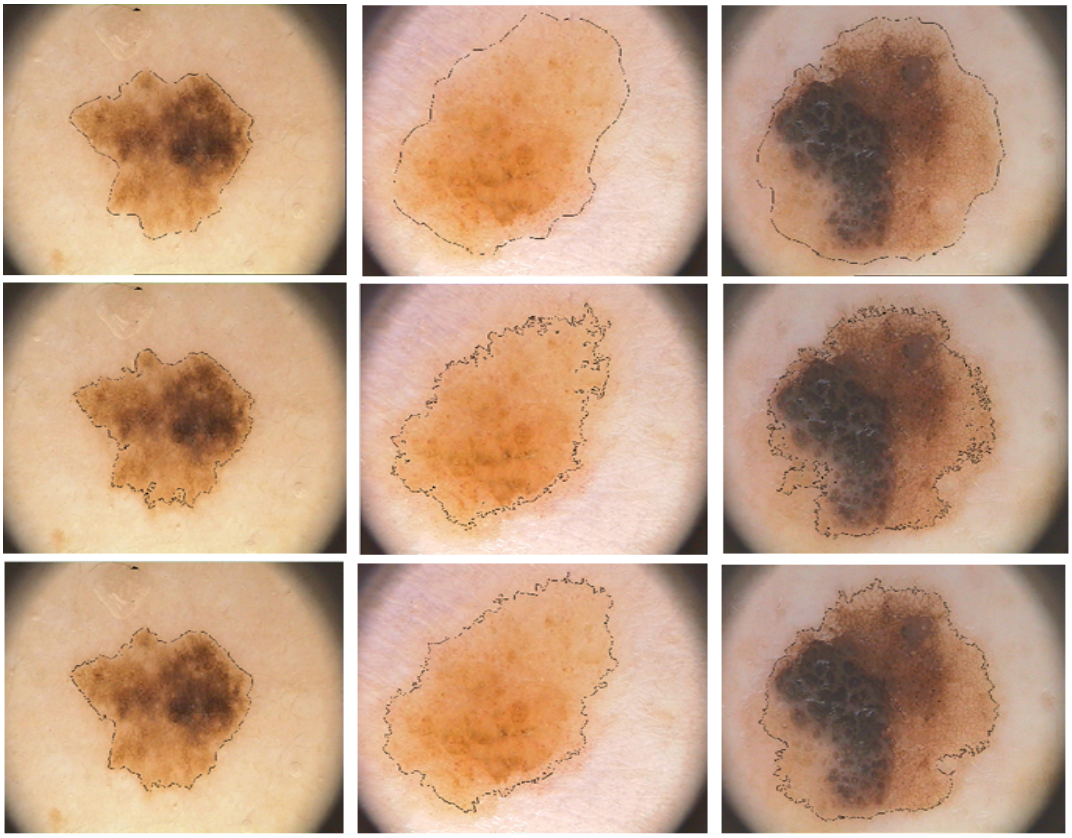


Figure 3.9: Results of lesion segmentation in dermoscopic images: lesion delineation derived from ground truth (first row), results derived from applying the FCM (second row) and the proposed method outcomes (third row).

Table 3.1: Segmentation performance on the complete dataset. SE-sensitivity, SP-specificity, AC-accuracy and DSC-Dice similarity coefficient.

Method	SE	SP	AC	DSC
FCM [32]	0.8800	0.9510	0.9200	0.9040
JSEG [54]	0.7108	0.9714	0.8947	0.7554
SRM [110]	0.1035	0.8757	0.6766	0.1218
Otsu [114]	0.5221	0.7064	0.6518	0.4293
Level Set [51]	0.7188	0.8003	0.7842	0.6456
ASLM [117]	0.8024	0.9722	0.8966	0.8257
Proposed Method	**0.9320**	**0.9800**	**0.9400**	**0.9105**

performs all other methods in terms of all evaluation metrics. For example and in comparison with the results obtained by applying the FCM, the sensitivity increases form 0.8800 to 0.9320, the specificity raises from 0.9510 to 0.9800, the accuracy increases from 0.9200 to 0.9400 and the dice similarity coefficient becomes 0.9105 as opposed to 0.9040. We have observed that the same increases can be produced if the comparison with the JSEG or the ASLM methods was performed, since their results are quite close to the FCM's results. Thus, we can say that the method performed well and achieved promising results in segmenting the affected lesion areas.

As mentioned above, the images are assigned according to their medical diagnosis, Thus, we performed a finer analysis by implementing the validation of the three diagnostic classes (common moles, atypical nevi and melanomas) separately. Table 3.2 illustrates the experimental results of 80 common mole images only. It can be seen that, for our method, the sensitivity raises from 0.9320 to 0.9531, the accuracy increases form 0.9400 to 0.9680. and the dice similarity coefficient raises from 0.9105 to 0.9242. In terms of the accuracy, the proposed method results is higher than the performance results of FCM [32], JSEG [54], SRM [110], Otsu [114], Level Set [51] and ASLM [117] methods, it increases approximately 0.10%, 3.30%, 33.51%, 40.06%, 21.06% and 2.14% respectively. We also observed that the specificity ratio of our method is slightly less than the one obtained by the FCM [32] method; meanwhile, the sensitivity is the highest compared with the alternative methods that, in turn, refer to the segmented lesions. An overall of the proposed results, the performance of the method is still

Table 3.2: Segmentation performance on 80 common healthy images. SE-sensitivity, SP-specificity, AC-accuracy and DSC-Dice similarity coefficient.

Method	SE	SP	AC	DSC
FCM [32]	0.9182	**0.9835**	0.9670	0.9083
JSEG [54]	0.6977	0.9783	0.9370	0.7265
SRM [110]	0.0751	0.9332	0.7250	0.0611
Otsu [114]	0.4777	0.7832	0.6911	0.3658
Level Set [51]	0.7069	0.8262	0.7996	0.5856
ASLM [117]	0.8717	0.9760	0.9477	0.8690
Proposed Method	**0.9531**	0.9783	**0.9680**	**0.9242**

the highest against the other methods in terms of almost all evaluation metrics. This indicates that the proposed approach achieves higher results in segmentation images of common moles.

The segmentation performance of 80 atypical nevi images only is shown in Table 3.3. It is also very clear that the proposed approach outperformed the alternative methods in all used criteria. in particular, the rise from 0.9320 to 0.9457, 0.9400 to 0.9631 and 0.9105 to 0.9295 of the sensitivity, accuracy and dice similarity coefficient respectively. In addition, we observed that our implemented algorithm (FCM [32]) is the only comparable method, since its results are close to the current method results. With respect to the four criteria including sensitivity, specificity, accuracy and dice similarity coefficient, the percentage of increase obtained were 3.09%, 0.43%, 0.38% and 0.39% respectively. Moreover, the same raises can be produced if the comparison with the JSEG or the ASLM methods was performed, since their results are quite close to the FCM's results. As for the specificity, it was slightly decreased compared to the one presented in Table 3.1 and 3.2, but it was higher than other methods that we compared to. Consequently, even in the case of atypical nevus, which is considered as benign lesions, the proposed approach continues to obtain higher segmentation results.

Conversely, a huge decrease can be observed in the quality of segmentation results when only melanoma images (40 images) are processed. The overall segmentation performance results of 40 melanoma images are presented in Table 3.4. In particular, the proposed method presents a large decrease in terms of three criteria: the sensitivity decreases to 0.8203 compared with that presented in Table 3.1 (0.9320), the accuracy becomes 0.8389 as opposed

Table 3.3: Segmentation performance on 80 atypical mole images. SE-sensitivity, SP-specificity, AC-accuracy and Dice similarity coefficient.

Method	SE	SP	AC	DSC
FCM [32]	0.9173	0.9739	0.9594	0.9258
JSEG [54]	0.7435	0.9708	0.9236	0.7768
SRM [110]	0.1042	0.8954	0.6812	0.0919
Otsu [114]	0.5515	0.7579	0.6779	0.4372
Level Set [51]	0.7364	0.8237	0.7985	0.6532
ASLM [117]	0.8640	0.9733	0.9271	0.8689
Proposed Method	**0.9457**	**0.9781**	**0.9631**	**0.9295**

to 0.9400 and the dice similarity coefficient decreases from 0.9105 to 0.8588. Although the large decrease in terms of almost all criteria, the method still achieved the highest results and outperformed all six methods. For instance, by comparing the performance results of our method with the results derived from the FCM [32] method, we can demonstrate the percentage difference between these two methods in terms of the all evaluation metrics as 1.53% sensitivity, 15.06% specificity, 1.87% accuracy and 1.16% dice. Furthermore, another comparison can be performed with results obtained by the JSEG [54] method. The increase ratio is approximately 21.59% sensitivity, 2.14% specificity, 10.51% accuracy and 11.38% dice similarity coefficient. This means the method continues to yield promising results and outperforms the other methods, even in cases of melanoma images.

It is worth mentioning that the specificity ratio in terms of all images type (common moles, atypical nevi and melanomas) was approximately 98%, which means that the method able to classify the background pixels (normal skin) as a healthy normal skin. Thus, the all results confirm that the proposed approach outperforms the six methods with respect to all evaluation criteria. The experimental results indicated that the proposed method achieved a high accuracy of skin lesion segmentation, it successfully achieved 93.2% sensitivity, 98.0% specificity and 94.0% accuracy.

In summary, the proposed approach achieved very good results in terms of the four criteria, when dealing with the whole dermoscopy images (200). In addition, the accuracy of all evaluation metrics increased with regards to the benign lesions (common moles and atypical

Table 3.4: Segmentation performance on 40 melanoma images. SE-sensitivity, SP-specificity, AC-accuracy and Dice similarity coefficient.

Method	SE	SP	AC	DSC
FCM [32]	0.8079	0.8516	0.8235	0.8489
JSEG [54]	0.6746	0.9593	0.7591	0.7710
SRM [110]	0.2234	0.7512	0.4148	0.2852
Otsu [114]	0.5971	0.4870	0.5524	0.6064
Level Set [51]	0.7073	0.7015	0.7249	0.7503
ASLM [117]	0.5404	0.9597	0.6615	0.6524
Proposed Method	**0.8203**	**0.9799**	**0.8389**	**0.8588**

nevi), while a decline was observed in the method's performance when melanoma images were processed.

3.5 Summary

This chapter described an automated segmentation process of lesion areas in dermoscopy images using Fuzzy C-means and Markov Random Field methods. The algorithm began by enhancing the quality of the image by detecting and removing the noise, to perform that lesion area segmentation. The results of this process improve the robustness and the accuracy of the Markov Random Field segmentation. Expected regions are extracted by combining the Fuzzy C-Means and the Markov Random Field methods. The proposed method was tested on a dataset of 200 dermoscopy images including 40 melanoma, 80 common moles and 80 atypical moles. The results of the proposed method were compared with the ground truth lesions. Our experimental results indicated that the proposed method provided high accuracy of skin lesion segmentation, and also achieved exceptional performance against alternative methods we used in comparison.

There is a need for fully automated segmentation techniques, and we have described one possible way in this chapter, with promising results. However, due to the impact and sensitive nature of the systems, as it deals with the health status of individuals, the need for improvement arises as always. This is beneficial in assisting dermatologists to acquire mean-

ingful information from the image and properly identify diseased lesions, since the actual lesion should be determined and extracted with neither under-segmentation or over-segmentation. The next chapter explores an alternate and slightly more efficient approach to improve the performance of fully automated lesion segmentation methods.

Chapter 4

Lesion Segmentation in Dermoscopy Images Using Particle Swarm Optimization and Markov Random Field

4.1 Introduction

The standard approach in dermoscopic image analysis often comprises four phases: artifacts detection and removal, lesion segmentation, feature extraction and lesion classification. The most important stage is image segmentation, since the subsequent steps usually rely on its performance for their accuracy. It helps acquire meaningful information from images and assists the dermatologists to interpret skin cancer images properly; as well as identify melanoma. The segmentation process should be accurate and robust, because it plays an important role in determining melanoma disease at its early stage. For instance, the borders structure information, as mentioned in Chapter 3, is very important as it plays a key role in identifying melanoma diagnosis early. It is well known that the over segmentation and under segmentation both affect the lesion segmentation accuracy, which leads to the wrong diagnosis later. Consequently,

lesion segmentation must be correct and accurate. However, due to the existence of the hair and reflection artifacts on images as described earlier. In addition, many melanoma borders are often invisible or fuzzy, which makes visual identification very difficult for dermatologists. Furthermore, the interpretation of the images is time consuming and subjective. The lesion border delineation is a daunting task and very challenging.

This chapter presents a novel method for detecting lesion borders on dermoscopy images. First, the image is pre-processed to remove the noise and enhance its quality. This process improves the accuracy of the image segmentation stage. Second, the Particle Swarm Optimization (PSO) and the Markov Random Field (MRF) are combined to segment the lesion area from images. The image segmentation is formulated as an optimisation problem of the energy function with MRF theory. The PSO method is used to perform the initial labeling based on the optimal threshold value, which was obtained by maximizing the fitness function. Then, an additional local search is performed for each segmented image by integrating it with MRF method. Accordingly, the pixels are reassigned to different classes based on the minimum cost. This process continues until a stopping criterion is met. Upon comparison, the proposed approach provides high performance in automatic image segmentation on dermoscopic images. This chapter is structured as follows: Section 4.2 provides an overview of the PSO method. The proposed method is described in details in Section 4.3. Section 4.4 illustrates the results and discussions. Finally, a summary of the chapter is presented in Section 4.5.

4.2 Overview of the PSO Method

The original PSO proposed by Eberhart and Kennedy [56] is a computational optimisation method based on swarm intelligence theory. It is initialised with a group of random particles to discover the optimal solution through the search space; each particle represents a candidate solution of the problem based on a fitness function. In the PSO approach, the whole swarm is modelled as multidimensional space S; therefore, each particle $P_i = \{F_i, V_i\} \in S$ has two components: position F_i and velocity V_i. The best previously visited position of particle i is denoted as its individual best position, while the best position of all these individual bests

is denoted as the global best position. In the beginning, the position and the velocity of each particle (solution of the problem) are initialised randomly. Then, the problem is being optimised by flying each particle through the search space, and updating its individual best position and global best position. Therefore, the performance of each particle is evaluated using a fitness function. This process is repeated until a stopping criterion is met. The stopping criterion could be that all particles positions do not change any more than a certain threshold, or the maximum number of iterations is met. At each step, the velocity and the position of particle f_i are updated using (4.2.1) and (4.2.2), respectively.

$$v_i = w \times v_i + c_1 \times r_1 \times (fb_i - f_i) + c_2 \times r_2 \times (G - f_i) \tag{4.2.1}$$

$$f_i = f_i + v_i \tag{4.2.2}$$

where w is the inertia weight which controls the interaction power between the particles, v_i is the current velocity and f_i, fb_i and G are the current position of the particle, the best position which particle has achieved so far and the location of overall best value respectively. The r_1 and r_2 are random values generated in the range between 0 and 1. The positive values c_1 and c_2 are constants referring to the acceleration coefficients in order to guide particles into good directions, and control the maximum step size.

4.3 The Proposed Method

In the pre-processing stage, the image enhancement is carried out by detecting hairs and reflection artifacts. The Image inpainting method is applied to remove the pixels which indicate the hairs and reflection artifacts mask. The second step of the proposed method is to delineate the border of the lesion area, which is achieved by combining the PSO with the MRF method. The image segmentation is formulated as an optimization problem of the energy function with MRF theory. The PSO approach assigns each image pixel to a cluster based on the highest fitness function; then the PSO and the MRF are integrated, in order to do a local search and refine the segmented image. Accordingly, the pixels are reassigned to different classes based

on the minimum cost. This process continues until a stopping criterion is met. The full details of individual steps are described as follows:

4.3.1 Dermoscopic Image Pre-processing

The pre-processing stage is responsible for detecting and reducing the number of artifacts from images, with the purpose of enhancing their quality. Therefore, this step is mandatory, since as aforementioned that the dermoscopic images contain several types of noise (hairs and reflection artifacts), which is covering most of the lesion areas. Incorrect segmentation of pigmented lesions can be obtained if hairs covering the images and lightening reflection are not detected and removed. Therefore, we used the same algorithm presented in Chapter 3 Section 3.3.1 to eliminate all these issues and obtain clearer images.

4.3.2 Skin Lesion Segmentation

Image segmentation is the process of adequately separating pixels into few groups, whose pixels share similar characteristics, such as texture, colour, and shape. Indeed, the performance of segmentation should be fast and accurate, since all the subsequent steps, such as feature extraction, feature selection and classification phase are dependent on its performance. In dermoscopy images, the segmentation stage is one of the most important and challenging steps, due to several reasons, such as the lesions have large variations in size and colour, low contrast between the lesion area and surrounding healthy skin, as well as the presence of the hairs with several artifacts as mentioned earlier.

Our approach aims primarily to construct an efficient, robust and automatic segmentation tool for melanoma lesion detection. The image segmentation phase is implemented, in order to separate the lesion area from the healthy skin. Useful results can be acquired by applying appropriate segmentation techniques. Consequently, the PSO and the MRF methods were combined to perform the final segmentation of the images by minimising the energy function.

The image segmentation is formulated as an optimisation problem of the energy function with the MRF method. For this, we use the PSO method to perform the initial labelling. The underlying idea of our approach is a cooperative search of the best class label for each pixel

in the image, using a population of artificial particles. Each particle assigns pixels to a class iteratively based on fitness function. In the beginning, the number of particles is determined, and then the position value of each particle is randomly set within the boundaries of the search space, while the velocity of each particle is set to zero. The search space will rely on the maximum intensity value L, which means the number of particles are distributed randomly between 0 and 255. One particle in the swarm represents one solution for clustering the image based on fitness function. Therefore, the whole swarm represents a number of candidate clustering solutions for the whole image. In each step of the proposed algorithm, each particle compares its current fitness value with the value of its own *Pbest* solution; as well as the fitness value of the whole swarm *Gbest* solution. The fitness function is defined as the between-class variance σ_b^2 of the intensity distributions of the image [67] by:

$$\sigma_b^2 = \sum_{j=1}^{n} w_j(\mu_j - \mu_t)^2 \tag{4.3.1}$$

where j represents a specific class in such a way that μ_j and w_j are the average and the probability of occurrence of class j respectively. The probabilities of occurrence w_j of class $D_1, ..., D_n$ are given by:

$$w_j = \sum_{i=1}^{t_j} p_i, j = 1$$

where p_i is the probability of occurrence of each pixel in the image, which can be obtained by:

$$p_i = \frac{h_i}{N}, \sum_{i=1}^{N} p_i = 1$$

where i represents intensity level, $i.e., 0 <= i >= L-1, N$ represent the total number of the pixels in the image and hi refers to the image histogram.

And the average of each class μ_j can be calculated as:

$$\mu_j = \sum_{i=1}^{t_j} \frac{ip_i}{w_j}, j = 1$$

where ip_i presents the total mean (combined mean) and can be obtained by:

$$\mu_t = \sum_{i=1}^{L} ip_i$$

In other words, the segmentation problem is reduced to an optimisation problem to search for the optimal threshold value by maximising the fitness function. Higher fitness means there is a greater probability of the particle being successful.

$$ft = \max \sigma_b^2(t) \tag{4.3.2}$$

Therefore, the optimal threshold value can be used for dividing the image into 2 classes: foreground and background. Through this process, we can obtain the membership value easily for each pixel in the image by:

$$g(x,y) = \left\{ \begin{array}{lll} 1, & \text{if} & \text{f(x,y)} > \text{T} \\ 0, & \text{if} & \text{f(x,y)} \leq \text{T} \end{array} \right\}$$

The main drawback of the PSO approach is that it addresses each pixel individually using its intensity value. An additional local search must be performed for each segmented image by combining it with the MRF method in order to address this issue and enhance the performance of the PSO method.

As a statistical method, the MRF provides a tool for Bayesian modelling using spatial continuity and has been widely used in image segmentation with numerous applications [66, 129, 130]. Typically, it is based on a local calculation of probability and potential functions. As we described the process in Chapter 3 Section 3.2.2, the pixels of the image are indexed by a rectangular patch S and each image pixel s is characterised by the grey level y_s from the set $y = \{y_s : s \in S\}$. The labelling process consists of accurately labelling each image pixel $s \in S$ with a class label representing the pattern class in the image. A label set is defined as $\Lambda = 1, 2, ..C$ where C is the number of classes. A labelling is indicated by $x = x_s : x_s \in \Lambda, s \in S$ where $x_s = l$ denotes that the class label l is assigned to the pixel s. The goal is to discover

the labelling $\hat{x}$ of the image, which maximises $P(x|y)$ based on the Bayes decision theorem.

The primary aim of using the MRF to perform image segmentation is minimising the energy function or maximizing the probability of pixel allocation to a cluster by using Maximum A Post Priority (MAP) [66]. According to the MAP estimate and the Hammersley-Clifford theorem [30] with assumption of existence of Gaussian noise in the images, the energy function can written as

$$U(x) = [\sum_{s \in S} \frac{(y_s - \mu_{xs})^2}{2\sigma_{xs}^2} + \sum_{s \in S} \log(\sigma_{xs}) + \sum_{c \in C} V_c(x)]$$

The overview of the MRF method can bee seen in Chapter 3 Section 3.2.2.

Our algorithm used the membership value instead of the conditional probability to correctly assign each image pixel and refine the segmented image. Equation (4.3.3) is used to calculate the conditional probability.

$$V_c(y_i|x_j) = g(x, y) \tag{4.3.3}$$

The Iterated Conditional Modes (ICM) method is used widely in MRF applications, which achieves optimal labelling with minimum energy. Actually, ICM is an iterated algorithm, which maximises local conditional probabilities by propagating messages along nodes in the MRF. Therefore, we use it to obtain optimal clusters. Our method assumes that one pixel has 8-neighbours. Therefore, the second order clique potential is defined on pairs of neighbouring pixels:

$$V_c(xi, xj) = (1 - i_{xi,xj})$$

where $i_{xi,xj} = 0$ if $xi \neq xj$ and 1 if $xi = xj$. Therefore, the energy function is defined as the sum of conditional probability and the second order energies as in (4.3.4).

$$U(X) = \sum_{i \in S} V_c(y_i|x_j) + \sum_{j \in S} \sum_{N_i} V_c(xi, xj) \tag{4.3.4}$$

By applying the PSO approach, each pixel in the image is labeled to a cluster based on the highest fitness function. Thus, to perform the local search and refine the segmented image, we integrate the PSO with the MRF approach. Accordingly, all pixels will be labeled again to different classes by getting the minimum cost (4.3.4). This process continues until a stopping

criterion is met or no further change occurs between clusters. The general schema of our approach is given in Algorithm 1. As a final step of the method, the morphological operations were used to fill in and remove several existing holes inside the segmented lesion; as well as few outlier pixels in the background.

Algorithm 1 PSO-MRF image segmentation

procedure MyProcedure
  Result: Regions of interest (objects)
  Initialization:
  Place images pixels in a matrix;
  Initialize randomly the position of each particles;
  The velocity of each particle is set to zero $v_i = 0$;
  Repeat
  for each particle i in S do
  Compute the fitness function based on equation (4.3.1);
  Get the optimal threshold value (best fitness) by maximising the fitness function equation (4.3.2);
  Construct the candidate segmentation according to optimal threshold value ;
  Calculate the membership value using equation (4.3.3);
  //* Local search *//
  Improve the solution by using the steepest descent method (local minimum) equation (4.3.4);
  Update the particle's best position
  Update the global best position
  Update the particle's velocity and position
  Update the velocity of particle f_i using equation (4.2.1);
  Update the position using equation (4.2.2);
  end for
  $it = it + 1$; // Number of iteration
  until The number of iterations (total max) is reached
end procedure

It is worth noting that the implementation of new segmentation method for affected lesions in dermoscopic images is desperately needed for several purposes, as described below:

- Due to the effect and sensitive nature of the existing systems, as it deals with the health status of individuals.
- Real lesions provide enough information without neither less or more irrelevant features, such as in under-segmentation and over-segmentation.

- The actual lesions have the key role in melanoma diagnosis since the ABCD-E rule is dependent on its accurate segmentation.
- An effective segmentation of the actual lesions can improve the productivity of dermatologists during diagnosis process.

In addition to the extraction of the actual lesion boundary (true positive) and improvement the segmentation accuracy, there is also the need for new segmentation techniques to handle the major drawback by using the previous method proposed in Chapter 3. The method's major drawback lies with the FCM approach as it addresses pixels individually based solely on their intensity values. Therefore, it lacks the capacity to model the overall appearance of a local neighbourhood region. Moreover, the FCM was used for the purpose of estimating the initial parameters (mean and variance) in the first iteration, which were needed in the MRF method. Moreover, in the FCM the initial points (centroids) started randomly, which means sometimes no results provided at all. Meanwhile, in this chapter, the PSO algorithm generated the best solution (always have a solution) to provide the initial parameters and share its solution with the MRF approach to perform a local search in each iteration. Thus, each pixel is reassigned to an optimal cluster; therefore pixels are divided into two clusters: cancerous pixels as the foreground (lesion) and normal skin pixels as the background.

4.4 Results and Discussions

The proposed method was tested on a publicly-available database PH2 [101], the same dataset used in Chapter 3. Four different criteria have been selected to evaluate the performance of the segmentation results: sensitivity (SE), specificity (SP), accuracy (AC) and the Dice similarity coefficient (DSC) [90]. Segmentation of the skin lesion was implemented using PSO [67], Fuzzy C-Means (FCM) [32] and the proposed method respectively. Thus, the comparison of these three methods was performed with the lesions acquired by an expert dermatologist, in order to evaluate the performance of the proposed method.

Examples of manual segmentation by dermatologists together with the results by three methods of skin lesion segmentation are shown in Figure. 4.1. The ground-truth is shown in the

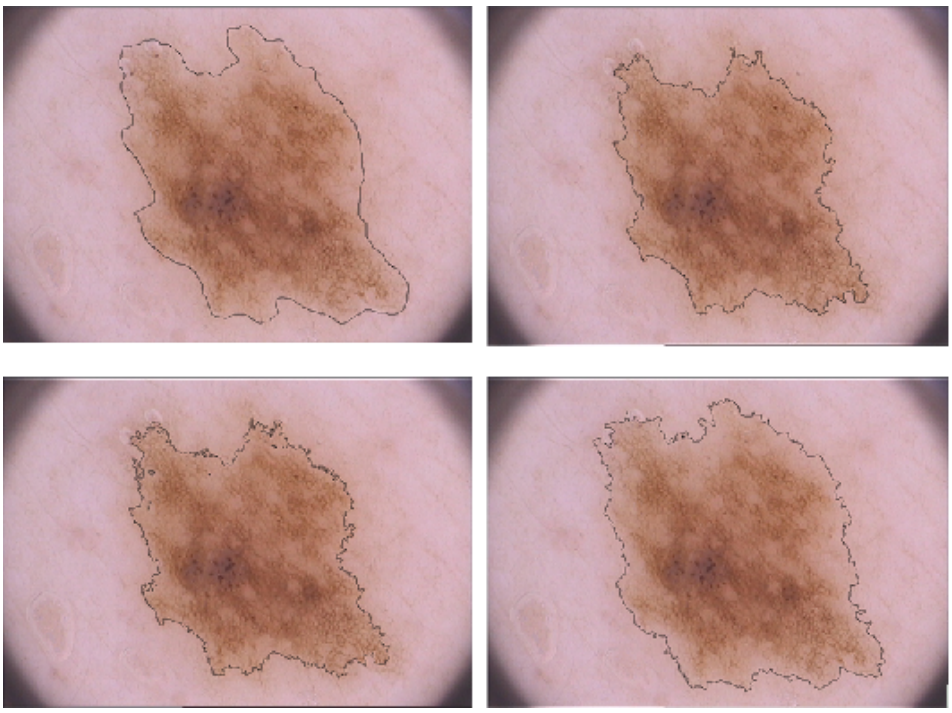


Figure 4.1: Results of lesion delineation: lesion border detection derived from ground truth (top left), results of lesion delineation derived from applying the PSO (top right), lesion boundary detection obtained from applying the FCM (bottom left) and the proposed method outcomes (bottom right).

top left image and results obtained using the PSO, FCM and the proposed method are in the top right, bottom left and bottom right images respectively. In general, the proposed method has the best performance in terms of the accuracy. For instance, as we can see in Figure. 4.1, the image in the top right (PSO method) did not succeed in delineating the entire lesion, with part of the lesion misclassified as a healthy skin (background). Moreover, the detected edge is not close to the actual boundary of the lesion; thereby leading to misclassification. This was the same when the FCM method was applied, as seen in the bottom left image. The result from the proposed method reveals that the detected edge is quite close to the real boundary of the skin lesion, while almost the whole lesion was detected and delineated.

In addition to the qualitative evaluation (visual observations), a quantitative evaluation was performed by comparing the performance of our method with 8 alternatives; namely, J-image segmentation (JSEG) [54], Statistical Region Merging (SRM) [110], Otsu [114], Level Set [51], Automatic Skin Lesion Method (ASLM) [117], PSO [67], FCM [32] and Y. Li et al

[94], which have been already used for the same dataset, and considered for lesion segmentation. The total experimental results on the same dataset of 200 dermoscopy images are presented in Table 4.1. The performance results of J-image segmentation (JSEG) [54], Statistical Region Merging (SRM) [110], Otsu [114], Level Set [51] and the Automatic Skin Lesion Method (ASLM) [117] were generated from their original manuscripts, while the performance results of FCM [32] and PSO [67] methods were obtained through our implementation. The performance of the dierent methods was generated using all the 200 dermoscopic images. We observed that the method presented by Y. Li et al [94], FCM [32] and PSO [67] are the only comparable methods, since their results are close to the current method results. Therefore, withe respect to Y. Li et al [94] method, it can be seen that the sensitivity, accuracy and dice are slightly increased, while the specificity value remained unchanged. The percentage of increase obtained are 0.72% sensitivity, 0.78% accuracy and 1.38% dice. Conversely, if we compare our performance results with the results obtained by the PSO [67] and the FCM [32] methods, we can get approximately 3.12% and 1.44% increase respectively, in terms of the accuracy alone. With respect to sensitivity and specificity, the percentage of increase approximately 19.95%, 0.93% and 5.72%, 2.97% compared with PSO and FCM methods respectively. Consequently, all results confirm that the proposed approach outperforms the eight methods, in terms of all evaluation criteria. It performs the best in delineating the lesion area, and the experimental results indicated that the proposed method successfully achieved approximately 94.00% sensitivity, 98.00% specificity, 95.00% accuracy and 92.30% dice. Therefore, we believe that it can be used for classification stage and provide greater accuracy for melanoma detection as an alternative to ground truth labelling.

The same analysis to that reported in Section 3.4 was conducted to implement a finer analysis of all images. Three diagnostic classes (common moles, atypical nevi and melanomas) were evaluated separately. Table 4.2 illustrates the experimental segmentation results of 80 common mole images only. In our method, the sensitivity raises from 0.9388 to 0.9651, the specificity slightly increases from 0.9758 to 0.9784, the accuracy increases form 0.9474 to 0.9740. and the dice similarity coefficient raises from 0.9231 to 0.9361. By comparing our results with an alternative method, such as the PSO [67] approach, we can observe that the

Table 4.1: Segmentation performance on the complete dataset. SE-sensitivity, SP-specificity, AC-accuracy and DSC-Dice Similarity Coefficient.

Method	SE	SP	AC	DSC
PSO [67]	0.7826	0.9709	0.9187	0.8481
FCM [32]	0.8880	0.9517	0.9339	0.9040
JSEG [54]	0.7108	0.9714	0.8947	0.7554
SRM [110]	0.1035	08757	0.6766	0.1218
Otsu [114]	0.5221	0.7064	0.6518	0.4293
Level Set [51]	0.7188	0.8003	0.7842	0.6456
ASLM [117]	0.8024	0.9722	0.8966	0.8257
Y. Li et al. [94]	0.9320	**0.9800**	0.9400	0.9105
Proposed Method	**0.9388**	**0.9800**	**0.9474**	**0.9231**

proposed method still achieves higher performance in terms of all evaluation metric. The sensitivity increases from 0.7535 to 0.9651, the specificity raises from 0.9711 to 0.9784, the accuracy becomes 0.9740 as opposed to 0.9479 and the dice similarity coefficient increases from 0.8156 to 0.9361. Moreover, another comparison can be made of the results obtained by FCM [32] method. Similarly, the sensitivity raises from 0.9182 to 0.9651, the accuracy increases from 0.9670 to 0.9740 and the dice similarity coefficient becomes 0.9361 as opposed to 0.9083. However, the specificity slightly decreases from 0.9835 to 0.9784, this means the percentage of decrease approximately 0.5%, which is quite low. Whereas, The increase ratio of the sensitivity, accuracy and the dice similarity coefficient are about 5.10%, 0.72% and 3.06%, which are too high. We believe that the performance of our method is still the highest against the other methods, in terms of all the evaluation metrics. This indicates that the proposed approach achieves higher results in segmenting images of common moles.

The region segmentation performance of 80 atypical nevi images only is presented in Table 4.3. Moreover, it is very clear that the proposed approach outperformed all other methods in almost all used criteria. In particular, the sensitivity rises from 0.9388 to 0.9589, the accuracy increases form 0.9474 to 0.9653. and the dice similarity coefficient raises from 0.9231 to 0.9321. In terms of the comparison with other methods, our results are higher than those obtained from PSO [67] method. The rise from 0.8334 to 0.9589, 0.9702 to 0.9709, 0.9432 to 0.9653

Table 4.2: Segmentation performance on 80 common healthy images. SE-sensitivity, SP-specificity, AC-accuracy and DSC-Dice similarity coefficient.

Method	SE	SP	AC	DSC
PSO [67]	0.7535	0.9711	0.9479	0.8156
FCM [32]	0.9182	**0.9835**	0.9670	0.9083
JSEG [54]	0.6977	0.9783	0.9370	0.7265
SRM [110]	0.0751	0.9332	0.7250	0.0611
Otsu [114]	0.4777	0.7832	0.6911	0.3658
Level Set [51]	0.7069	0.8262	0.7996	0.5856
ASLM [117]	0.8717	0.9760	0.9477	0.8690
Y. Li et al. [94]	0.9531	0.9783	0.9680	0.9242
Proposed Method	**0.9651**	0.9784	**0.9740**	**0.9361**

and 0.8867 to 0.9321 of the sensitivity, specificity, accuracy and dice similarity coefficient respectively. Furthermore, its results much better than the results achieved by FCM [32]. The sensitivity, accuracy and dice similarity coefficient respectively are increased from 0.9173 to 0.9589, 0.9594 to 0.9653 and 0.9258 to 0.9321. Another comparison can be performed with the method presented in ASLM [117] work. Likewise, the increases of the sensitivity accuracy and dice similarity coefficient is approximately of 10.98%, 4.12% and 7.27% respectively. With respect to the specificity, we observed that it reduces to 0.9709 compared to FCM [32] and ASLM [117] results, but in the meanwhile, our segmentation accuracy is the highest, compared with the other methods. Thus, even in the case of atypical nevus, which are considered benign lesions, the proposed approach continues to obtain higher segmentation results.

Conversely, a huge decrease can be observed in the quality of segmentation results when only melanoma images (40 images) are processed. The overall segmentation performance results of 40 melanoma images are presented in Table 4.4. In particular, the proposed method presents a large decrease in terms of four criteria: the sensitivity decreases to 0.8459 compared with that in Table 4.1 (0.9388), the specificity reduced from 0.9800 to 0.9595, the accuracy becomes 0.8605 instead of 0.9474 and the dice similarity coefficient decreases from 0.9231 to 0.8821. Despite, the large decrease in all evaluation criteria, the method still achieved the highest results and outperformed all other methods. For example, comparing the performance

Table 4.3: Segmentation performance on 80 atypical mole images. SE-sensitivity, SP-specificity, AC-accuracy and Dice similarity coefficient.

Method	SE	SP	AC	DSC
PSO [67]	0.8334	0.9702	0.9432	0.8867
FCM [32]	0.9173	0.9739	0.9594	0.9258
JSEG [54]	0.7435	0.9708	0.9236	0.7768
SRM [110]	0.1042	0.8954	0.6812	0.0919
Otsu [114]	0.5515	0.7579	0.6779	0.4372
Level Set [51]	0.7364	0.8237	0.7985	0.6532
ASLM [117]	0.8640	0.9733	0.9271	0.8689
Y. Li et al. [94]	0.9457	**0.9781**	0.9631	0.9295
Proposed Method	**0.9589**	0.9709	**0.9653**	**0.9321**

results of the proposed method with those derived from PSO [67] approach, we can note that our method yielded better results. I was able to increase the sensitivity, accuracy and dice similarity coefficient by 12.05%, 5.88% and 5.66% respectively. Moreover, it produced higher results than the ones obtained from the FCM [32] method. With respect to the all evaluation metric, the method raises the sensitivity form 0.8079 to 0.8459, specificity from 0.8516 to 0.9595, accuracy from 0.8235 to 0.8605 and the dice similarity coefficient from 0.8489 to 0.8821. This indicates that, even in the case of melanoma images, the method successfully achieved promising results and its performance is better than the other alternative methods that we compared to.

Despite, the obtained promising results of skin lesion segmentation, the method was not able to delineate the whole actual lesion in many cases of melanoma images as provided by dermatologists. The example in Figure 4.2 illustrates three melanoma images with proposed obtained result and ground truth respectively. It is clear that the method provided a binary image containing under-estimated lesions, as illustrated in the second column, which means not all of the affected lesion was segmented. The reason of that is because many cases have variegate colouring inside the lesion area such as the one in the first row, which makes the detection of the whole lesion quite difficult, since in our approach, the number of clusters was set as 2. Moreover, in cases of presence very poor contrast between the lesion edges and the

Table 4.4: Segmentation performance on 40 melanoma images. SE-sensitivity, SP-specificity, AC-accuracy and Dice similarity coefficient.

Method	SE	SP	AC	DSC
PSO [67]	0.7549	0.9763	0.8127	0.8348
FCM [32]	0.8079	0.8516	0.8235	0.8489
JSEG [54]	0.6746	0.9593	0.7591	0.7710
SRM [110]	0.2234	0.7512	0.4148	0.2852
Otsu [114]	0.5971	0.4870	0.5524	0.6064
Level Set [51]	0.7073	0.7015	0.7249	0.7503
ASLM [117]	0.5404	0.9597	0.6615	0.6524
Y. Li et al. [94]	0.8203	**0.9799**	0.8389	0.8588
Proposed Method	**0.8459**	0.9595	**0.8605**	**0.8821**

surrounding healthy skin in many parts of the lesion as appears in the images in the second and the third row, the dermatologists usually delineate the lesion area with extra parts in order to be more accurate. Therefore, the algorithm could not classify these parts correctly; furthermore, this process affects the segmentation evaluation results, since the manual segmentation is used as a ground truth.

In summary, the proposed approach yielded excellent results regarding the four criteria when dealing with the whole dermoscopy images (200). In addition, the accuracy of all evaluation metrics was increased when dealing with the benign lesions (common moles and atypical nevi), while decreasing when melanoma images are processed.

It is worth noting that the comparison with the method presented in Chapter 3 (FCM with MRF) has been performed, for the purpose of clarifying the difference between their results, and demonstrating the advantages of combining the PSO and the MRF methods. As mentioned previously that integrating these two methods to do a local search gave the algorithm robustness and high accuracy in detecting the real lesion area as seen in Figure (4.3). The example in Figure (4.3) illustrates the outcomes derived from skin cancer experts (ground truth) in the first column, the previous method (FCM with MRF) in the second column, and the current method (PSO-MRF) presented in the third column. It is noticeable that the lesions illustrated in the second column images are not detected properly, compared

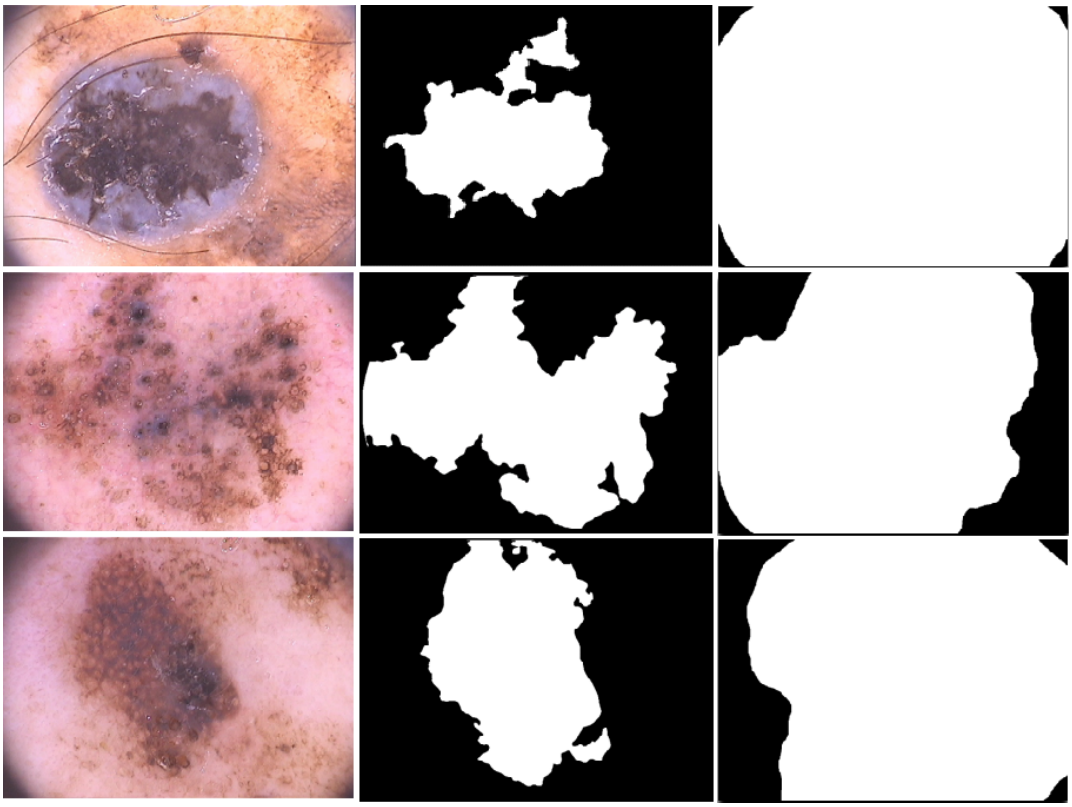


Figure 4.2: Example shows some cases of dermoscopic images: original images (Source:[101]) (first column), results derived from our method [94] (second column) and the ground truth (third column).

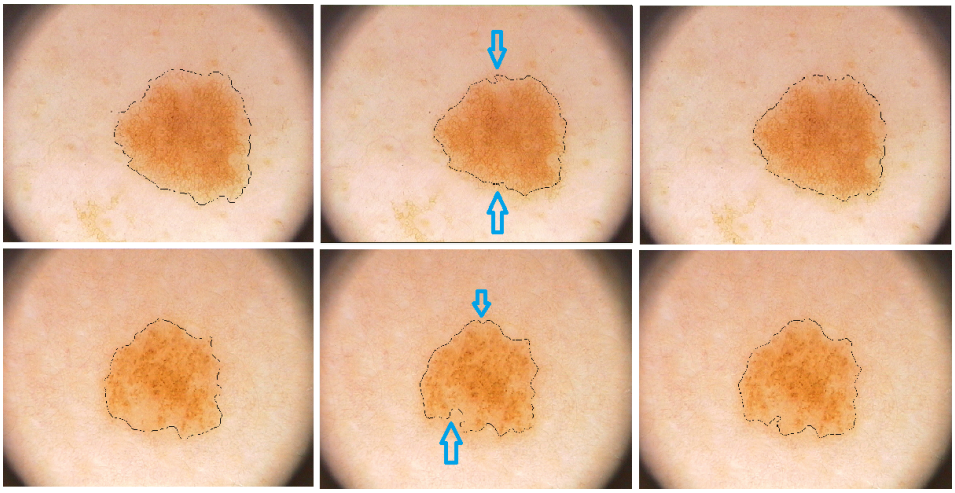


Figure 4.3: Results of lesion delineation in dermoscopic images: lesion border detection derived from ground truth (first column), results of lesion delineation obtained from applying the FCM with MRF "our previous method" (second column), and the proposed (PSO-MRF) method outcomes (third column).

with the ground truth images. In the part of low contrast edges between the lesion area and its surrounding skin, the method classified part of the lesion as a background (normal healthy skin), which in turn led to obtaining under-segmentation (see the blue arrows). This issue occurred because incorrect initial parameters (mean and variance) were obtained and used for MRF method, since the FCM deals with pixels individually by their intensity values only as aforementioned. Whereas the lesions presented in the third column images are almost fully detected. This is due to assigning each pixel to classes properly by sharing the obtained best solution from PSO method with the neighbours of each pixel. Furthermore, the detected lesions presented in the third column are very close to the actual boundary, which is quite similar to the one acquired from the dermatologists (ground truth). In terms of the segmentation of the whole dataset images, as we described earlier that the current method achieved approximately 94.74% accuracy, while the method presented in Chapter 3 obtained accuracy of 94.00%, which means the percentage of increase obtained is about 0.78%. We acknowledge that it is slightly increased of the accuracy, but achieving greater accuracy, however slight, is useful and meaningful.

4.5 Summary

This study presents an automated approach for region segmentation in dermoscopy images. The images are pre-processed using the aforementioned method presented in Chapter 3 Section 3.3.1. The results of this process improve the robustness and the accuracy of the lesion segmentation. Lesions are segmented by integrating the PSO and the MRF methods. The method proved accurate and robust, leading to successful segmentation results of the region boundaries used to assess and monitor skin cancer disease.

The proposed method was tested on a dataset of 200 dermoscopy images including 40 melanoma, 80 common moles and 80 atypical moles. The results of the proposed method were compared with the ground truth lesions. The experimental results indicated that the proposed method provided a high accuracy of skin lesion delineation. In addition, a comparison against a selection of several alternative methods shows that the proposed method performs the best in terms of all evaluation metric. It also exhibited better performance opposed to our previous method proposed in Chapter 3, as it is capable of segmenting almost the whole actual lesion. All the above factors indicate that this approach can cope with the image noise, and provide a high accuracy of skin lesion detection.

According to the obtained unaffected borders of the pigmented regions when images are pre-processed, we believe that our pre-processing approach can be useful when extracting pigment network structures since it was able to extract the hair without affecting the lesion boundaries. Therefore, it can be applied when dealing with pigment network detection.

From the promising results obtained in the identification of the pigmented region areas, we deem it necessary to provide a new method of melanoma detection. This will improve the productivity of dermatologists by reducing their time required while diagnosis process. Consequently, the next chapter describes a semi-automated method for pigmented lesion detection.

Chapter 5

Detection of Pigment Network Structures in Dermoscopy Images

5.1 Introduction

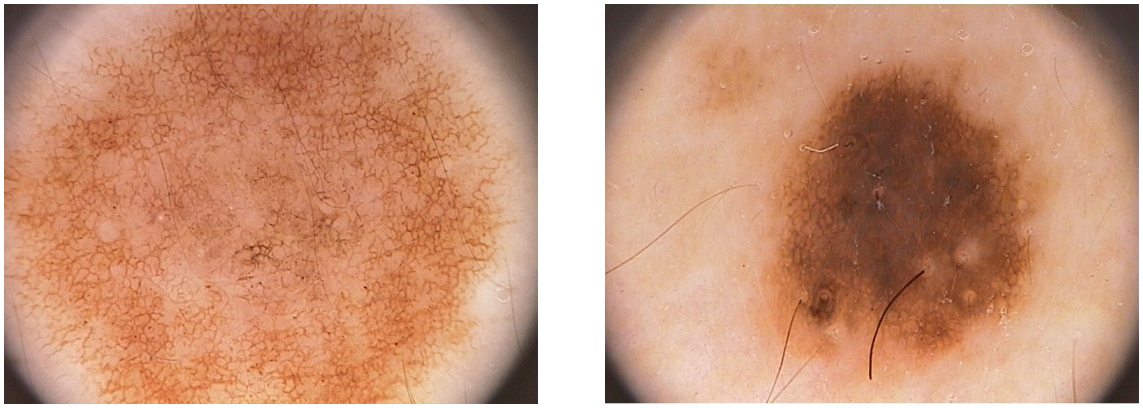
Based on the ABCD-E criteria, the pigment networks structures correspond with the (D) parameter, which refers to the differential structures criteria detailed in Appendix A Section (A.2). The segmentation of dermoscopic image structures, such as dots, pigment network, streaks and blue-whitish veil, has been of great interest since it could be used as a non-invasive diagnosis in modern dermatology. The pigment network is regarded as one of the most relevant dermoscopic structures, its appearance on the body is an indicator of the existence of melanin in deep layers of the skin. It appears as a grid of thin brown lines over a lighter background with a shape very similar to a honeycomb called reticular [25] [20]. Pigment network can cover almost the whole lesion area as illustrated in Figure 5.1(left), or only several small parts of it, as illustrated in Figure 5.1(right). The shape of the pigment network can be classified as either Typical or Atypical: typical when the pattern is regularly with a light-to-dark-brown network small uniform space network holes and thin network lines distributed more or less regular throughout the lesion; while atypical is characterised by black, brown, or gray and thickened lines distributed irregularly throughout the lesion. Consequently, an atypical pigment network

is considered as a melanoma[19] [25]. Examples are illustrated in Figure 5.2.

In the classification of skin lesions, the pigment network plays an essential role in identifying the lesions as melanoma, because it facilitates the distinction between various classes in the lesions. Therefore, the presence of pigment network on lesion areas in dermoscopy images represents an important feature, which can be used for melanoma detection in its early stage.

Computer-aided diagnosis (CAD) systems have been proposed by many different groups to identify malignant melanomas in dermoscopy images. These systems use several features, such as colour, shape and texture, to classify the images as either normal or abnormal.[19] [25] [105]. In the detection of melanoma, dermatologists typically use the ABCD rule to analyse four parameters (Asymmetry, Border, Colors and Diameter) and distinguish skin lesions on the 7-points checklist. This is a scoring method for a set of different classifiers according to colour, shape and texture [147] [105]. However, it must be clarified that this procedure is time consuming and requires significant effort. Thus, the use of an automated system decreases the time consumed and fatigue. In addition, it assists physicians in interpreting the image correctly and supports their diagnosis. Besides, storing the diagnostic information could be very useful in future investigation or developing a new method. It is noteworthy that the automatic detection of pigment networks is a challenging task because there is a smooth transition between the colour of the pigment network lines and the background. Moreover, the presence of dark hair covering the lesions and the existence of specular reflection. In addition, the circular structures, such as dots, are detected wrongly as pigment networks or vice-versa [25].

This chapter outlines the four steps implemented to construct an automatic algorithm for the detection of pigment network lines in dermoscopy images, with a view to the early diagnosis of melanoma skin cancer. First, the image is pre-processed to remove the noise; for example, hairs and reflection detection. We focus on this process since it improves the accuracy of the subsequent steps such as segmentation and classification. A sample threshold is used to extract the lightening reflection, and then the Gabor filters using a bank of 64 directional filters with image sharpening and Sobel filter are implemented to detect and extract the hair. Second, based on the intensity values and geometrical properties of the pigment networks, a



(a) (b)

Figure 5.1: Two examples of pigment network distribution in dermoscopy images. Left: the pigment network is diffusing in almost the whole lesion. Right: it is existing in separate parts of the pigmented lesion area. (Source:[101]).

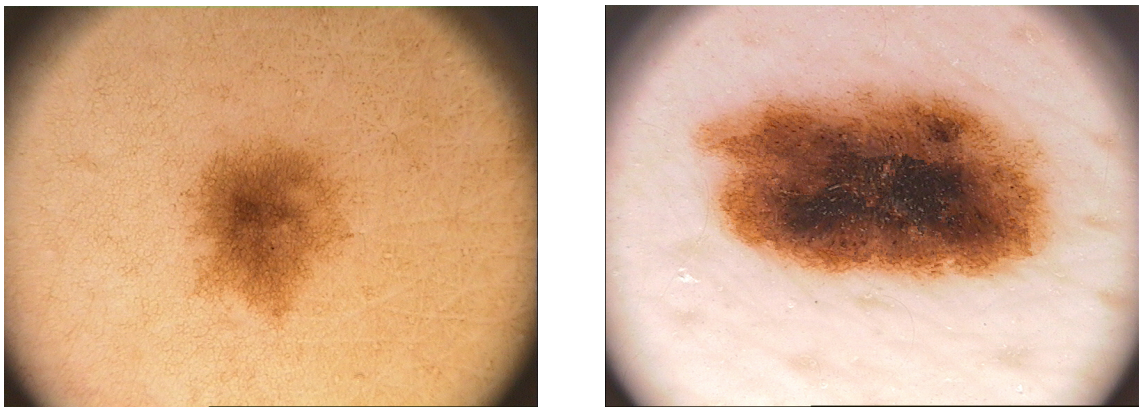


Figure 5.2: Examples of pigment network shapes in dermoscopy images. Left: a typical pigment network. Right: an atypical pigment network. (Source:[101]).

set of tentative pigment network regions are detected using connected component analysis. Then, features are extracted from the given lesion and by using a trained classifier, the given lesion is classified in one of two classes (with or without pigment network). This chapter is structured as follows. The proposed system is described in details in Section 5.2. Section 5.3 illustrates the experimental results. Finally, a summary of the chapter is presented in Section 5.4.

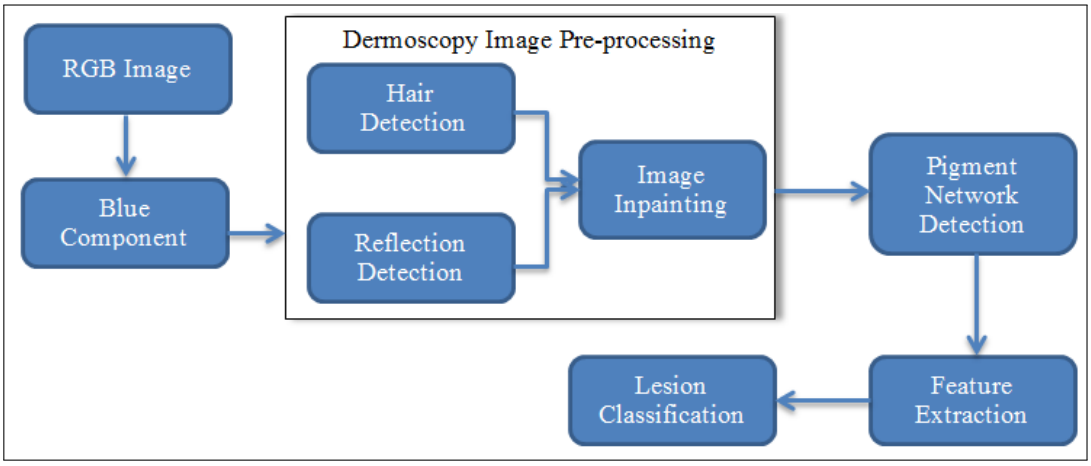


Figure 5.3: Overview of the detection system.

5.2 Methods

This section describes and discusses the proposed system for skin lesion detection. Initially, image quality is enhanced by detecting and removing the light reflection and hairs. Next, the lines of pigment network are detected using a bank of directional filters with a connected component analysis. Thus, several features are extracted from the segmented network and used to train an Artificial Neural Network and classify each lesion as normal or abnormal skin. The scheme of melanoma detection is illustrated in Figure 5.3.

5.2.1 Image Pre-processing

As mentioned previously that the hairs are presence in most of the dermoscopy images, which in turn could affect the lines of pigment networks during segmentation process, since their shapes are quite similar. Hairs covering the lesion areas could also hide the pigment network lines and make them invisible, which make its detection a challenge. In addition, incorrect detection of the lesion (pigment network features) could lead to the wrong classification later. Accordingly, the whole hair should be detected and removed, in order to obtain effective segmentation and classification results. Due to the similarity between the pigment network lines and the hair, it is challenging to extract hairs without any effect of the pigment network lines, but at the same time, we can use the same approach to detect both. Thus, we used

Figure 5.4: Examples of hair detection process and inpainting images: original image (Source:[101]) (first column), hair detection mask (second column) and inpainted image (third column).

the same algorithm presented in Chapter 3 Section 3.3.1.2 to obtain images that were more understandable. One such example can be seen in Figure. 5.4.

It should be noted here that the hair which was not detected using previous methods, such as "Barata's method" [25], was identified and classified as a noise. Three examples are illustrated in Figure.5.5.

5.2.2 Pigment Network Detection

Pigment network is considered one of the features that used by experts dermatologists to perform the diagnosis of melanoma. In general, dermatologists use the net hole size and the thickness of lines to assess the network structure and classify it as either typical or atypical [25]. Therefore, we tried to implement an algorithm to extract the skeleton of pigment network from dermoscopy images, which could be quite useful for the specialist.

Pigment network and hair artifacts both have a very similar line shape, and they appear with more than one orientation in dermoscopy images. Therefore, the same technique (Gabor Filter) with different parameters can be used to detect the pigment network lines. Two primary steps have been applied for this purpose: network enhancement and network detection.

5.2.2.1 Network Enhancement

The output image that was obtained from pre-processing stage (cleaned image) is filtered with a bank of directional filters {see equations (3.3.1)-(3.3.1.2} in Chapter 3. Due to the difference of the length and the width of each line stroke of the hairs and pigment network,

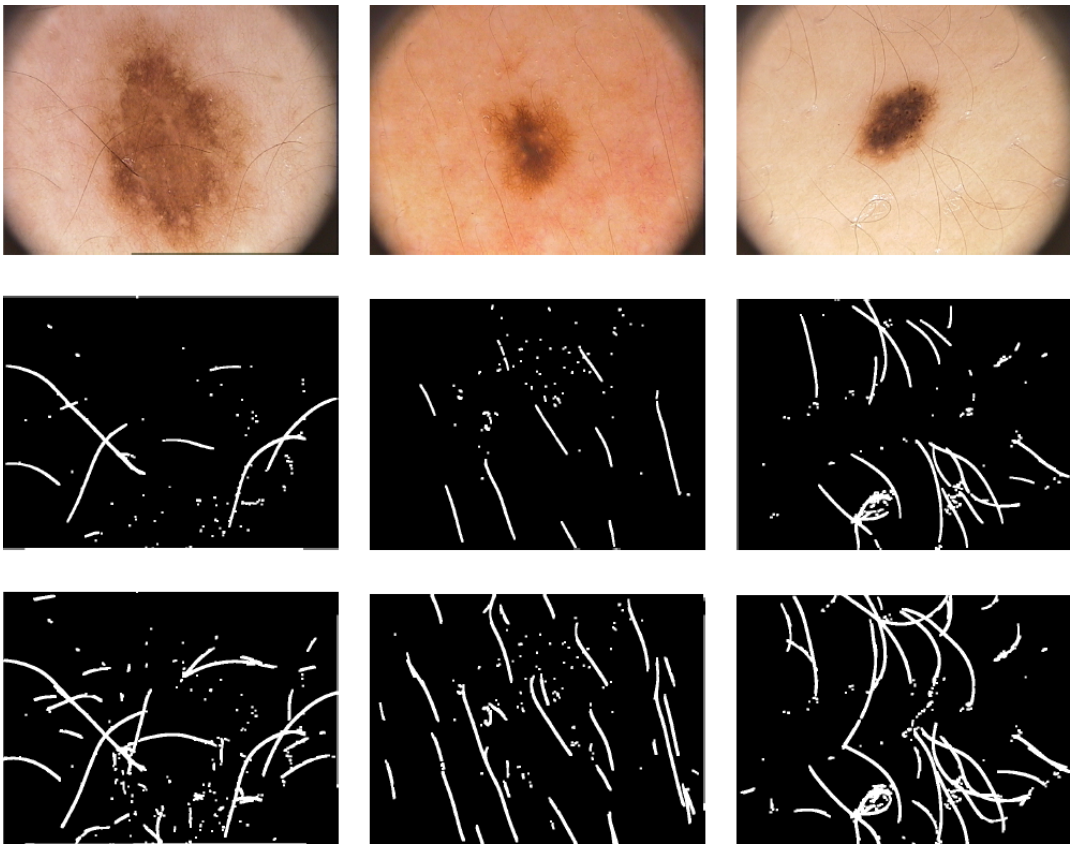


Figure 5.5: Examples of hair detection process: original images (Source:[101])(first row), results derived from Barata's method [25](second row) and our results (third row).

the parameters used to detect the pigment network lines are different from those used to detect the hairs. Theses parameters were set as: σ_{x1} =40, σ_{y1}=40, σ_{x2}=3 and σ_{y2}=0.5 with mask filter dimension of 11 x 11 and the number of rotations for θ_i was set as $N = 18$.

To detect the pixels which belong to the net of pigment networks, a threshold value is implemented and set as T_H=0.0185. Thus, the process replaces each pixel in the segmented image with a white pixel (1) if the image intensity value greater than threshold $I(x, y) > T$ and black otherwise (0). It is noteworthy that all these parameters were determined experimentally and the obtained results were promising for all images [25].

5.2.2.2 Network Detection

This step is to extract large linked lines of pigment networks, since it is assumed that the pigment network consists a set of connected lines. Therefore, a connected component analysis is used to identify the connectivity between these pixels. Connected component analysis (labeling) scans the segmented image (from top to bottom and left to right) and groups its pixels into components (objects), based on pixel connectivity. We used 8-connectivity (8 neighbours pixels), since it provides the information of vertical, horizontal, and diagonals connections. The pixels whose have similar value and connected to each other become one connected component (one group). Once all groups have been defined, each pixel is labeled with the same colour (colour labeling) according to the component to which it was assigned. The 8-connectivity of a pixel p(x,y) is defined as:

$$\begin{aligned} N_8(p) =\{&(x+1,y),(x-1,y),(x,y+1),(x,y-1),\\ &(x+1,y+1),(x-1,y-1),(x+1,y-1),(x-1,y+1).\} \end{aligned}$$

To extract the large linked group or object, a threshold value is applied for each connected components, each connected component in the binary image is extracted and classified as a pigment network if its area is greater than the threshold value, and it is given by the following condition.

$$A(R_c) > A_{min}$$

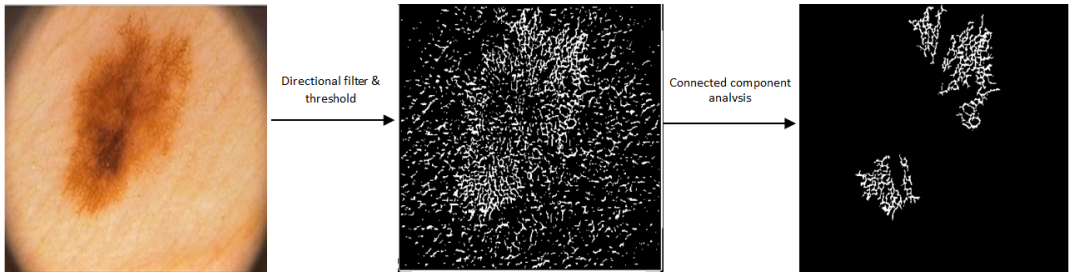


Figure 5.6: Example shows the process of pigment network detection. Left: the pre-processed image. Medial: the output of the directional filters bank and threshold . Right: Final outcomes of the connected component analysis (the largest connected groups of pigment networks).

where R_c is the cth connected region, $A(R_c)$ is its area, and A_{min} is a threshold value which was set as 900 experimentally. By applying the previous condition, all connected components with a small area can be classified as noise. An example of pigment network detection is shown in Figure.5.6. The obtained image is a binary "net mask "of all connected regions according to following equation:

$$R = \bigcup_{c:A(R_c)>A_{min}} R_c$$

5.2.3 Feature extraction

It is well known that the diagnosis of melanoma in dermoscopy images using CAD systems usually requires a selection of suitable features, which in turn could assist the dermatologists in their diagnosis. As mentioned previously in Chapter 2, the features could be extracted from images as either globally or locally; however, consideration must be given to the fact that the extracted features should contain sufficient information to distinguish between two classes (melanoma or non-melanoma). On this basis, our proposed algorithm aims to extract the relevant features from the final binary images. The density and regular distribution of the pigment networks are the main properties, and they can be used to identify the images, where pigment network is present. By characterising these two properties of pigment network, we defined five features, and we believe that they will be quite good in distinguishing between all images.

Figure 5.7: Examples of lesion segmentation: Left: original dermoscopy image. Right: binary segmentation mask. (Source:[101])

All the five features that were identified to be useful in melanoma diagnosis require both: pigment networks detection and the border of the lesion areas. Therefore, it was necessary to segment the lesions, in order to extract the discriminant features. Since the lesion segmentation is not the main focus of this method, this phase was not automated. Consequently, the ground truth images were used as segmented lesions, which is actually a binary mask. One example can be seen in Figure 5.7. The extracted features from all dermoscopy images, which will be used in the final stage for melanoma classification are described as:

Network/Lesion Ratio: This value computes the area of the network R (see Section 5.2.2) with the area of the whole lesion

$$Network/Lesion = \frac{A(R)}{A(L)}$$

where A(R) is the area of the detected pigment network and A(L) is the area of the segmented lesion, which has been segmented by an experienced dermatologist (Figure.5.7 shows an example).

Network/Region Ratio: This value computes the area of the pigment network lines with the area of pigment network regions, because the pigment network consists of holes surrounded by pigmented lines. A simple morphological filling process is applied for all connected regions, to get the pigment network region B_c (see Figure.5.8(c)). Then a pigment network regions

mask is defined as the union of all the regions.

$$B = \cup_c B_c$$

The network/regions ratio is then obtained as follows:

$$Network/Region = \frac{A(R)}{A(B)}$$

where $A(R)$ is the area of the detected network R and $A(B)$ is the area of all the pigment network regions present in B.

Number Of Holes: As above mentioned, the pigment network consists of holes surrounded by pigment network lines. Therefore, this value computes the total number of the holes of the detected net. It can be simply obtained by subtracting R from B, where R is the detected image and B is the union of all regions. An example can be observed in Figure.5.8(d).

Holes/Lesion Ratio: This value compares the number of the holes with the area of the lesion.

$$Holes/lesion = \frac{H_n}{A(L)}$$

where H_n is the total number of the holes in the detected image and $A(L)$ is the area of the segmented lesion L.

Holes/Region Ratio: This value compares the number of the holes with the area of the detected region.

$$Holes/Region = \frac{H_n}{A(B)}$$

where H_n is the total number of the holes in the detected image and $A(B)$ is the area of all the pigment network regions B.

The same number of features was acquired from all images and used as image descriptor to classify the new images using machine learning techniques. These are described in the following section.

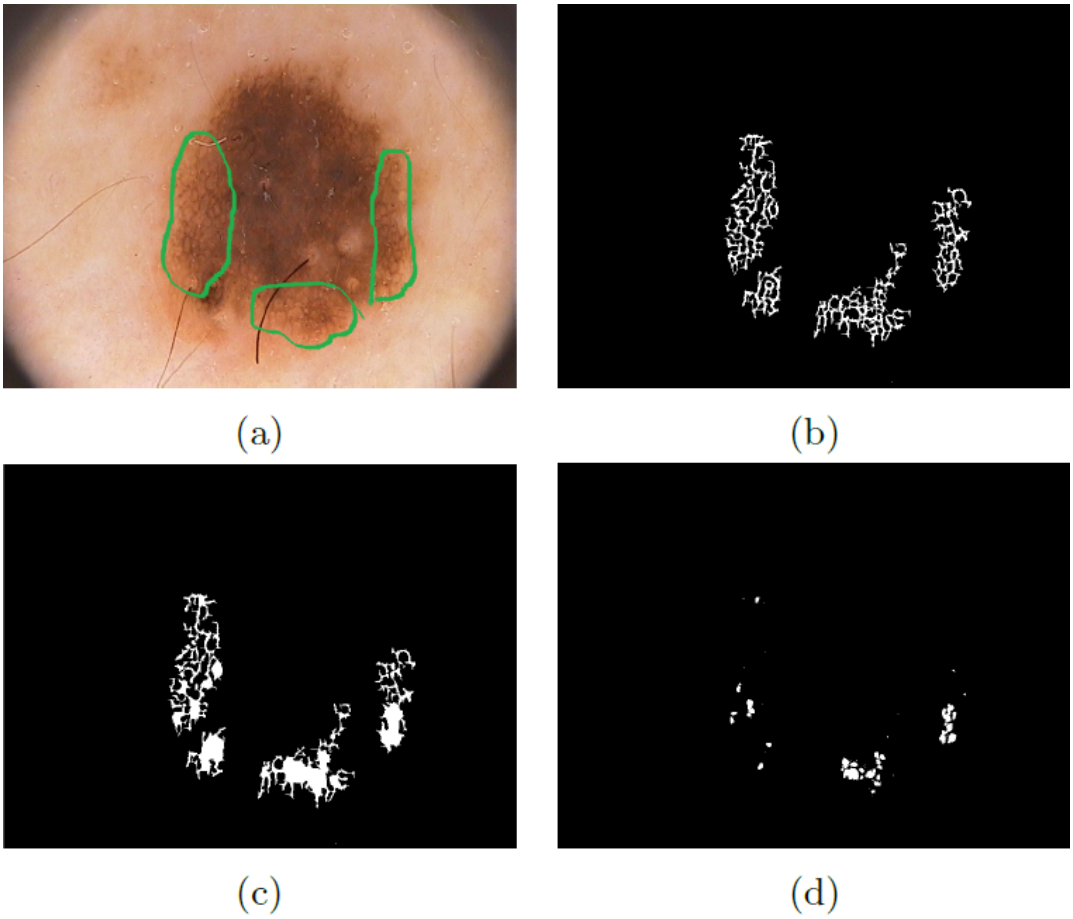


Figure 5.8: Pigment network masks. (a) original image with pigment network regions highlighted; (b) pigment network detection ;(c) pigment network regions-mask and (d) holes mask.

5.2.4 Lesion Classification

For the automatic classification of the pigmented skin lesions, many different classication methods were used in literature such as Support Vector Machine (SVM), K-Nearest Neighbour (KNN), AdaBoost, Artificial Neural Network (ANN), etc for the purpose of classifying a given lesion into one of two classes (melanoma or non melanoma). Before describing the lesion classification process, a brief definition is provided of several algorithms used for classification.

- Support Vector Machine (SVM): Support Vector Machines proposed by Cortes and Vapnik in 1995 [50]. It has been introduced recently in the statistical learning theory domain for classication and regression problems. Based on statistical learning, SVM provides an optimal design criterion for linear classier. The technique consists in nding the optimal separation surface between classes thanks to the identication of the most representative training samples of the side of the class. These samples are called support vector [102]. The separation hyperplane must be the one that has the largest distance to the nearest training pattern of each class. This distance is called margin, hence one can formulate the learning problem of SVM as that of nding the optimal hyperplane, that maximizes the margin of the training data. Occasionally, it is not possible to separate the training data using a hyperplane, since the two clouds of training features overlap. A brief description of the classification process is given as:

  A two class classication problem can be stated the following way: N training sample are available and can be represented by the set pairs $(y_i, x_i), i = 1, 2, ..., N$ with y_i a class label of value ś1 and $x_i n$ feature vector with n components. The classier is represented by the function $f(x;)ßy$ with the parameters of the classier. The SVM method consist in nding the optimum separating hyperplan so that:

  - Samples with labels $y =$ ś1 are located on each side of the hyperplane.
  - The distance of the closest vectors to the hyperplane on each side of the maximum. These are called support vectors and the distance is the optimal margin (see Figure 5.9).

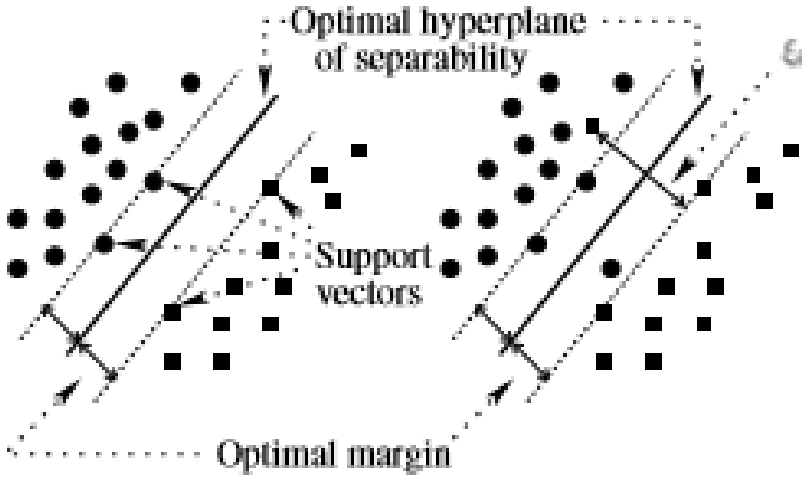


Figure 5.9: Generalized optimal separation hyperplane: Left: Linear separability. Right: non linear separability. (Source:[102])

The hyperplane is dened by $w.x+b = 0$ where (w, b) are the parameters of the hyperplane. The vectors that are not on this hyperplane lead to: $w.x + b0$ and allow the classier to be dened as: $f(x;) = sign(w.x + b)$. The support vectors lie on two hyperplanes, which are parallel to the optimal hyperplane, of the equation: $w.x + b =$ ś1. The maximisation of the margin with the equations of the two support vector hyperplanes leads to the following constrained optimisation problem:

$$min\frac{1}{2} \parallel w \parallel^2 with y_i(w.x + b) >= 1, i = 1, ..., N.$$

If the training data set cannot be separated linearly, a kernel method is used to simulate a non-linear projection of the data in a higher dimension space, where the classes are separated linearly. Different kernel functions can be used and the most popular is the Gaussian radial basis function (RBF), dened as follows:

$$Kernel(x_i, x_j) = exp(-\gamma \parallel x_1 - x_j \parallel^2),$$

where x_i, x_j are feature vectors associated with patterns i and j, and γ is a parameter

that must be optimised and dene the width of the kernel.

- K-Nearest Neighbour (KNN): The algorithm has a very simple formulation, given a training set of patterns (feature vectors) for which the classes are known, each new pattern will be classied in the same class as that of the closest training pattern (called nearest neighbour). The comparison between patterns is performed by computing the distance between feature vectors. It is possible to determine the class of the test pattern by taking into account not only one, but the k closest training patterns (neighbours). In this case, the class of the pattern is that which is more common among the selected training patterns. Two parameters must be optimised for this algorithm: the number of neighbours k and the distance metric used to compare the feature vectors of training and test images.

- AdaBoost: The AdaBoost classier was originally proposed by Freund and Schapire [63]. It is a parallel classier combined with many linear weak classiers. every weak classier only focuses on the classication of one dimension in the input feature vector. During the entire training process, after the goal is given to the classier, the algorithm is able to self-adaptively increase the number of weak classiers, so as to improve the overall accuracy rate of the classication and focus on key features. Following the addition of a weak classifier, the algorithm uses the minimum error to calculate its weight value and readjust the weight value of every training example. The value is then passed to the next newly-added weak classier. Based on the newly-added weak classier, the effect of the overall parallel classier is improved. The summary of the algorithm and how to obtain the strong classifier is illustrated below.

  - Assign weights to each training pattern.
  - For W iterations:
    - Learn a stump (weak) classier for each feature component and determine its performance.
    - Choose the weak classier with the lowest classication error.

    - Update the weights of the training patterns: increase the weight if a pattern is misclassied and decrease it otherwise.
  - Dene the strong classier as a linear combination of the W weak classiers.

- Artificial Neural Network (ANN): An artificial neural network is a mathematical model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of several artificial neurons that are highly interconnected. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information, that flows through the network during the learning phase. The neurons are arranged in layers, the neurons in the first layer called the input layer and receive the feature vector of an object. Information will flow from this layer until the output layer. If it is a multilayer network, there will be intermediate layers called hidden layers, which include an activation function. The most commonly used and the simplest type of ANN is the feed-forward neural network, it named like this because the information flows in one direction only (forward) from the neurons in the input layer to the neurons in the output layer (see Figure 5.10). They are also called supervised networks because they require a desired response in order to be trained. In order to train a network, the most common algorithm used is the back propagation algorithm [57]. At the training stage, the feature vectors are applied to the input of the network, and the desired output classes are known. Once the information reaches the output layer, where the input features result in a class label, the back propagation algorithm runs the network in the opposite direction, updating the weights and biases. These are often initialised randomly between the connected neurons until all examples are correctly classified or a stopping criterion is reached.

Our intention was to use the ANN algorithm to determine lesions type, since it is simple and easy to implement; moreover, it achieves competitive performances when the size of the dataset and/or feature space is small, as is the case of this work.

It is well known that after extracting features from images and storing them in a database, machine learning techniques are needed in order to learn the system, which in turn could

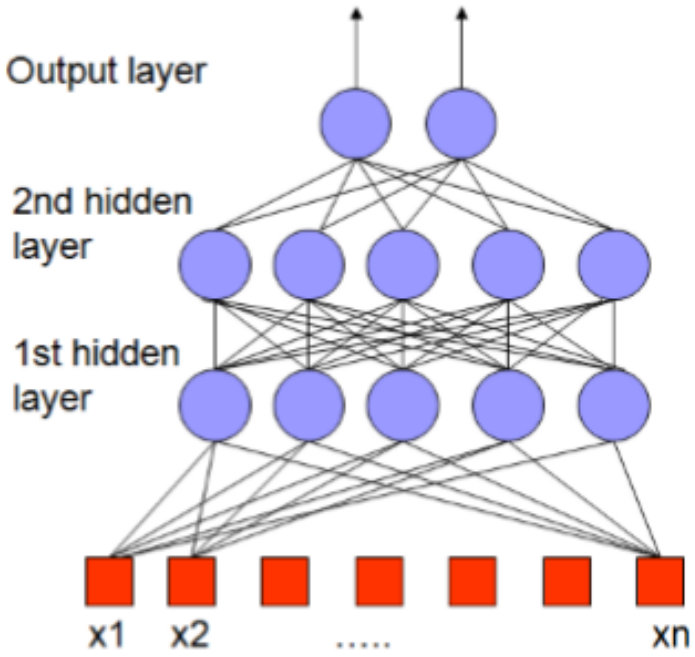


Figure 5.10: Feed forward neural network. (Source:[57])

classify new cases properly. Classifier performance always depends on the characteristics of the data to be classified, and no single classifier works efficiently on all problems. Thus, in melanoma detection problem, the feed forward neural network based upon the scaled conjugate gradient back-propagation methods with a sigmoid function in hidden layers as a transfer function has been selected, since it is one of the most popular classifiers and has been used in previous works related with malignant melanoma detection.

Scaled conjugate gradient back-propagation is a network training function, that updates weight and bias values according to the scaled conjugate gradient method. Back-propagation is used to calculate derivatives of performance with respect to the weight and bias variables. The advantages of using this function in lesion classification include not performing a line search at each iteration, which is computationally expensive [108] [53].

In our approach, the discriminant features of the pigment networks were extracted from images and stored in a database as numerical data, which is then used as input to a classifier, in order to determine the image type as a normal image or abnormal. Accordingly, the data

is divided into two cases: training and testing, then the training data with the known target are used to feed an Artificial Neural Network (ANN) as a binary classifier. As a target: **(0)** indicates an abnormal skin and **(1)** indicates a normal skin. Two-layers feed-forward network (hidden layer and output layer) are used with 10 neurons in a hidden layer. The sigmoid transfer function is used to map the sum of the weighted inputs to the output of the neuron. Therefore, the network is trained with supervision using a gradient descent training technique called back-propagation, which minimises the squared error between the actual outputs and the desired outputs of the whole network. That means the hidden and output layer nodes adjust the value of the weight depending on the classification error to predict the correct class. Once the network is trained with the known target, the testing data (unknown images) are used as input to the network to assess its performance in classifying new cases as normal or abnormal.

5.3 Experimental Results

The system described in the previous sections was evaluated using a public database PH2 [101], which provides 200 dermoscopy images (84 with pigment network and 116 without pigment network). Each image was classified by an experienced dermatologist as either melanoma or non-melanoma, which in turn used as ground truth. These images were taken from the database of Hospital Pedro Hispano, Matosinhos, and were obtained by dermatologists during clinical exams using a dermatoscope with a magnification of 20×. They are acquired in 8-bit RGB colour and stored in bitmap (BMP) formats with an average resolution of 573x765. Three criteria are used to evaluate the performance of our proposed method; namely, sensitivity, specificity, and accuracy.

The aim of this step is to evaluate the ability of our approach to classify images of skin lesions as malignant melanoma or benign. In order to fulfill this purpose, the performance of the network has to be tested. Consequently, new samples (set of similar data unused during training) are fed to the network and then classified. Then, the predicted classes are compared to the known classes. Thus, to be more accurate for lesion classification, the training and

Table 5.1: The average results of lesion classification.

No	Sensitivity	Specificity	Accuracy
1	83.3%	95.8%	93.3%
2	95.0%	96.3%	85.3%
3	100.0%	90.3%	90.0%
4	100.0%	96.0%	90.0%
5	88.7%	92.0%	86.7%
6	100.0%	100.0%	93.3%
7	90.0%	92.6%	86.7%
8	86.0%	100.0 %	83.3%
9	87.6%	100.0 %	96.7%
10	100.0%	96.0%	94.0%
11	100.0%	89.3%	90.0%
12	83.3%	87.5%	86.7%
13	100.0%	96.2%	93.7%
14	100.0%	92.9%	86.7%
15	83.3%	96.0%	93.3%
16	80.0%	96.7%	90.0%
17	100.0%	94.5%	93.3%
18	83.3%	91.7%	86.3%
19	100.0%	92.3%	93.3%
20	86.0%	100.0 %	90.0%
Average	92.3%	94.8%	90.1%

testing operations were repeated 20 times, and the three evaluation criteria namely sensitivity, specificity and accuracy were obtained as can bee seen in Table 5.1. The final average of all the above mentioned criteria shows that the method performed well in identifying lesions type and achieved promising results, it successfully achieved 92.3% sensitivity, 94.8% specificity and 90.1% accuracy. This refers to that the method is able to determine the lesion types with high accuracy.

Quantitative comparison between various methods is difficult because different datasets and criteria have been used. However, we were able to evaluate our method against Barata et al [25], as they have the same objectives and are based on the same dataset. Table 5.2 and Figure 5.11 present the comparative results. They indicate that our method achieved better results and outperforms Barata et al [25] by three evaluation criteria namely: sensitivity, specificity and accuracy. For instance, the sensitivity and specificity increase from 91.1 to

Table 5.2: Results of the lesion classification SE-sensitivity, SP-specificity and AC-accuracy.

Work	SE	SP	AC
Barata et al[25]	91.1%	82.1%	86.2%
Proposed method	**92.3**%	**95.0**%	**90.1**%

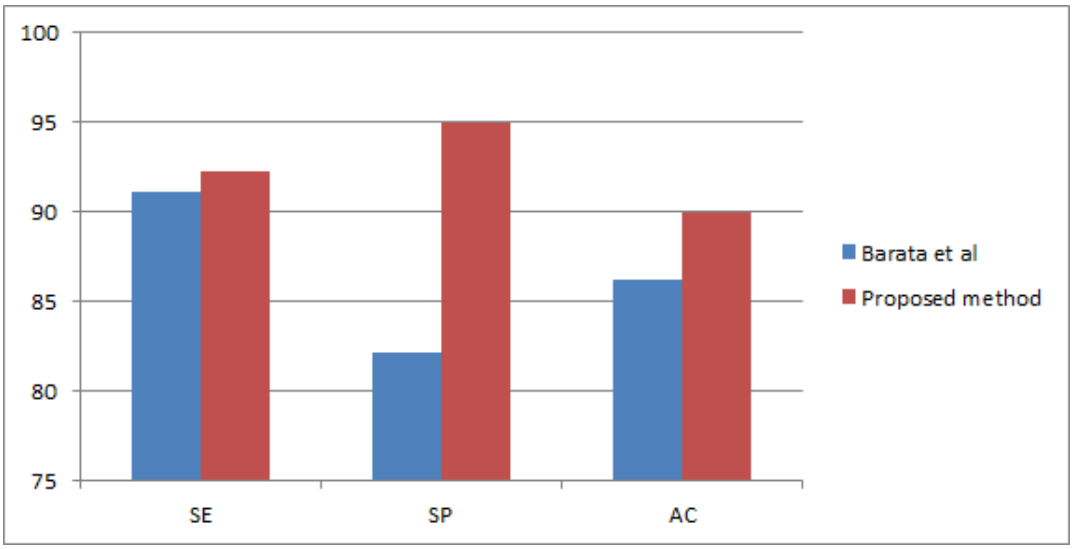


Figure 5.11: Results of lesion classification SE-sensitivity, SP-specificity and AC-accuracy

92.3 and from 82.1 to 95.0 respectively, which means the correct classification of lesions were increased by 1.31% sensitivity and 15.71% specificity. This refers to that melanoma images were classified as melanoma cases correctly, and non-melanoma images were identified as normal skin. With respect to the accuracy, it raises from 86.2 to 90.1, this means an increase of approximately 4.5% percent. As a conclude, removing hairs from images using Sobel filter after applying Gabor filter made the method accurate and robust; also able to classify the given lesions as either normal or abnormal skin. In addition, increase all the above criteria and obtained high results; as well as located the pigment network structures, which can assist dermatologists in making visual observations and the right diagnostic decisions.

It is important to highlight that the manual segmentation images are required to extract the desired features from actual lesions. Since the lesion segmentation phase is not the primary focus of this method and the method presented in Chapter 4 failed to segment the actual whole lesions in many cases, especially melanoma images, as presented in Figure 4.2. This led

subsequently to a huge decrease in terms of sensitivity, specificity, and accuracy, as mentioned in Section 4.4. Moreover, the extracted features are dependent on the affected lesions in three types (common moles, atypical nevi and melanoma images), the ground truth images are adopted to obtain the required features from the real affected lesions.

5.4 Summary

The work reported in this chapter described a process of automated enhancement and segmentation of the pigment network structures in dermoscopy images, in addition, feature extraction and classification. In order to perform the segmentation of the pigment network features, and determination of lesions type as melanoma or normal skin, the algorithm started by enhancing the appearance of several noises in the image. Thus, the enhanced images were obtained by applying the same algorithm presented in Chapter 3 Section 3.3.1.2. The outcomes of these process improve the robustness and the accuracy of the segmentation step and classification result. Due to the similarity between the hair and the pigment network lines, the segmentation step was also applied by using Gabor filter with different parameters. Consequently, connected component analysis with threshold approach was applied to detect the large connected objects, which in turn indicate the pigment network lines. Several features were extracted and used to feed the Artificial Neural Network as classifier approach. The proposed method was tested on a dataset of 200 dermoscopy images including 84 with pigment network and 116 without pigment network. The evaluations also show that our method achieved exceptional performance against alternative methods we compared to.

Although the final results of melanoma detection were quite promising, user intervention was necessary to determine and segment the affected lesions; leading ultimately to the construction of a semi-automatic system. Thus, to construct an integrated diagnostic system without user intervention, an appropriate algorithm should be selected and implemented. Therefore, a new fully automated approach for lesion detection is presented in the next chapter.

Chapter 6

Skin Cancer Detection in Dermoscopy Images Using Sub-Region Features

6.1 Introduction

The early detection of skin cancer is extremely important since it can often be cured with a simple excision[148]; thereby reducing the mortality rate. Dermoscopy is one of the major tools used in its diagnosis and is used widely by dermatologists, due to its value in the early detection of malignant melanoma. The tool provides better visualisation of several pigmented structures that are invisible to the naked eye, such as streaks, dots, pigment networks and blue-white areas [159].

In the last two decades, computer-aided diagnosis (CAD) systems of melanoma detection have diminished the gap between the medical and engineering knowledge, since these systems try to mimic the performance of dermatologists when diagnosing a skin lesion area. Thus, they help differentiate more quickly between melanoma and benign lesions [24] [134].

In dermoscopy image analysis, images and segments the lesion areas (ROI) are typically enhanced first. This is followed by extracting several features, which could be local or global

and using them to learn an appropriate classifier, in order to predict the lesion label (melanoma or non-melanoma). Each step of above process depends on the previous one. For instance, the classification stage depends on the performance of all previous steps. Therefore, to achieve a high classification rate and increase the accuracy of the diagnosis of skin cancer, all or most of the previous steps should be implemented with the best strategy.

It is well-known that the skin lesion classification methods are based typically on the feature extraction. Therefore, the extraction of representative features of the lesions under analysis is a very important stage for efficient classification [113]. In the ABCD-E rule, the colour features which indicate the (C) parameter describes the colour properties inside the lesion thus; many different features can be obtained such as statistical descriptors (mean, standard deviation, skewness, entropy, etc) and colour histograms. For this purpose, we focus on the best way to extract the desired colour features from images, using an improved method to segment the lesion.

In this chapter, four main steps have been implemented to build an automatic process for detection of malignant melanoma in dermoscopy images. First, a pre-processing step was applied on each image, for the purpose of removing undesirable artifacts. Second, the Particle Swarm Optimization (PSO) [56] and the Markov Random Field (MRF) [66] methods are integrated to segment the pigmented lesion areas. Therefore, the k-means was applied on the segmented image (lesion), in order to separate each homogeneous set of pixels in one group (cluster). Consequently, several features are extracted from the given lesion based on existing clusters. By using a trained classifier, the lesion is classified as either benign or melanoma. The proposed method provides a high performance in automatic melanoma detection on dermoscopic images. This chapter is structured as follows: Section 6.2 explains the proposed method including image pre-processing, image segmentation, feature extraction and lesion classification. Section 6.3 illustrates the experimental results. Finally, a summary of the chapter is presented in Section 6.4.

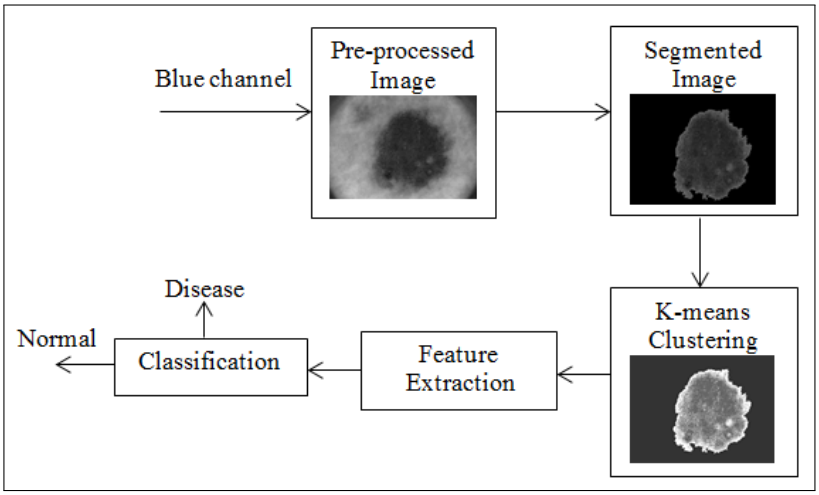


Figure 6.1: Overview of the proposed approach for skin lesions classification.

6.2 The Proposed Method

The proposed method for skin cancer detection is described and discussed in this section. As an initial step, the quality of the image is improved by detecting and removing several artifacts, which have been described in previous chapters. Then, the skin lesion which is expected to be a melanoma or non-melanoma is segmented from the surrounding healthy skin by applying PSO with MRF methods. Therefore, the k-means approach was applied on the segmented image (lesion) to divide it into sub region (clusters). Thus, a number of features, such as the colour moments and colour histogram, are extracted from each sub region and used as input to an Artificial Neural Network (ANN) classifier, in order to be trained and determine the skin lesions type. The overview of the proposed approach is illustrated in Figure. 6.1.

6.2.1 Pre-Processing of Dermoscopic Images

As mentioned in previous chapters the dermoscopic images do not have the expected quality to perform the diagnostic analysis. Moreover, the pre-processing stage is quite important and mandatory, since the noise presence in images affect the subsequent stages. The improvement of the image quality should be performed as a first step. Thus, to improve the quality of images and prevent false positive detection; as well as obtain a high accuracy of melanoma

detection, we adopt the same algorithm described in Chapter 3 Section 3.3.1 to detect and remove most of the hair from images.

6.2.2 Image Segmentation

The purpose of image segmentation is to separate the homogeneous lesions from the surrounding healthy skin. It is the most important phase for analysing images properly, since it affects the subsequent steps accuracy. However, appropriate segmentation in dermoscopy images is a daunting task and very challenging, because the lesions have large variations in size, shape and colour. As well as, the existence of low contrast between the lesions and surrounding healthy skin. In this work the image segmentation stage was implemented in two steps, the lesion area was segmented in the first step, while the sub region inside the segmented lesion was segmented in the second.

6.2.2.1 Skin Lesion Segmentation

The automatic segmentation approach was implemented to extract the whole lesion areas. The PSO and the MRF methods were combined, in order to minimise the energy function. The image segmentation is formulated as an optimization problem of the energy function with MRF theory. The PSO method is used to perform the initial labeling based on the optimal threshold value, which was obtained by maximising the fitness function. Then, an additional local search is performed for each segmented image by integrating it with MRF method. The purpose of this is to refine the segmented image by minimising the energy function or maximising the probability of pixel allocation to a cluster, by using Maximum A Post Priority (MAP). [66]. Accordingly, the pixels are reassigned to different classes based on the minimum cost. This process continues until a stopping criterion is met. More details of the approach can be found in Chapter 4 4.3.2.

6.2.2.2 Sub-Region Clustering

Our aim is to divide the segmented lesion into a few clusters, with a more homogeneous distribution of pixels for each cluster. A simple K-means approach is used for this purpose.

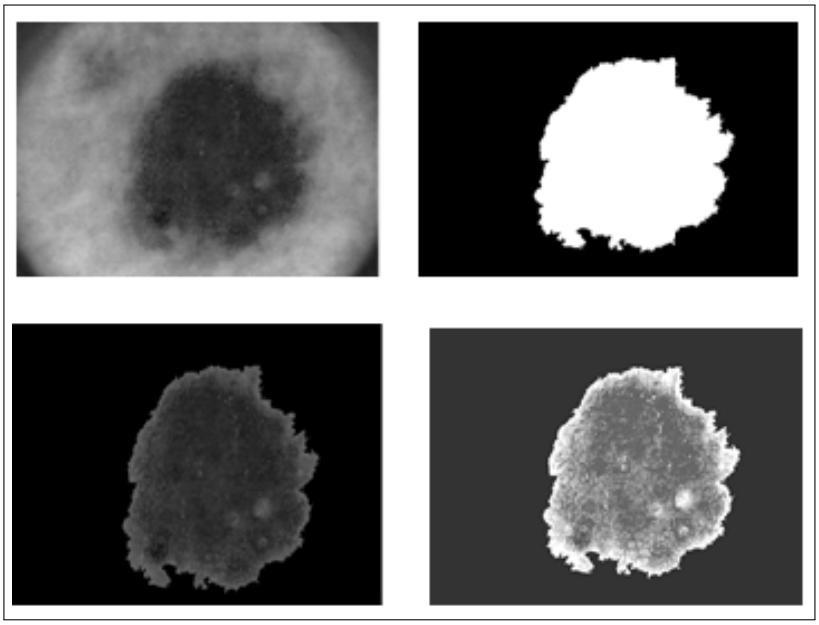


Figure 6.2: Segmentation and clustering: Blue scale image (pre-processed)(top left), segmented binary image (top right), segmented blue scale image, i.e. the lesion (bottom left) and result of clustering (bottom right).

The outcomes (binary masks or segmented lesions) from the previous step are multiplied by grey scale images (see the bottom left image in Figure. 6.2), to separate the pixels located inside the lesion into several groups. The K-means clustering method is used since it is very simple and has low computational complexity. Moreover, the number of clusters (k) could be determined easily. Hence, the number of clusters has been experimentally obtained and set as k=5, (discussions in section 6.3).

The Euclidean distance is used to calculate the distance between the image pixels and the centroids of the clusters. Every single pixel was assigned to the appropriate cluster, based on its distance. Thus, the location of each cluster was updated and the pixels were re-assigned. This process continues until no more changes to cluster membership. The final result of this step is several homogeneous clusters, which can be used for the subsequent step of feature extraction. An example is given in Figure. 6.2.

6.2.3 Feature Extractions

Hundreds of features can be acquired from each dermoscopy image and used as an image descriptor. However, not all are appropriate for lesion classification. Too many irrelevant features make the classifier complicated and require more computational time, which also reduces the classification accuracy. The best features have to be able to represent the characteristics of the regions in skin cancer images. Therefore, the suitable number of features should be extracted, with the best way possible to distinguish between images. Accordingly, melanoma images can be discriminated significantly from the benign images by a classifier. On this basis, the segmented lesion images can be used to extract several numbers of features as the best way to address the region in isolation. The density and the regular distribution of the blue color are the main properties and can be used to identify the images. In this work, three colour moments with colour histogram are used as features, in order to determine the skin lesion type.

6.2.3.1 Colour Moments

Colour moments can be used to distinguish images based on their colour distribution. Typically, probability distributions are characterised by the number of unique moments [13]. Consequently, they can be used as colour features. The first colour moment can be interpreted as the average colour in each sub region inside the lesion, and can be calculated using the following equation:

$$E_i = \sum_{i=1}^{N} \frac{1}{N} Pi \ .$$

where N is the total number of pixels inside the sub region and Pi is the pixel value.

The second colour moment used as a feature is the standard deviation, which can be obtained by taking the square root of the variance of the colour distribution.

$$\sigma_i = \sqrt{(\frac{1}{N} \sum_{i=1}^{N} (pi - E_i)^2)} \ .$$

where E_i is the average value and N is the total number of the pixels inside the sub region.

The third and final colour moment used in our approach is the skewness; in other words the colour distribution asymmetry. Therefore, this provides useful information pertaining to the shape of the colour distribution. Skewness can be calculated as:

$$S_i = \sqrt[3]{(\frac{1}{N}\sum_{i=1}^{N}(pi - E_i)^3)} \ .$$

6.2.3.2 Colour Histogram

Colour histogram is a way to represent the distribution of the composition of colours in images. It reveals the number of pixels in each type of colour in the image. The histogram associated with the blue colour component $I_c, c \in \{3\}$ is given by:

$$h_c(i) = \frac{1}{N}\sum_{x,y} b_c(I_c(x,y))i = 1, ..., B_c \ .$$

where N is the number of pixels inside the sub region, i is histogram bin, B_c in the number of bins and $b_c(.)$ is the characteristic function of ith.

$$b_c(I_c(x,y)) = \left\{ \begin{array}{ll} 1 & I_c(x,y) \quad \in \text{ith colour bin} \\ 0 & \text{otherwise} \end{array} \right\} \ .$$

The bins are defined by dividing the colour component range into intervals with the same width. For all histograms, the number of bins is given by:

$$B_c \in \{5, 10, 15, 20\}.$$

It was found that the best performance was achieved when the number of bins set as 10 (discussions in section 6.3). Thus, the elements of each sub region are divided into 10 equally spaced bins.

The colour moments and colour histogram features are concatenated into one vector, to represent the image. It is noteworthy that all the extracted features are obtained from only the blue colour moment, since it provides the best discriminatory performance. The same

number of features was obtained from all images, three features related to the colour moment and ten features from the colour histogram. These features are used as an image descriptor and the images are classified using machine learning techniques.

6.2.4 Skin Lesion Classification

Based on the extracted features from images. The classification process of skin lesions must be performed to a very high standard, because it supports the decision of the dermatologists when diagnosing melanoma. Therefore, the Artificial Neural Network (ANN) as a classifier is used to determine the image type as normal or abnormal. The output of the classifier is (1) for melanoma and (0) for non-melanoma. The Radial Basis Function (RBF) is used as an activation function and given by: [36] [14].

$$h(x) = exp\left(\frac{-(x-c)^2}{r^2}\right)$$

where c is the center of bell-shaped Gaussian and r is the width.

A radial basis function (RBF) network is a special type of neural network, that uses a radial basis function as its activation function. It is a type of feed forward neural network composed of three layers, namely input, hidden and output layer [89]. It adds new neurons to the hidden layer of a radial basis network until the specified mean squared error goal is met or a maximum number of neurons is reached. The error of the new network is checked if it is low enough, the network is finished. Otherwise, the next neuron is added. The advantages of using this function in lesion classification images include designing a radial basis network often takes less time than training a sigmoid network, and can sometimes result in the use of fewer neurons [98]. Therefore, the network was constructed by two-layers (hidden layer and output layer) with 100 neurons in the hidden layer. All the discriminant features were extracted and stored in a database as numerical data, which was divided into training and testing. For the purpose of training the classifier, the training data with its target are used as input to the Radial Basis Function Network to make a prediction. As a target: **(1)** refers to melanoma and **(0)** indicates a normal skin. Once the network is trained with the known target, the testing

data, which is the data that is not yet trained (unknown images), are used as input to the network to assess its performance.

6.3 Experimental Results

The proposed approach was evaluated on a public database PH2 [101] which provides 200 dermoscopic images. Each image in the database was classified by the dermatologists as either normal or abnormal (the ground truth labels). The extracted features from all images were used as an input to the ANN classifier, with the ground truth labels as the output. The training and testing process was performed by using a 5-fold stratified cross-validation method. The images were divided into five subsets, each of which contained approximately the same number of melanoma and non-melanoma images.

We used five common evaluation criteria i.e. Sensitivity or True Positive Rate (TPR), Specificity or True Negative Rate (TNR), Accuracy (AC), False Positive Rate (FPR), and Mean Square Error (MSE). The AC is defined as the sum of the true positives (images correctly classified as melanoma) and the true negatives (normal skin correctly identified as a non-melanoma), divided by the total number of images. The TPR is defined as the total number of true positives divided by the total number of images marked in the ground truth as melanoma. The TNR is defined as the total number of true negatives divided by the total number of images marked in the ground truth as normal skin. The FNR is calculated as the total number of false negatives divided by the number of images marked as non-melanoma in the ground truth image. It is noteworthy that a perfect classification would have a sensitivity (TPR) of 1 and FNR and MSE of 0.

The obtained results of our method are presented in Table 6.1. The training and testing operations were repeated 5 times for greater accuracy in the lesion classification results; thus, the average of all the above mentioned criteria was calculated. It is clear that the method yielded promising results regarding all aforementioned evaluation criteria. The average of the sensitivity, specificity and accuracy were approximately 99.00%, 85.00% and 96.00% respectively. This means the method was able to distinguish between lesions type with high accuracy.

Table 6.1: Results of lesion classification in dermoscopy images.

No	SE (TPR)	SP (TNR)	AC	FNR	MSE
1	0.9874	0.9000	0.9698	0.01257	0.0470
2	1.0000	0.8750	0.9748	0.0000	0.0420
3	0.9748	0.8250	0.9447	0.0251	0.0592
4	0.9937	0.8250	0.9597	0.0062	0.0474
5	0.9811	0.8250	0.9497	0.0188	0.0514
Mean	0.9874	0.8500	0.9597	0.0125	0.0494

Moreover, the FNR and MSE averages were quite low, which means a few melanoma images were classified incorrectly as non-melanoma.

Making a quantitative comparison of various methods is difficult since different datasets and criteria have been used. However, we evaluated the performance of our approach against Barata et al [25] [24] [22], Alfed et al[13], Eltayef et al[58], Marques et al[99], Barata et al[26], Barata et al[21], Riaz et al[122], Barata et al[27] and Lei et al [33] as they have the same objectives, and they are based on the same database. Table 6.2 presents the comparison results.

To facilitate the performance comparison between our method and the alternate approaches, parameters, such as sensitivity (SE), specificity (SP), accuracy rates (AC) and balance accuracy (average of sensitivity and specificity, denoted as BAC) are computed for each method against the ground truth. Table 6.2 presents the results of performance comparison. Out of the four criteria, the proposed method performed better than all other methods by SE and AC; however, it is not as good as Eltayef [58] and Barata [22] in terms of SP and BAC, respectively. However, it is worth mentioning that the ground truth was used as the segmented images in both methods, while we did both segmentation and classification from the original images. Moreover, the work presented in [22] requires more computational time, because a number of image patches are used for feature extraction, and this could be expensive when the patch size is large. Furthermore, the late fusion strategy was used in their work, which requires additional time to classify each type of features separately. Consequently, our method achieved higher accuracy in identifying melanoma skin cancer compared with the

Table 6.2: Performance comparison with several methods.

Method	SE	SP	AC	BAC
Barata et al[25]	91.10%	82.10%	86.20	86.60%
Barata et al[24]	93.00%	85.00%	——	89.00%
Barata et al[22]	98.00%	90.00%	——	**94.00**%
Alfed et al[13]	91.00%	85.00%	——	88.00%
Eltayef et al[58]	92.30%	**95.00**%	90.00	93.65%
Marques et al[99]	94.10%	77.40%	79.10%	85.75%
Barata et al[26]	85.00%	87.00%	87.00%	86.00%
Barata et al[21]	92.5%	76.30%	84.30%	84.40%
Riaz et al[122]	84.00%	94.00%	——	89.00%
Barata et al[27]	96.00%	80.00%	——	88.50%
Lei et al [33]	87.50%	93.13%	92.00%	90.31%
Proposed method	**98.74**%	85.00%	**95.97**%	91.87%

state-of-the-art methods. This provides a clear indication that the method can be used in the early detection of melanoma.

It is worth mentioning that the comparison with the method presented in Chapter 5 (pigment network detection) was undertaken, in order to present that which obtained greater accuracy in the classification of lesions. The method presented in Chapter 5 used five extracted features to identify the given lesions type as melanoma, and it successfully achieved 92.30% sensitivity, 95.00% specificity and 90.00% accuracy. Whereas, the current method used colour moments and colour histogram to extract relevant required features. The method produced exactly 98.74% sensitivity, 85.00% specificity and 95.97% accuracy. It is clear that the current method yielded better results than the previous one in terms of sensitivity and accuracy. The sensitivity and accuracy are increased from 92.30 to 98.74 and from 90.00 to 95.97 respectively, which means the increases average of sensitivity is approximately 6.97% and specificity is about 6.63% percentage. This indicates that many melanoma images were wrongly classified as non-melanoma cases in the previous method, and correctly detected in this method. Specificity decreased significantly from 95.00% to 85.00%, which means the percentage decrease is about 11.76%. However, the method presented in Chapter 5 is a semi-automatic method, since it used the ground truth as a segmented image to extract various of features, while the

Table 6.3: Results of determining the number of clusters.

No of clusters	SE	SP	AC	FNR	FPR
3	0.9836	0.8100	0.9487	0.01635	0.1900
4	0.9849	0.7700	0.9417	0.01509	0.2300
5	0.9874	**0.8500**	**0.9597**	**0.01132**	**0.1500**
7	**0.9886**	0.7800	0.9467	0.01635	0.2200
8	0.9861	0.8300	0.9547	0.01383	0.1700
10	0.9761	0.8150	0.9437	0.02389	0.1850

current one is a fully automatic method. It performed the full process including image pre-processing, segmentation, feature extraction and classification automatically and without user intervention. Moreover, it determined the lesion types with high accuracy. Thus, we believe the method has great potential to detect malignant melanoma at an early stage, and support the clinical diagnosis.

Furthermore, to determine the number of clusters, we selected 6 numbers randomly and then apply k-means method inside the segmented lesions based on these numbers separately. At each stage, where the k-means is applied, we extract the mentioned features and use them as input to the ANN classifier. Thus, we evaluate the performance of the lesion classification by the five fore mentioned criteria. As can bee seen in Table 6.3, it is clear that the best performance was achieved when the number of clusters (k) was set to 5, since it gives the highest accuracy and specificity, alongside the lowest false negative and false positive rate. Therefore, it is selected as the best number of clusters.

The same strategy has been adopted to define the best number of bins. Table 6.4 illustrates the experimental results of determining the best number of bins value from four random numbers. It was observed that the higher accuracy and specificity were obtained when the number of bins was set to 10, beside that the lower false negative and false positive rate was acquired. Thus, the value of bins was set to 10.

It is important to state that the experimental results in Table 6.3 reveal that the results for each single metric are very close, which means the method is not impacted significantly by a different number of clusters. For instance, the sensitivity metrics are quite similar and high

Table 6.4: Results of determining the number of bins.

No. bins	SE	SP	AC	FNR	FPR
5	**0.9886**	0.8350	0.9577	0.01132	0.1650
10	0.9874	**0.8500**	**0.9597**	**0.01132**	**0.1500**
15	0.9886	0.8250	0.9557	0.01257	0.1750
20	0.9861	0.8400	0.9567	0.01383	0.1600

in terms of all clusters, and the same applies to the accuracy metrics. Moreover, the same thing occurred in Table 6.4 when the number of bins was defined. This is a clear indication that our method does not require complex parameters or specific numbers of bins or clusters to classify the lesions with a high accuracy, which is an additional advantage of the proposed method.

Conversely, the combination of the features derived from pigment network detection presented in Chapter 5 with the current extracted features was performed in an attempt to boost the accuracy of lesion classification. The overall classification performance results of 200 images are presented in Table 6.5. Moreover, the training and testing operations were repeated 5 times to ensure the result of lesion classification is correct. The average of all evaluation metrics reveals that the method performed well and obtained higher results. It achieves approximately 97.00% sensitivity, 83.00% specificity, and 94.10% accuracy; thereby indicating that the method can be used in the detection of melanoma lesions. However, a slight decrease can be observed in the classification results with respect to all evaluation criteria, compared with the current method results. In particular, the sensitivity decreases to 0.9698 compared to that presented in Table 6.1 (0.9874), the specificity reduced to 0.8300, the accuracy becomes 0.9417 as opposed to 0.9596. Moreover, the average of FNR and MSE are decreased from 0.0125 to 0.0301 and from 0.0494 to 0.0522 respectively. The reason of that is because, in the pigment network detection, the method extracted the net mask from inside and outside the lesion, which in turns led to producing irrelevant features. These features have reduced the average value of the classification accuracy. Moreover, undetected thin/grey hairs as mentioned in Chapter 3 Section 3.3.1.2 may produce wrong values of extracted features, such as

Table 6.5: Performance results of lesion classification in dermoscopic images using pigment network and sub-region features.

No	SE (TPR)	SP (TNR)	AC	FNR	MSE
1	0.9811	0.8000	0.9447	0.0188	0.0499
2	0.9685	0.8250	0.9396	0.0314	0.0509
3	0.9622	0.8500	0.9396	0.0377	0.0546
4	0.9685	0.8500	0.9447	0.0314	0.0564
5	0.9685	0.8250	0.9396	0.0314	0.0492
Mean	0.9698	0.8300	0.9417	0.0301	0.0522

number of holes, which led to a decrease in the accuracy of lesion identification. Despite, the reduction of all aforementioned criteria, the method still yielded very promising results and can be used as a tool for non-invasive diagnostic technique in dermatology, in addition to supporting dermatologists in their diagnosis.

6.4 Summary

In this chapter, a fully automatic approach to melanoma detection in dermoscopy images was performed. The input images are first pre-processed by detecting and removing the noise. Then the lesions are segmented by applying PSO and MRF methods. Therefore, the segmented lesion was used to extract a number of features, the K-means is applied on the segmented image, and the features such as the colour moments and colour histogram, are extracted at the sub-region (cluster) level. Consequently, The features are used as input to an ANN for final melanoma classification. The proposed approach achieved approximately 96.0% accuracy, 99.00% sensitivity and 85.00% specificity on a dataset of 200 images. A comparison of several alternative methods reveals that the proposed method achieved overall superior performance in terms of sensitivity and accuracy.

The main contributions of the work are as follows.

1. A comprehensive method including the whole process of image enhancing, segmentation of lesions and melanoma classification is developed.

2. A method for lesion segmentation is proposed by combining the PSO and MRF methods.

3. Feature extraction at the sub-region level is performed by separating the segmented lesions into homogeneous clusters.

The method proved to be flexible, robust, accurate, and fast; thereby leading to successful classification results. However, there remains scope for improvement in certain aspects of this study. This includes the removal of thin and grey hair from images without affecting lesion borders, improving lesion segmentation accuracy especially with melanoma images alone, and extracting more relevant features, which can increase the rate of the accuracy in diagnosing skin cancer.

Chapter 7

Conclusion and Future Work

7.1 Introduction

In recent years, incidences of skin cancer have become more widespread in countries including the USA, Australia and the UK; especially among white-skinned people who are exposed to the sun. Dermoscopy is considered one of the major tools in its diagnosis and is used widely by clinicians due to its value in early detection. Medical research has been accelerating through the use of computers and software to interpret medical data and provide meaningful information. The motivation for this is obvious; when performed manually, the analysis of medical images can be time consuming, tedious and subjective.

The introduction of computer-aided diagnosis (CAD) in dermoscopy image analysis signified the arrival of a new platform for the early detection of melanoma. Using computers in melanoma diagnosis has become a major research area in recent years with the aim of helping the dermatologists through the provision of useful information on the skin lesions. The structures, such as pigment network and region area, are considered key indicators of melanoma disease. Thus, effective segmentation of these significant structures can improve the productivity of dermatologists by reducing their time required while diagnosis process. This chapter is organized as follows: Section 7.2 reflects the overall aim/objectives of this study, Section 7.3 demonstrates the research limitations, and the future work is described in Section 7.4.

7.2 Research Summary

The main aim/objectives of this thesis as summarised in Chapter 1, were as follows:

- provide accurate, robust and reliable automated dermoscopy image analysis technique, for the purpose of facilitating early detection of melanoma skin cancer disease.
- Develop a new method to reduce the noise from dermoscopic images without corrupting image information, and produce images more suitable for further processing.
- Implement a new algorithm to extract the actual lesion boundaries in dermoscopic images.
- Develop an automatic algorithm for the detection of pigment network structures.
- Identify skin lesion types based on the presence or absence of the pigment network structures.
- Develop a fully automatic approach for melanoma detection in dermoscopy images based on the suitable number of features with an ideal extraction process.

To address the aforementioned research aim/objectives, we conducted a review of the literature in Chapter 2 to identify existing ndings alongside gaps in the state-of-the-art methods. Then, we proposed four innovative methods, including two for lesion segmentation, one for pigment network detection, and one for lesion classification. The whole experimental results were obtained on a public database PH2, alongside a comparison of existing methods in the literature, which revealed that our proposed approach is accurate, robust and efficient in the segmentation and classification of the pigmented lesions in dermoscopy images. The aforementioned objectives of this study are outlined in the following sections.

7.2.1 Region Segmentation

Determining the actual regions in dermoscopy images is quite important and necessary since it is responsible for identifying melanoma in the early stage by applying (ABCD) rule, which is an essential role for melanoma diagnosis. Over segmentation and under segmentation both

affect the lesion segmentation accuracy and lead to the wrong diagnosis later. Thus, region segmentation must be correct and accurate to obtain exact and meaningful information for diagnosing melanoma correctly. Chapters 3 and 4 provided two methods of region border detection in dermoscopy images, which can be used in modern dermatology to assist in the diagnosis of melanoma.

In Chapter 3, the automatic segmentation process of regions in dermoscopy images began by removing the noise from all images, since this increases the accuracy of the segmentation results. Therefore, the robust pre-processing approach was performed for the purpose of enhancing the quality of images and rendering them more understandable. It involves reflection detection using adaptive threshold, hair extraction using directional Gabor filter, image sharpening and Sobel filter and lightning reflection and hair repairing using image inpainting approach. The experimental results indicate that the method was able to detect and inpaint the noise pixels with highly accurate. The visual observations indicate that the pre-processed images were clearer than the original one, which led to making the region borders more visible. In terms of the region segmentation, the FCM and the MRF methods were used to detect and extract the region border, the FCM was executed for the purpose of segmenting the region area and providing the initial parameters. Therefore, the MRF method was implemented based on the estimated parameters to refine the segmented images by reassigning each image pixel to one cluster based on higher probability. A morphological operation was performed, to fill in the holes and remove the small few objects to present the segmented region as one connected region.

Another fully automatic method of region border extraction in dermoscopy images was presented in Chapter 4. Images must be pre-processed to enhance the quality of images and increase the accuracy of the segmentation results. Therefore, the same algorithm presented in Chapter 3 section 3.3.1 was applied to detect and remove most of the noise from images, as an image pre-processing phase. The obtained pre-processed images displayed the region area clearly, which helped significantly in the process of region segmentation. Combining the PSO with the MRF methods was performed to segment the region area from surrounding healthy skin. Based on the highest fitness function, the PSO method was able to segment the image

into the foreground (region) and the background. Therefore, to perform a local search and obtain a robust segmentation, the segmented image was integrated with the MRF approach. The process relabeled each pixel to a different class based on its neighbours. The main aim of this process was to reassign each pixel to optimal cluster based on its neighbors by getting the minimum cost of the energy function in each iteration. The morphological operation such as fill in function was used to remove or fill in the holes in the obtained segmented images.

It is worth mentioning that this method exhibited better performance opposed to our previous method proposed in Chapter 3, as it is capable of segmenting almost the whole actual region. This method used the best solution obtained form PSO algorithm, based on the highest fitness and shared it with MRF approach to performing a local search in each iteration. While in the method presented in Chapter 3, the FCM was used only to estimate the initial parameters. Moreover, in FCM, the initial points began randomly, which means sometimes no results were yielded; however, there is always a solution with PSO, since each particle provide one solution.

In terms of the qualitative and quantitative evaluation, the two proposed methods were tested on a dataset of 200 dermoscopy images, including 40 melanoma, 80 common moles, and 80 atypical moles. As a qualitative evaluation, the obtained segmented images of the proposed methods were compared with the ground truth lesions. The experimental results of the two methods indicated that they provided a high accuracy of skin lesion delineation. While with respect to the quantitative evaluation, the experimental results were also compared against a selection of several other methods in the literature on the same dataset and objects. Their comparison indicated that the two proposed methods achieved exceptional performance and outperformed the all selected methods. It is worth mentioning that, the proposed approach presented in Chapter 4 performed well in terms of all validation criteria. We believe it can address the image noise and provide high accuracy of skin lesion segmentation. Due to the great potential for detecting the region border correctly, the method can be used to help the dermatologists for visual observation of the lesion area.

It is important to state that the two proposed approaches outperformed all the other methods in terms of the accuracy of all the evaluation metrics in processing the benign images,

while the accuracy decreased with respect to melanoma images.

7.2.2 Pigment Network Detection

Pigment network is considered one of the significant features that can be used to identify melanoma in the early stage. Therefore, Chapter 5 proposed a novel approach to pigment network detection in dermoscopy images. The algorithm was started by applying a robust proposed approach described in chapter 3 section 3.3.1 for the purpose of enhancement the image quality, because it was able to remove most of the noise and provide very clear images, without impacting the pigment network features. This process known as pre-processing phase involves directional Gabor filter, image sharpening, Sobel filter, and image inpainting methods. The pre-processed images worked very well with segmentation process; thereby boosting its accuracy, since the noise (hair) covering the pigmented lesions was mostly removed, which led to the lines of pigment network becoming more visible. Therefore, the outcomes of this process increase the segmentation and the classification accuracy. With respect to detecting and segmenting the pigment network lines, directional Gabor filter was also executed with different parameters from those used to detect the hairs. Thus, the large mesh of pigment network lines was extracted by implementing the connected component analysis with adaptive threshold method. The Artificial Neural Network is trained by several features which were extracted from the segmented image to identify the image type as normal or abnormal. The proposed method was tested on a dataset of 200 dermoscopy images, including 84 with pigment network and 116 without pigment network. The evaluations also reveal that our method yielded promising results and outperformed the alternative method with which it was compared

The key achievement of this chapter is that the method was able to extract the hair from images without affecting the pigment network lines, which led subsequently to segmenting the pigment network correctly. In addition, we classified the lesions as melanoma or non-melanoma based on the extracted features, which were obtained from the mesh of the pigment network. Consequently, this method can be used as a part of an automatic diagnosis system for lesion classification.

7.2.3 Sub-Region Feature Extraction

A comprehensive approach to determining the skin lesions type was presented in Chapter 6. As the first stage of image processing technique and for the purpose of removing undesirable objects, the original input images are pre-processed by detecting and removing the noise. Therefore, the region areas are extracted by applying the PSO and the MRF methods. As opposed to direct feature extraction from the segmented lesions, K-means is applied and the desired features, such as the colour moments and colour histogram, are extracted at the sub-region (cluster) level. These features are fed into an ANN with Radial Basis Function as an activation function for final melanoma classification. The proposed approach proved to be robust and accurate, leading to successful classification results of lesions; subsequently, aiding in the assessment of melanoma disease.

Regarding the qualitative evaluation, the experimental results of 200 dermoscopy images of the proposed approach were compared with the ground truth. The method achieved approximately 96.0% accuracy, 99.00% sensitivity, and 85.00% specificity, which are considered very promising results. In addition to the qualitative evaluation, a comparison of several alternative methods was performed. The evaluation metric indicated that the method achieved the highest results. Thus, we believe that our efficient and robust method can help dermatologists reach better diagnoses, and in a fast and accurate way. Once the code is loaded onto dermoscopy machines, the method will enhance these doctors abilities, and the outcomes for skin cancer patients.

7.3 Research Limitations

This work has certain limitations that should be acknowledged. For the purpose of enhancing the quality of images, the algorithm was mostly focused on hair detection and removal. However, it may be possible to focus on detecting other noises as well such as air bubbles and gel, to render the images even more understandable and gain clear images.

The accurate segmentation of lesion boundaries (actual lesions) can provide meaningful and fundamental information for malignant melanoma detection in an early stage. However,

our segmentation methods performed well in case of benign images. However, with regard to melanoma images, the algorithms failed to delineate whole lesions in many cases, which could be considered a limitation of this study.

With respect to feature extraction, the study focused solely on extracting several features from the pigment network structures or the lesion areas. This also can be a limitation of this work, since those features depended only on the regular distribution of the pigment network, colour and intensity value. We could have considered more features, such as lesion shape, texture etc.

Additionally, another possible limitation might be that all four methods were applied to one dataset (200 images), obtained from (dermatology service of Hospital Pedro Hispano Portugal [101]). Consequently, the generalisability of findings to other sources of datasets may be limited.

7.4 Future Work

One of the primary aims of the process of dermoscopic image analysis using CAD systems is to increase physicians' confidence in differentiating between normal and abnormal skins. Therefore, future work will focus primarily on implementing computer algorithms that are capable of analysing accurately and more efficiently, for the purpose of diagnosing the disease correctly. In this section, we propose and discuss ideas that will drive the future direction of this research.

Firstly, in this thesis, we have performed a segmentation of dermoscopic images with lesion classification using one dataset. In planning for future work, we intend using a different dataset of skin cancer images to determine whether our methods are limited to dermoscopic images. Moreover, different classiers like AdaBoost, KNN or SVM could be tested and the one with the higher accuracy adopted.

Secondly, in the three diagnostic classes (common moles, atypical nevi and melanomas), the segmentation methods were able to detect the actual lesions with very high accuracy rate, in terms of the common moles and atypical nevi. However, the accuracy rate was decreased

when melanoma images are processed. Thus, it will be interesting to improve our approach and focus more on melanoma images, and images with low contrast between lesion edges and surrounding healthy skin.

Thirdly, we introduced a method for detecting and removing the hair from images. as plans for future work, we should extend or improve the method to detect and remove the thin hair and the hair which tends to be grey in colour, since it causes a significant degradation to the pigment network lines and regions border segmentation. Furthermore, in Chapter 5, we have proposed a pigment network detection approach. This approach is currently semi-automate since the ground truth images are used as segmented lesions. Therefore, it is imperative to implement an algorithm capable of segmenting the actual lesions of any affected area. We plan to develop our algorithm by integrating it with another method able to delineate the real lesions. Moreover, dermoscopic features such as dots/globules and streaks are crucial, since their presence in pigmented skin lesions are a sign of malignancy. We intend to work out to extract these important structures for the purpose of increasing the accuracy diagnosis of melanoma disease.

In addition to the proposed extracted features presented in Chapter 6, researchers could work towards extracting more topological features using different detect descriptors such as Speeded Up Robust Features (SURF), Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradient (HOG),etc, in order to provide some extra information regarding the intensity of the region. Moreover, they can extract several features from the three colour channels (RGB), rather than using only grey scale images. Thus, these features could be combined with those derived from the proposed method and used as an image descriptor. This might provide high accuracy in differentiating between lesions types.

Bibliography

[1] https://challenge.kitware.com/challenge/n/isic$_2$017

[2] https://livelovefruit.com/10-skin-cancer-signs-you-should-never-ignore/.

[3] https://www.dermnetnz.org/cme/dermoscopy-course/introduction-to-dermoscopy/.

[4] https://www.dermnetnz.org/cme/dermoscopy-course/pattern-analysis/.

[5] http://www.dermoscopy.org/atlas/base.htm.

[6] http://www.dermoscopy.org/consensus/2b.asp.

[7] Q. Abbas, M. E. Celebi, I. Fondón García, and M. Rashid. Lesion border detection in dermoscopy images using dynamic programming. *Skin Research and Technology*, 17(1):91–100, 2011.

[8] Q. Abbas, I. Fondón, and M. Rashid. Unsupervised skin lesions border detection via two-dimensional image analysis. *Computer methods and programs in biomedicine*, 104(3):e1–e15, 2011.

[9] Q. Abbas, I. F. Garcia, M. Emre Celebi, W. Ahmad, and Q. Mushtaq. A perceptually oriented method for contrast enhancement and segmentation of dermoscopy images. *Skin Research and Technology*, 19(1):e490–e497, 2013.

[10] Q. Abbas, I. F. Garcia, M. Emre Celebi, W. Ahmad, and Q. Mushtaq. Unified approach for lesion border detection based on mixture modeling and local entropy thresholding. *Skin Research and Technology*, 19(3):314–319, 2013.

[11] N. R. Abbasi, H. M. Shaw, D. S. Rigel, R. J. Friedman, W. H. McCarthy, I. Osman, A. W. Kopf, and D. Polsky. Early diagnosis of cutaneous melanoma: revisiting the abcd criteria. *Jama*, 292(22):2771–2776, 2004.

[12] J. F. Alcón, C. Ciuhu, W. Ten Kate, A. Heinrich, N. Uzunbajakava, G. Krekels, D. Siem, and G. De Haan. Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. *IEEE journal of selected topics in signal processing*, 3(1):14–25, 2009.

[13] N. Alfed, F. Khelifi, and A. Bouridane. Improving a bag of words approach for skin cancer detection in dermoscopic images. In *Control, Decision and Information Technologies (CoDIT), 2016 International Conference on*, pages 024–027. IEEE, 2016.

[14] N. Alfed, F. Khelifi, A. Bouridane, and H. Seker. Pigment network-based skin cancer detection. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 7214–7217. IEEE, 2015.

[15] G. Argenziano, C. Catricalà, M. Ardigo, P. Buccini, P. De Simone, L. Eibenschutz, A. Ferrari, G. Mariani, V. Silipo, I. Sperduti, et al. Seven-point checklist of dermoscopy revisited. *British Journal of Dermatology*, 164(4):785–790, 2011.

[16] 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. *Archives of dermatology*, 134(12):1563–1570, 1998.

[17] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara, et al. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. *Journal of the American Academy of Dermatology*, 48(5):679–693, 2003.

[18] B. Aribisala and E. Claridge. A border irregularity measure using a modified conditional entropy method as a malignant melanoma predictor. *Image analysis and recognition*, pages 914–921, 2005.

[19] J. L. G. Arroyo and B. G. Zapirain. Detection of pigment network in dermoscopy images using supervised machine learning and structural analysis. *Computers in biology and medicine*, 44:144–157, 2014.

[20] J. L. G. Arroyo and B. G. Zapirain. Detection of pigment network in dermoscopy images using supervised machine learning and structural analysis. *Computers in Biology and Medicine*, 44:144 – 157, 2014.

[21] C. Barata, M. E. Celebi, and J. S. Marques. Improving dermoscopy image classification using color constancy. *IEEE journal of biomedical and health informatics*, 19(3):1146–1152, 2015.

[22] C. Barata, M. E. Celebi, and J. S. Marques. Melanoma detection algorithm based on feature fusion. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 2653–2656. IEEE, 2015.

[23] C. Barata, M. A. Figueiredo, M. E. Celebi, and J. S. Marques. Color identification in dermoscopy images using gaussian mixture models. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 3611–3615. IEEE, 2014.

[24] C. Barata, J. S. Marques, and T. Mendonça. Bag-of-features classification model for the diagnose of melanoma in dermoscopy images using color and texture descriptors. In *International Conference Image Analysis and Recognition*, pages 547–555. Springer, 2013.

[25] C. Barata, J. S. Marques, and J. Rozeira. A system for the detection of pigment network in dermoscopy images using directional filters. *Biomedical Engineering, IEEE Transactions on*, 59(10):2744–2754, 2012.

[26] C. Barata, J. S. Marques, and J. Rozeira. Evaluation of color based keypoints and features for the classification of melanomas using the bag-of-features model. In *International Symposium on Visual Computing*, pages 40–49. Springer, 2013.

[27] C. Barata, M. Ruela, M. Francisco, T. Mendonça, and J. S. Marques. Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal*, 8(3):965–979, 2014.

[28] C. A. Z. Barcelos, M. Boaventura, and E. Silva. A well-balanced flow equation for noise removal and edge detection. *IEEE Transactions on image processing*, 12(7):751–763, 2003.

[29] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 417–424. ACM Press/Addison-Wesley Publishing Co., 2000.

[30] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 259–302, 1986.

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

[32] J. C. Bezdek, R. Ehrlich, and W. Full. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2-3):191–203, 1984.

[33] L. Bi, J. Kim, E. Ahn, D. Feng, and M. Fulham. Automatic melanoma detection via multi-scale lesion-biased representation and joint reverse classification. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1055–1058. IEEE, 2016.

[34] R. P. Braun, H. S. Rabinovitz, M. Oliviero, A. W. Kopf, and J.-H. Saurat. Dermoscopy of pigmented skin lesions. *Journal of the American Academy of Dermatology*, 52(1):109–121, 2005.

[35] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980.

[36] M. D. Buhmann. *Radial basis functions: theory and implementations*, volume 12. Cambridge university press, 2003.

[37] C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.

[38] H. Bustince, E. Barrenechea, and M. Pagola. Image thresholding using restricted equivalence functions and maximizing the measures of similarity. *Fuzzy Sets and Systems*, 158(5):496–516, 2007.

[39] H. Bustince, M. Pagola, E. Barrenechea, J. Fernández, P. Melo-Pinto, P. Couto, H. R. Tizhoosh, and J. Montero. Ignorance functions. an application to the calculation of the threshold in prostate ultrasound images. *Fuzzy sets and Systems*, 161(1):20–36, 2010.

[40] G. Campos-do Carmo and M. Ramos-e Silva. Dermoscopy: basic concepts. *International journal of dermatology*, 47(7):712–719, 2008.

[41] G. Capdehourat, A. Corez, A. Bazzano, R. Alonso, and P. Musé. Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions. *Pattern Recognition Letters*, 32(16):2187–2196, 2011.

[42] H. Castillejos, V. Ponomaryov, L. Nino-de Rivera, and V. Golikov. Wavelet transform fuzzy algorithms for dermoscopic image segmentation. *Computational and mathematical methods in medicine*, 2012, 2012.

[43] M. E. Celebi, S. Hwang, H. Iyatomi, and G. Schaefer. Robust border detection in dermoscopy images using threshold fusion. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 2541–2544. IEEE, 2010.

[44] M. E. Celebi, H. Iyatomi, G. Schaefer, and W. V. Stoecker. Lesion border detection in dermoscopy images. *Computerized medical imaging and graphics*, 33(2):148–153, 2009.

[45] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, and R. H. Moss. A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics*, 31(6):362–373, 2007.

[46] M. E. Celebi and A. Zornberg. Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification. *IEEE systems journal*, 8(3):980–984, 2014.

[47] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277, 2001.

[48] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE transactions on pattern analysis and machine intelligence*, 17(8):790–799, 1995.

[49] K. Clawson, P. Morrow, B. Scotney, D. McKenna, and O. Dolan. Computerised skin lesion surface analysis for pigment asymmetry quantification. In *Machine Vision and Image Processing Conference, 2007. IMVIP 2007. International*, pages 75–82. IEEE, 2007.

[50] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[51] R. Crandall. Image segmentation using the chan–vese algorithm. *Project report from ECE*, 532, 2009.

[52] A. Criminisi, P. Perez, and K. Toyama. Object removal by exemplar-based inpainting. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–721. IEEE, 2003.

[53] H. Demuth, M. Beale, and M. Hagan. Neural network toolbox™ 6. *User's guide*, pages 37–55, 2008.

[54] Y. Deng and B. Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE transactions on pattern analysis and machine intelligence*, 23(8):800–810, 2001.

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

[56] R. Eberhart and J. Kennedy. A new optimizer using particle swarm theory. In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, pages 39–43. IEEE, 1995.

[57] M. Elgamal. Automatic skin cancer images classification. *IJACSA) International Journal of Advanced Computer Science and Applications*, 4(3):287–294, 2013.

[58] K. Eltayef, Y. Li, and X. Liu. Detection of pigment networks in dermoscopy images. In *Journal of Physics: Conference Series*, volume 787. IOP Publishing, 2017.

[59] M. Emre Celebi, H. A. Kingravi, H. Iyatomi, Y. Alp Aslandogan, W. V. Stoecker, R. H. Moss, J. M. Malters, J. M. Grichnik, A. A. Marghoob, H. S. Rabinovitz, et al. Border detection in dermoscopy images using statistical region merging. *Skin Research and Technology*, 14(3):347–353, 2008.

[60] M. Emre Celebi, Q. Wen, S. Hwang, H. Iyatomi, and G. Schaefer. Lesion border detection in dermoscopy images using ensembles of thresholding methods. *Skin Research and Technology*, 19(1), 2013.

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

[62] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41. Society for Imaging Science and Technology, 2004.

[63] Y. Freund, R. E. Schapire, et al. Experiments with a new boosting algorithm. In *Icml*, volume 96, pages 148–156. Bari, Italy, 1996.

[64] R. J. Friedman, D. S. Rigel, and A. W. Kopf. Early detection of malignant melanoma: The role of physician examination and self-examination of the skin. *CA: a cancer journal for clinicians*, 35(3):130–151, 1985.

[65] R. Garnavi, M. Aldeen, M. E. Celebi, A. Bhuiyan, C. Dolianitis, and G. Varigos. Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels. *International Journal of Medicine and Medical Sciences*, 1(2):126–134, 2010.

[66] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741, 1984.

[67] P. Ghamisi, M. S. Couceiro, J. A. Benediktsson, and N. M. Ferreira. An efficient method for segmentation of images based on fractional calculus and natural selection. *Expert Systems with Applications*, 39(16):12407–12417, 2012.

[68] A. Gijsenij, T. Gevers, and J. Van De Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 2011.

[69] J. Glaister, A. Wong, and D. A. Clausi. Segmentation of skin lesions from digital images using joint statistical texture distinctiveness. *IEEE transactions on biomedical engineering*, 61(4):1220–1230, 2014.

[70] A. G. Goodson and D. Grossman. Strategies for early melanoma detection: Approaches to the patient with nevi. *Journal of the American Academy of Dermatology*, 60(5):719–735, 2009.

[71] C. Grana, R. Cucchiara, G. Pellacani, and S. Seidenari. Line detection and texture characterization of network patterns. 2006.

[72] G. Hamerly and C. Elkan. Learning the k in k-means. In *Advances in neural information processing systems*, pages 281–288, 2004.

[73] L.-K. Huang and M.-J. J. Wang. Image thresholding by minimizing the measures of fuzziness. *Pattern recognition*, 28(1):41–51, 1995.

[74] J. Humayun, A. S. Malik, and N. Kamel. Multilevel thresholding for segmentation of pigmented skin lesions. In *Imaging Systems and Techniques (IST), 2011 IEEE International Conference on*, pages 310–314. IEEE, 2011.

[75] S. Hwang and M. E. Celebi. Texture segmentation of dermoscopy images using gabor filters and g-means clustering. In *IPCV*, pages 882–886, 2010.

[76] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, pages 448–456, 2015.

[77] A. G. Isasi, B. G. Zapirain, and A. M. Zorrilla. Melanomas non-invasive diagnosis application based on the abcd rule and pattern recognition image processing algorithms. *Computers in Biology and Medicine*, 41(9):742–755, 2011.

[78] H. Iyatomi, H. Oka, M. E. Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka, and K. Ogawa. An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. *Computerized Medical Imaging and Graphics*, 32(7):566–579, 2008.

[79] P. Jaccard. The distribution of the flora in the alpine zone. *New phytologist*, 11(2):37–50, 1912.

[80] M. Jafari, N. Karimi, E. Nasr-Esfahani, S. Samavi, S. Soroushmehr, K. Ward, and K. Najarian. Skin lesion segmentation in clinical images using deep learning. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, 2016.

[81] J. Jaworek-Korjakowska and R. Tadeusiewicz. Determination of border irregularity in dermoscopic color images of pigmented skin lesions. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 6459–6462. IEEE, 2014.

[82] R. H. Johr. Dermoscopy: alternative melanocytic algorithms—the abcd rule of dermatoscopy, menzies scoring method, and 7-point checklist. *Clinics in dermatology*, 20(3):240–247, 2002.

[83] J. N. Kapur, P. K. Sahoo, and A. K. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer vision, graphics, and image processing*, 29(3):273–285, 1985.

[84] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[85] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder. Diagnostic accuracy of dermoscopy. *The lancet oncology*, 3(3):159–165, 2002.

[86] H. Kittler, M. Seltenheim, M. Dawid, H. Pehamberger, K. Wolff, and M. Binder. Morphologic changes of pigmented skin lesions: a useful extension of the abcd rule for dermatoscopy. *Journal of the American Academy of Dermatology*, 40(4):558–562, 1999.

[87] J. Kittler and J. Illingworth. Minimum error thresholding. *Pattern recognition*, 19(1):41–47, 1986.

[88] M. Kruk, B. Świderski, S. Osowski, J. Kurek, M. Słowińska, and I. Walecka. Melanoma recognition using extended set of descriptors and classifiers. *Eurasip Journal on Image and Video Processing*, 2015(1):43, 2015.

[89] T. Kurban and E. Beşdok. A comparison of rbf neural network training algorithms for inertial sensor based terrain classification. *Sensors*, 9(8):6312–6329, 2009.

[90] V. Labatut and H. Cherifi. Accuracy measures for the comparison of classifiers. *arXiv preprint arXiv:1207.3790*, 2012.

[91] E. H. Land. The retinex theory of color vision. *Scientific American*, 237(6):108–129, 1977.

[92] S. Lankton and A. Tannenbaum. Localizing region-based active contours. *IEEE transactions on image processing*, 17(11):2029–2039, 2008.

[93] H. Lee and Y.-P. P. Chen. Skin cancer extraction with optimum fuzzy thresholding technique. *Applied intelligence*, 40(3):415–426, 2014.

[94] Y. Li et al. Lesion segmentation in dermoscopy images using particle swarm optimization and markov random field. 2017.

[95] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.

[96] N. V. Lopes, P. A. M. do Couto, H. Bustince, and P. Melo-Pinto. Automatic histogram threshold using fuzzy measures. *IEEE Transactions on Image Processing*, 19(1):199–204, 2010.

[97] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[98] A. R. Maouche and M. Attari. Nonlinear adaptive rbfnn control of a one-link flexible manipulator. In *Machine and Web Intelligence (ICMWI), 2010 International Conference on*, pages 165–170. IEEE, 2010.

[99] J. S. Marques, C. Barata, and T. Mendonça. On the role of texture and color in the classification of dermoscopy images. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4402–4405. IEEE, 2012.

[100] E. Mendi, C. Yogurtcular, Y. Sezgin, and C. Bayrak. Automatic mobile segmentation of dermoscopy images using density based and fuzzy c-means clustering. In *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium on*, pages 1–6. IEEE, 2014.

[101] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira. Ph 2-a dermoscopic image database for research and benchmarking. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5437–5440. IEEE, 2013.

[102] G. Mercier and M. Lennon. Support vector machines for hyperspectral image classification with spectral-based kernels. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International*, volume 1, pages 288–290. IEEE, 2003.

[103] S. Merler, C. Furlanello, B. Larcher, and A. Sboner. Tuning cost-sensitive boosting and its application to melanoma diagnosis. In *International Workshop on Multiple Classifier Systems*, pages 32–42. Springer, 2001.

[104] M. Messadi, A. Bessaid, and A. Taleb-Ahmed. Extraction of specific parameters for skin tumour classification. *Journal of medical engineering & technology*, 33(4):288–295, 2009.

[105] M. Messadi, H. Cherifi, and A. Bessaid. Segmentation and abcd rule extraction for skin tumors classification. *Journal of Convergence Information Technology*, 9(2):21, 2014.

[106] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *International journal of computer vision*, 60(1):63–86, 2004.

[107] N. K. Mishra and M. E. Celebi. An overview of melanoma detection in dermoscopy images using image processing and machine learning. *arXiv preprint arXiv:1601.07843*, 2016.

[108] M. F. Møller. A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, 6(4):525–533, 1993.

[109] F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig. The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30(4):551–559, 1994.

[110] R. Nock and F. Nielsen. Statistical region merging. *IEEE Transactions on pattern analysis and machine intelligence*, 26(11):1452–1458, 2004.

[111] L. A. Nowak, M. J. Ogorzalek, and M. P. Pawlowski. Pigmented network structure detection using semi-smart adaptive filters. In *Systems Biology (ISB), 2012 IEEE 6th International Conference on*, pages 310–314. IEEE, 2012.

[112] 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(7):971–987, 2002.

[113] R. B. Oliveira, J. P. Papa, A. S. Pereira, and J. M. R. Tavares. Computational methods for pigmented skin lesion classification in images: review and future trends. *Neural Computing and Applications*, pages 1–24, 2016.

[114] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

[115] N. Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

[116] G. Pellacani, C. Grana, and S. Seidenari. Algorithmic reproduction of asymmetry and border cut-off parameters according to the abcd rule for dermoscopy. *Journal of the European Academy of Dermatology and Venereology*, 20(10):1214–1219, 2006.

[117] A. Pennisi, D. D. Bloisi, D. Nardi, A. R. Giampetruzzi, C. Mondino, and A. Facchiano. Skin lesion image segmentation using delaunay triangulation for melanoma detection. *Computerized Medical Imaging and Graphics*, 2016.

[118] I. Pirnog, R. O. Preda, C. Oprea, and C. Paleologu. Automatic lesion segmentation for melanoma diagnostics in macroscopic images. In *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pages 659–663. IEEE, 2015.

[119] J. Qi, M. Le, C. Li, and P. Zhou. Global and local information based deep network for skin lesion segmentation. *arXiv preprint arXiv:1703.05467*, 2017.

[120] M. M. Rahman and P. Bhattacharya. An integrated and interactive decision support system for automated melanoma recognition of dermoscopic images. *Computerized Medical Imaging and Graphics*, 34(6):479–486, 2010.

[121] M. Rastgoo, R. Garcia, O. Morel, and F. Marzani. Automatic differentiation of melanoma from dysplastic nevi. *Computerized Medical Imaging and Graphics*, 43:44–52, 2015.

[122] F. Riaz, A. Hassan, M. Y. Javed, and M. T. Coimbra. Detecting melanoma in dermoscopy images using scale adaptive local binary patterns. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 6758–6761. IEEE, 2014.

[123] P. Rubegni, G. Cevenini, M. Burroni, R. Perotti, G. Dell'Eva, P. Sbano, C. Miracco, P. Luzi, P. Tosi, P. Barbini, et al. Automated diagnosis of pigmented skin lesions. *International Journal of Cancer*, 101(6):576–580, 2002.

[124] M. Ruela, C. Barata, J. S. Marques, and J. Rozeira. A system for the detection of melanomas in dermoscopy images using shape and symmetry features. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 5(2):127–137, 2017.

[125] D. Ruiz, V. Berenguer, A. Soriano, and B. SáNchez. A decision support system for the diagnosis of melanoma: A comparative approach. *Expert Systems with Applications*, 38(12):15217–15223, 2011.

[126] M. Sadeghi, M. Razmara, T. K. Lee, and M. S. Atkins. A novel method for detection of pigment network in dermoscopic images using graphs. *Computerized Medical Imaging and Graphics*, 35(2):137–143, 2011.

[127] M. Sadeghi, M. Razmara, P. Wighton, T. K. Lee, and M. S. Atkins. Modeling the dermoscopic structure pigment network using a clinically inspired feature set. In *Medical Imaging and Augmented Reality*, pages 467–474. Springer, 2010.

[128] A. Sáez, B. Acha, and C. Serrano. Pattern analysis in dermoscopic images. In *Computer Vision Techniques for the Diagnosis of Skin Cancer*, pages 23–48. Springer, 2014.

[129] A. Salazar-Gonzalez, D. Kaba, Y. Li, and X. Liu. Segmentation of the blood vessels and optic disk in retinal images. *IEEE journal of biomedical and health informatics*, 18(6):1874–1886, 2014.

[130] A. Salazar-Gonzalez, Y. Li, and D. Kaba. MRF reconstruction of retinal images for the optic disc segmentation. In *International Conference on Health Information Science*, pages 88–99. Springer, 2012.

[131] C. Schmid. Constructing models for content-based image retrieval. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 2, pages II–II. IEEE, 2001.

[132] S. Seidenari, G. Pellacani, and C. Grana. Pigment distribution in melanocytic lesion images: a digital parameter to be employed for computer-aided diagnosis. *Skin Research and Technology*, 11(4):236–241, 2005.

[133] M. A. Sheha, M. S. Mabrouk, and A. Sharawy. Automatic detection of melanoma skin cancer using texture analysis. *International Journal of Computer Applications*, 42(20):22–26, 2012.

[134] M. Silveira, J. C. Nascimento, J. S. Marques, A. R. Marçal, T. Mendonça, S. Yamauchi, J. Maeda, and J. Rozeira. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *Selected Topics in Signal Processing, IEEE Journal of*, 3(1):35–45, 2009.

[135] N. Situ, X. Yuan, J. Chen, and G. Zouridakis. Malignant melanoma detection by bag-of-features classification. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3110–3113. IEEE, 2008.

[136] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In *null*, page 1470. IEEE, 2003.

[137] N. Smaoui and S. Bessassi. A developed system for melanoma diagnosis. *International Journal of Computer Vision and Signal Processing*, 3(1):10–17, 2013.

[138] P. Soille. *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.

[139] S. Sookpotharom. Border detection of skin lesion images based on fuzzy c-means thresholding. In *Genetic and Evolutionary Computing, 2009. WGEC'09. 3rd International Conference on*, pages 777–780. IEEE, 2009.

[140] H. P. Soyer, G. Argenziano, I. Zalaudek, R. Corona, F. Sera, R. Talamini, F. Barbato, A. Baroni, L. Cicale, A. Di Stefani, et al. Three-point checklist of dermoscopy. *Dermatology*, 208(1):27–31, 2004.

[141] R. J. Stanley, W. V. Stoecker, and R. H. Moss. A relative color approach to color discrimination for malignant melanoma detection in dermoscopy images. *Skin Research and Technology*, 13(1):62–72, 2007.

[142] S. Suer, S. Kockara, and M. Mete. An improved border detection in dermoscopy images for density based clustering. *BMC bioinformatics*, 12(10):S12, 2011.

[143] S. C. Trotter, N. Sroa, R. R. Winkelmann, T. Olencki, and M. Bechtel. A global review of melanoma follow-up guidelines. *Journal of Clinical & Aesthetic Dermatology*, 6(9), 2013.

[144] K. Van De Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, 32(9):1582–1596, 2010.

[145] J. Van De Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *IEEE Transactions on image processing*, 16(9):2207–2214, 2007.

[146] M. Vestergaard, P. Macaskill, P. Holt, and S. Menzies. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *British Journal of Dermatology*, 159(3):669–676, 2008.

[147] F. M. Walter, A. T. Prevost, J. Vasconcelos, P. N. Hall, N. P. Burrows, H. C. Morris, A. L. Kinmonth, and J. D. Emery. Using the 7-point checklist as a diagnostic aid for pigmented skin lesions in general practice: a diagnostic validation study. *Br J Gen Pract*, 63(610):e345–e353, 2013.

[148] P. Wighton, T. K. Lee, H. Lui, D. I. McLean, and M. S. Atkins. Generalizing common tasks in automated skin lesion diagnosis. *IEEE Transactions on Information Technology in Biomedicine*, 15(4):622–629, 2011.

[149] A. Wong, J. Scharcanski, and P. Fieguth. Automatic skin lesion segmentation via iterative stochastic region merging. *IEEE Transactions on Information Technology in Biomedicine*, 15(6):929–936, 2011.

[150] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing*, 7(3):359–369, 1998.

[151] L. Xu, M. Jackowski, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, and A. Huntley. Segmentation of skin cancer images. *Image and Vision Computing*, 17(1):65–74, 1999.

[152] X. Yang, Z. Zeng, S. Y. Yeo, C. Tan, H. L. Tey, and Y. Su. A novel multi-task deep learning model for skin lesion segmentation and classification. *arXiv preprint arXiv:1703.01025*, 2017.

[153] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE transactions on medical imaging*, 36(4):994–1004, 2017.

[154] Y. Yuan, M. Chao, and Y.-C. Lo. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. *IEEE Transactions on Medical Imaging*, 2017.

[155] Y. Yuan, M. L. Giger, H. Li, K. Suzuki, and C. Sennett. A dual-stage method for lesion segmentation on digital mammograms. *Medical physics*, 34(11):4180–4193, 2007.

[156] M. E. Yüksel and M. Borlu. Accurate segmentation of dermoscopic images by image thresholding based on type-2 fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 17(4):976–982, 2009.

[157] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning—i. *Information sciences*, 8(3):199–249, 1975.

[158] I. Zalaudek, J. Kreusch, J. Giacomel, G. Ferrara, C. Catricalà, and G. Argenziano. How to diagnose nonpigmented skin tumors: a review of vascular structures seen with dermoscopy: part i. melanocytic skin tumors. *Journal of the American Academy of Dermatology*, 63(3):361–374, 2010.

[159] H. Zhou, M. Chen, L. Zou, R. Gass, L. Ferris, L. Drogowski, and J. M. Rehg. Spatially constrained segmentation of dermoscopy images. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 800–803. IEEE, 2008.

[160] H. Zhou, G. Schaefer, M. E. Celebi, H. Iyatomi, K.-A. Norton, T. Liu, and F. Lin. Skin lesion segmentation using an improved snake model. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 1974–1977. IEEE, 2010.

[161] H. Zhou, G. Schaefer, M. E. Celebi, F. Lin, and T. Liu. Gradient vector flow with mean shift for skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 35(2):121–127, 2011.

[162] H. Zhou, G. Schaefer, A. H. Sadka, and M. E. Celebi. Anisotropic mean shift based fuzzy c-means segmentation of dermoscopy images. *IEEE Journal of Selected Topics in Signal Processing*, 3(1):26–34, 2009.

Appendix A

A.1 Dermoscopy Images

Dermoscopy or epiluminescence light microscopy (ELM) is a non-invasive method that enables the skin cancer experts to perform direct microscopic examination of diagnostic clinical features, for the purpose of detecting melanoma disease an early stage. It is also one of the major tools allows for a better visualization of pigmented skin lesion structures; therefore permits the identification and recognition of several dermoscopic structures, which are not visible to the naked eye examination. This technique is widely used by physicians, due to its value in providing a meaningful clinical features from pigmented skin lesions [70]. Examples of dermoscopy imaging devices are shown in Figure 1.

It has been shown that the naked eye test for the diagnosis of melanoma in image skin lesions is less accurate than dermoscopy image examination. A comparison between dermoscopy

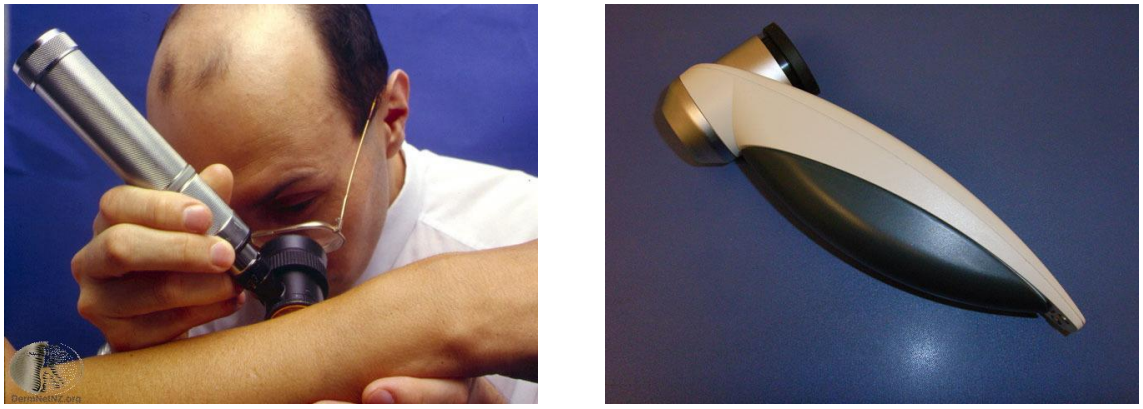


Figure 1: Examples of dermoscopy imaging tools. (Source:[3])

images and naked eye examinations has been performed by M. Vestergaard et al [146]. There report showed that the mean sensitivity of 95.00% was achieved with dermoscopy images, which is higher than the obtained sensitivity with naked eye examination 74.00%.

In addition, the macroscopic images which are obtained by using a magnifying glass, provide less details compared to dermoscopy images. Figure 2 illustrates a visual comparison between two dermatoscopic and macroscopic images of the same lesion. In terms of normal blue nevus, the images in the first row show two images (macroscopic and dermatoscopic) of the same skin region area. It has been observed that the only one single color (steel-blue areas) is appeared, and no clinical feature structures are visible in both images, which can easily be classified as a normal blue nevus . The second row presents images of an affected skin lesion (melanoma). It is very clear that the pigment network and streaks structures are appeared in dermoscopy image (bottom right), while they are completely invisible in macroscopic image (bottom left). This indicates the high performance of dermoscopy images in detecting melanoma at early stage.

The expert dermatologists are able to perform a clinical diagnosis of melanoma disease based on several number of dermoscopic clinical features, which are obtained from pigmented skin lesion images. These features have to be assessed in the most widely used medical diagnosis procedures, such as the ABCDE rule and 7-point checklist. The presence of these specific structures in different regions of the same skin lesion contributes to make a diagnosis of melanocytic lesions. Most dermoscopic features such as pigment network, dots/globules, streaks,regression area, and blue-whitish veil structures, whose used by the dermatologists to detect or distinguish between benign and malignant melanoma lesions are described below.

A.1.1 Pigment Network

The pigment network structure is a grid composed of pigmented lines, which tend to be a brown or a black, and hypopigmented holes. The appearance of it on the lesion can be either typical or atypical. A typical network is relatively uniform, regularly distributed throughout the lesion, homogeneous in color(light-to-dark-brown), and usually thinning out at the periphery. An atypical network is non-uniform thickened lines, with grey, brown, or black broadened

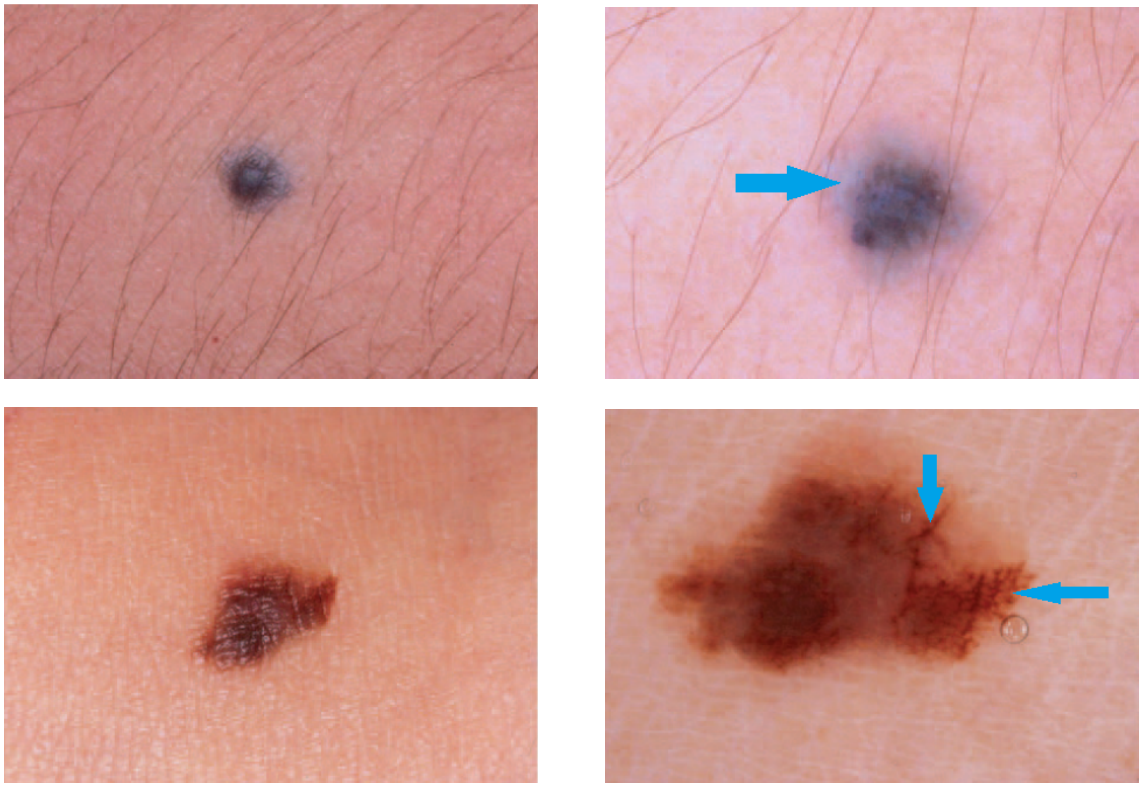


Figure 2: Examples of dermoscopy and macroscopic images: macroscopic image of a blue nevus (top left), dermoscopy of the same lesion shows steel-blue areas (top right), macroscopic image of a superficial spreading malignant melanoma (bottom left) and dermoscopy image of the same lesion illustrates (atypical) pigment network and branched streaks (bottom right). (Source:[34]).

lines and holes, which are inhomogeneous in shape and area. The pigmented lines may end abruptly at the periphery [34]. The presence of an atypical network in the pigmented skin lesions always denotes to melanoma disease, which makes it play a key role in distinguishing between benign and melanoma lesions.

A.1.2 Dots/Globules

Dots appear in the affected lesion as small round structures, with diameter less than 0.1 mm (smaller than globules). Their color can be brown, grey, blue or black. In the pigmented benign lesions (non-melanoma), dots locate in the center of the lesion and they are rather regular in size and shape, while in case of melanoma, usually they present in the periphery and tend to vary in size and shape. A multiple grey-bluish occurrence indicates melanophages in the dermis, and a spotted pattern or pattern resembling "black pepper grains" suggests melanoma [40] [34].

Globules are round structures and they may present different colors (brown, black, or red) depending on the degree of aggregation of melanin. Globules are larger than dots, where their diameter usually exceed 0.1 mm. Milky red globules often corresponding to an elevated part of the lesion [158] [40] [34].

Both dots and globules may occur in benign or melanoma skin lesions.Their presence in lesion area is particularly useful for the distinction between melanoma and non-melanoma lesions.

A.1.3 Streaks

Streaks are defined as finger like projections of the pigment network from the periphery of the lesion. They can appear as regular (symmetrical radial arrangement over the whole lesion) or irregular, when they are unevenly distributed. The existence of streaks in a skin lesions is by itself a signal of malignancy [101].

A.1.4 Regression or Depigmentation Areas

Regression area appears as white scar-like de-pigmentation, usually it is lighter than the surrounding skin, or it appears as "peppering" (speckled multiple blue-grey granules within a hypopigmented area). Not uncommonly, malignant melanomas can have both color patterns of regression in a skin lesion. [34] [82].

A.1.5 Blue-whitish veil

The blue white veil can be defined as an irregular, indistinct, blue pigmentation areas with an overlying white color, pale, and ground-glass haze [34] [17]. The pigmentation can not cover the entire skin lesion.

These significant clinical features can be shown in the lesion with an regular/irregular or typical/atypical nature, implying malignancy or not. Example of several dermoscopic features appear throughout the lesion can be seen in Figure 3.

In the literature section we will briefly explain numerous works that are focused on the automatic detection of several local features and used for melanoma identification.

A.2 Melanoma Diagnosis Criteria

To be able to distinguish melanoma versus benign melanocytic lesions, the expert dermatologists are adopted four main diagnosis methods namely; the ABCD rule, Menzies method, pattern analysis, and 7-Point Checklist. All these approaches were evaluated during the 2,000 Consensus NetMeeting on Dermoscopy (CNMD) [17] by skin cancer experts from all over the world. A two-step process was applied to facilitate the diagnosis:

- To distinguish the lesion as either melanocytic or non-melanocytic.
- To determine whether the melanocytic lesion is suspect, benign, or malignant.

In addition to the previous methods, there is one more method named 3-point checklist which also used for the same purpose, due to its simplification. Brief description of all these methods is presented below.

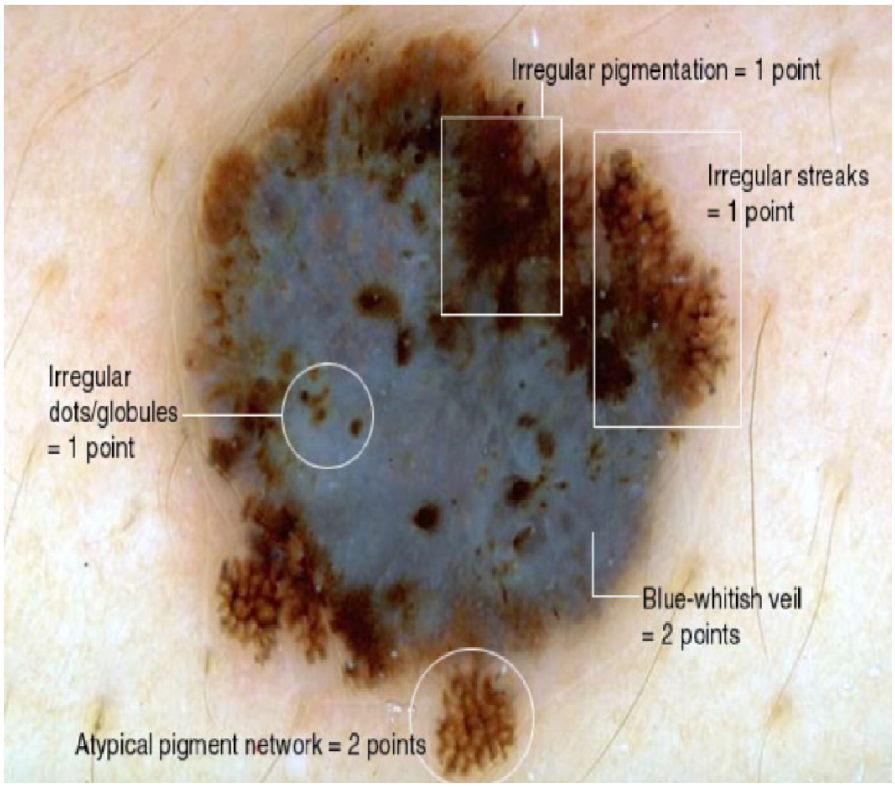


Figure 3: Example shows the dermoscopic structures with its score. (Source:[6])

A.2.1 Pattern Analysis

Pattern analysis process is considered as one of the preferred method by many skin cancer experts to distinguish between skin lesions (benign melanocytic or malignant melanoma). They also perform the final clinical diagnoses of the disease. Pattern analysis indicates the simultaneous assessment of the diagnostic value of all dermoscopic clinical structures shown by the pigmented skin lesion. In other word, it seeks to detect specific patterns from the lesions (locally and globally). The global features usually are appeared as arrangements of textured patterns covering most of skin lesion area. Whereas, The local dermoscopic features are presented as individual groups of characteristics throughout the lesion. The multicomponent pattern (three patterns or more and usually non-concentric) was considered as the most predictive for the diagnosis of melanoma melanocytic lesions. While the cobblestone, globular, starburst and homogeneous patterns were considered as the most predictive for the diagnosis of benign lesions. Regarding to the local dermoscopic features, it has been shown that the presence of irregular streaks, atypical pigment networks, and regression structures on the lesion, followed by irregular blotches, irregular dots/globules, and blue-whitish veil indicate the melanoma. On the contrary, the existence of regular streaks, typical pigment networks, regular blotches, and regular dots/globules are mostly referred to benign melanocytic lesions [128] [4].

As a summary, benign tumours have limited number of colours, symmetrical in pattern, and a regular structure. Whereas, malignant malignant lesions often have several colours, asymmetry of pattern, and disordered structure. The brief description of the pattern analysis in dermoscopy images is illustrated below:

- Local features of benign lesion
  - Pigment network diffused regularly and fading out.
  - Dots and globules distributed regularly.
  - Streaks distributed regularly.
- Global features of benign lesions

  - Reticular pattern (spread or patchy network).
  - Globular pattern (globules within brown color shades).
  - Homogeneous pattern in colour.
  - Starburst pattern (uniform peripheral radial streaks, dots or globules).
  - Parallel pattern (along furrows, palms and soles only).
  - Multicomponent pattern which can be three patterns, and usually concentric).
  - Complex pattern, which can be two patterns within one lesion, and usually symmetrically or regularly distributed.

- Local features of melanoma lesions
  - Multicomponent pattern, which can be three or more and usually non-concentric.
  - Unspecific pattern (mainly structureless or two patterns, irregular).
  - Parallel pattern (along ridges, palms and soles only).

- Global features of melanoma lesions
  - Atypical pigment network structures (thickened, asymmetrical)
  - Dots and globules diffused irregularly throughout the lesion, and of different shapes and sizes.
  - Asymmetrical blotches (featureless colours).
  - Mostly five or six colours (white, grey, red, tan, brown, black).
  - Blue-white veil over part of the lesion.
  - White scar like depigmentation (Regression area).
  - Blue pepper like granules.

A.2.2 ABCD-E Rule

The ABCD criteria was proposed by Friedman et al [64] in 1985, which has been widely used in clinical practice, for the purpose of distinguishing between benign and malignant melanoma

lesions. In The diagnosis of a pigmented skin lesion based on its Asymmetry (A), Border irregularity (B), Color variegation (C), and Diameter (D) $\geqslant$ 6mm across (approximately ij inch), although malignant lesions can sometimes be smaller than this. On other hand, (D) also could refers to Differential structures such as pigment network, structureless area, dots, globules, and streaks. Later in 2004, Abbas et al [11] proposed expanding the ABCD criteria to ABCDE rule by combining the evolving (E) lesion during the time [70]. The ABCDE rule was the second algorithm developed after the pattern analysis method, and the first attempt to simplify the process of melanoma early detection. It was very helpful for clinicians who are not fully experienced in dermoscopy images observation, since of its lower complexity vs pattern analysis method. The choice of these five parameters is based on dermatology criteria which is including color, shape, and symmetry. In order to diagnosing pigmented skin lesions as melanoma or not, These significant criteria have to be assessed by the dermatologists [82]. The images in Figure 4 illustrate the typical normal nevus and melanomas, which are quite useful in showing the people the variations between lesions, as a start point to detect the disease. The method is based on check the skin lesion as shown below:

A.2.2.1 Asymmetry

Asymmetry is a significant parameter that most contributes to the final diagnosis of melanoma. This clinical features is considered in terms of shape and color. With respect to the form, the lesion is divided into halves, along the principal axes, and folded one half of the lesion over the other half. Therefore, asymmetry can be defined as one half of the lesion does not match the other half in 2 perpendicular axes (see Figure 5). Whereas in color, asymmetry is performed by assessing the overall color similarity between symmetric patches or blocks over the pigmented skin lesion [124]. Example of color symmetry analysis is shown in Figure 6.

According to the work presented in [34], Asymmetry has a weighting factor of 1.3, which is higher than the other weighting factors of all criteria, as displayed in Table 1. In terms of the evaluation process, the a symmetry is assessed as follow:

- If it is asymmetric in both axes (vertical and horizontal), it will score 2.6.

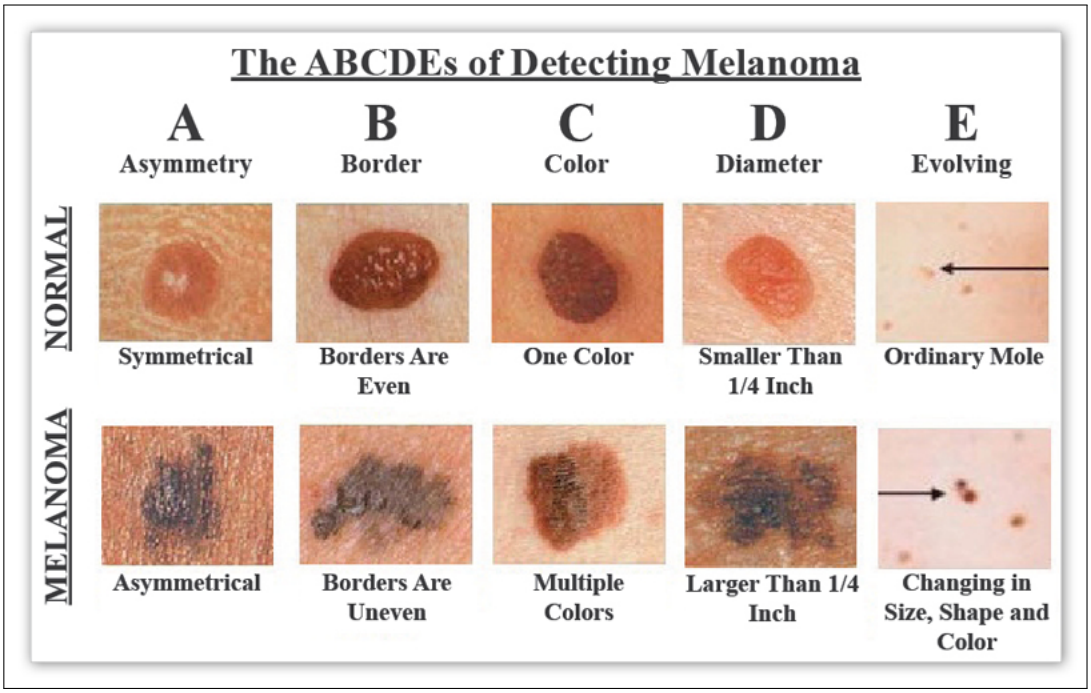


Figure 4: Example of Asymmetry, Border, Color, Diameter, and Evolution (ABCD-E) rule for Melanoma Diagnosis. (Source:[2]).

- if it is asymmetric in just one of the axes, it will score 1.3.
- if it is fully symmetric, it will score zero.

A.2.2.2 Border

Melanoma boundaries are usually irregular, ragged, or blurred, while benign lesions are clearly defined. For the evaluation of the border score, the pigmented skin lesion is visually divided into 8 sectors (pie-shaped) as shown in Figure 7. For each sector in which abrupt cutoff of pigment pattern is present, one point is added to the score leading to a minimum border score of zero and maximum border score of eight. The calculation is based on the euclidean distance and the standard derivation in each sector [137] [86].

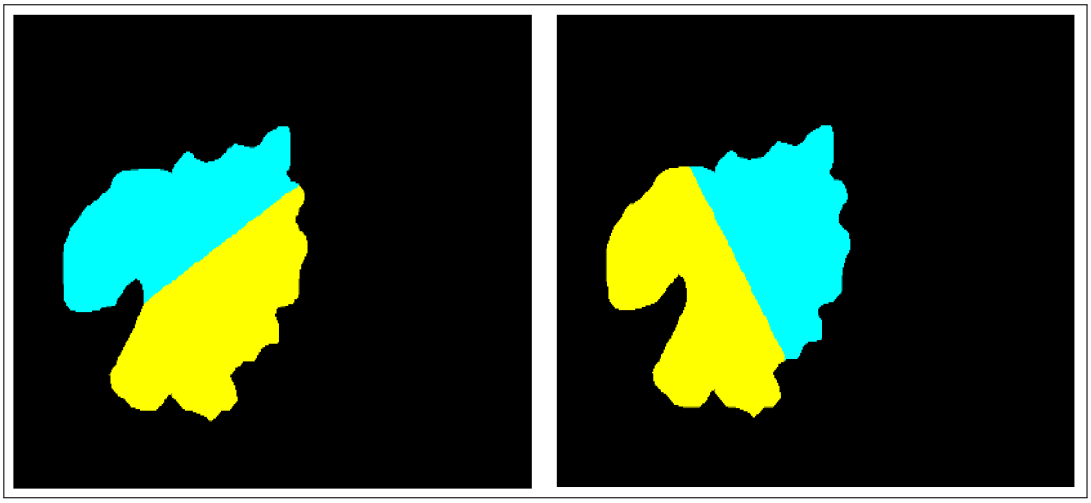


Figure 5: Example of lesion divided into halves according to the principal axes (2 perpendicular axes) of the image. (Source: [124])

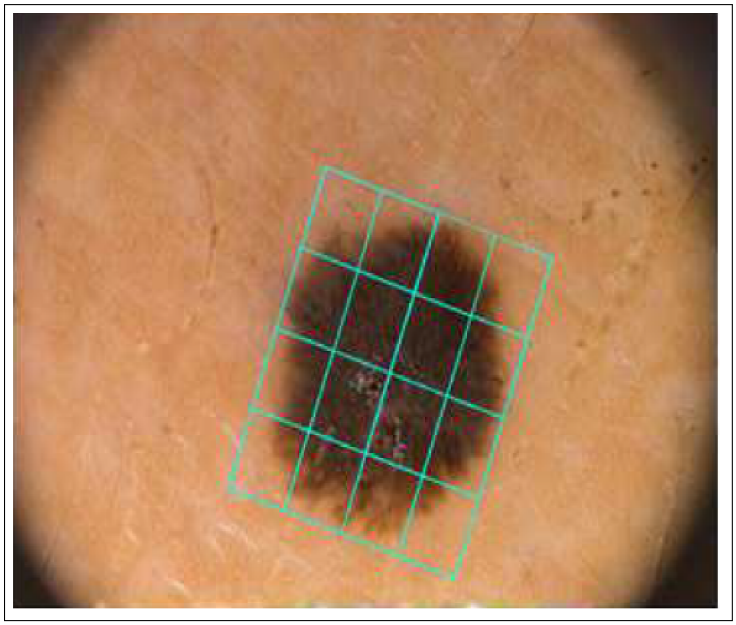


Figure 6: Example of lesion divided into regular grid of blocks. (Source: [124])

Table 1: ABCD rule criteria of dermoscopy according to [34].

Parameter	Defination	Points	Weight factor
Asymmetry	Complete symmetry	0	1.3
	Asymmetry in 1 axis	1	
	Asymmetry in 2 axis	2	
Border	8 sectors	0-8	0.1
Color	1 point for each color	1-6	0.5
	White		
	Red		
	Light brown		
	Dark brown		
	Black		
	Blue-grey		
Differential structures	1 point for every structure	1-5	0.5
	Pigment network		
	Structureless area		
	Dots		
	Globules		
	Streaks		

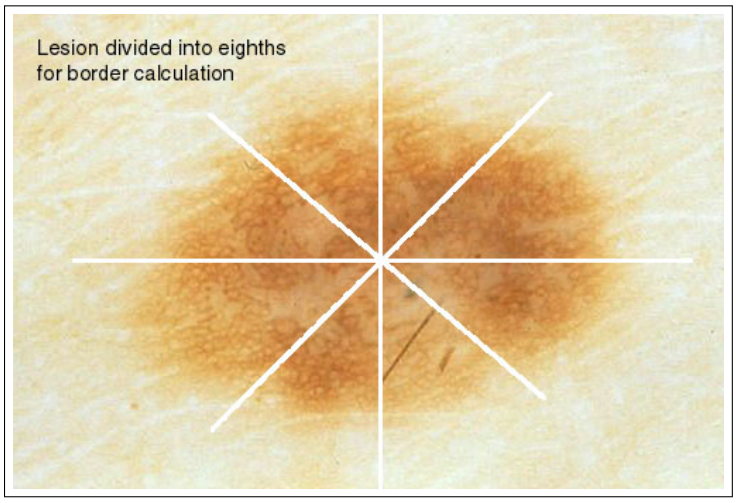


Figure 7: Example of lesion divided into eight sectors. [6]

A.2.2.3 Color

During the diagnosis of a melanocytic lesion, the skin cancer experts are focused on several number of different colors over the skin lesion, because malignant melanoma is characterized by the appearance of these colors throughout the lesion. These different colors comprises the white, red, light brown, dark brown, blue-grey, and black. For the color score, the number of colors appear within the skin lesion are counted, and therefore one point for each color is added, this leads to a maximum score of six when the six colors are present, and the minimum score is one [86].

A.2.2.4 Differential structures

The differential structures criteria such as pigment network, structureless area, dots, globules, and streaks have a weighting factor of 0.5 and score range from one to five, which means one pint for each structure [34].

A.2.2.5 Evolving

The evolving criteria was introduced by Abbas et al [11] to increase the diagnosis accuracy of melanoma disease. They defined the evolving of lesions as the changing of size, shape, symptoms (eg, itching, tenderness), surface (eg, bleeding), or shades of color, which are indicator to be malignant melanoma.

As a final step of the ABCD-E rule, and with the purpose of differentiating between benign and melanoma lesions, the obtained values of the above parameters are used to calculate the Total Dermatoscopic Value (TDV) based on the equation A.2.1, where each one of the presented characteristics is multiplied by its corresponding weighting factor as descried in the work presented by Stolz et al [109] and appeared in Table 1.

$$TDS = 1.3 * A + 0.1 * B + 0.5 * C + 0.5 * D. \tag{A.2.1}$$

This score contributes to identify melanoma or benign lesions as follows

- TDS < 4.75 refers to benign lesion.

- $4.75 \leqslant \text{TDS} \leqslant 5.45$ indicates suspicious.
- TDS>5.45 considered to be melanoma.

According to the work presented in [86], the enhanced ABCD-E score was calculated by adding 1.2 to the standard ABCD score for changing lesions and subtracting 0.8 from the standard ABCD score for nonchanging lesions according to the results of their multivariate analysis.

A.2.3 3-point checklist

In the detection of skin cancer at an early stage, the dermoscopy 3-point checklist method has a high sensitivity for melanoma [140]. The main purpose of it is to define whether the lesion being examined has to undergo a biopsy or not. Therefore, It does not need accurate diagnosis to be made. There a high probability of malignancy (melanoma or basal cell carcinoma) if an affected skin lesion has any two of the following criteria:

- Asymmetry: asymmetry of colour and structure in one or two perpendicular axes
- Atypical network: pigment network with irregular holes and thick lines
- Blue-white structures: any type of blue and/or white colour, i.e. combination of blue-white veil and regression structures

The 3-point checklist method is quite easy to learn and has been designed to permit non-expert dermatologists not to miss detection of melanomas. However, it is not as specific as pattern analysis method.

A.2.4 7-point checklist

Based on the the analysis of 342 skin images, which was performed by Argenziano and colleagues in 1998 [16] [17] [34], the 7-point checklist was adopted by focusing on seven dermoscopic features in the pigmented skin lesion, which have to be assessed by the clinicians. These dermoscopic clinical structures are divided into two groups namely major and minor criteria.

Table 2: 7-point checklist criteria of dermoscopy with related individual score according to [16].

7-point checklist criteria	Score	
Major criteria	1. Atypical pigment network	2
	2. Blue-whitish veil	2
	3. Atypical vascular pattern	2
Minor criteria	4. Irregular streaks	1
	5. Irregular pigmentation	1
	6. Irregular dots/globules	1
	7. Regression structures	1

The major criteria have score of 2 points, while the minor criteria have 1 point, as can be illustrated in Table 2. Based on 7-point checklist, a minimum total score of 3 is required for the detection of melanoma skin cancer [34]. The pigmented skin lesion is considered to be melanoma or not dependent on the following conditionals:

- If a total score of three or more is given, the lesion is classified as malignant melanoma.
- If a total score of less than three is given, the lesion is diagnosed as normal mole or nevus.

The major criteria consists of an atypical pigment network, blue-whitish veil, and atypical vascular pattern. While, the minor criteria includes radial or irregular streaks, irregular pigmentation, globules and irregular spots, and regression patterns.

A.2.5 Menzies Scoring Method

According to Menzies method, eleven dermoscopic features are adopted for the purpose of distinguishing melanoma melanocytic lesions. The characteristics were divided into two groups (negative and positive). The negative group which is consisting of a symmetry of the lesion and presence of a single color is used to define the lesions as benign. While, the blue-whitish veil, multiple brown dots, pseudopods, radiated streaks, scar-like depigmentation, black dots and globules in the periphery of the lesion, multiple different colors (5 or 6), multiple blue-grey dots, and enlarged pigmentary net are considered to be in the positive group, and utilized to determine the lesion as melanoma. The existence of a positive features (single color and a

Table 3: The dermoscopic characteristics adopted by Menzies scoring method according to [34].

Negative features
1. Point and axial symmetry of pigmentation
2. presence of a single color
Positive features
3. Blue-white veil
4. Multiple brown dots
5. Pseudopods
6. Radial streaks
7. Scar-like depigmentation
8. Peripheral black dots-globules
9. Multiple different colors (5 or 6)
10. Multiple blue/gray dots
11. Broadened network

symmetry lesion), added to the absence of all other negative features structures, is extremely sufficient for the diagnosis of cutaneous melanoma [34] [40]. Table 3 shows the 11 dermoscopic characteristics used in Menzies scoring method.

Appendix B

B.1 Inpainting Approach

Consider the user determine a target region, Ω, to be filled. The source region, Φ, may be defined as the whole image mines the target region ($\Phi = \mathcal{I} - \Omega$) as a dilated band around the target region. The user also determine the size of the template window Ψ. Once all these parameters are defined, the algorithm iterates the following steps until all pixels in the target region are filled:

B.1.1 Computing patch priorities

The task is performed by computing the priority value of each patch.The priority computation is biased toward those patches which are on the continuation of strong edges and which are surrounded by high-confidence pixels. For instance, patches that include corners of the target region will be filled first, since they are surrounded by more pixels from the original image. Given a patch Ψ_p centered at point p for some $p \in \delta\Omega$ (see Figure 8), its priority $P(p)$ is defined as the product of two terms:

$$P(p) = C(p)D(p)$$

where $C(p)$ call the confidence term and $D(p)$ the data term, and they are defined as follows:

$$C(p) = \frac{\Sigma_{q \in \Psi p \bigcap \bar{\Omega}} C(q)}{|\Psi_p|} \quad , \quad D(p) = \frac{|\nabla I_{\frac{\perp}{p}}.n_p|}{\alpha}$$

where $|\Psi_p|$ is the area of Ψ_p , α is a normalization factor ($e.g., \alpha = 255$ for a gray level image), $\delta\Omega$ is a target region border, and n_p is a unit vector orthogonal to the front $\delta\Omega$ in p point (perpendicular). The priority of every border patch is computed with distinct patches for each pixel on the boundary of the target lesion.

The confidence term C(p) can be considered as a measure of the amount of reliable information surrounding the patch center point (pixel p). The main aim is to fill in first those patches, since they have more pixels. The data term $D(p)$ is a function of the strength of isophotes hitting the front $\delta\Omega$ at each iteration. This term boosts the priority of a patch that an isophote "flows" into.

B.1.2 Propagating texture and structure information

Once all the priorities have been computed on the boundary of target region $\delta\Omega$, the patch $\Psi_{\hat{p}}$ with the highest priority is found and filled by image texture data extracted from the source region Φ. The following equation is used to search for the patch which is most similar to $\Psi_{\hat{p}}$ in the source region.

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_{\hat{p}}, \Psi_q)$$

where the distance $d(\Psi_{\hat{p}}, \Psi_q)$ between two generic patches $\Psi_{\hat{p}}$ and Ψ_q is defined as the sum of squared differences (SSD) of the already filled pixels in the two patches. Having found the source exemplar $\Psi_{\hat{q}}$, the value of each pixel to be fill, $p'|p' \in \Psi_{\hat{p}} \bigcap \Omega$ is copied from its corresponding position inside $\Psi_{\hat{q}}$. This process is sufficient to achieve the propagation of both: structure and information from Source Φ to target Ω.

B.1.3 Updating confidence values

After the patch $\Psi_{\hat{p}}$ is filled with new pixel values, the confidence term $C(p)$ is updated in the area delimited by $\Psi_{\hat{p}}$ as follows:

$$C(q) = C(\hat{p}) \forall q \in \Psi_{\hat{p}} \bigcap \Omega$$

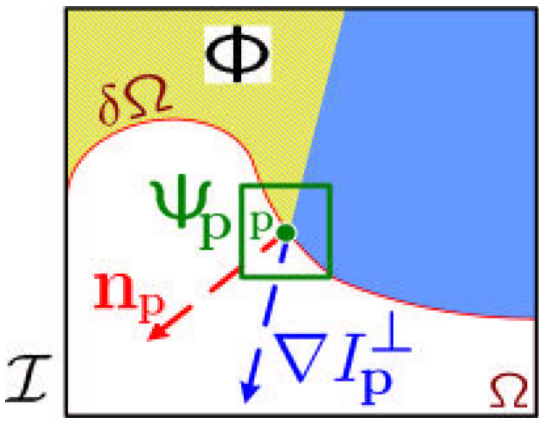


Figure 8: Notation diagram. Given the patch Ψ_p , np is normal to the contour $\delta\Omega$ of the target region or inpainted region Ω, the boundary region is denoted by $\delta\Omega$ and $\nabla I^{\perp}_{p}$ is the isophote (direction and intensity) at point p. The source region is denoted by Φ while the entire image is referred to $\mathcal{I}$.