ABSTRACT

AUTOMATIC MELANOMA DETECTION IN DERMATOLOGICAL IMAGES

By

Javier Eslava Rios

January 2015

Malignant melanoma is one of the most dangerous types of skin cancer. A very important aspect of this type of cancer is that, if detected early, it can be completely removed from the body. These characteristics make the research on automated melanoma detection systems a field with high potential. In this thesis, a system for automatic detection of melanoma is designed, developed and studied. The system is composed of five stages: image acquisition, illumination correction, lesion segmentation, feature extraction and classification. The techniques implemented in illumination correction are based in morphological operators and the Retinex algorithm. The four proposed methods for lesion segmentation include Otsu's method thresholding, GVF Snakes, and two novel methods based in Mean Shift clustering using color and texture information. The classification stage makes use of linear discriminant analysis and SVMs. In addition, a GUI tool that takes advantage of the mentioned techniques is created and presented.

AUTOMATIC MELANOMA DETECTION IN DERMATOLOGICAL IMAGES

A THESIS

Presented to the Department of Electrical Engineering

California State University, Long Beach

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Electrical Engineering

Committee Members:

Christopher Druzgalski, Ph.D. (Chair)
Fumio Hamano, Ph.D.
James Ary, Ph.D.

College Designee:

Antonella Sciortino

By Javier Eslava Rios

B.S., 2013, Universidad Autonoma de Madrid

January 2015

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Dr. Chris Druzgalski for providing directions, insight and tools in order to make this thesis possible. Thank you for introducing me to the field of biomedical engineering and for always talk about the thesis, approaching challenges and providing guidance in such a friendly attitude.

I would also like to thank all the people working at the College of Engineering especially the graduate advisor Dr. James Ary, for helping me and all the students on their way to obtain the title with astounding efficiency.

A big thank you goes to my family for all the unconditional support that they gave throughout the whole program and for always encouraging me to improve myself. None of this extraordinary experience would have been possible without your complete support; you can feel proud of being part of something that has changed my life in a very positive way.

Thanks to all my friends that, while really far away from each other, somehow make it feel like we are very close. Thanks to Rafael Martin in particular for providing insightful and relevant comments for the thesis.

Finally, I would like to say thank you and dedicate this thesis to Leyre. Thank you for your determination in the pursuit of a dream and for making it come true, this thesis is, in many ways, ours.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iii

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS x

CHAPTER

1. INTRODUCTION 1

    Objectives 1
    Thesis Structure and Timeline 2
    Dermatologic Background 3

2. STATE-OF-THE-ART TECHNIQUES IN IMAGE PROCESSING FOR MELANOMA DETECTION 9

    Image Acquisition 10
    Illumination Correction 10
    Lesion Segmentation 17
    Feature Extraction 22
    Classification 24

3. SYSTEM FOR AUTOMATIC MELANOMA DETECTION 30

    Illumination Correction 30
    Lesion Segmentation 38
    Feature Extraction 47
    Feature Selection 51
    Classification 51
    Conclusion 53

CHAPTER Page

4. EXPERIMENTS AND RESULTS 54

Dataset 54
Experiments Protocol 56
Illumination Experiments 57
Segmentation Experiments 65
Classification Experiments 77

5. GRAPHICAL USER INTERFACE TOOL 82

Features 82
Program Structure 85
Conclusion 87

6. CONCLUSIONS 88

Summary of Work 88
Future Work 90

REFERENCES 92

LIST OF TABLES

TABLE Page

1. Mean Values of Variance in Original, Morpho and Retinex Datasets 60

2. Average of Cumulative Mean Differences 63

3. Segmentation Results..... 66

4. Averaged Mean Difference between Lesion and Skin Areas in each Image.... 72

5. Standard Deviation Average in Lesion Area 76

6. Classification Results..... 79

7. Specificity, Sensitivity and Accuracy of Classifiers..... 80

LIST OF FIGURES

FIGURE Page

1. Timeline of the work relevant to this thesis 2

2. General melanoma detection system 9

3. Example of uneven illumination in melanoma images 11

4. Original image (left) and dilated image (right) 12

5. Original image (left) and eroded image (right). 13

6. Original image (left) and morphologically opened image (right) 14

7. Original image (left) and morphologically closed image (right) 15

8. Retinex algorithm results for different iterations 17

9. Original image (left) and thresholded image (right). 18

10. Edge detection using a Sobel operator 20

11. GVF snakes performed over a shape. 22

12. 2-dimensional set of vectors separated by the SVM hyperplane 26

13. Data separated by a non-linear Gaussian radial basis kernel 28

14. Example of indexed image where each pixel points to a position in the colormap 31

15. Intensity image where every pixel refers to a single value in a scale 32

16. RGB image where each pixel is defined by three values stored in three matrices. 33

17. Original melanoma image (left) and extracted background (right). 35

FIGURE Page

18. Original melanoma images (top row) and corrected results by morphological operator (bottom row) 36

19. Original images (top row) and corrected results with Retinex method (bottom row) 38

20. Original image (left) and thresholded result with Otsu's method (right). 40

21. Thresholded image (left) and result (right) by performing preprocessing actions that remove noise and separate the lesion area from other thresholded areas 41

22. Original images (top row) and segmented results (bottom row) using Otsu's method 42

23. Original image (left) and calculated edge (right) using Canny method..... 43

24. Original images (top row) and segmented results (bottom row) using GVF snakes..... 44

25. Mean Shift convergence calculation for a feature space based on intensity values. 46

26. Segmentation of a) original image by performing b) Mean Shift with color information or c) Mean Shift with color and texture information to obtain lesion d) 47

27. Horizontal asymmetry performed by 'folding' the mask and obtaining the non-matching pixels..... 49

28. First level coefficients after calculating the wavelet transform 50

29. Sample of melanoma images from DermQuest database 54

30. Sample of non-melanoma images acquired for this thesis..... 55

31. Variances of melanoma and non-melanoma sets for the Original, Morpho and Retinex sets. 58

32. Variance difference between Original and the two corrected sets..... 59

FIGURE Page

33. Mean values of the red, green and red channels for the Original, Morpho and Retinex sets (melanoma)..... 61

34. Mean values of the red, green and red channels for the Original, Morpho and Retinex sets (non-melanoma) 62

35. Cummulative mean difference between Original and both Morpho and Retinex sets. 64

36. Mean differences between lesion area and skin area for melanoma (left column) and non-melanoma (right column) subsets. 69

37. Mean differences between lesion area and skin area for melanoma (left column) and non-melanoma (right column) subsets. 70

38. Mean differences between lesion area and skin area for melanoma (left column) and non-melanoma (right column) subsets 71

39. Standard deviation of the lesion area for melanoma (left column) and non-melanoma (right column) subsets. 73

40. Standard deviation of the lesion area for melanoma (left column) and non-melanoma (right column) subsets. 74

41. Standard deviation of the lesion area for melanoma (left column) and non-melanoma (right column) subsets. 75

42. Actual value vs predicted value diagram 78

43. ROC curve for SVM classifier..... 81

44. Database preview in the melanoma detection tool 83

45. Illumination correction performed over loaded image 84

46. Otsu's method incorrect segmentation of original image 84

47. Otsu's method properly segmenting corrected image 85

48. Classification of a melanoma image 86

LIST OF ABBREVIATIONS

MATLAB	MATrix LABoratory
GUI	Graphical User Interface
PCA	Principal Component Analysis
UV	Ultraviolet
DNA	Deoxyribonucleic Acid
UVA	Ultraviolet A
UVB	Ultraviolet B
TNM	Tumor, Node, Metastasis
OCT	Optical Coherence Tomography
GVF	Gradient Vector Flow
LDA	Linear Discriminant Analysis
SVM	Support Vector Machines
ANN	Artificial Neural Networks
RGB	Red, Green, Blue
HSV	Hue, Saturation, Value
CMYK	Cyan, Magenta, Yellow, black
PDF	Probability Density Function
MS	Mean Shift
SD	Standard Deviation

ROC	Receiver Operating Characteristic
GUIDE	Graphical User Interface Development Environment
PNG	Portable Network Graphics
BMP	BitMaP
JPEG	Joint Photographics Expert Group
TIFF	Tagged Image File Format
OS	Operating System

CHAPTER 1

INTRODUCTION

In this chapter, an introduction to the work accomplished, structure of the thesis, and a brief medical background of the problem are provided. This thesis focuses on the automatic detection of melanoma lesions which leads to the creation of a full system that is able to detect and predict if a certain lesion has a risk of being a melanoma.

Objectives

Melanoma, while not the most common, is the most dangerous form of skin cancer [1]. It is the skin cancer that causes the most deaths with an increasing number of new cases per year doubling the figures of 1978 [2]. The treatment of this cancer usually involves surgery to remove the affected area and has a very high rate of success of completely eliminating the cancer from the body when diagnosed in its early stages [3]. Early detection is, thus, a critical step in saving a patient's life. These facts are the main motivations for the research on ways to enhance the work of dermatologists by providing tools that help in diagnosing this type of cancer.

The main objectives of this thesis are the research on state-of-the-art techniques in image processing related to melanoma detection, the design, development and implementation of a system that is able to make use of said techniques and the development of a tool that provides a graphical user interface that takes full advantage of of the system. In order to achieve these goals, the software tool MATLAB® will be the

preferred platform to code the algorithms. In addition, experiments that measure the accuracy and performance of the techniques implemented will be carried out by acquiring a database of images that will serve as the basis for these experiments.

Thesis Structure and Timeline

The different sections of this thesis describe to the work developed in a 6 month period. The time-structured diagram of the work is depicted by Figure 1.

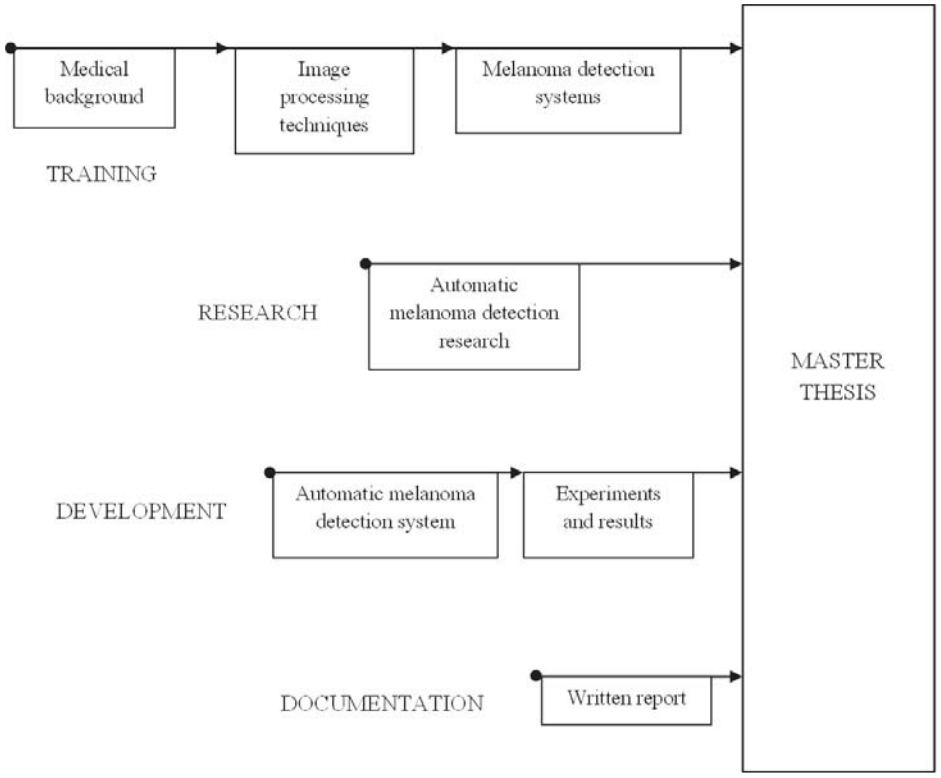


FIGURE 1. Timeline of the work relevant to this thesis.

This thesis is organized in six chapters. Chapter 1 serves as the introduction of the problem giving medical background on melanoma and providing the motivation and objectives set for the accomplishment of the thesis.

Chapter 2 gives a review in existing techniques and methods that compose a melanoma detection system. Background of the selected techniques in image processing is presented in detail.

In chapter 3, the main techniques developed for the thesis are described giving general information on the methodology followed to compose the full system. Specific details are given on each algorithm that was used in every task of the melanoma detection system.

Chapter 4 describes the experiments conducted to test each of the techniques' performance and the whole system accuracy. Comparisons between techniques are presented in order to have a general vision of how a method performs regarding the others.

Chapter 5 outlines the tool developed by making use of all the techniques and methods presented in chapter 3. The tool contains a graphical user interface (GUI) that shows results on screen. A description of its features and structure is also provided.

Chapter 6 concludes the thesis by giving a summary of the work done and adding future work suggestions based on the results obtained in this work.

Dermatologic Background

Melanoma is the general name given to a type of skin cancer that affects the melanocytes, the cells that produce the dark pigment melanin, responsible for the color of skin. This cancer is highly invasive due to its ability to generate metastasis. Melanomas are more common in pale skinned people that live in sunny locations. The most frequent area to develop a melanoma is the back for men and the legs for women [4].

Causes

The risk of contracting melanoma depends on two factors, the family genetic history and, more importantly, the exposure to ultraviolet (UV) light that causes DNA damage [3].

The source of the light radiation is not significant, just the wavelength of said light. It can come from a natural source, such as the sun; or an artificial source, for example a tanning table. UVB light which has a wavelength between 315 and 280 nm, is absorbed by the skin cell DNA and results in DNA damage called cyclobutane pyrimidine dimers. UVA light is also absorbed by skin cell DNA but due to its longer wavelength (400–315 nm) the efficiency for the absorption rate is much lower [5].

Some genes and mutations increase the risk of developing melanoma. A mutation of the gene CDKN2A destabilizes p53 which is related with half of the human cancers. Another mutation results in a nonfunctional inhibitor of CDK4, a cyclin-dependent kinase that promotes cell division. Any mutation that causes xeroderma pigmentosum, a disease that reduces the ability to repair damage caused by UV light, also increases the risk of developing melanoma [6].

Melanoma Diagnosis

Melanoma detection is based mostly on visual inspection by a dermatologist. The main technique to decide if a lesion is a potential melanoma is the use of the ABCD rule devised by Friedman et al. [7]. The ABCD acronym refers to the Asymmetry, Border, Color and Diameter of the lesion.

Regarding asymmetry, melanomas often appear as highly asymmetric lesions. The asymmetry is assessed when one half of the lesion doesn't match in shape the other. In addition, melanomas are also asymmetric in color, meaning that the pigmented regions inside the lesion halves are different [7].

Melanomas usually present irregular borders having uneven, ragged or notched sections. The borders are often highly defined in terms of color gradient with the skin, having abrupt edges that change from lesion to normal skin color without a soft progression [7].

The colors found in melanomas are typically shades of dark brown, black, red, white and blue. Melanomas often present more than one color in the same lesion. In addition, the pigmentation of one color is not uniform with variable degrees [7].

All lesions that are bigger than 6 mm in diameter have higher chances of being a melanoma. This characteristic doesn't mean that lesions smaller than 6 mm could not be melanoma, it just states the risk factor based on the diameter size. There are melanomas that start below 6 mm in diameter therefore all lesion sizes must be taken into consideration [7].

Other characteristics have also been proven decisive in melanoma detection. Steep changes in size over a short period of time are one of the situations where a lesion can be classified as a potential melanoma.

The method that completely determines if a lesion is a melanoma is the skin biopsy. However, this method requires minor surgery and local anesthesia as part of the lesion is extracted for testing [8].

Stages of Melanoma

Melanoma stages are classified following the TNM system. TNM describes the thickness of the lesion (T), the possibility of spreading to the nearby lymph nodes (N) and the presence of metastasis to other parts of the body (M).

Stage 0, also called melanoma *in situ*, is the appearance of abnormal melanocytes in the epidermis. In this stage the cancer has not yet appeared, but there is a potential risk that the abnormal melanocytes may become cancer [3].

In stage 1 the cancer has already formed. There are two substages called stage IA and stage IB. Stage IA refers to a lesion without ulceration that is less than 1 mm thick. For a lesion with ulceration and a thickness of less than 1 mm, or a lesion without ulceration and a thickness between 1 and 2 mm, the stage is defined as stage IB [3].

Stage II is, in this case, divided into three substages; stage IIA, stage IIB and stage IIC. In stage IIA, the lesion presents either a thickness between 1 and 2 mm with signs of ulceration, or a thickness between 2 and 4 mm without signs of ulceration. Regarding stage IIB, the lesion has either reached 4 or more mm of thickness without ulceration, or it is more than 2 mm but less than 4 mm thick when ulceration is present. Finally, stage IIC describes a lesion with ulceration that is more than 4 mm thick [3].

Stage III describes the melanoma's first spreading events outside the primary lesion. The thickness and ulceration are not taken into account in this stage. Alternately, one or more of the following statements are true: Cancer has spread to one or more lymph nodes. Lymph nodes may be joined together. Cancer may be in a lymph vessel

between the primary lesion and nearby lymph nodes. Small tumors appear not more than 2 cm away from the primary lesion [3].

The final stage of the melanoma is stage IV, where the cancer develops distant metastasis and spreads to other parts of the body affecting other organs. Diagnosing the cancer in this stage means a very low rate of survival due to the fact that it is very difficult to effectively treat the cancer and eradicate it as a whole [3].

Treatment

The type of treatment required for melanoma varies depending on the stage in which the cancer is found. In most favorable cases, treatment involves surgery with the occasional use of adjuvant therapy [8].

Regarding stage 0 melanomas, the common practice involves surgery (wide excision) to remove the melanoma. A margin of 0.5 cm around the healthy skin is used and if the edges of the removed area present cancer cells, the process is repeated with a wider margin [8].

Similarly to stage 0, stage I melanomas are treated by wide excision. In this case, the margin to be used depending on the thickness of the melanoma however margins of over 2 cm are not used, because make healing more difficult and doesn't improve the treatment [8].

In stage II melanomas, the same procedure as in stage I melanomas is used with the exception of adding a treatment using interferon (adjuvant therapy) after surgery for patients that may have lymph nodes with cancer. This treatment is advised in order to reduce the chance for the cancer to come back [8].

Cancers that have reached the lymph nodes as stage III melanomas require wide excision of the primary lesion as in earlier stages, along with lymph node dissection. Besides adjuvant therapy with interferon, radiation therapy on the areas where the lymph nodes were removed is also an option. In the case of the presence of multiple melanomas, all of them should be removed [8].

Stage IV melanomas are very hard to cure mostly due to the fact that the cancer has spread to other parts of the body. Skin melanomas and affected lymph nodes are removed as well as metastases in other organs when it is possible to do so. In the event of the impossibility to remove them, they may be treated with radiation, immunotherapy, targeted therapy or chemotherapy [8].

CHAPTER 2

STATE-OF-THE-ART TECHNIQUES IN IMAGE PROCESSING FOR MELANOMA DETECTION

In this chapter, a comprehensive description of the most commonly used techniques for melanoma detection is given. The image processing techniques available can be algorithms that require some form of user input in addition to the target image or fully automatic systems that only require the input image to process everything. This chapter focuses on automatic techniques since the specific scope of this thesis is building an automatic detection system.

A basic melanoma detection system is typically composed of the stages depicted in Figure 2. Most of these stages are often shared with other detection systems in terms of high level functionality.

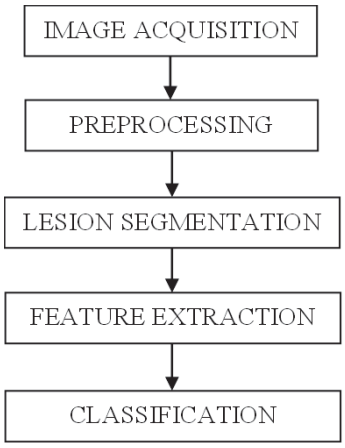


FIGURE 2. General melanoma detection system.

Image Acquisition

The first stage in a detection system is how the image is acquired. Most of the time, the images are already acquired and are, thus, obtained from a database. However, the process of obtaining the image and the devices used differ depending on the target to acquire.

Regarding dermatologic images, the most common and simple procedure is acquiring images with a device based on a digital camera capable of macro photography [9]. Since illumination is an important factor in this process, the device usually carries a polarized light that highlights the target area to have the best possible illumination and color characteristics, without undesired artifacts like reflections. Other acquisition devices include Optical Coherence Tomography (OCT) [10] capable of displaying a three dimensional representation of the lesion including its thickness.

Illumination Correction

Dermatologic images often present uneven illumination that is commonly present as a shadow gradient that goes from bright to dark across the image [11]. Figure 3 describes a common case of uneven illumination in dermatologic images. An image with illumination issues makes the subsequent stages of the detection system to loose precision. In addition, the real colors of the image are affected by these problems, making decisions based on lesion color more difficult.

The need for an illumination correction stage is evident given the issues previously discussed. The problem of correcting uneven illumination is a challenging one in image processing, as most approaches give inconsistent results. One of the most

widely known methods for illumination correction is the one that makes use of morphological operators. There also are specific algorithms that try to estimate the illumination from an illumination-reflectance model. This thesis describes one of the first algorithms in illumination correction using an estimation of lightness called Retinex [12].

FIGURE 3. Example of uneven illumination in melanoma images.

Illumination Correction Using a Morphological Operator

This technique comes from the idea that an image's background can be extracted by applying a morphological operator [13]. Morphological operators are often used in binary images and consist on moving a structuring element pixel by pixel across the image, making decisions based on the operation and the neighborhood covered by the structuring element. The structuring element is a predefined shape that covers a certain area of pixels. Some examples of widely used structuring elements include a 3x3 square or a disk of radius *r*.

The two basic operations which the rest of the morphological operators are based upon are called dilation and erosion [14].

Dilation. Dilation is an operation whose visual effect in a binary image is to thicken the parts of the image that are set to a logical 1. The mathematical definition of the dilation of a set A with a structuring element B, denoted $A \oplus B$ is:

$$A \oplus B = \left\{ z | \left(\hat{B} \right)_z \cap A \neq \emptyset \right\} \tag{2.1}$$

$\left(\hat{B} \right)_z$ is the reflection of the structuring element, translated by z. The direction and extension of the thickening is defined by the shape of the structuring element chosen. Figure 4 illustrates an example of dilation in a binary image. Graphically, every time the reference pixel of the structuring element (if set to the center pixel as commonly done) is a logical 1 in the image, the rest of the neighborhood defined by the structuring element's shape is 'filled' with logical 1's [14].

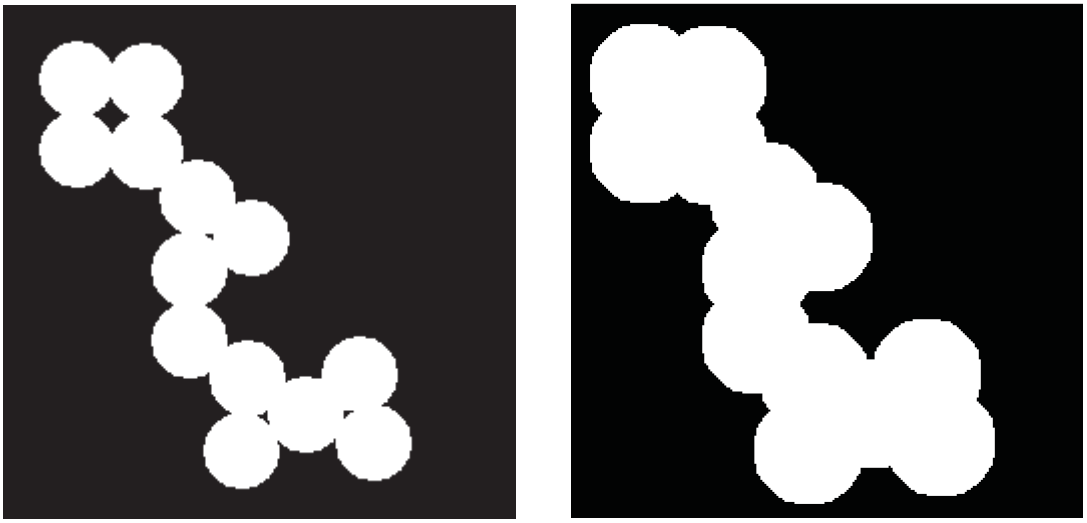


FIGURE 4. Original image (left) and dilated image (right).

Erosion. Erosion is the opposite operation of dilation. In this case the image areas that are set to logical 1's are shrunk or, for the sake of equivalence, the logical 0's are dilated following the structuring element's shape restrictions. The mathematical definition for the erosion of a set A with a structuring element B, denoted $A \ominus B$ is:

$$A \ominus B = \{z|(B)_z \subseteq A\} \tag{2.2}$$

Figure 5 illustrates an example of erosion in a binary image. Complementary to the dilation process, the structuring element's neighborhood is set to logical 0 for every pixel in the image that is a logical 0. The structuring element's reference pixel is aligned with each pixel that is 0 in order to make the calculation [14].

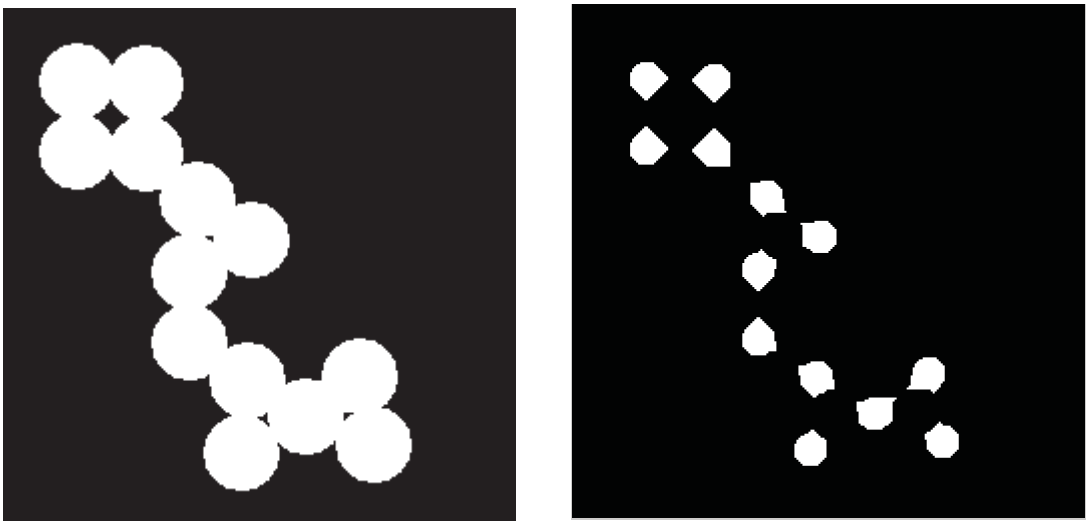


FIGURE 5. Original image (left) and eroded image (right).

Opening. The morphological opening is a compound operation that involves performing an erosion followed by a dilation. The mathematical expression for a morphological opening of A by B is:

$$A \circ B = (A \ominus B) \oplus B \tag{2.3}$$

In the example provided by Figure 6, it is easy to see that one of the effects of performing a morphological opening is that the big areas set to logic 1 remain mostly unaltered and the smaller areas (smaller than the size of the structuring element) are deleted. Hence, opening is a useful operation to remove small noise in an image without altering the main image [14].

FIGURE 6. Original image (left) and morphologically opened image (right).

Closing. The morphological closing is similar to the opening however, the order of the basic operations is inverted. In this case, a dilation is performed first and then it is followed by an erosion. The mathematical expression is:

$$A \bullet B = (A \oplus B) \ominus B \tag{2.4}$$

Regarding closing, the results from this operation tend to leave unaltered big and small areas of logical 1's and set small areas of logical 0's to logical 1's. This operation effectively fills the 'holes' of an image [14] (see Figure 7).

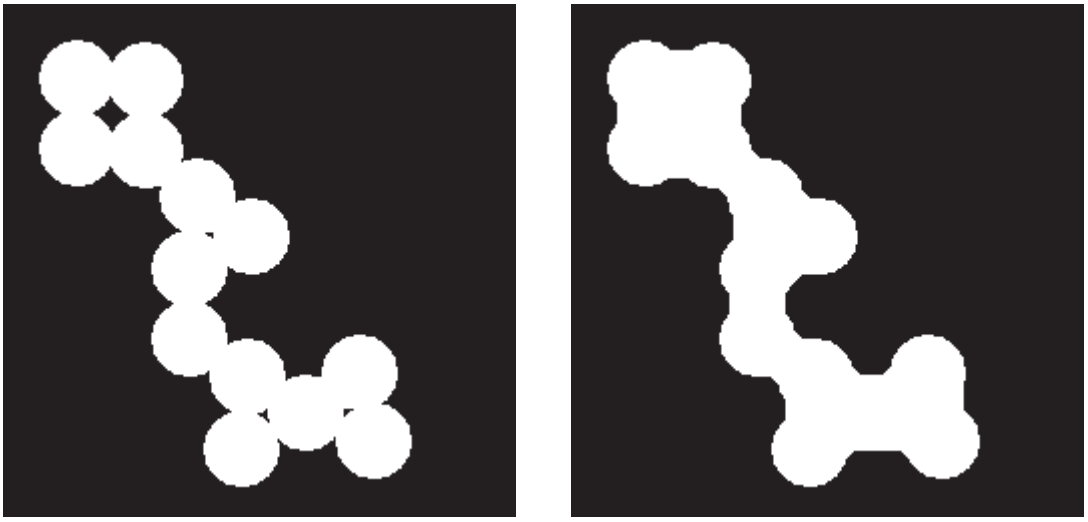


FIGURE 7. Original image (left) and morphologically closed image (right).

Extracting the background image using a morphological closing. The main concept on this technique is the subtraction of the background image, which is the calculated illumination gradient [13]. In the case of dermatological images, the background refers to the shadows and uneven illumination present in melanoma images. The erosion and dilation operators are performed over a grayscale image; therefore the definition of the operators is, in this case, slightly different. The erosion of *f* by a structuring element *b* at any location (*x, y*) is defined as the minimum value of the image in the region coincident with *b* when the reference of *b* is at (*x, y*). The dilation of *f* by a structuring element *b* at any location (*x, y*) is defined as the minimum value of the image in the region coincident with $\hat{b}$, where $\hat{b}$ is the reflection of the structuring element *b*. Dilation and erosion in a given location are, thus, mathematically defined with the following equations:

$$f \ominus b(x,y) = \min_{(s,t)\in b}\{f(x+s, y+t)\} \tag{2.5}$$

$$f \oplus b(x,y) = \max_{(s,t)\in b}\{f(x-s, y-t)\} \quad (2.6)$$

where *x* and *y* are incremented through all values required so that the reference of *b* visits every pixel in *f*. By using a large structuring element, the elements extracted by the closing operator are the ones which are bigger than the structuring element and darker than their surroundings; effectively extracting the shadows in the image. The structuring element should be large enough to not alter the lesion intensity values. The corrected image is obtained by subtracting the original image to the values extracted by the closing operator. The mathematical expression for this illumination correction is:

$$T_b(f) = f \bullet b - f \quad (2.7)$$

where f denotes a grayscale image and b the structuring element [13].

Illumination Estimation Models

One of the first algorithms that calculate the estimation of lightness is the Retinex model introduced by Land and McCann [12]. The Retinex model distinguishes between physical reflectance, the sensation of lightness, and perceived reflectance. The model aims to calculate the sensation of lightness trying to obtain the true colors of an image with illumination variations [15].

This model's calculations are based on four operators; ratio, product, reset and average [15]. The basic principle for the computation of lightness at a given pixel is the comparison of that pixel's value with others following a specific path. A sequential product is accumulated along the path and it gets multiplied by the ratio of each pair of compared pixels and averaged. The length of the path controls how each pixel is affected by others. Shorter paths mean that the pixels are influenced by localized regions of the

image. Extremely long paths produce an output that is the same as the input. Figure 8 depicts an example of the Retinex model applied with different iterations that directly correlate with path length.

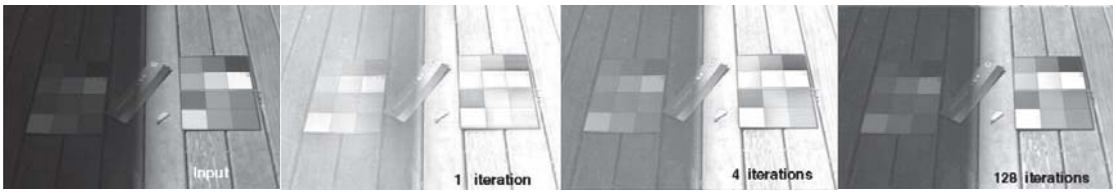


FIGURE 8. Retinex algorithm results for different iterations.

Lesion Segmentation

The lesion segmentation step is the main process in a melanoma detection system. Segmenting a lesion is a process that aims to define the area which is regarded as melanoma and to extract the region of interest. Extracting the region of interest gives valuable information such as borders, pixel values, area, etc.

There are several segmentation methods available in the image processing literature based on different image features [16]. The methods studied are the ones that offer the best results given the characteristics of a dermatologic image, where there is a large light area considered skin and a typically darker area referring to the lesion. Some of them include global thresholding of a grayscale image, color clustering techniques or edge detection algorithms [16].

Segmentation Using Global Thresholding

This method of segmentation focuses on calculating a black and white thresholded image based on the values of the pixels of an intensity image. Figure 9 illustrates an example of a thresholded image.

FIGURE 9. Original image (left) and thresholded image (right).

The basic method is selecting a threshold based on the average of the pixel values. The result image is calculated by setting all the pixels that have a value below the threshold to logical 0 and all the pixels whose value is above or equal the threshold to logical 1 as:

$$T_{global}(g) = \begin{cases} 0 & if < t \\ 1 & if \geq t \end{cases} \tag{2.8}$$

where the image is depicted as T_{global}, the pixel value is depicted as g, and the threshold is depicted as t. Other more effective methods include calculating the threshold by finding an optimal value based on an image's histogram [13].

Once the image has been thresholded, the biggest element belonging to the region of interest is selected and the rest are removed to retain the lesion.

Segmentation Using Clustering Techniques

Clustering techniques represent a given image using a reduced set of colors, effectively clustering areas that share similar colors with only one. The main concept of the segmentation is that the difference between the skin and lesion colors allows the clustering techniques to define the lesion with one color and the rest with other colors.

One of the most popular methods is *k*-means clustering [17]. *K*-means is a clustering method that aims to partition a set of *n* observations in *k* clusters. Each observation belongs to the cluster closest to the mean.

The most basic procedure starts by randomly partitioning the space in *k* sets $S = \{S_1, S_2, \cdots, S_k\}$, then calculating the mean of the observations inside each cluster. This mean is defined as the centroid *c* and calculated as:

$$c_i = \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j \tag{2.9}$$

Once the new centroids are calculated, a new partition of the space is performed and the process is repeated until the centroids converge.

Another method is Mean Shift clustering [18]. This is a novel method in the application of dermatological images and one of the techniques developed during the work of this thesis. Therefore, a detailed description is included on the next chapter.

Segmentation Using Edge Detection

Given the characteristics of the dermatological images to work on, edge detection is a viable option to segment a melanoma since there is a change of intensity between skin and lesion [16]. Many edge detector calculations use a gradient operator to obtain the gradient magnitude $g(x, y)$ and direction $\theta(x, y)$ which are defined as:

$$g(x,y) \cong (\Delta x^2 + \Delta y^2)^{\frac{1}{2}} \tag{2.10}$$

$$\theta(x,y) \cong \tan^{-1}\left(\frac{\Delta y}{\Delta x}\right) \tag{2.11}$$

where:

$$\Delta x = f(x+n,y) - f(x-n,y) \tag{2.12}$$

$$\Delta y = f(x,y+n) - f(x,y-n) \tag{2.13}$$

Denoting $f(x,y)$ as an image function and *n* as a small integer. The gradient operators are typically a mask that is convolved with the image. A basic example of an operator is $[1 \quad 0 \quad 1]$ which allows calculating the horizontal gradient and, thus, the vertical edges.

Popular variations extended the neighborhood to a 3x3 matrix which resulted in an averaging of the adjoining pixel values. Examples include the Prewitt and Sobel operators [19] (see Figure 10).

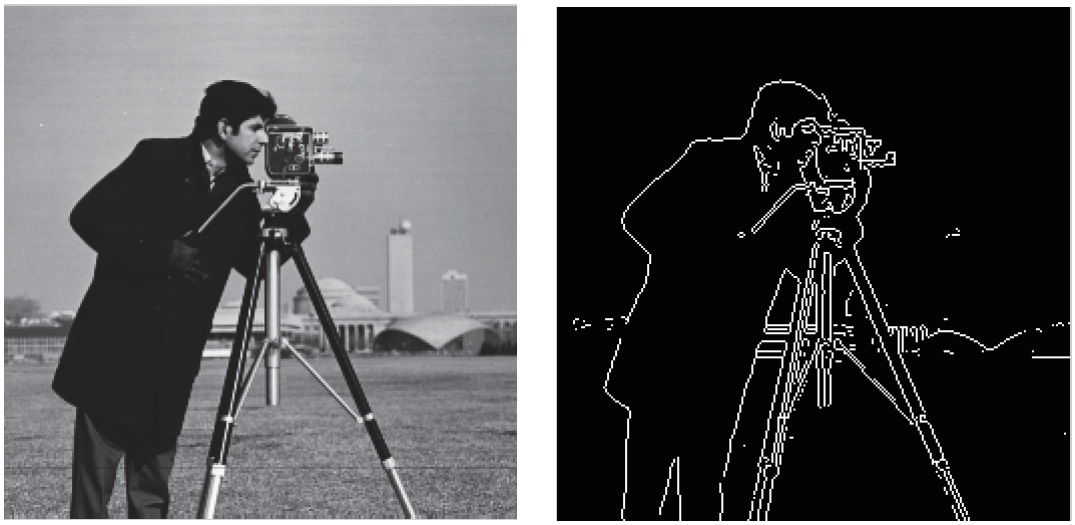


FIGURE 10. Edge detection using a Sobel operator.

One of the most popular edge detectors in this category is the Canny algorithm [20]. This algorithm involves a four step process. First, the image is smoothened by applying a Gaussian filter in order to lower potential high frequency noise. The gradient is then calculated as previously explained and an edge thinning technique called non-maximum suppression is applied. The non-maximum suppression estimates which edges are more likely to be noise or part of a wider edge and removes them. Finally, the edges are calculated by thresholding the gradient image with hysteresis. A high threshold is set to get the strong edges that form a line removing strong values that don't follow a line. The low threshold is then used while tracing the strong edges and completing a line as long as it is connected to a strong one.

An edge detection method that has proven to yield good results is Gradient Vector Flow (GVF) snakes [21]. Snakes are curves that can move thanks to the influence of internal forces intrinsic to the curve and external forces derived from the image values.

The general process of this algorithm is composed of four steps. First of all, a starting curve $x(s,t)$ is drawn close to the boundary to detect. An edge map $f(x,y)$ is then calculated from the original image. The GVF field is then described as the vector field $\mathrm{v}(x,y) = \big(u(x,y), v(x,y)\big)$ that minimizes the energy functional

$$\varepsilon = \iint \mu\big(u_x^2 + u_y^2 + v_x^2 + v_y^2\big) + |\nabla f|^2 |\mathrm{v} - \nabla f|^2 dxdy \tag{2.14}$$

The GVF field is then calculated by solving the following Euler equations:

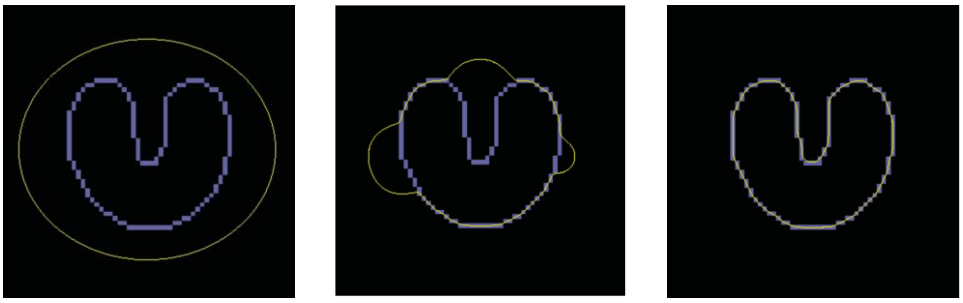
$$\mu\nabla^2 u - (u - f_x)\big(f_x^2 + f_y^2\big) = 0 \tag{2.15}$$

$$\mu\nabla^2 u - \big(u - f_y\big)\big(f_x^2 + f_y^2\big) = 0 \tag{2.16}$$

where ∇^2 is denoted as the Laplacian operator. Finally, the new curve $x_t(s,t)$ is calculated as:

$$x_t(s,t) = \alpha x''(s,t) - \beta x''''(s,t) + \mathrm{v} \tag{2.17}$$

where α and β are denoted as weighting parameters that control the curve's tension and rigidity, respectively. This last step is then iterated to obtain the best curve fitting of a boundary [21] (see Figure 11).



a) b) c)

FIGURE 11. GVF snakes performed over a shape. a) Initial contour b) Iteration 25 c) Iteration 125.

Feature Extraction

The feature extraction stage involves obtaining a set of features that has high intraclass correlation and low interclass correlation. In other words, the aim is to search for features that are commonly found in melanomas and are uncommon to other types of lesions. The extracted features are accumulated in a vector called feature vector.

Extracting ABCD Rule Features

A widely used approach for feature extraction in melanoma images is the obtainment of the ABCD rules that are available in the image. While asymmetry, border

and color are easily obtainable; diameter is often skipped due to the fact that most dermatological images don't have any reference to determine the lesion's size [22, 23]. In order to obtain a score for each of the lesions, a TDS system is employed where each feature is given a weight depending on their importance and is described as

$$TDS = 1.3 \cdot A + 0.1 \cdot B + 0.5 \cdot C + 0.5 \cdot D \tag{2.18}$$

where the asymmetry A is a value between 0 and 2, 0 when there is no asymmetry, 1 when there is asymmetry in one axis and 2 when there is asymmetry in both axes. The border B is divided in eight partitions and each one is determined separately. If the border of one partition is irregular, one point is added to the score up to a maximum of 8 points when all the border partitions are regarded as irregular. Color C considers 6 common colors found in melanoma; white, red, light brown, dark brown, gray, blue or black. If any of these colors is found, one point is added to a maximum of 6. The D is not for diameter but for different structural elements like pigmented networks, points, areas without structure, etc. In this case, the maximum value is 5. When the total score is calculated, a lesion is classified as benign if the score is lower than 4.76 and suspicious of potential melanoma if the score falls between 4.76 and 5.45. A score above 5.45 is classified as melanoma [22].

Extracting Texture Features

Other feature that is relevant to the classification of melanoma is the information stored in the image's texture. The texture describes the arrangement of intensity values in the image in regards of the position of each pixel.

There are different methods to obtain an image's texture, and each method provides a different texture highlighting a specific characteristic. For instance, an edge detector previously introduced in the preceding section, is a way to obtain a texture based on the edges. Other methods of effectively extracting the texture include Gabor filtering [24], local binary patterns (LBPs) or wavelet transformations such as tree-structured [25] or pyramid-structured wavelets.

Extracting Color Features

Apart from the set of colors described by the ABCD rule, general color features have also been proven to perform well in melanoma classification systems. A color based feature extraction is the calculation of color histograms and measurement of the correlation between non melanoma and melanoma images [26]. Other measures include the calculation of variance and mean of the segmented lesion, since a melanoma commonly presents more color variations than a non-melanoma lesion.

Classification

The last step in a detection system is the classification of an input and the consequent decision. The classifier usually needs a set of positive (melanoma images) and negative (non-melanoma) images in order to use the feature vectors, described in the previous section, for training.

Linear Classifiers

Linear classifiers make the classification based on a linear combination of the values in the feature vector. One of the most popular linear classifiers is Linear Discriminant Analysis (LDA) [27], closely related to PCA. LDA aims to find a vector $\vec{w}$

which best separates the classes and it's normal to the class hyperplane where the features $\vec{x}$ are projected by performing $\vec{x}_{proj} = \vec{x} \cdot \vec{w}$. In the case of two classes, which is the situation often faced in melanoma detection, LDA assumes normally distributed probability density functions $p(\vec{x}|y = 0)$ and $p(\vec{x}|y = 1)$ where $\vec{x}$ is denoted as the feature vector and y as the class label [27]. Normal distribution implies that each class has mean and covariance denoted as $(\vec{\mu}_0, \Sigma_0)$ and $(\vec{\mu}_1, \Sigma_1)$ respectively. The means and covariances are predicted using a training set of data and, thus, a vector $\vec{w}$ is found with:

$$S = \frac{(\vec{w} \cdot (\vec{\mu}_1 - \vec{\mu}_0))^2}{\vec{w}^T \cdot (\Sigma_0 + \Sigma_1) \cdot \vec{w}} \tag{2.19}$$

where S denotes the ratio of the variance between the classes to the variance within the classes, which describes the separation between classes. The maximum separation is found when:

$$\vec{w} \propto (\Sigma_0 + \Sigma_1)^{-1}(\vec{\mu}_1 - \vec{\mu}_0) \tag{2.20}$$

Once the projection vector is found, any feature vector can be projected on the hyperplane and a decision can be made based on the established threshold *c,* found as:

$$c = \vec{w} \cdot \frac{1}{2}(\vec{\mu}_0 + \vec{\mu}_1) \tag{2.21}$$

Support Vector Machines

SVMs are supervised learning models that are used for classification and regression analysis [28]. SVMs aim to find a hyperplane that best separates a set of *n*-dimensional classes. As Figure 12 depicts, in the basic example of two sets of two-dimensional classes, the hyperplane found by the SVM model is a line, using as a margin reference some of the dataset points denoted as the support vectors.

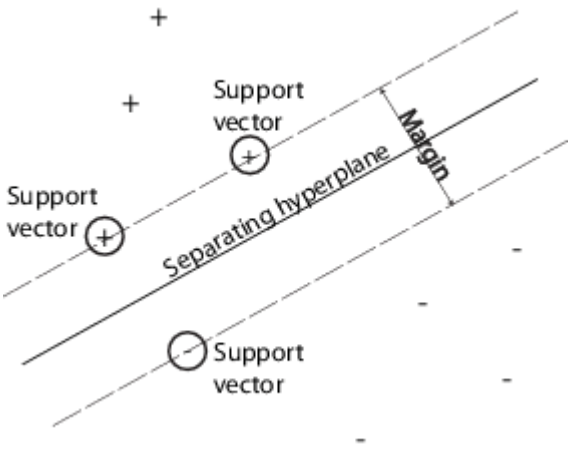


FIGURE 12. 2-dimensional set of vectors separated by the SVM hyperplane.

Formally, a hyperplane H is defined as the set of points *x* that satisfy:

$$w \cdot x + b = 0 \tag{2.22}$$

where · is denoted as the dot product, *w* the normal vector to the hyperplane (also known as weight vector) and *b* the bias. A linear classifier is characterized by the set of pairs (*w*,*b*) that satisfies the following inequalities for any pattern x_i in the training set:

$$w \cdot x_i + b > +1 \quad if \quad y_i = +1 \tag{2.23}$$

$$w \cdot x_i + b < -1 \quad if \quad y_i = -1 \tag{2.24}$$

denoting y_i as each class of the data set.

For all points from the hyperplane H ($w \cdot x + b = 0$), the distance between origin and the hyperplane H is $\frac{b}{||w||}$. The patterns from the class -1 that satisfy the equality $w \cdot x + b = -1$ are considered, and the hyperplane H_1 thus, determined; the distance between origin and the hyperplane H_1 is equal to $\frac{|-1-b|}{||w||}$. Similarly, the patterns from the class +1 satisfy the equality $w \cdot x + b = +1$, and determine the hyperplane H_2; the

distance between origin and the hyperplane H_2 is equal to $\frac{|+1-b|}{||w||}$. Hyperplanes H, H_1, and H_2 are parallel and no training patterns are located between hyperplanes H_1 and H_2. Based on the above considerations, the distance between hyperplanes (referred as the margin) H_1 and H_2 is $\frac{2}{||w||}$.

From these considerations, the identification of the optimum separation hyperplane is performed by maximizing $\frac{2}{||w||}$. The problem of finding the optimum separation hyperplane is represented by the identification of (w,b) which satisfies:

$$\begin{cases} w \cdot x_i + b \geq +1 & if \quad y_i = +1 \\ w \cdot x_i + b \leq -1 & if \quad y_i = -1 \end{cases} \tag{2.25}$$

for which ||w|| is minimum [29].

The hyperplane found by SVM depends on the kernel function used to compute it. The most basic procedure is the use of linear kernels but the SVM models can also perform non-linear classification by using a different kernel (see Figure 13). Some of the most common non-linear kernels include; polynomial, Gaussian radial basis function and hyperbolic tangent.

Decision Trees

Decision tree learning uses a decision tree model as a predictive model. The structure of a decision tree is composed of leaves that represent class labels and branches that represent features that lead to the class labels. One of the most popular algorithms to create decision trees for classification is the ID3 algorithm [30]. The algorithm calculates the entropy of each feature in the training set and selects the one with the smallest entropy and creates a node that splits the data into more subsets. This algorithm keeps

iterating through the data selecting the next lowest entropy feature and creating nodes that will end in a leaf of one of the two classes when all of the features under that branch fall into a class.

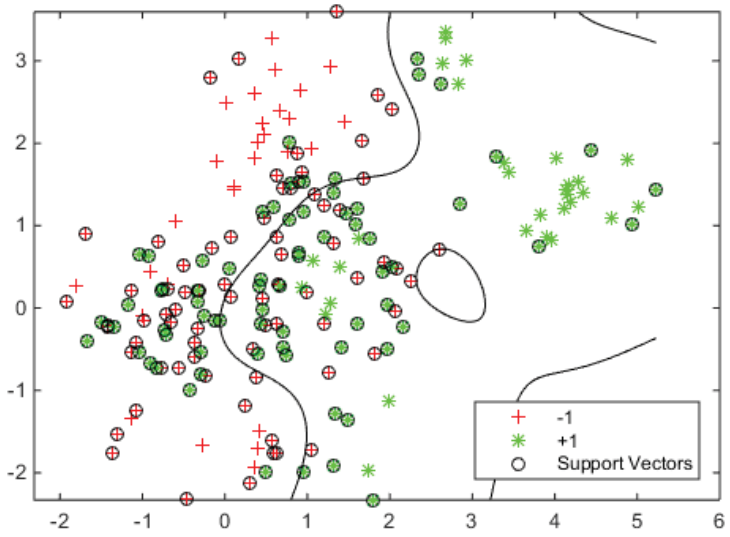


FIGURE 13. Data separated by a non-linear Gaussian radial basis kernel.

Neural Networks

The Artificial Neural Networks, typically known as ANNs, were first described by Rosembaltt in 1949 [31], as an algorithm for pattern recognition. ANNs are computational models of learning and automatic processing inspired by the animal's nervous system. It is an interconnected system which collaborates to produce an output.

A neural network is composed of several parallel units named neurons. Each of them receives a series of inputs throughout other interconnected neurons, and sends an output.

The output is determined by different parameters, such as the summation of the inputs multiplied by their interconnection weight, the activation function that converts a neuron's weighted input into its output activation, and the function chosen by the author, which applies to the output of the activation function.

In most cases, the first layer of the system has input neurons, and they send the corresponding data to the neurons of the second layer, and then those to the third layer of output neurons [32]. Each connection between neurons has its own parameters; hence they can be manipulated separately regarding the interconnection patterns, the weights of the interconnections and the activation function.

The objective of ANNs is to obtain an automatic learning process of the desired properties with an input significant enough, taking into account that the selection of the right network model, the correct variables and the preprocessing of the information are the most important parameters.

CHAPTER 3

SYSTEM FOR AUTOMATIC MELANOMA DETECTION

In this chapter, the techniques and algorithms implemented for the creation of the automatic melanoma detection system are detailed. Following a similar structure as in the previous chapter, the stages of the system include; image acquisition, illumination correction, lesion segmentation, feature extraction, feature selection and classification. In addition to the definition of specific methodology and development procedures, a discussion over the results obtained from the set of images will be provided.

All techniques and methods have been fully implemented in MATLAB.

Illumination Correction

In the illumination correction stage, two algorithms were selected for implementation; based in morphological operators [14] and in the Retinex [12] algorithm respectively. The set of images typically present linear uneven illumination from one side of the image to the other. The aim of these algorithms is to remove the shadow caused by the uneven illumination in order to improve the segmentation rate.

Morphological Operators

Using a morphological operator for illumination correction yields a computationally fast algorithm as the main operations are Gaussian filtering and a morphological erosion and dilation which don't require heavy processing calculations [13].

The type of image to be used and the way the data is stored is a very relevant factor in image processing. MATLAB supports three main types of images; indexed images, intensity images and RGB (Truecolor) images. Indexed images are composed of a data matrix *X* and a colormap matrix *map*. The colormap is an *m*-by-3 matrix that stores color values in RGB format; each row contains the red green and blue values respectively. The number of rows *m* of the colormap is variable depending on the number of colors used to represent it. The *X* matrix contains indices that point to the row number in *map*, each position in *X* belonging to a pixel in the image (see Figure 14).

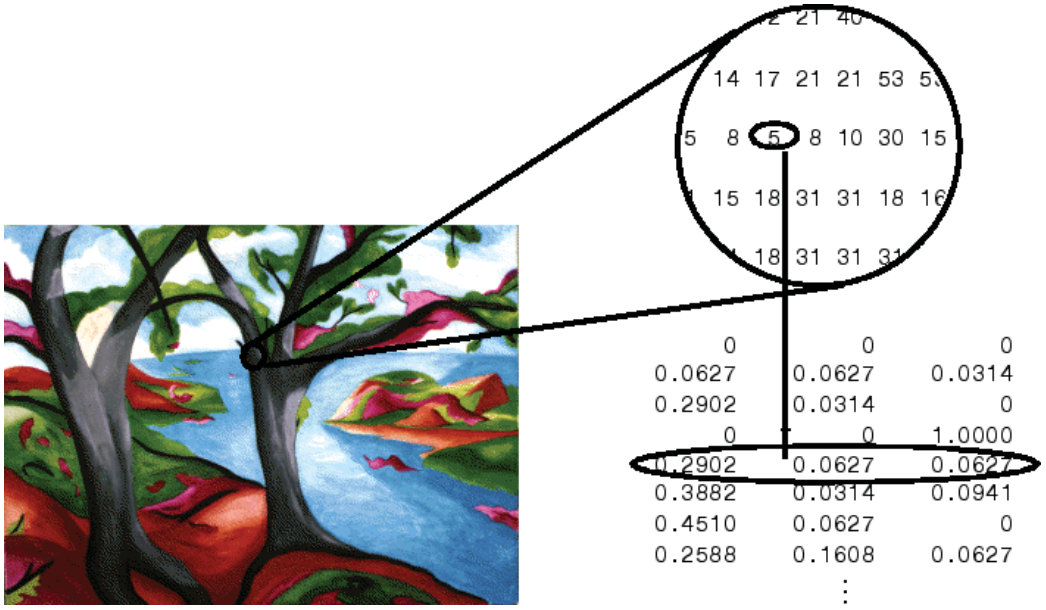


FIGURE 14. Example of indexed image where each pixel points to a position in the colormap.

Intensity images are composed of a single matrix with one value per pixel that represents an intensity value range. The most common way to represent these images is

by using the grayscale for the colormap as Figure 15 shows; hence, they are commonly referred as grayscale images. However, other colormaps can be used to represent the value range.

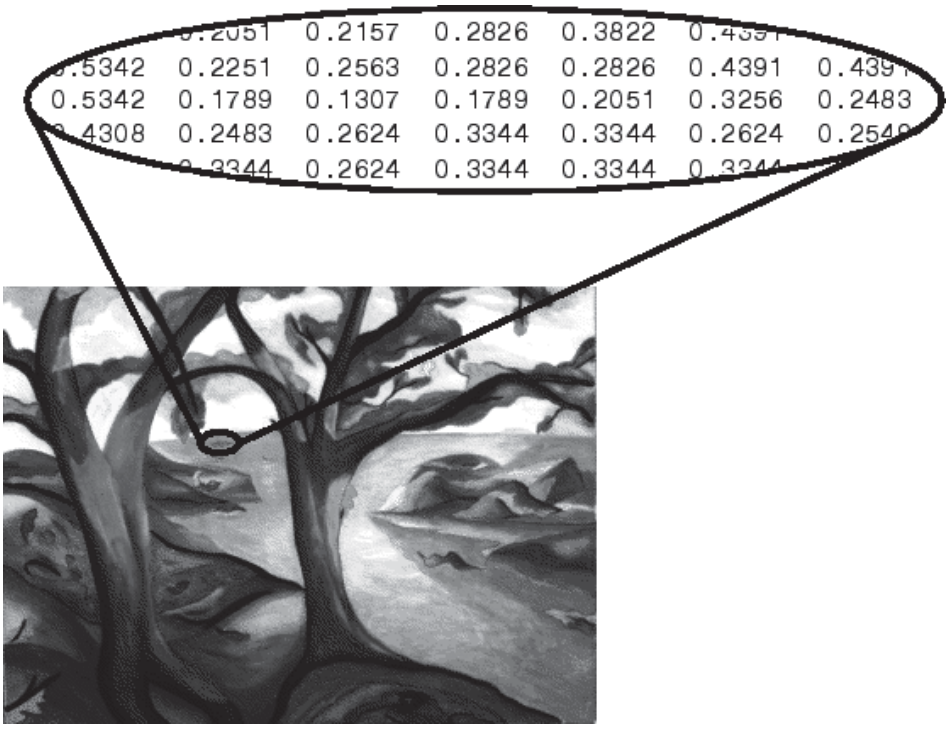


FIGURE 15. Intensity image where every pixel refers to a single value in a scale.

RGB images are *m*-by-*n*-by-3 arrays that contain three matrices, one matrix per color channel. As Figure 16 depicts, each matrix element defines the color component in a given pixel regarding red, green and blue values.

When an image is loaded by MATLAB, it will be computed as one of the three image types described above. However, there are more color space models that can be used to represent an image's information.

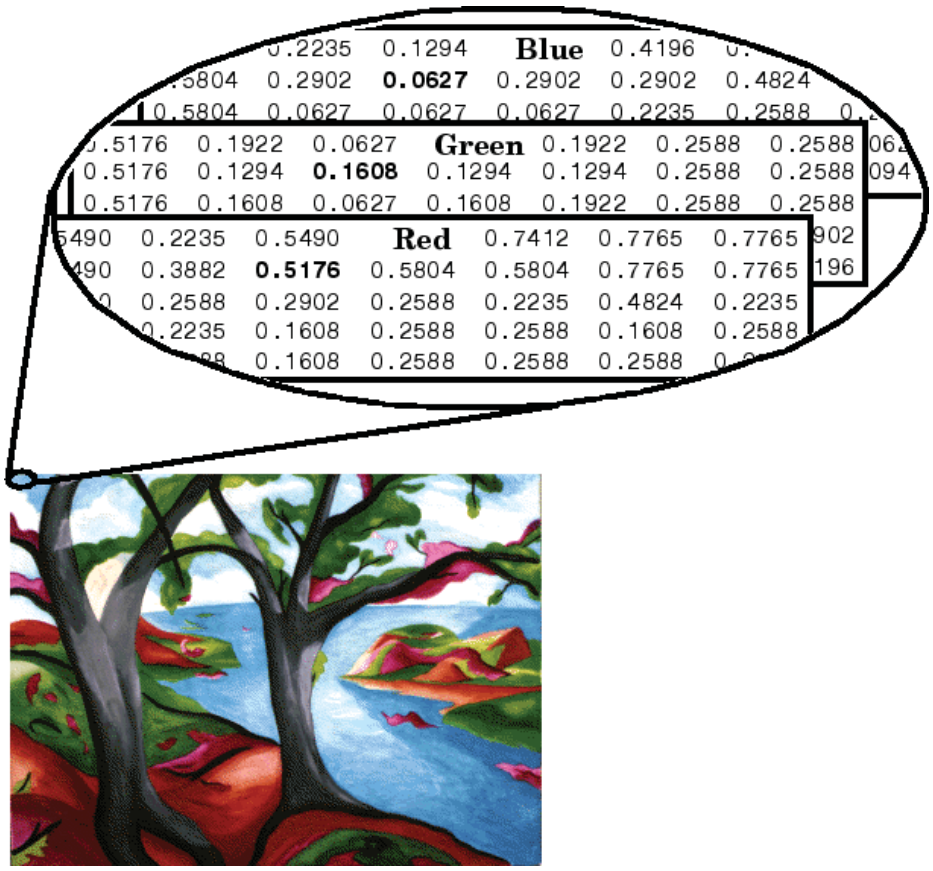


FIGURE 16. RGB image where each pixel is defined by three values stored in three matrices.

Other popular models include HSV which represents the hue saturation and value of an image; *Lab*, which represents lightness and two color-opponent dimensions; or CMYK, a subtractive color space used in color printing.

The illumination correction algorithm using morphological operators work with intensity images. Converting an RGB or indexed image to intensity, results in the color information to be lost. Consequently, a different color space must be used that allows altering the intensity values without removing the color information. Various popular

color spaces have characteristics that allow this, including HSV, *Lab* and YCbCr [13]. In this thesis, the color space HSV has been selected over other color spaces since the *V* channel provides a representation of the intensity values and the simplicity of converting an RGB image into HSV. Thus, the algorithm works with the *V* channel of a converted RGB image.

A morphological closing is performed in order to extract the background illumination [13]. The structuring element must be big enough to extract only the information related to the uneven illumination and to not alter the lesion area. A small structuring element extracts the lesion as part of the background and that leads to lesion altering. Since the size of the images is variable, the structuring element size is calculated to be proportional to the image's size. The selected shape for the structuring element is a disk as it is the shape that extracts a background closer to the shape of the illumination.

In addition, the algorithm tries to estimate if the selected structuring element's size is smaller than the lesion by computing a thresholded image and obtaining the biggest connected component. If the lesion is bigger, the structuring element size is increased to avoid altering the lesion. An example of an extracted background by using this method is shown in Figure 17.

The resulting background image is smoothened by a Gaussian filter of size three times the structuring element used to perform the closing. The standard deviation of the filter is seven times smaller than its size.

FIGURE 17. Original melanoma image (left) and extracted background (right).

Finally, the resulting image is subtracted by the original image in order to obtain an illumination corrected image. This operation is also known as black top-hat transform [13], defined mathematically as:

$$T_b(f) = f \bullet b - f \tag{3.1}$$

where $\bullet$ is the closing operation, b is the structuring element and f is the intensity image.

Figure 18 shows some examples of dermatologic images processed with the morphological operator. As it can be observed, color alterations are present in some of the lesion areas.

Retinex Algorithm

Regarding the Retinex algorithm, the implementation called McCann99 Retinex developed by Frankle and McCann is used [33]. Similar to the morphological operation method, the Retinex algorithm works with intensity images. Therefore, the need to convert the RGB images to HSV color space is required in order to extract the *V* channel.

FIGURE 18. Original melanoma images (top row) and corrected results by morphological operator (bottom row).

Instead of using path following as introduced in Chapter 2, this implementation creates a multi-resolution pyramid by averaging the image's values [33]. The algorithm starts from the lowest resolution image created at the top of the pyramid and computes the lightness. After this first estimation, the algorithm takes the next higher resolution image and uses the calculated lightness values as initial values for estimation. These values get refined with each resolution level until it reaches the highest resolution image [33].

To compute the number of layers that the algorithm will create, the greatest common divisor of the image's width and length that is a power of two is found, then the binary logarithm is calculated to get the final number of layers.

Each layer's input size is half the image of the next lower layer, being the size of the *n*th layer, the original size divided by 2^n. At each layer, the original image is divided in blocks of size 2^n-by-2^n and all the pixels inside each block are averaged and assigned

to that layer's image. For instance, the penultimate layer which its image is half the original image size, is assigned as its first pixel (1,1) the average value of the 2-by-2 block generated at the (1,1) position of the original image's matrix. The second pixel of the first row in this image (1,2) will be the average of the values covered by the 2-by-2 block at position (1,3) [33].

After all layers have been computed, the level of lightness is calculated for each pixel in each layer. In order to accomplish this, the 8-pìxel neighborhood is obtained in clockwise order. For each neighbor, the log luminance is subtracted which corresponds to the ratio step commented in the previous chapter and then it is added to the old product corresponding to the product step. The new product is obtained by averaging the pixel with the old product. This calculation is repeated a number of iterations, in the case of this thesis the number of iterations are set to 125. The higher the number of iterations makes pixels with higher distance to influence the target pixel [33].

Some of the results of using the Retinex algorithm with the dataset are shown in Figure 19. The lesions are affected by the illumination correction similar to what happened with the morphological operator method. However, even if the background color sometimes changes, the algorithm creates a more constant and less uneven illuminated image. These characteristics make the Retinex algorithm more suitable for applications that use the method for automatic segmentation and using the original image's features rather than a method for correcting images in order to obtain higher quality features.

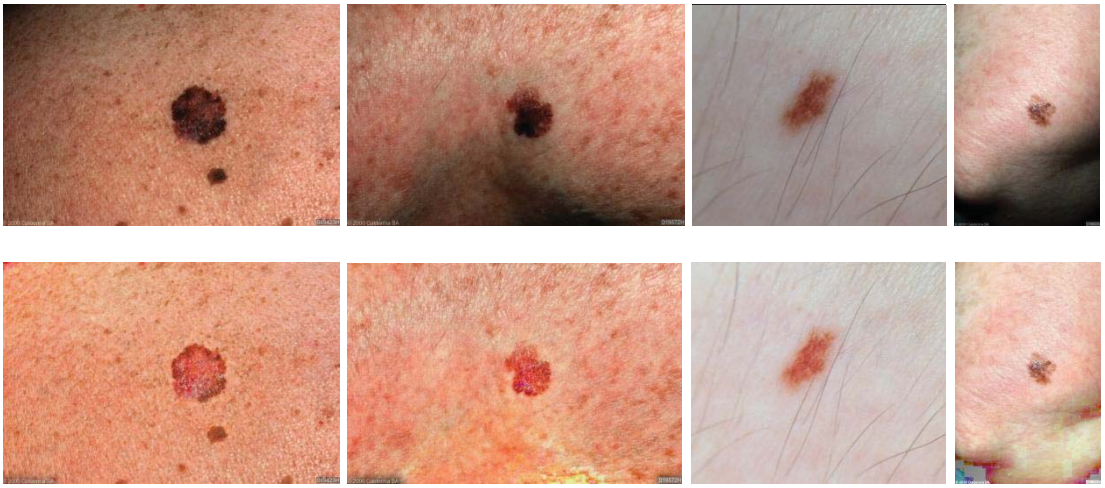


FIGURE 19. Original images (top row) and corrected results with Retinex method (bottom row).

Lesion Segmentation

In the lesion segmentation stage, three methods are proposed. The first method is a widely known global thresholding technique known as Otsu's method [34]. An active contour technique (GVF snakes) is also implemented [21]. Lastly, a novel technique in melanoma segmentation is developed by implementing Mean Shift filtering for image clustering [18].

One of the focuses of this thesis is the study of this novel technique in comparison with two well-known methods that have been proven to obtain good results in lesion segmentation.

Otsu's Method

In image thresholding, finding an optimal threshold is often a difficult task due to the image's characteristics. Otsu's method, developed by Nobuyuki Otsu [34], tries to

find the most optimal threshold by predicting which parts of the image belong to the foreground and which to the background.

The base of this algorithm lies in calculating the variances σ_0^2 and σ_1^2 of each of the partitions $K_0(t)$ and $K_1(t)$ separated by the threshold *t* in the image histogram. The point where the variances of each section are minimal is the optimal threshold. Denoting the arithmetic mean of the values in each section as $\bar{g}_0$ and $\bar{g}_1$, the variances are then calculated as:

$$\sigma_0^2(t) = \sum_{g=0}^{t}(g - \bar{g}_0)^2 p(g) \tag{3.2}$$

$$\sigma_1^2(t) = \sum_{g=t+1}^{G}(g - \bar{g}_1)^2 p(g) \tag{3.3}$$

The aim is to maintain the lowest possible variance inside each section while finding the highest variance between sections. The variance between segments is found by calculating:

$$\sigma_{zw}^2(t) = P_0(t) \cdot (\bar{g}_0 - g)^2 + P_1(t) \cdot (\bar{g}_1 - g)^2 \tag{3.4}$$

where P_0 and P_1 are the accumulated probabilities of each value that belongs to its corresponding section.

The variance inside the sections is the sum of both weighted by their respective probabilities:

$$\sigma_{in}^2(t) = P_0(t) \cdot \sigma_0^2(t) + P_1(t) \cdot \sigma_1^2(t) \tag{3.5}$$

The optimal threshold is, thus, obtained from the highest ratio obtained by dividing the variance between sections by the variance inside the sections:

$$Q(t) = \frac{\sigma_{zw}^2(t)}{\sigma_{in}^2(t)} \tag{3.6}$$

Otsu's method starts by selecting an initial threshold at the first element of the histogram and computing the ratio with the formulas described above. This step is repeated for all the elements of the histogram and the highest ratio is selected as the point where the threshold is optimal [34].

The first step in the lesion segmentation of a melanoma image is using the Otsu's method to obtain a binary image where the pixels set to 0 are the ones relevant to the lesion itself, and the ones set to 1 are the pixels that belong to the background (see Figure 20).

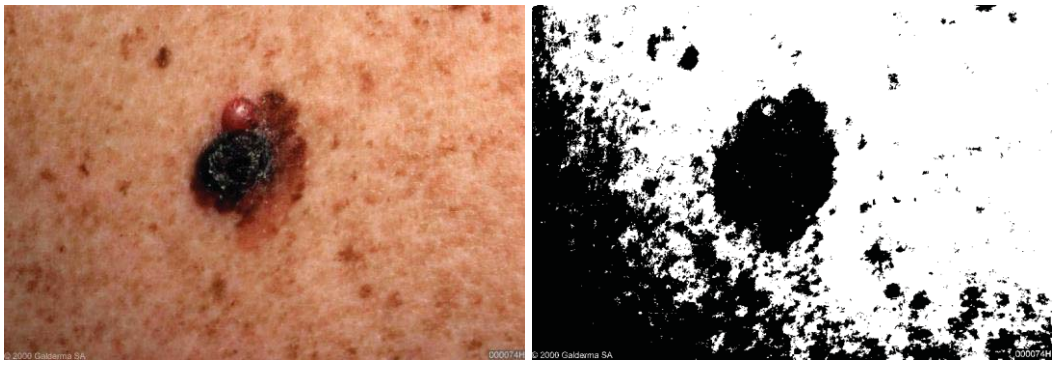


FIGURE 20. Original image (left) and thresholded result with Otsu's method (right).

After thresholding, a closing operation is performed in order to remove potential noise like hair or small dark spots. The closing operator uses a disk shaped structuring element that changes its size dynamically, depending on the input image size similar to the one described in the previous section. After the closing operation, a morphological opening is performed to fill the holes and join potentially separated regions of the lesion.

The regions that are connected to the borders are treated as noise and deleted, since the lesion area is assumed to be centered (see Figure 21).

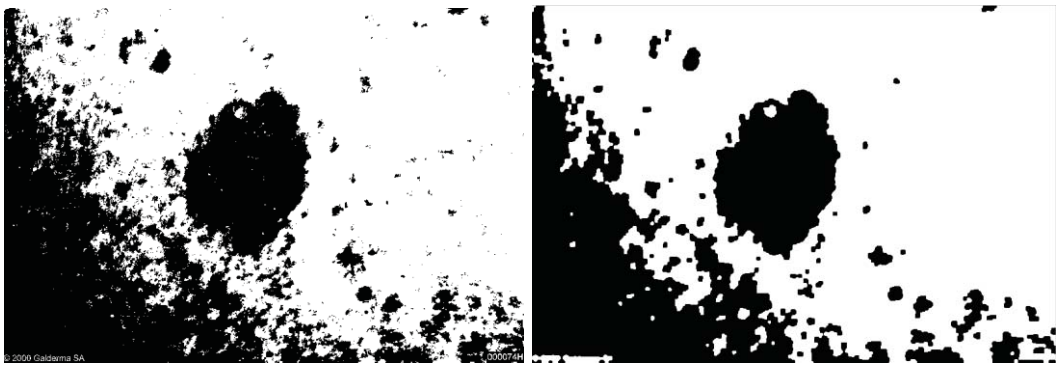


FIGURE 21. Thresholded image (left) and result (right) by performing preprocessing actions that remove noise and separate the lesion area from other thresholded areas.

The last step involves selecting the biggest connected component of the image that will be labeled as the lesion area. The boundary will then be calculated based in the mentioned area. Figure 22 shows some examples of segmentation using this method.

As the results show, this method is very susceptible to uneven illumination problems, obtaining irregular results when the illumination of an image is highly unbalanced.

Gradient Vector Flow

The method used for segmentation based in GVF is called GVF snakes, introduced in the previous chapter. This technique uses part of the implementation developed for MATLAB by Xu and Prince [35].

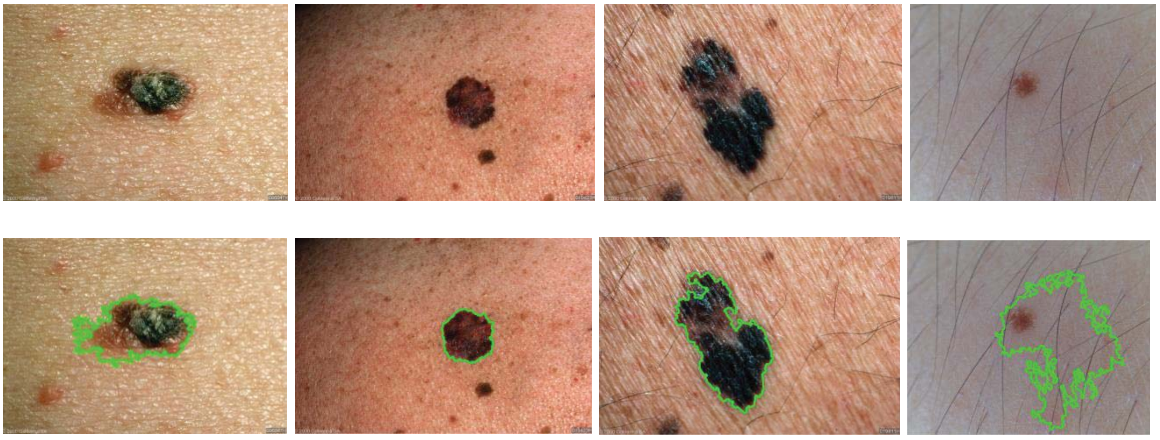


FIGURE 22. Original images (top row) and segmented results (bottom row) using Otsu's method.

GVF snakes need an initial boundary to start the calculations of each deformation. Commonly, the boundary is given manually by setting a number of points or drawing the whole line. Since the aim of this thesis is the design and development of an automatic detection system, the initial points are automatically calculated based on an estimation of the lesion shape. This estimation is done by calculating the binary mask of the lesion using the Otsu's Method described in the previous section. Once the lesion mask is calculated, eight equally angled radiuses are drawn from the centroid of the connected component until the boundary is intersected [36]. These eight intersection points are then given a five pixel margin in order to avoid the boundary itself. The initial eight points for the snake are, thus, defined this way. Other methods like obtaining the points from the bounding box or giving a margin proportional to the image size were discarded, since this technique doesn't perform as expected when the distance between the real border of the lesion and the initial snake is too great. This often is the case in large images where the

initial snake is set too far away and the gradient force field is not strong enough to pull the line towards the edge.

After the points have been set, the snake is constructed by interpolating them and creating a pseudo-ellipse of 126 values. The edge map of the image is obtained by performing a Canny edge detection with a threshold of 0.3 and sigma of 15, in order to get the most relevant borders and avoid noise (see Figure 23). With the obtained edge map, the GVF field image can be calculated by using Xu and Prince's implementation [35]. With all the parameters calculated, the last step involves iterating the deformation calculations that will pull the initial snake towards the edges. The iterations were set to 125 [36]. Figure 24 shows some of the results obtained by using GVF snakes in lesion segmentation.

In most unfavorable cases, the technique fails due to incorrect initial point setting (subject to Otsu's method performance) or inability to deform the line given the distance between the initial snake and the real edge.

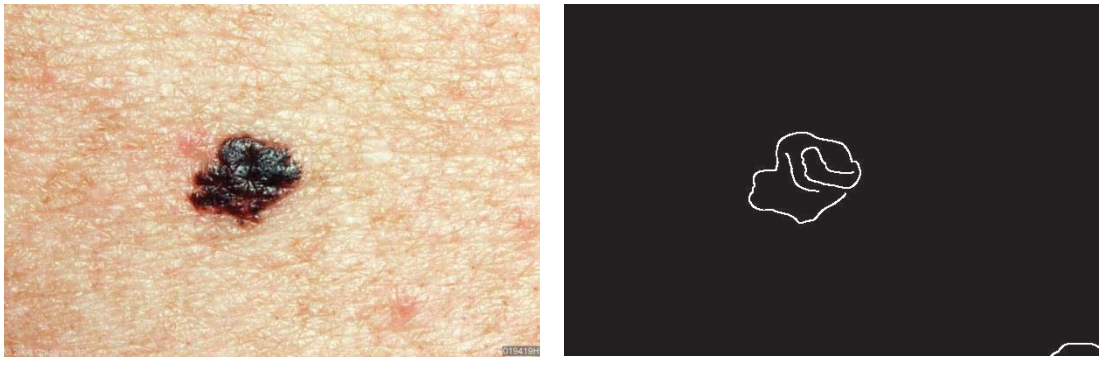


FIGURE 23. Original image (left) and calculated edge (right) using Canny method.

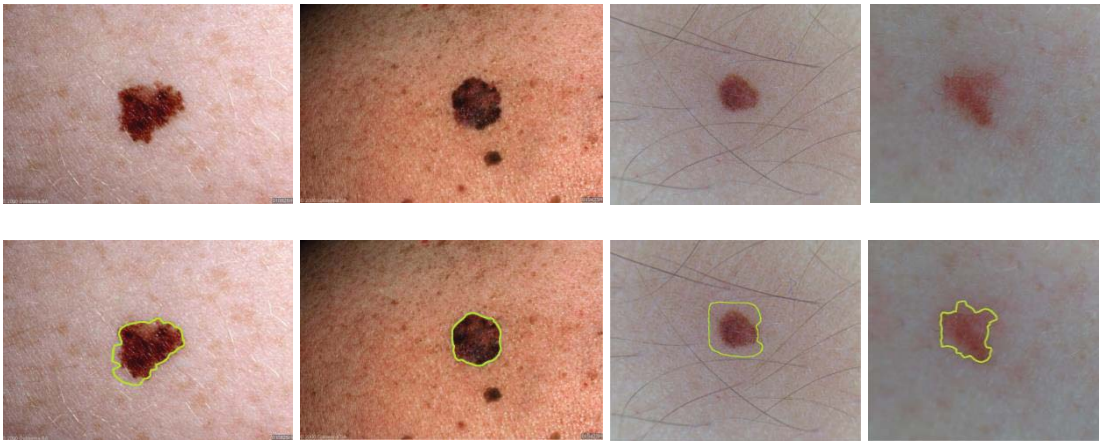


FIGURE 24. Original images (top row) and segmented results (bottom row) using GVF snakes.

Mean Shift Filtering

Mean Shift is a non-parametric iterative algorithm that can be used for various purposes such as mode finding, clustering or object tracking [18]. This algorithm was first introduced by Fukunaga and Hostetler [18] and it gained more attention when Cheng applied it in the field of computer vision [37]. One of the main differences of this technique over other clustering methods such as *k*-means is that the number of clusters is not known and tries to find an optimal number by finding convergence in its iterations.

Mean Shift uses a kernel function $K(x_i - x)$ to calculate the mean $m(x)$ inside the neighborhood defined by the kernel at an initial point *x* randomly estimated. Mean Shift treats the points of a feature space as a probability density function (*pdf*) and calculates the mean to locate the maxima of the *pdf* by using:

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x) \cdot x_i}{\sum_{x_i \in N(x)} K(x_i - x)} \tag{3.7}$$

Dense regions in the feature space correspond to local maxima. This new calculated point is set as the new x and the mean is calculated again until there is convergence. Another parameter that is relevant to the calculation is the bandwidth. The bandwidth defines the size of the selected neighborhood and thus, influences the mean in each calculation [38].

In image clustering, the feature space can be composed of only color values or both color values and texture. In the first case the feature space will be one dimensional and in the second case, two dimensional. The kernel windows are set across the feature space consecutively without overlapping. The mean will be calculated for each window and they will be moved accordingly until convergence. Figure 25 depicts the process of iteratively calculating means and moving the windows until convergence ending with three clusters. All the pixel values that a specific window passed will belong to the cluster the window ended at. It is easy to see that the size of the window will affect the number of clusters. A big window size will make the means converge at the same point creating very few clusters whereas a small window size will create a high number of clusters due to the different convergence points found.

In this thesis, the implemented Mean Shift clustering technique uses two modes. The first only takes color values as the feature space and the second takes both color values and texture information. The aim is to compare both methods on the next chapter to calculate their performance under the same dataset and determine which algorithm gives better results.

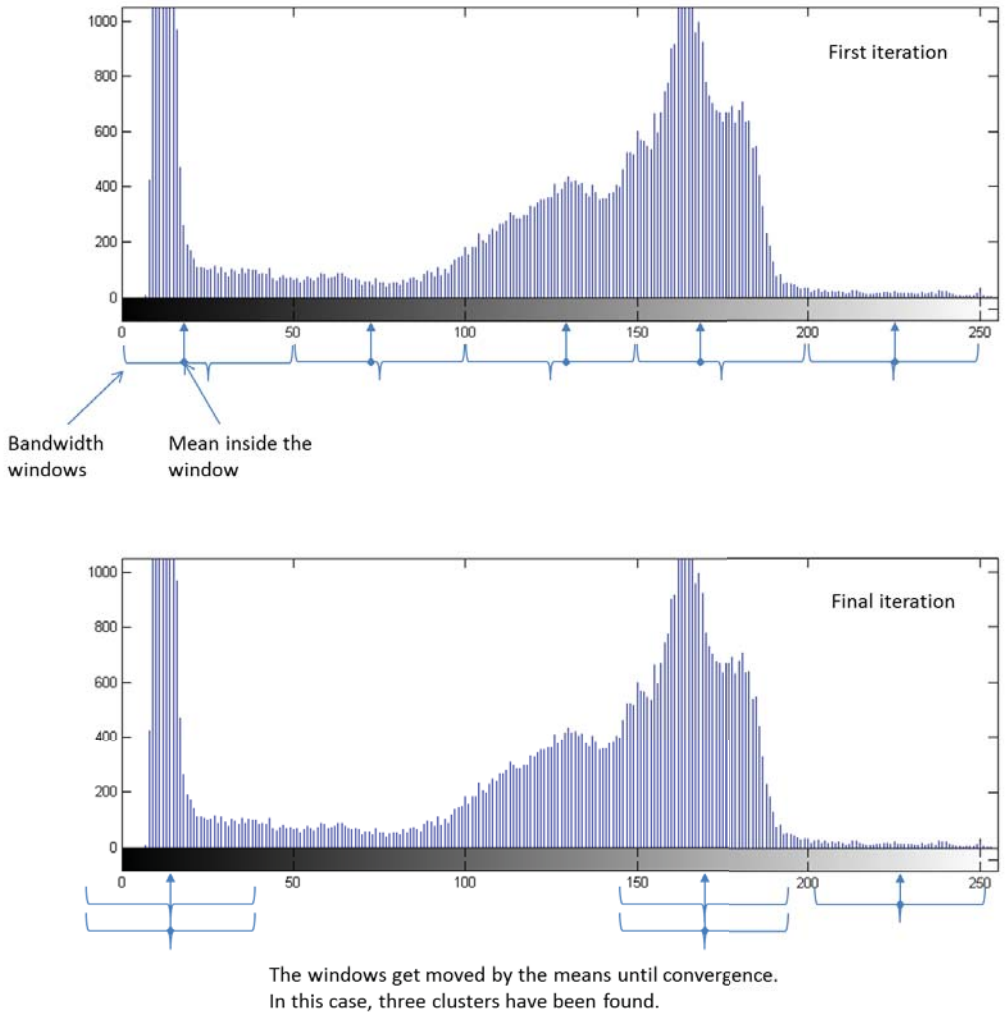


FIGURE 25. Mean Shift convergence calculation for a feature space based on intensity values.

Given the importance of the bandwidth, the algorithm developed uses an adaptive bandwidth that starts at a fixed value and if the number of clusters is below 2 or above 10, the bandwidth gets reduced or increased respectively. By doing this, some problems of not having enough clusters to define the lesion or too many clusters that the lesion is labeled with multiple clusters can be avoided. The lesion area is extracted as the biggest

cluster that is not connected to any of the image's borders. Figure 26 shows some results of clustered melanoma images and their consequent segmentation.

While this technique gives good results at segmenting melanoma images, it is the slowest of the three segmentation implementations. The Mean Shift clustering based on color values and texture is especially slow compared to the rest of techniques being up to five times slower than the one that only uses color values as feature space.

a) b) c) d)

FIGURE 26. Segmentation of a) original image by performing b) Mean Shift with color information or c) Mean Shift with color and texture information to obtain lesion d).

Feature Extraction

In this thesis, the feature extraction step calculates a number of geometric features mostly based on asymmetry and border characteristics taken from the ABCD rule [22]. For features based on texture and color, the former is obtained with a tree-structured wavelet transform [25] and the latter is implied in this calculation since the

transformation is not only calculated for the intensity values but for all the RGB channels and luminance values.

Geometric Features

The features related to geometry are divided into border characteristics and asymmetry index. As previously introduced, melanomas tend to present highly irregular borders and very asymmetric shapes.

Border Irregularity. There are several ways to measure the degree of irregularity that a given boundary presents. In this work, the methods selected are the circularity ratio and the irregularity index [25]. Both calculations estimate the irregularity as a function that takes the area A and perimeter P of the lesion, which has been extracted in the segmentation step.

The circularity ratio is defined as:

$$CRC = \frac{4 \cdot A \cdot \pi}{P^2} \tag{3.8}$$

The irregularity index is defined as:

$$IrA = \frac{P}{A} \tag{3.9}$$

Asymmetry Index. The asymmetry of a lesion is defined in the ABCD rule as the asymmetry in two axes; horizontal and vertical. The proposed implementation to calculate asymmetry takes the logical mask of the lesion (where the lesion area pixels are set to 1 and the background pixels are set to 0) and calculates its orientation. The lesion is then rotated and moved in order to make the bounding box's longer sides parallel with the x axis and to set the axes intersection at the center of the image. For the horizontal asymmetry, the image is divided in two halves by a horizontal line having one of the

halves flipped up to down. An XOR calculation is performed in order to see the pixels that do not match when the image gets 'folded' (see Figure 27). The same procedure is repeated for the vertical asymmetry by dividing the image vertically. Finally, the asymmetry ratio is obtained by dividing the number of non-matching pixels by the total number of lesion pixels [25].

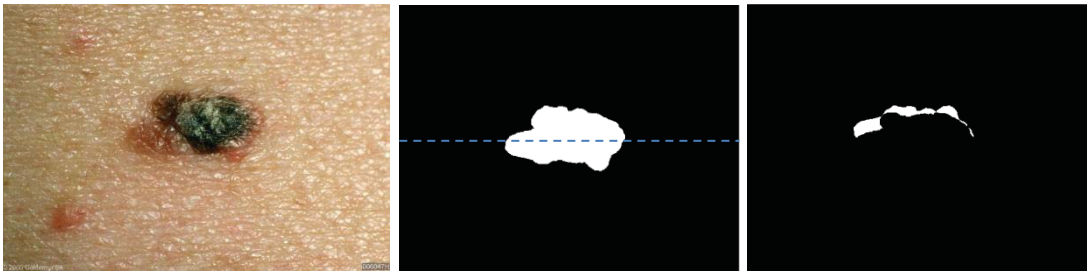


FIGURE 27. Horizontal asymmetry performed by 'folding' the mask and obtaining the non-matching pixels.

Features Based on Texture and Color

For texture and color features, a tree-structured wavelet transform is performed on the intensity image (grayscale), RGB channels and luminance image. The luminance image L can be calculated as a function of the RGB channels:

$$L = R \cdot 0.29 + G \cdot 0.59 + B \cdot 0.11 \tag{3.10}$$

The number of levels selected for the tree-structured wavelet is 4 and the type of filter is a Daubechies filter with one vanishing moment [25]. The image must be preprocessed in order to be able to calculate the wavelet transform. The image must be square and the sides must be divisible by 2^n where n is the number of wavelet levels, in this case, 2^4. The image is, thus, cropped by a square whose side length is the closest

number divisible by 2^4 from the smallest side of the original image. Once the wavelet transform is calculated, the three coefficients from each of the four levels are extracted (see figure 28). A number of measurements are performed for each extracted image that will become part of the feature vector. These measurements are; energy (*E*), standard deviation (*Std*), mean (*M*), entropy (*H*), skewness (*S*), kurtosis (*K*) and norm (*N*) [25].

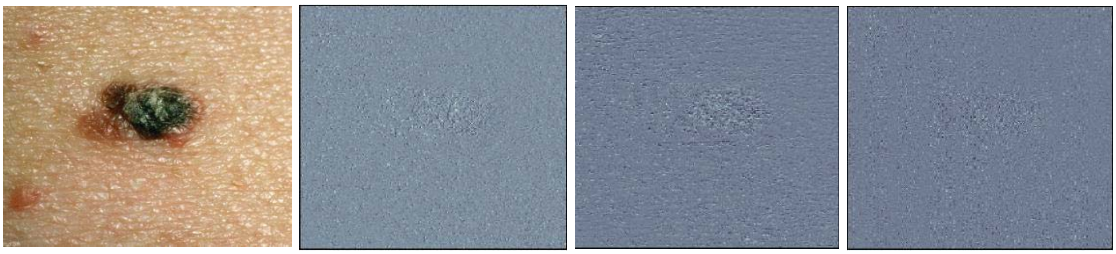


FIGURE 28. First level coefficients after calculating the wavelet transform.

$$E = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} x_{jk}^2}{J \times K} \tag{3.11}$$

$$M = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} x_{jk}}{J \times K} \tag{3.12}$$

$$Std = \sqrt{\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} (x_{jk} - M)^2}{J \times K}} \tag{3.13}$$

$$H = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} (x_{jk}^2 \times \log x_{jk}^2)}{J \times K} \tag{3.14}$$

$$S = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} \left(\frac{x_{jk} - M}{Std}\right)^3}{J \times K} \tag{3.15}$$

$$K = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} \left(\frac{x_{jk} - M}{Std}\right)^4}{J \times K} \tag{3.16}$$

$$N = \max \left(\sqrt{eig(X \times X')}\right) \tag{3.17}$$

where the image matrix is denoted by X and each element on the j-th and k-th row by x_{jk}. These 7 features are calculated on 4 image types that are further transformed in wavelets obtaining 3 images in a total of 4 levels. Hence, the total number of features obtained in texture and color is 336.

Feature Selection

Feature selection is an important part of the classification process because it reduces feature vector dimensionality, which makes the classification more efficient and sometimes improves the performance since it removes noisy information that doesn't add relevant data.

The feature selection method used in this thesis is called sequential feature selection. This technique creates candidate feature subsets by sequentially adding features not yet selected [39]. For each feature subset, a criterion is calculated by performing a 10-fold cross-validation with different train and test sets. Sequential feature selection sums the criterions and divides that sum by the total number of test observations. It then uses that mean value to evaluate each candidate feature subset.

The image set used for feature selection is not used in the classification process in order to obtain results not affected by reused sets.

Classification

For the last stage of the system, classification, two different methods have been selected, SVMs [29] and discriminant analysis with various types of functions. These

techniques require a training step prior to classification, for which the subsets used are detailed in chapter 4.

Support Vector Machines

SVMs have been proven to give good results at melanoma classification. As previously defined, SVMs take a dataset and tries to calculate the hyperplane that best separates the classes. For this calculation, a model has to be trained with a subset of the data, for information on how the data is partitioned, refer to chapter 4. The kernel used for the model training is a linear kernel since, given the characteristics of the feature vector (relatively low number of features and training observations), it gives better results than a non-linear kernel [40].

Once the SVM model is trained, it is used in the classification process where the system takes the test vectors and makes a decision based on the trained model parameters. The MATLAB implementation of SVMs is used for both training and classifying purposes. The result is a relatively fast classifier due to the use of a linear kernel and small feature vectors.

Discriminant Analysis

Discriminant analysis uses statistical analysis to predict a categorical dependent variable. It is also used in feature selection for dimensionality reduction purposes, since it is able to determine the effectiveness of a feature set.

Similar to SVM, discriminant analysis requires a train and a test set of images. Each observation of the train set is labeled to the class it belongs in order to perform the training. The type of discriminant function to be used in the training stage affects the

performance of the technique for classification. In this thesis, three different discriminant functions are used in order to test and study their performance in dermatologic images.

The first function is the linear function which fits a multivariate normal density to each class, with an estimate of covariance. The second function is quadratic, which fits multivariate normal densities with covariance estimates stratified by class. Lastly, mahalanobis distances are used as a discriminant function along with stratified covariance estimates.

The discriminant analysis implementation is faster than SVMs and it is proportionally faster the larger the feature vector is. However, it usually results in inferior classification rates.

Conclusion

In this chapter, a full automatic system for melanoma detection has been described. The chapter is divided by the systems most important stages; illumination correction, lesion segmentation, feature extraction, feature selection and classification. A detailed implementation of each stage is detailed along with the algorithms proposed. In each stage, the system makes use of different techniques that allows their comparison and study under the same dataset. In addition, a novel technique for melanoma segmentation based on Mean Shift clustering is proposed.

CHAPTER 4

EXPERIMENTS AND RESULTS

In this chapter, the dataset used for all experiments is introduced. The experiments and results are defined for each stage developed in the system. The experiments are performed for the illumination correction techniques, segmentation techniques and classification techniques. The results are studied and a discussion on the performance of the system is included.

Dataset

For this thesis, two types of dermatologic images are used; melanoma images and non-melanoma skin lesions. For the melanoma images the database DermQuest [41] is used. The database offers a variety of melanoma images from different subjects at different resolutions. This thesis uses 100 different images from the database. A sample of the images is shown in Figure 29.

FIGURE 29. Sample of melanoma images from DermQuest database.

The acquisition device for each image is different; the melanomas are mostly captured with digital cameras with flash. In all cases, each image presents a single lesion.

Regarding non-melanoma, the images were acquired specifically for this thesis making use of a macro-capable lens and a ring flash that tries to even the superficial illumination. The digital camera used in combination of the devices previously listed is an 8 Megapixel camera with an aperture of f/2.2. The images are cropped from the original acquisition; hence, the sizes vary but the height-to-width ratio remains constant. A total of 100 non-melanoma images were captured for the completion of the dataset. Figure 30 shows a sample of the non-melanoma images used in the experiments.

FIGURE 30. Sample of non-melanoma images acquired for this thesis.

Image Subsets

The total of 200 images in the dataset is further divided for the classification process. The classification techniques need a training set and a test set.

The train set is used for training the classifiers and learn the features. The test set is composed of the images that are actually being classified as melanoma or non-melanoma. Each image in a subset should not be contained in the other due to results being altered when test images have been previously used for training. The training subset is formed of 50 randomly selected melanoma images and non-melanoma images (25 images from each class). The test set is formed of the remaining 150 images (75 images from each class).

Experiments Protocol

The methodology behind the experiments aims to obtain performance measures and results from each individual system stage, with the aim of perform a general study of the system as a whole.

For the illumination experiments, the whole set of 200 images is used. The dataset will be labeled as 'Original' in the experiments. The illumination corrected datasets will be labeled 'Morpho' for the one obtained with morphological operators and 'Retinex' for the results of the Retinex algorithm.

Regarding segmentation, the study focuses on segmentation performance and the effect illumination correction has over the different measurements, thus, the Original set of 200 images along with the corrected sets Morpho and Retinex from the previous experiments are used. The segmentation results are labeled as 'Original_Otsu', 'Morpho_Otsu' and 'Retinex_Otsu' for the three sets segmented with Otsu's method. Subsequently, 'Original_Snakes', 'Morpho_Snakes' and 'Retinex_Snakes' are used to label the results of GVF snakes segmentation. Lastly, 'Original_MS', 'Morpho_MS' and

‘Retinex_MS’ belong to the segmented images using Mean Shift clustering as well as ‘Original_MSTexture’, ‘Morpho_MSTexture’ and ‘Retinex_MSTexture’ belong to the ones segmented with Mean Shift using both color and texture values for the subspace.

Finally, the classification experiments are performed with a training subset of 50 images and a test subset of 150 images. In order to obtain a measure of performance of each individual classification technique, the segmentation is performed manually so the results are not altered by any of the previous method’s own performance.

Illumination Experiments

Intensity correction measurements.

In order to measure illumination correction in each technique, the selected measurement for the experiment is the intensity variance. The variance is calculated for the original dataset and both the corrected dataset using morphological operators and Retinex algorithm. The variance of a corrected dataset is expected be lower since the image has its uneven illumination removed. Figure 30 represents the data separated by melanoma and non-melanoma sets. The results show that the Retinex algorithm seems to perform better since low variances are present throughout the dataset.

To further describe how well each system performs, a difference of variances between Original and Morpho is computed as well as a difference of variances between Original and Retinex (see Figure 31). A higher positive difference means that the corrected image has a lower variance in comparison with Original, a negative difference means that the corrected image has a higher variance than Original therefore, it performed poorly.

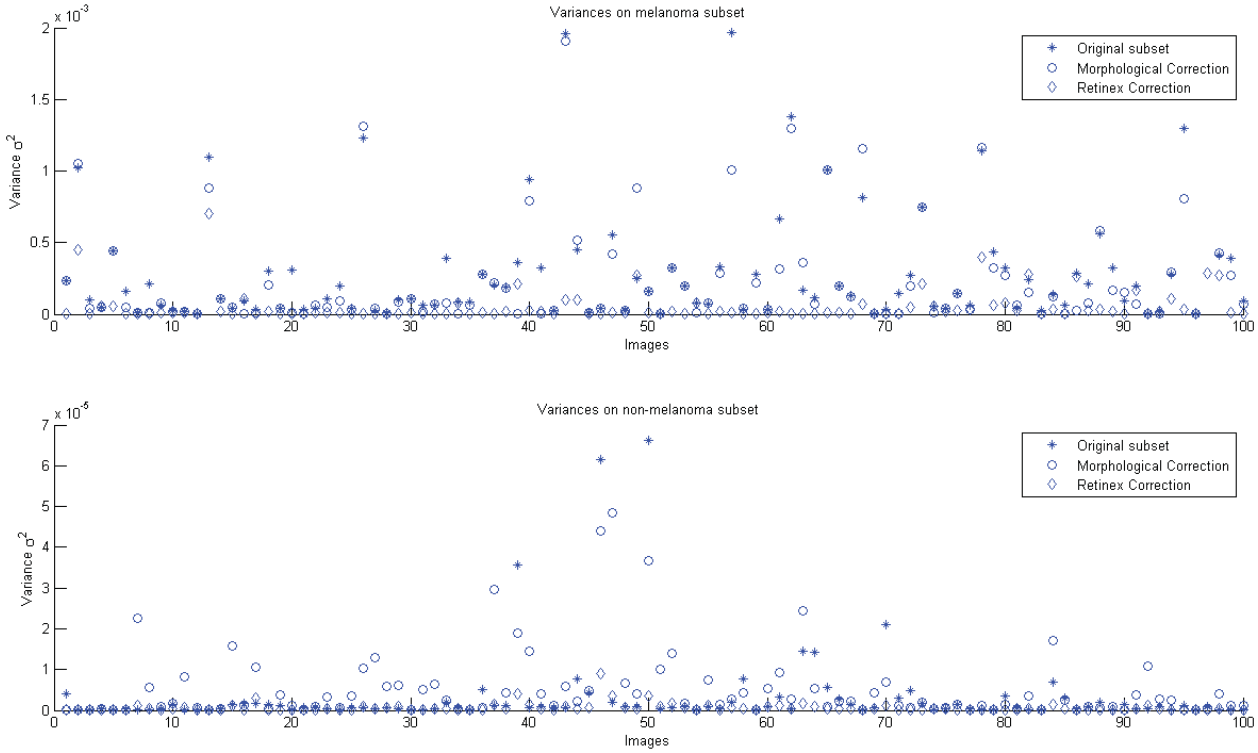


FIGURE 31. Variances of melanoma and non-melanoma sets for the Original, Morpho and Retinex sets.

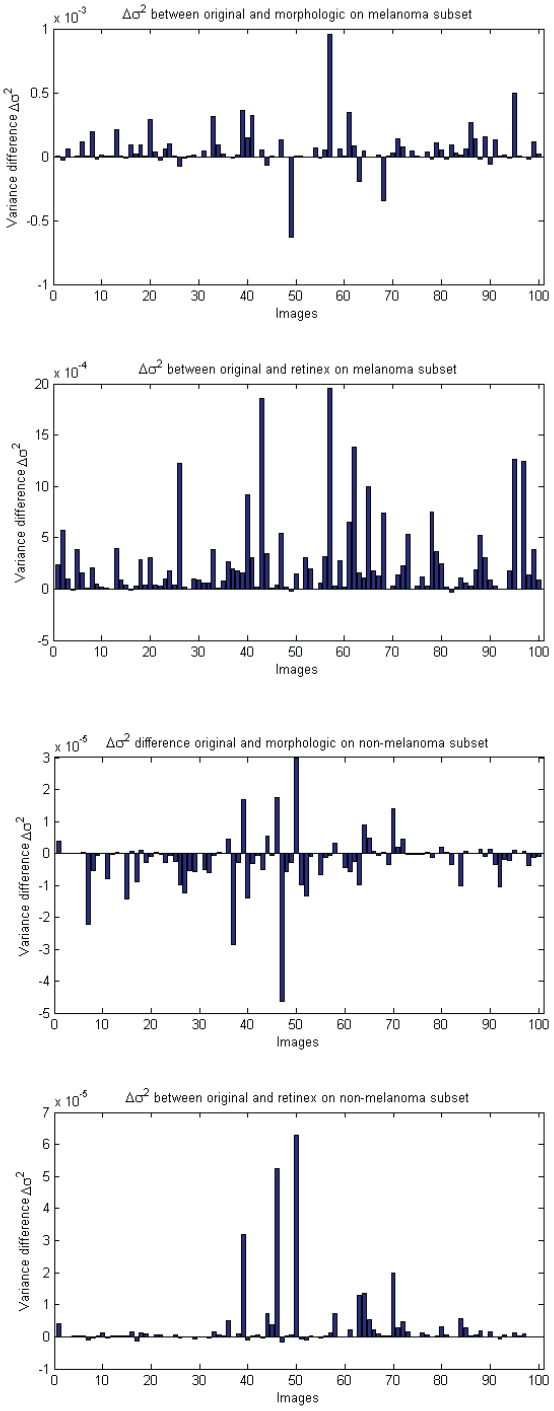


FIGURE 32. Variance difference between Original and the two corrected sets.

The results show that Retinex always perform much better than the morphological operator method given the higher variance differences present in the melanoma and non-melanoma subsets. The morphological operator performs especially worse on the non-melanoma subset creating images with higher variances than the originals.

Given all the results, the variance of the original non-melanoma dataset is lower presumably due to the use of a ring flash to avoid uneven illumination in the acquisition stage. Table 1 shows the mean values of variance for each dataset which corroborates the previous observation.

TABLE 1. Mean Values of Variance in Original, Morpho and Retinex Datasets

Dataset	Original	Morpho	Retinex
Mean Variance (Melanoma)	3.015×10^{-4}	2.527×10^{-4}	5.281×10^{-5}
Mean Variance (Non-melanoma)	3.278×10^{-6}	5.259×10^{-6}	6.036×10^{-7}

Lesion color alteration

Regarding the lesion, the means of each of the three RGB channels are calculated in order to compare the change of color of the corrected images with the Original set. In a similar way to the previous experiment, the means of the three channels of Original, Morpho and Retinex sets are shown in Figures 32 and 33.

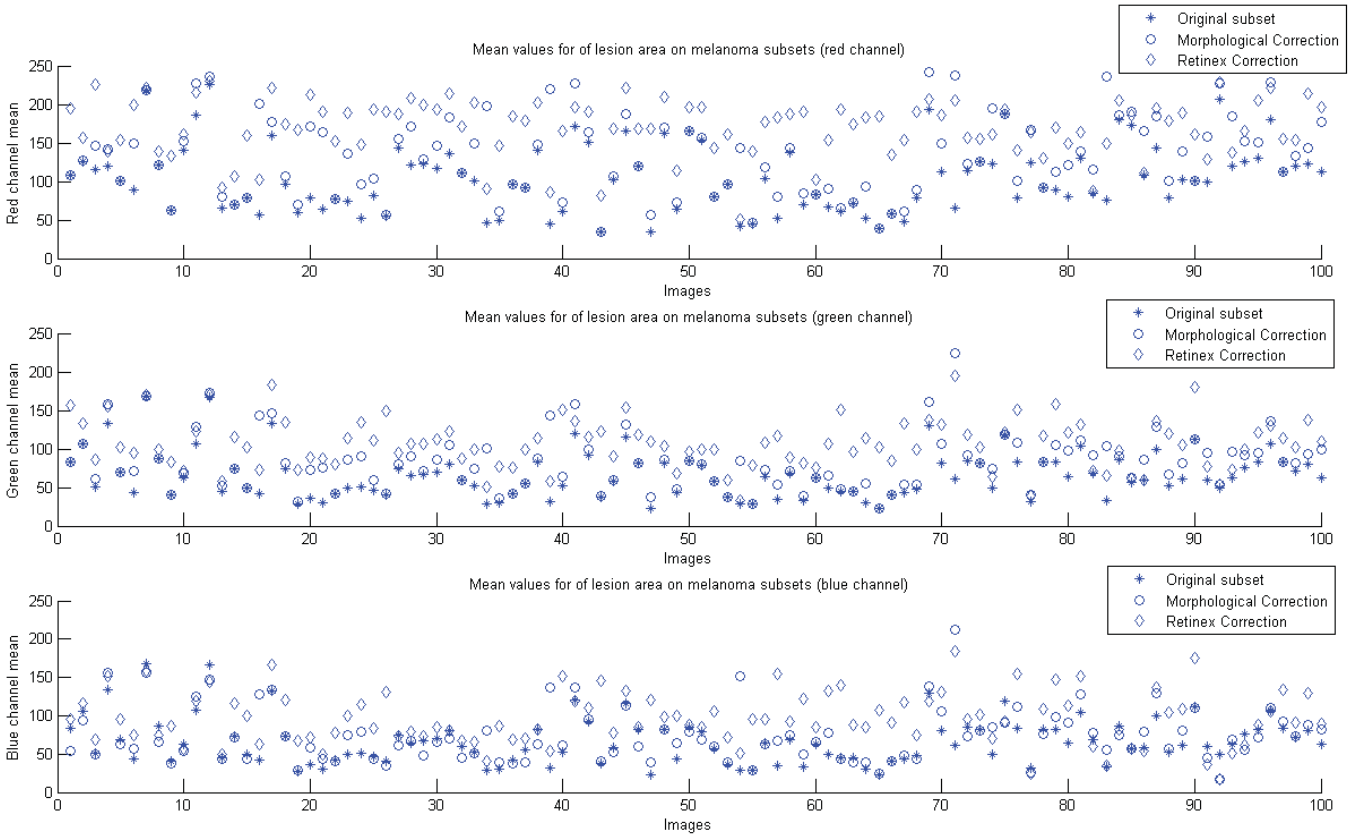


FIGURE 33. Mean values of the red, green and red channels for the Original, Morpho and Retinex sets (melanoma).

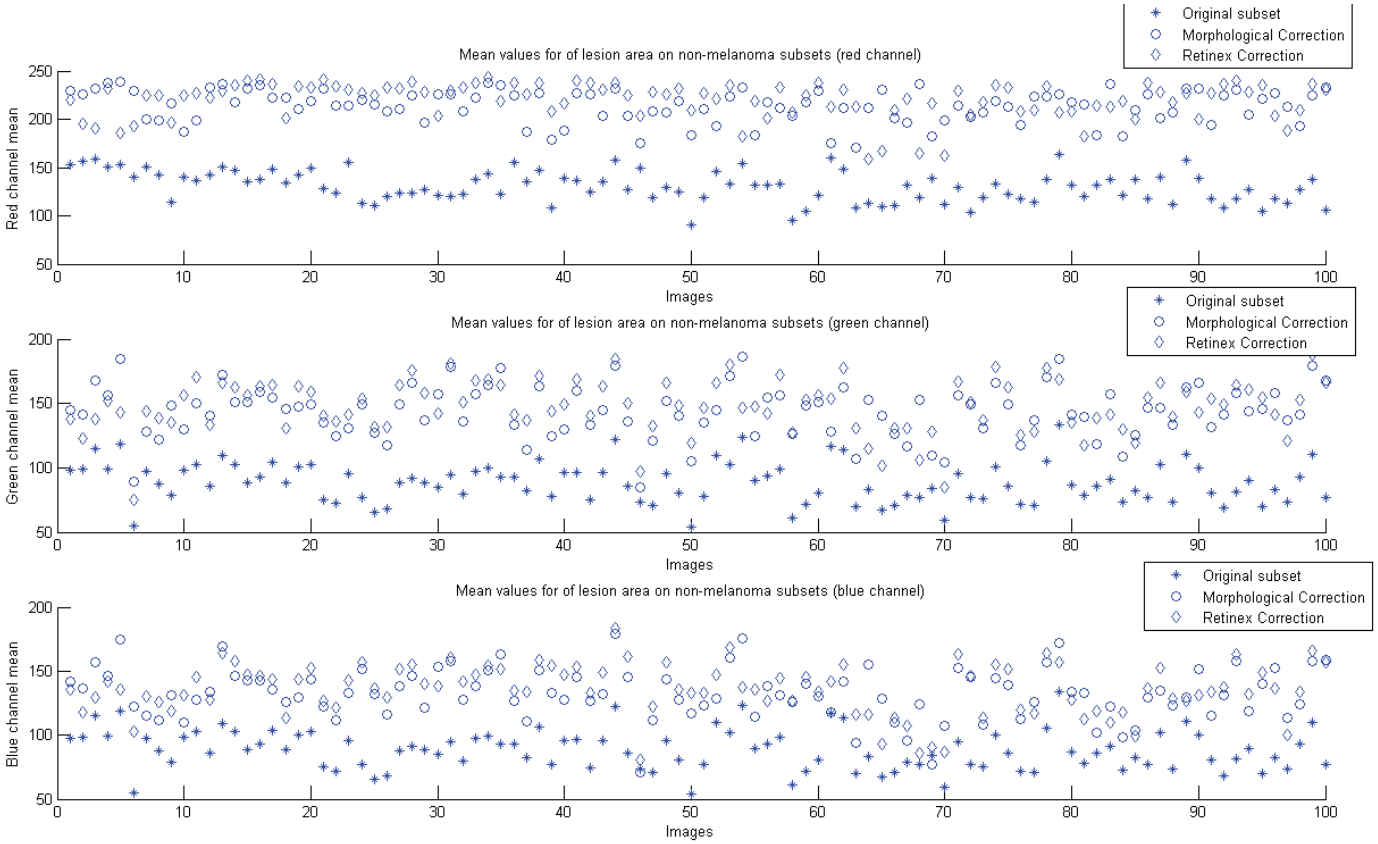


FIGURE 34. Mean values of the red, green and red channels for the Original, Morpho and Retinex sets (non-melanoma).

In this first result, the morphological operator leaves the melanoma lesions mostly unaltered when Retinex modifies the colors of almost all of them. In the non-melanoma set, both of the algorithms alter the lesion. In the case of the morphological operator, this is mainly due to the inability of correctly predicting the lesion size to resize its structuring element. For a more comprehensive view on the results, the differences of the means are calculated. Figure 34 shows the cumulative difference of means between the Original set and each of the corrected sets. In this case, the closest the difference is to 0, the less altered the lesion colors will be.

The results show that, for the melanoma subset, the morphological operator performs better due to a lower cumulative mean difference. However, both techniques seem to highly alter the colors of the lesions in the non-melanoma subset. Table 2 shows the calculated average of these values. It can be thus determined that the morphological operator is superior in all subsets.

TABLE 2. Average of Cumulative Mean Differences

Dataset	Morpho	Retinex
Average mean difference (melanoma)	61.006	138.932
Average mean difference (non-melanoma)	183.333	195.212

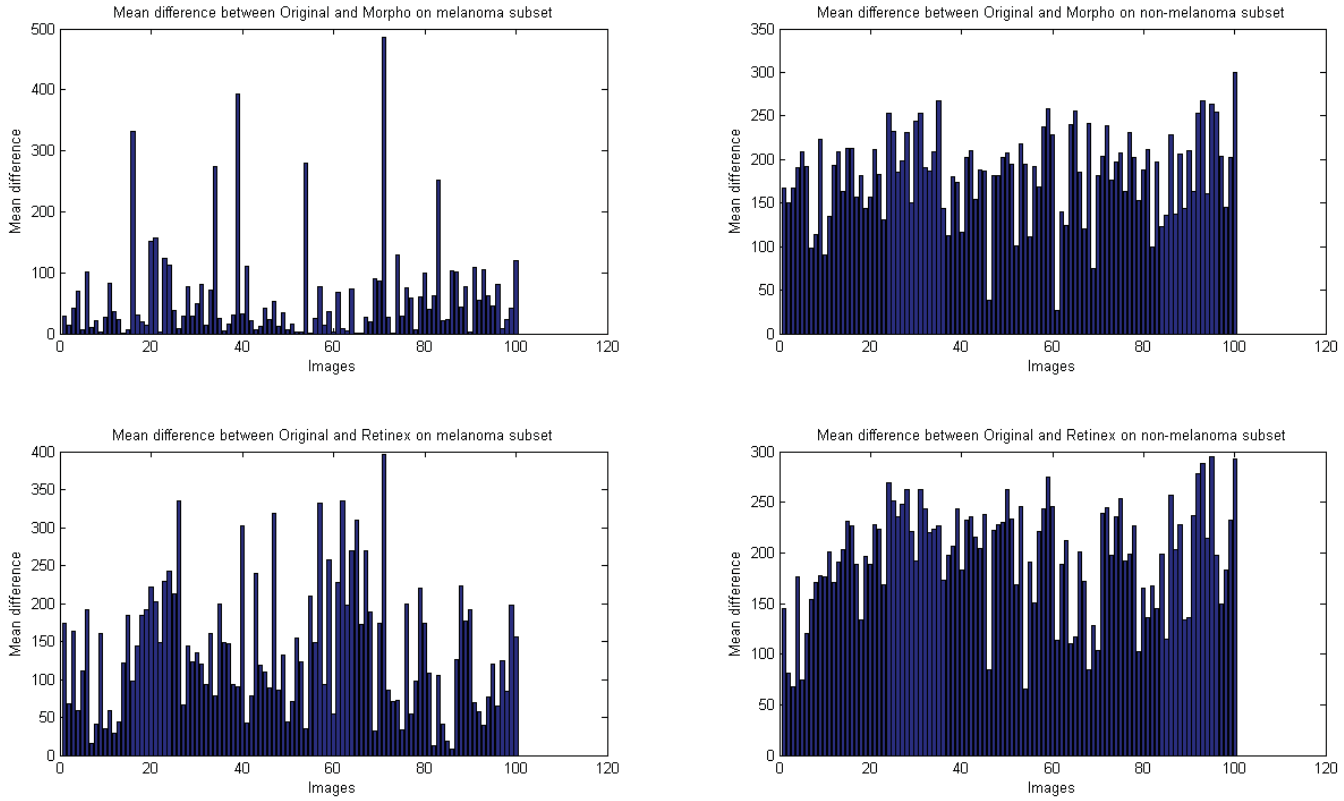


FIGURE 35. Cummulative mean difference between Original and both Morpho and Retinex sets.

Discussion.

The illumination experiments have shown that, while Retinex is superior in removing the uneven illumination, it is more prone to altering the lesion's color. While the segmentation techniques will probably work better with the Retinex algorithm, the resulting corrected images from morphological operation are better suited to use on regular diagnosis procedures.

Segmentation Experiments

As previously described, the segmentation experiments will use the Original, Morpho and Retinex datasets in order to measure the performance of the various segmentation techniques and how the illumination correction affects it.

General Lesion Segmentation Performance.

For this first experiment, the segmentation of all the datasets is performed by the four algorithms developed in this thesis. The results of the segmentation are labeled as 'Correct' when a lesion has been correctly segmented, 'Partial' when a lesion area is covered by the segmentation but not in its entirety and 'Failed' when the segmentation area doesn't contain the lesion at all.

These measurements are performed by empirical research on the segmented images and will be used to create subsets in subsequent experiments. Table 3 depicts the results of each segmentation technique for the Original, Morpho and Retinex sets. The four algorithms are Otsu's method, GVF snakes, Mean Shift clustering with color features and Mean Shift clustering with color and texture features.

TABLE 3. Segmentation Results

		Original	Morpho	Retinex
Otsu's Method	Correct	115	122	159
	Partial	6	20	28
	Failed	79	58	13
GVF snakes	Correct	111	114	152
	Partial	9	27	29
	Failed	80	59	19
Mean Shift Color	Correct	165	167	138
	Partial	11	13	28
	Failed	24	20	34
Mean Shift Texture	Correct	110	125	117
	Partial	23	28	29
	Failed	67	47	54

From the results, various conclusions can be drawn. First of all, the morphologic operator slightly improves the segmentation performance on all the techniques. On the other hand, the Retinex algorithm vastly improves Otsu's method and GVF snakes but decreases the performance of Mean Shift clustering with color features. This is due to Mean Shift being a color based algorithm and Retinex highly altering the color of the lesions as was illustrated in the previous experiment.

Regarding Mean Shift clustering, the algorithm that only takes color features into consideration yields better results in every dataset compared to the one that takes both color and texture features. Given the results, it can be seen that Mean Shift clustering provides a certain degree of robustness against uneven illumination since the corrected images don't really improve its performance. In addition, Mean Shift clustering offers better results on uncorrected images rather than the other algorithms using corrected ones.

The results provided by Otsu's method and GVF snakes are similar mainly due to GVF snakes requirement of an initial curve to start the algorithm. This curve is commonly done manually; however, in order to make it automatic, the Otsu's method is used to predict the initial points as described in Chapter 3. This means that there is a failed initial curve every time Otsu's method fails to predict the location of the lesion resulting in segmentation errors on the same images.

The results of this experiment will be used in the next one in order to measure which algorithm calculates the best fit of the lesion area. Only the images that were labeled as Correct in every technique will be used to measure the segmentation accuracy in order to avoid noisy results.

Lesion Segmentation Accuracy

The second segmentation experiment involves pixel values calculations aimed at estimating how well each technique segments the lesion, in other words, if the segmented area contains the most number of pixels relative to the lesion and the least number of pixels relative to healthy skin.

The measure used to determine the segmentation accuracy is the calculation of pixel means for the lesion area and the skin area. These means are then subtracted in order to obtain a difference. The higher the difference means that the lesion area has been better separated from the skin area.

A second measurement which involves the calculation of standard deviations is performed to further study the segmentation process. The standard deviation of the lesion area is expected to present a low standard deviation when comparing techniques.

Intensity mean value difference. Figure 35 depicts the difference of color means between lesion and skin areas of the melanoma and non-melanoma sets for the Original dataset by using the four different techniques. Likewise, Figure 36 shows the results related to the Morpho dataset and Figure 37 illustrates the results related to the Retinex dataset. As previously discussed, the images selected for this experiment belong only to the 'Correct' group in all techniques. The means are calculated for the intensity values.

In addition, Table 4 shows the average mean difference of each technique in order to have a clear understanding of how each technique performs in comparison with the rest and how the datasets affect them.

Studying the results, it can be seen that the melanoma subset has an overall higher mean difference than the non-melanoma subset. This is expected since the melanoma lesions are typically darker than the non-melanoma lesions. The overall difference of a subset doesn't imply that it is segmented better since the comparison must be done between each image's differences individually. This is depicted by Table 4 and it shows the average performance of each technique.

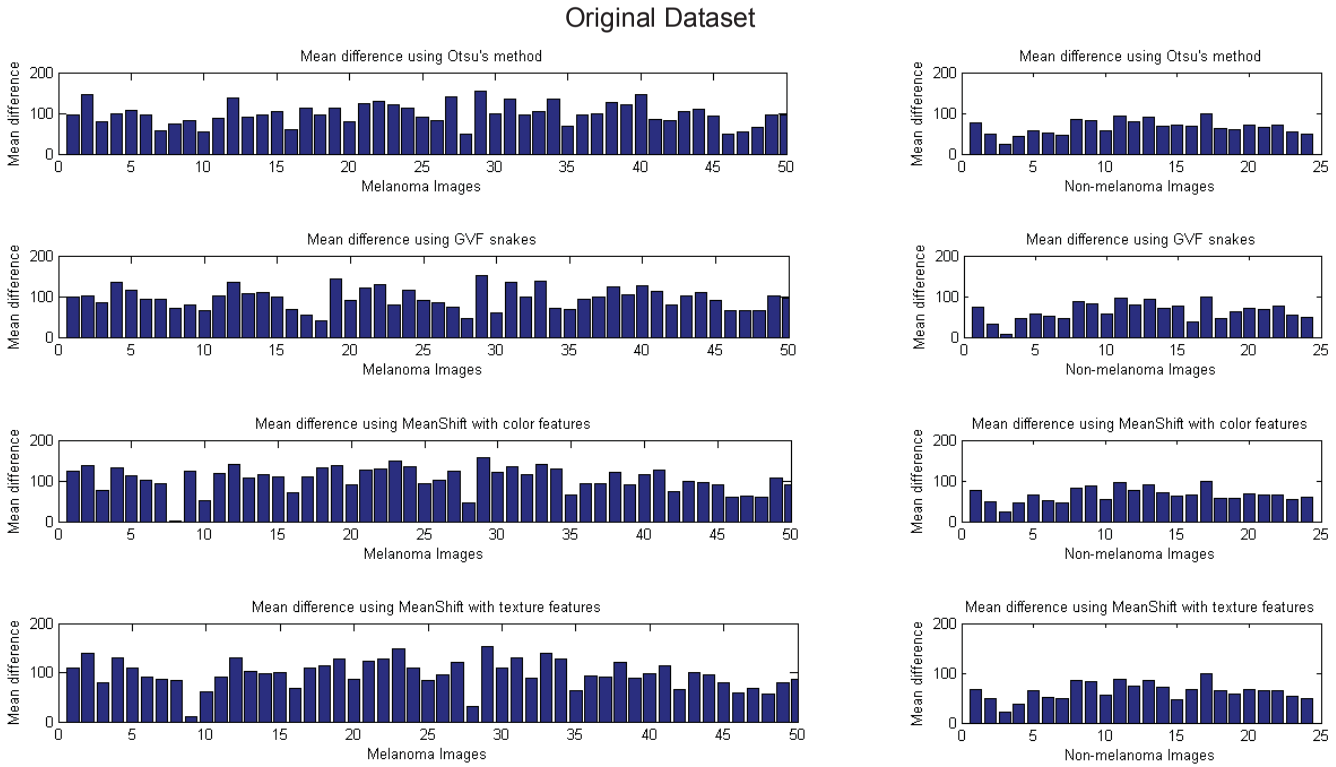


FIGURE 36. Mean differences between lesion area and skin area for melanoma (left column) and non-melanoma (right column) subsets. Results depicted for Otsu's method (first row) GVF snakes (second row) Mean Shift clustering with color features (third row) and Mean Shift clustering using color and texture features (fourth row) on the Original dataset.

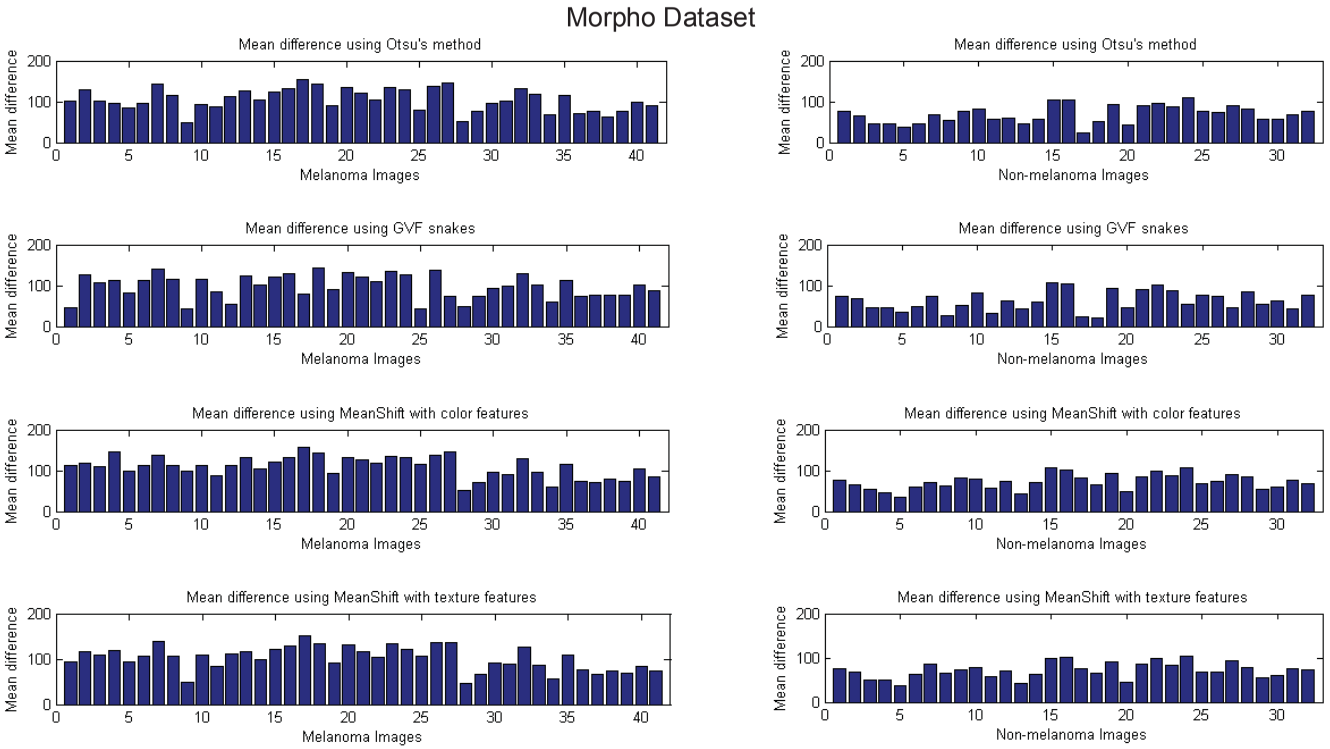


FIGURE 37. Mean differences between lesion area and skin area for melanoma (left column) and non-melanoma (right column) subsets. Results depicted for Otsu's method (first row) GVF snakes (second row) Mean Shift clustering with color features (third row) and Mean Shift clustering using color and texture features (fourth row) on the Morpho dataset.

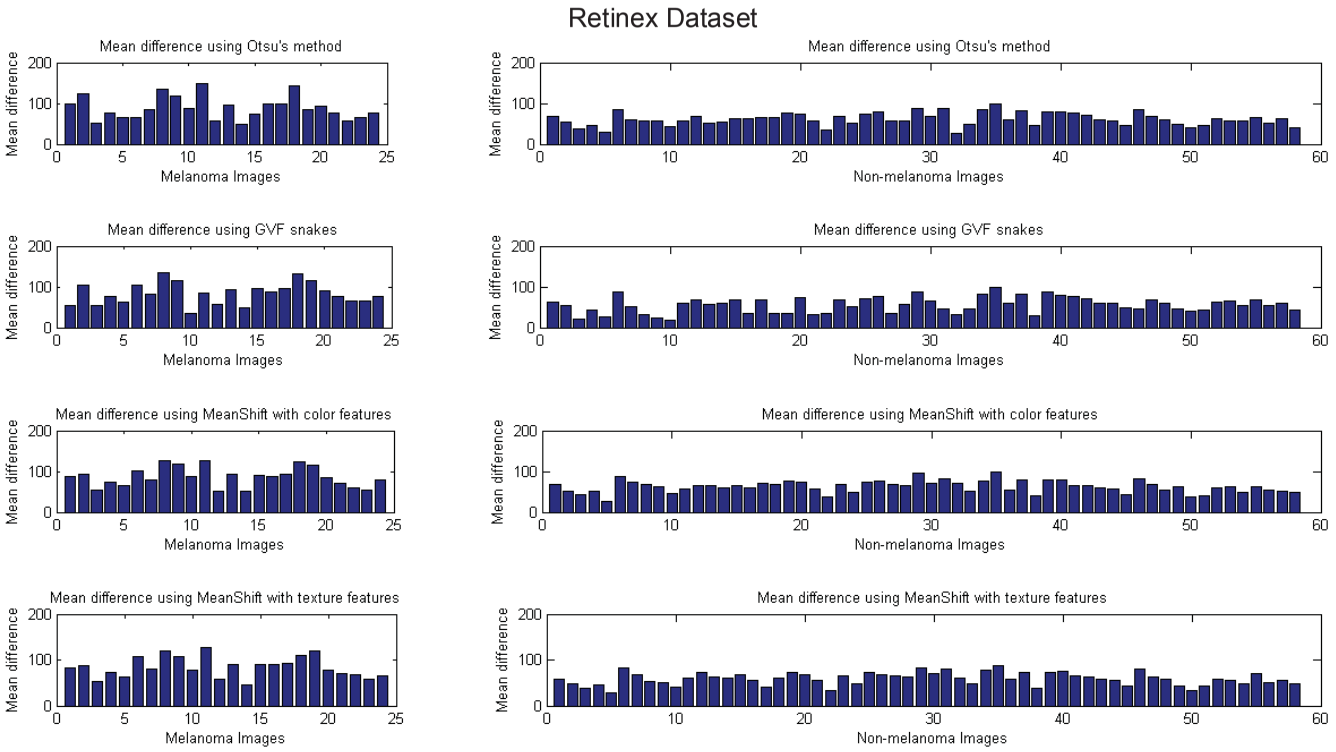


FIGURE 38. Mean differences between lesion area and skin area for melanoma (left column) and non-melanoma (right column) subsets. Results depicted for Otsu's method (first row) GVF snakes (second row) Mean Shift clustering with color features (third row) and Mean Shift clustering using color and texture features (fourth row) on the Retinex dataset.

TABLE 4. Averaged Mean Difference between Lesion and Skin Areas in each Image

		Otsu's method	Snakes	MS color	MS texture
Original	Melanoma	98.514	95.465	104.234	97.404
	Non-Melanoma	65.439	63.112	65.295	63.029
	Total	87.787	84.972	91.605	86.255
Morpho	Melanoma	104.786	97.997	109.008	101.853
	Non-Melanoma	68.594	62.164	72.681	71.8692
	Total	88.921	82.289	93.084	88.709
Retinex	Melanoma	88.297	83.483	86.129	84.152
	Non-Melanoma	73.812	67.247	75.895	71.27
	Total	69.05	63.799	69.635	66.349

Regarding Mean Shift, the algorithm that uses only color features as subspace is superior in all the datasets compared to the one that uses both color and texture. Moreover, Mean Shift has the highest total performance in all datasets compared to the rest of techniques. While the datasets from Original, Morpho and Retinex present very different average results, it is important to remember that the datasets for each of them are only the ones labeled as 'Correct' hence, the images from each dataset are not the same and this alters the results from set to set.

Lesion intensity standard deviation. The standard deviation (SD) also offers insight in the segmentation process. Figure 38, Figure 39 and Figure 40 illustrate the SD of the segmented lesions for Original, Morpho and Retinex sets respectively.

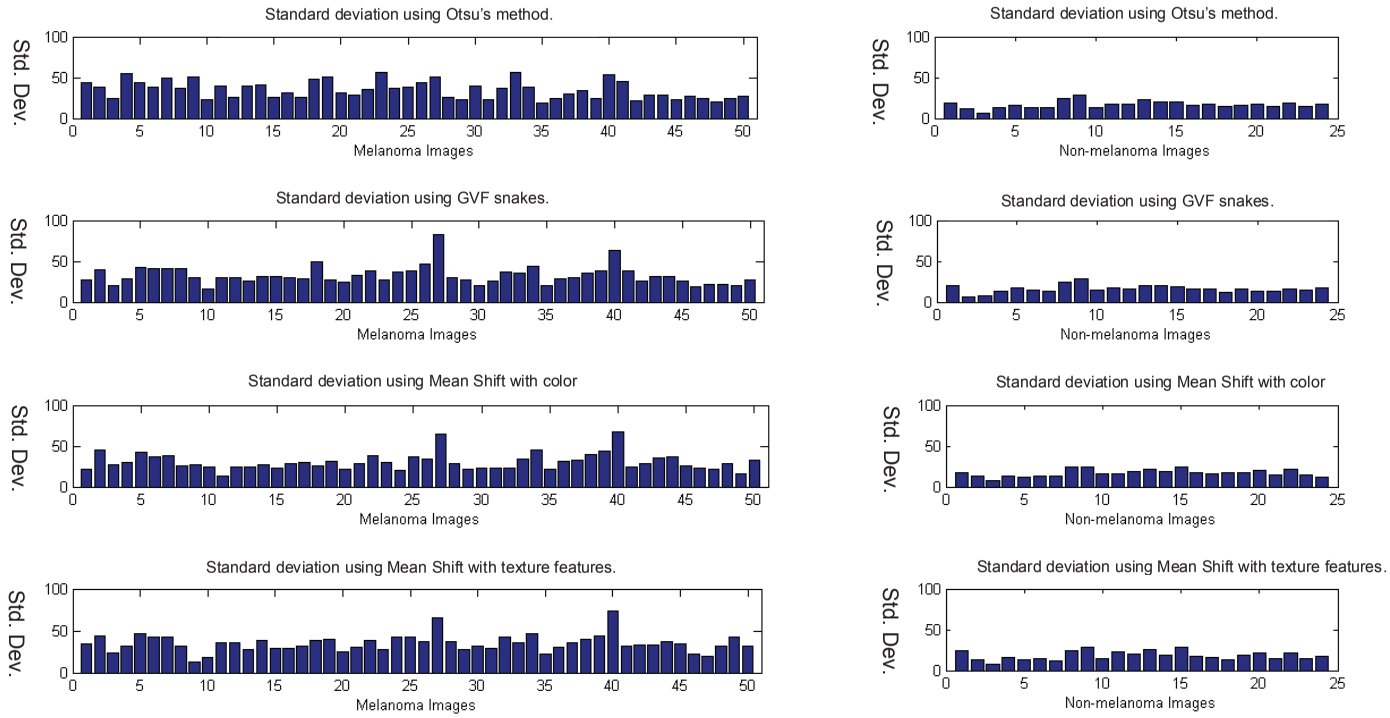


FIGURE 39. Standard deviation of the lesion area for melanoma (left column) and non-melanoma (right column) subsets. Results depicted for Otsu's method (first row) GVF snakes (second row) Mean Shift clustering with color features (third row) and Mean Shift clustering using color and texture features (fourth row) on the Original dataset.

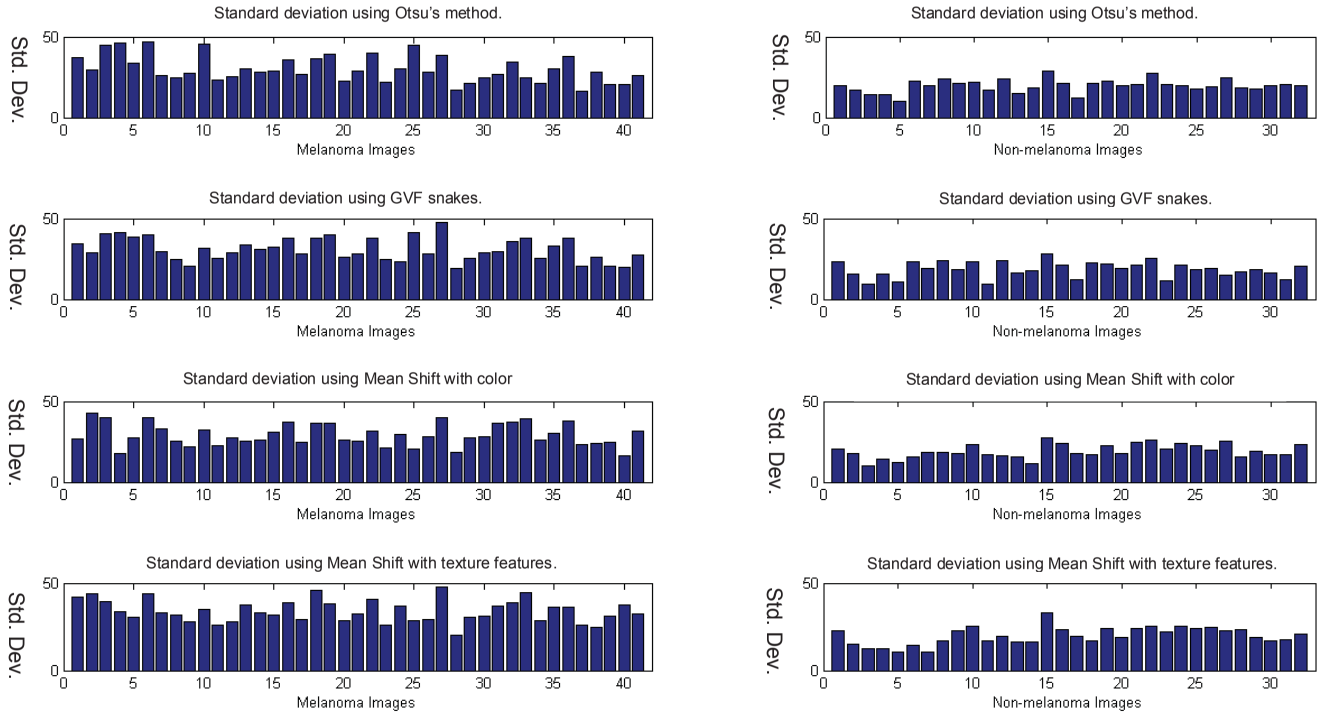


FIGURE 40. Standard deviation of the lesion area for melanoma (left column) and non-melanoma (right column) subsets. Results depicted for Otsu's method (first row) GVF snakes (second row) Mean Shift clustering with color features (third row) and Mean Shift clustering using color and texture features (fourth row) on the Morpho dataset.

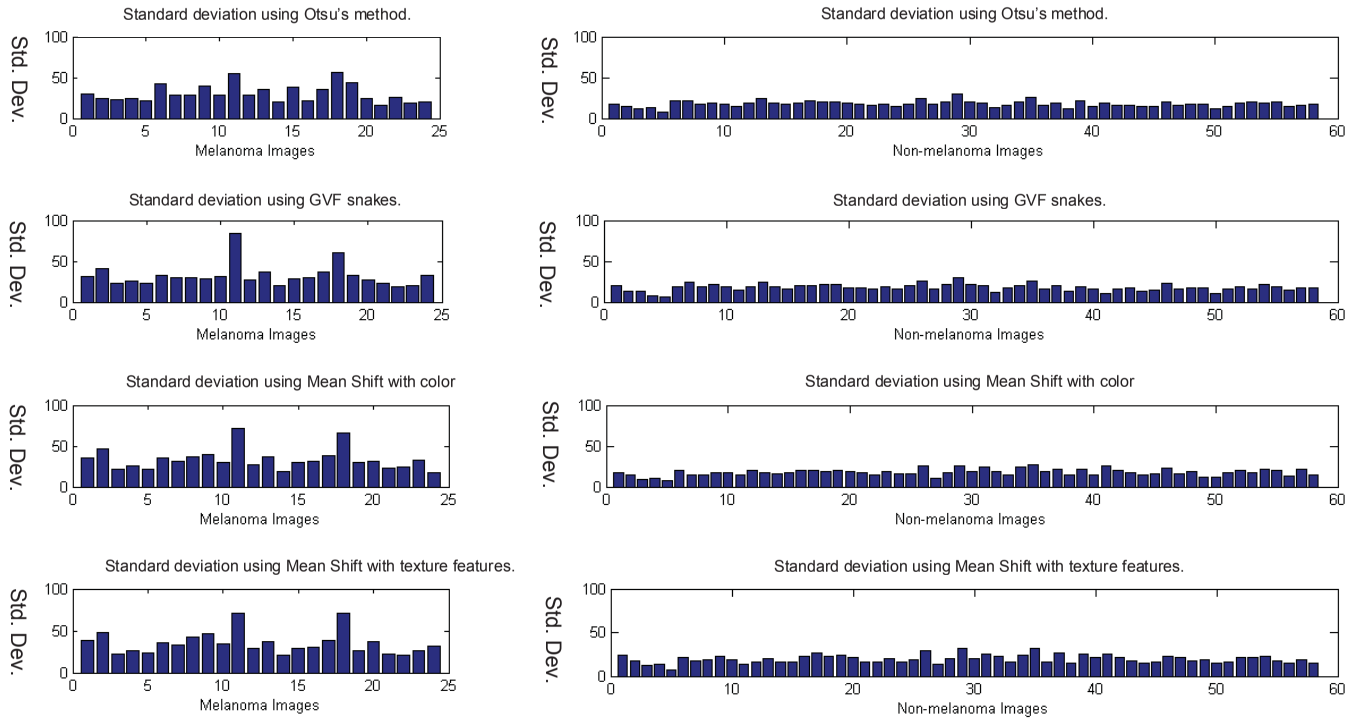


FIGURE 41. Standard deviation of the lesion area for melanoma (left column) and non-melanoma (right column) subsets. Results depicted for Otsu's method (first row) GVF snakes (second row) Mean Shift clustering with color features (third row) and Mean Shift clustering using color and texture features (fourth row) on the Retinex dataset.

In this case, the melanoma subset presents a higher standard deviation than the non-melanoma subset due to the expected changes in lesion color as opposed to the mostly constant color palette present in non-melanoma images.

Table 5 shows the averaged standard deviations in order to get a comparative view of the four segmentation techniques.

TABLE 5. Standard Deviation Average in Lesion Area

		Otsu's method	Snakes	MS color	MS texture
Original	Melanoma	35.007	32.493	30.454	35.151
	Non-Melanoma	16.76	16.224	16.987	18.321
	Total	29.089	27.216	26.0862	29.6927
Morpho	Melanoma	30.225	30.908	29.204	33.929
	Non-Melanoma	19.688	18.439	19.027	19.883
	Total	25.606	25.442	24.742	27.772
Retinex	Melanoma	30.329	32.295	33.297	35.099
	Non-Melanoma	21.305	21.458	21.332	23.733
	Total	21.348	22.0134	22.233	24.165

The results show that Mean Shift with color features is again superior having a lower standard deviation than the texture counterpart. Mean Shift has also lower standard deviations than the rest of the techniques in the Original and Morpho sets, being Otsu's method the algorithm that obtains the lowest standard deviation in the Retinex set.

Discussion. In the segmentation experiment, the results have shown that Mean Shift is an algorithm that matches or improves the performance over other well-known techniques that are widely used in melanoma detection. On the other hand, the Mean Shift algorithm is greatly slower than the rest of the techniques making its use under some environments (e.g. real time applications) unviable.

Regarding algorithm comparisons, while the implementation using color features has been proven to be the most successful, the clustering results for the texture based Mean Shift show that most of the failed segmentations come from an oversized initial bandwidth window, leading to a high number of clusters which result in incorrect lesion segmentations.

Classification Experiments

The last set of experiments is related to the classification stage of the system. For the experiments, test and a train vectors were used as previously described at the beginning of the chapter. Four classifiers were trained, an SVM model and three discriminant analysis classifiers; linear, quadratic and based on mahalanobis distances.

During classification, the classes to be determined are labeled as 'positive' and 'negative'. In this case, the system is trying to detect melanoma images so the melanoma ones are the positive class and the non-melanoma lesions are the negative class. In a classification experiment with two classes, four possible outcomes can occur denoted true positive, true negative, false positive and false negative. A true positive happens when an image which is a real melanoma is classified by the system as a melanoma. A true negative happens when an image which is not a melanoma is classified as non-melanoma.

On the other hand, a false positive happens when an image which is not a melanoma is classified as melanoma and a false negative happens when an image which is a melanoma is classified as non-melanoma (see Figure 41).

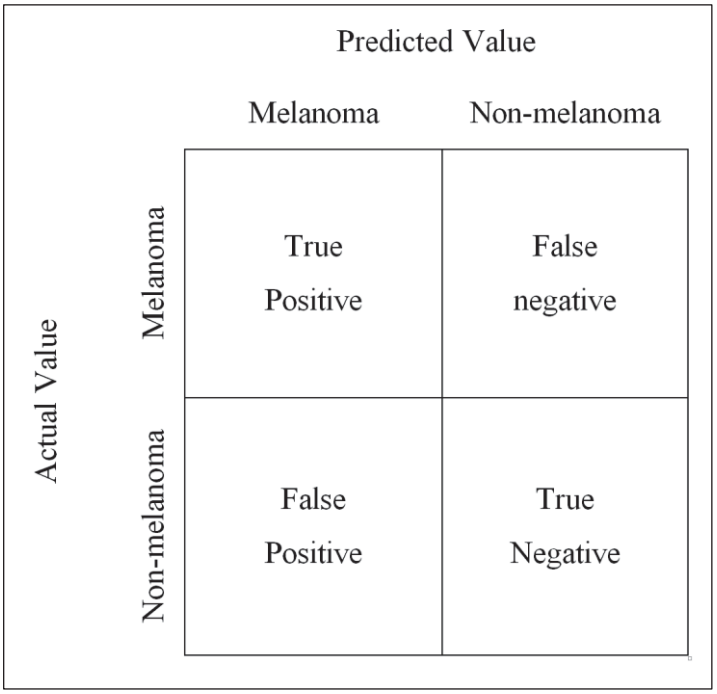


FIGURE 42. Actual value vs. predicted value diagram.

With these values, performance measures of specificity, sensitivity and accuracy can be calculated. These measures are defined as:

$$specificity = \frac{\sum true\ negative}{\sum true\ negative + \sum false\ positive} \quad (4.1)$$

$$sensitivity = \frac{\sum true\ positive}{\sum true\ positive + \sum false\ negative} \quad (4.2)$$

$$accuracy = \frac{\sum true\ positive + \sum true\ negative}{\sum All\ test\ images} \quad (4.3)$$

Table 6 shows the classification results of the four classifiers using 150 images as testing set.

TABLE 6. Classification Results

	SVM	Linear	Quadratic	Mahalanobis
True Positive	75	39	75	75
True Negative	71	75	67	38
False Positive	4	0	8	37
False Negative	0	36	0	0
Total Correct	146	114	142	113
Total Failed	4	36	8	37

The results show that the classifiers are able to robustly define a class while the other sometimes gets classification errors. This is mainly due to the characteristics of the datasets. The melanoma dataset is inherently varied given the different patients, acquisition devices and illumination environments while the non-melanoma dataset has been acquired in a controlled environment resulting in highly correlated images. This means that the classifier trains a strict model for the non-melanoma images and when an image is slightly different from the common characteristics, the classifier decides it belongs to the melanoma class.

Regarding performance, Table 7 depicts the specificity, sensitivity and accuracy of each technique that will define how well each classifier works.

TABLE 7. Specificity, Sensitivity and Accuracy of Classifiers

	SVM	Linear	Quadratic	Mahalanobis
Specificity	94.67 %	100 %	89.33 %	50.67 %
Sensitivity	100 %	52 %	100 %	100 %
Accuracy	97.33 %	76 %	94.67 %	75.33 %

In this case, sensitivity refers to the ability of the classifier to correctly detect a melanoma and specificity to the ability of correctly excluding a melanoma from an image. Given the preventive applications of these kinds of systems, it is more critical correctly detecting a melanoma than making fewer mistakes when determining that an image is not a melanoma. Thus, the best overall technique is the SVM classifier followed by the Quadratic function of discriminant analysis.

Another way to represent the performance is the receiver operating characteristic (ROC) curve which displays the true positive rate against the false positive rate. Figure 42 represents the ROC curve of the SVM classifier. The SVM high performance makes the curve reach the top left corner which means that it can achieve high melanoma detection rates even when the false positive error rates are low.

Discussion.

In this section, the implemented segmentation algorithms have been measured and studied in order to obtain the best performing one. The SVMs have been proven to be the best performing classifiers given their high accuracy rates. It is important to note that SVMs are also the slowest to process so this must be taken into consideration depending

on the application. One of the most important aspects to highlight in these experiments is the importance of the dataset to be used. As previously discussed, both melanoma and non-melanoma subsets have different characteristics that influence the way the classifier is trained. Highly correlated sets make the training process choose models that may not work for the whole dataset. In addition the bigger the database, the better the results can represent the functionality.

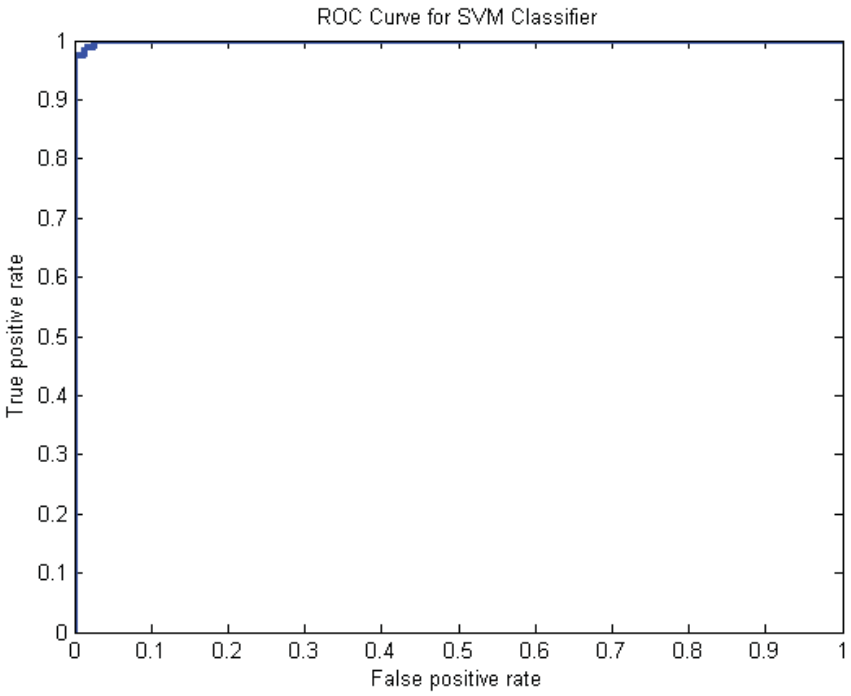


FIGURE 43. ROC curve for SVM classifier.

CHAPTER 5

GRAPHICAL USER INTERFACE TOOL

This chapter describes a tool developed using the system described and implemented in Chapter 3. The tool objectives are to provide an interactive and educational view on melanoma detection systems with ease of use thanks to a graphical user interface that shows results onscreen.

This chapter covers the structure of the program, its main features and functionality.

Features

The tool provides a set of buttons that allow the user to perform detection tasks by using different techniques. In addition, the partial results of each stage are shown in an area designated to provide feedback.

The database of images to be processed can be selected by the user thanks to the ability of the tool to load any folder by prompting a browse dialog. The tool automatically search for any image files inside the selected folder and loads them. The user can then preview each of the loaded images in the results area and start performing detection actions (see Figure 43). The left menu provides the detection actions divided by illumination correction, lesion segmentation and classification in a similar way the developed system defined in Chapter 3 was structured.

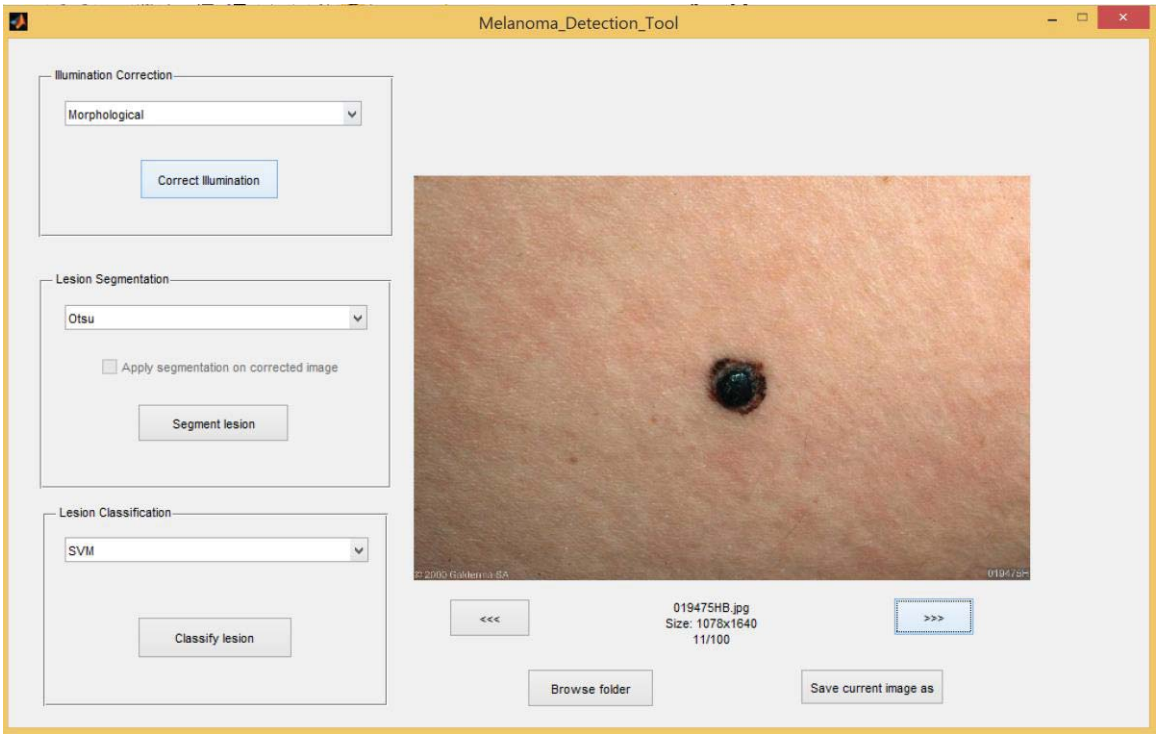


FIGURE 44. Database preview in the melanoma detection tool.

In every image, the user can perform illumination correction actions selecting the preferred algorithm. The software performs the correction and shows the results in the preview area (see Figure 43). At any moment, the user can save to a folder the image that is currently displayed on the preview area.

For lesion segmentation actions, the tool offers the four segmentation algorithms that were developed for the system although some of them display a warning due to their heavy process requirements and long waiting times. The user can select if the segmentation is to be performed on the original loaded image (see Figure 44) or on the illumination corrected one (see Figure 45). This provides a good understanding on how illumination correction influences segmentation techniques.

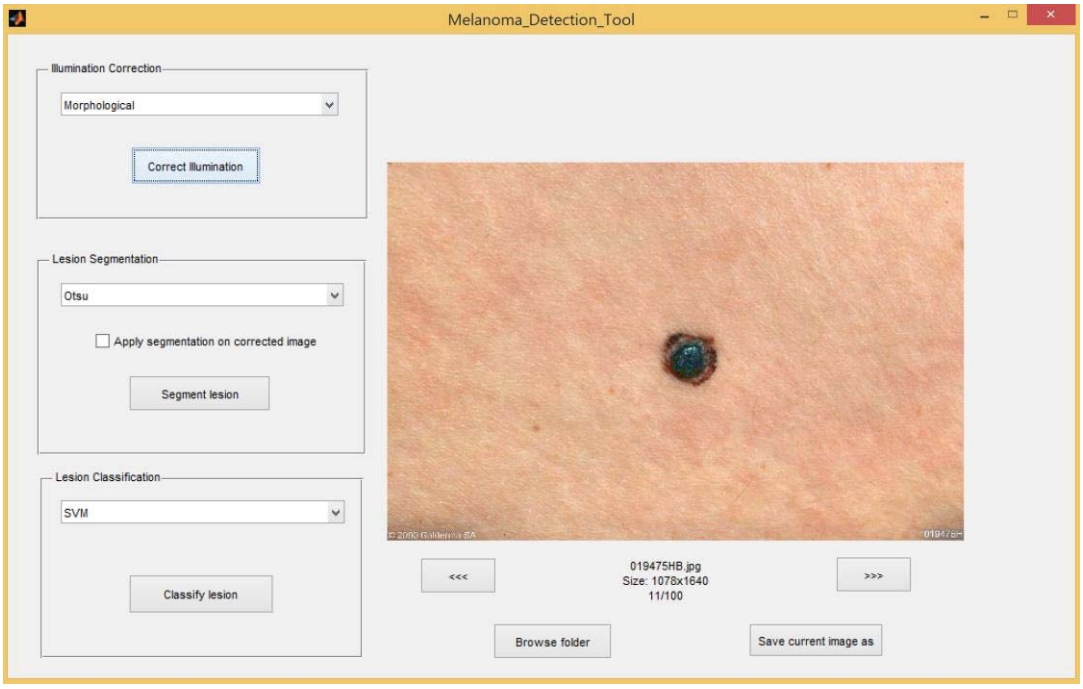


FIGURE 45. Illumination correction performed over loaded image.

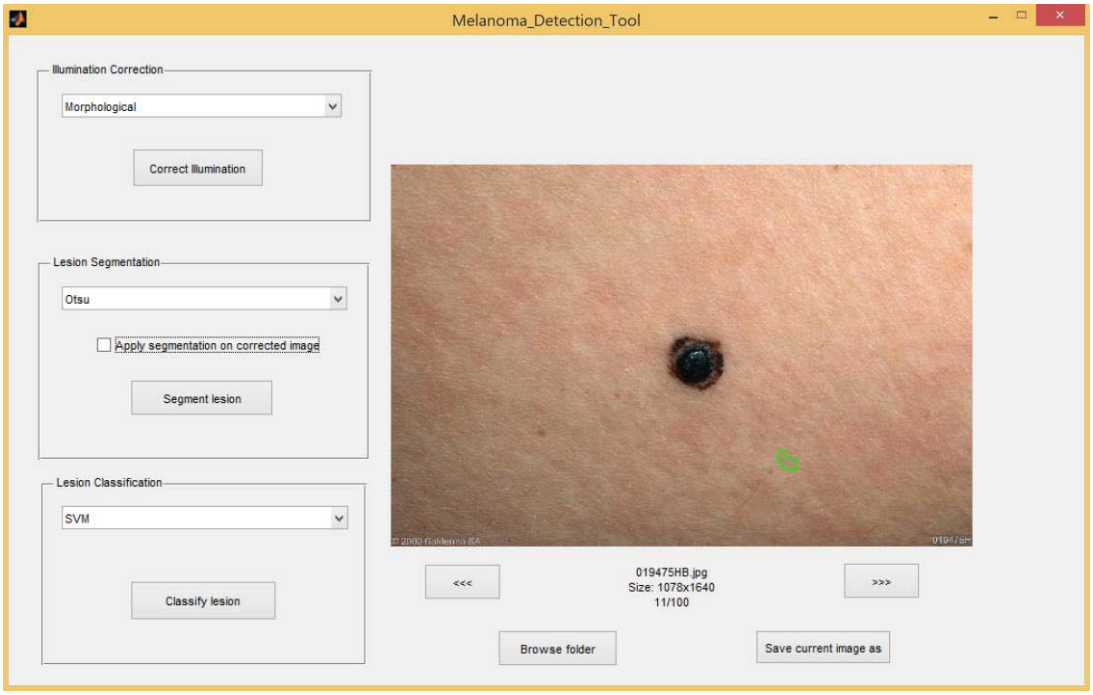


FIGURE 46. Otsu's method incorrect segmentation of original image.

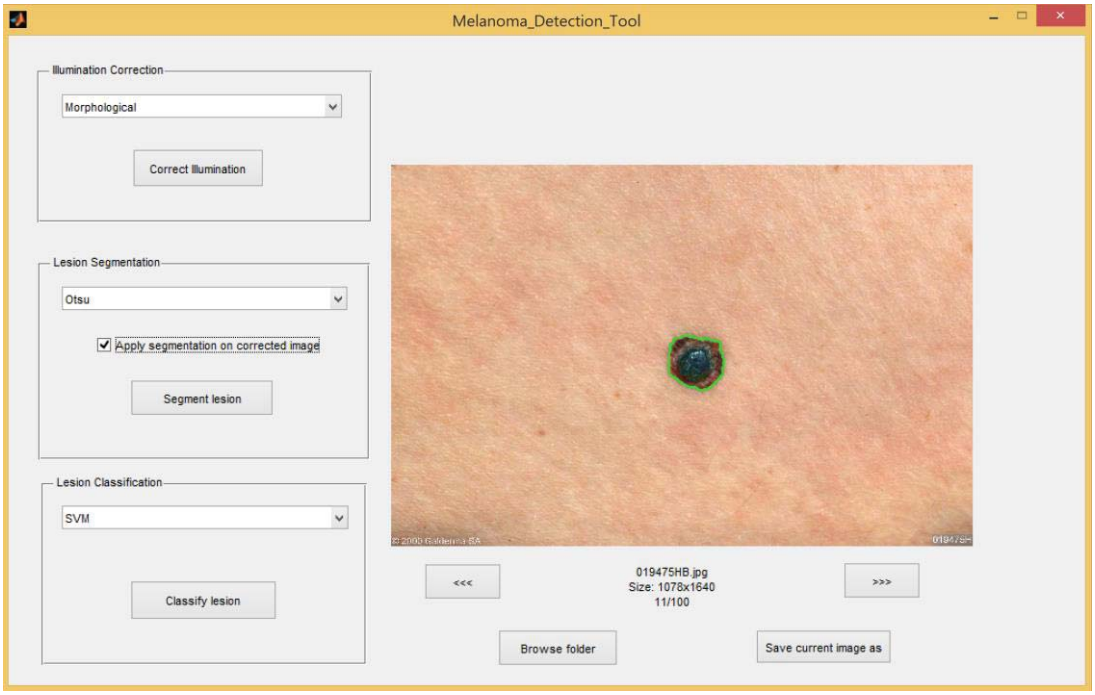


FIGURE 47. Otsu's method properly segmenting corrected image.

Finally, the classification can be performed using SVMs or any of the discriminant analysis techniques implemented. The program will take a trained model and make a decision by extracting the features of the segmented lesion and show the result on the preview area (see Figure 46). The train set can be modified by changing the images on the folder labeled as 'Train' with new ones. Apart from training images, the user must also provide masked images that define the lesion area in order to generate a robust classifier.

Program Structure

The program is developed using MATLAB GUIDE platform which provides tools for designing user interfaces for custom apps. Most of the code used to implement the techniques in previous chapters is used for the tool with some variations.

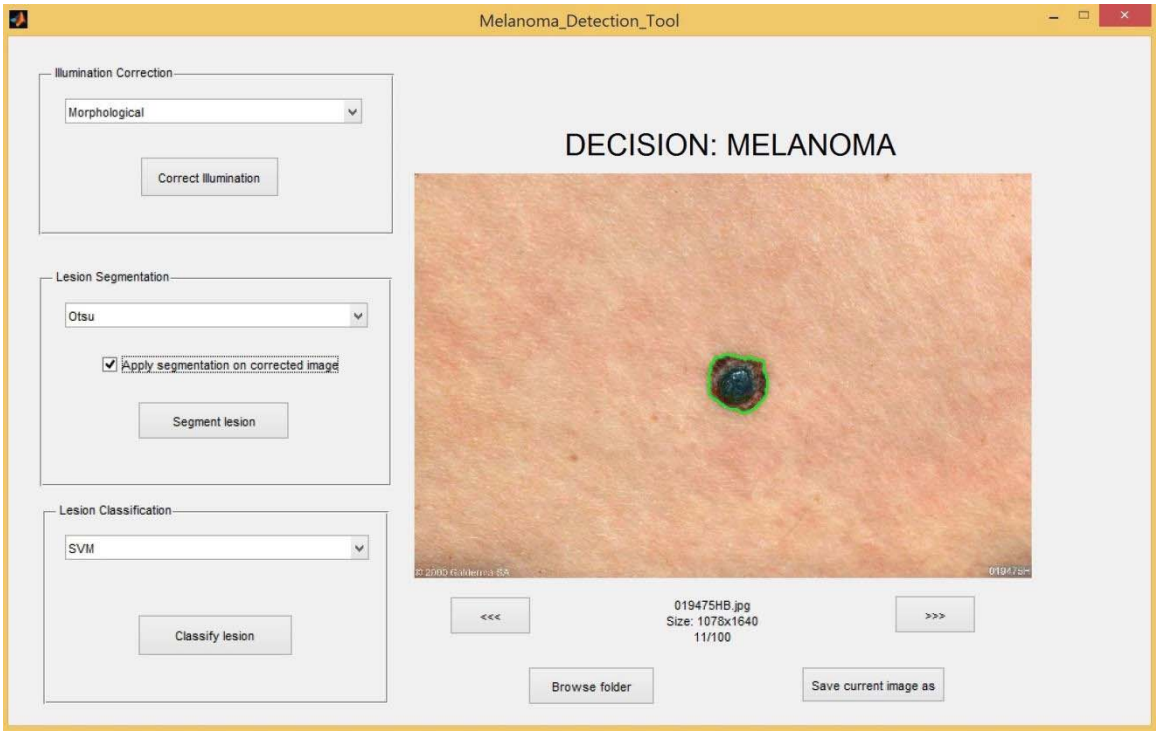


FIGURE 48. Classification of a melanoma image.

The GUI is divided in two main parts, preview and process. The preview area contains a MATLAB axes object where all the feedback is displayed. The controls include two pushbuttons to navigate a loaded database and a browse pushbutton to load new databases. The file types that the program searches for when loading a database are restricted to PNG, BMP, JPEG and TIFF. If no images have been found, the preview area displays a warning message. The size of each image is obtained and displayed since large images are resized to fit inside the axes.

Regarding the processing part, the controls in the left menu control the actions that involve the detection process. Each stage and technique has been rewritten as a function that will be called when the associated pushbutton is clicked. Instead of reading

a set of images in a loop, the images are stored in an array when the browse function is used and the array is the input passed to each function along with the index pointing to the image that needs to be processed. The functions output will be the processed image in the case of illumination correction techniques, the processed image and the logical lesion mask for the segmentation techniques and the decision variable for classification.

The tool considers the sequential aspect of the system by setting a series of constraints disabling the options that cannot be used given the circumstances. When there are no images in a specific folder, the whole processing menu is disabled in order to avoid errors. The classification can only be performed if the image has previously been segmented, and the segmentation of the corrected image can only be done when an image has previously been through the illumination correction stage.

Conclusion

In this chapter the tool for melanoma detection developed for this thesis has been presented and detailed. The tool is able to dynamically make use of the detection techniques that were implemented in previous chapters and provide visual feedback of the process. By allowing different datasets, this tool can be used for other research needs and objectives. It is important to note that the program cannot run by itself in any computer since it needs the software MATLAB that provides the framework and functions required to properly work, MATLAB is a multi-platform software that requires a license to be used.

CHAPTER 6

CONCLUSIONS

As a concluding chapter for this thesis, a summary of the related work along with contributions and some guidelines for potential future work are provided.

Summary of Work

In this thesis, a full automatic system for melanoma detection has been designed, developed and studied providing existing and novel techniques in the field.

A comprehensive study of the medical background on melanoma and dermatological procedures for detection has been performed. In addition, a research on state-of-the-art techniques in melanoma detection and image processing algorithms has been accomplished with comparative studies on each technique.

Each stage of a melanoma detection system is defined and techniques for each of them selected and studied before implementing them.

For illumination correction, a critical issue in most detection systems, two algorithms are implemented; one based in morphological operators and another based in the Retinex algorithm. Regarding segmentation, two existing algorithms are used; Otsu's method thresholding and GVF snakes. In addition, a novel technique in melanoma segmentation is developed based in Mean Shift clustering. Two different Mean Shift algorithms are developed to study their performance based on their feature subspace; only regarding color or regarding color and texture values.

Regarding feature extraction, existing techniques based on the ABCD rule mostly concerning lesion geometry and texture based features are extracted in order to construct a feature vector for classification. The features are further selected obtaining a vector that contains the most relevant information.

Two different techniques are selected and implemented for the segmentation stage; SVMs, and discriminant analysis making use of three different functions to contrast their impact on detection rate.

The development of the system is followed by a study of every technique implemented in a series of experiments that will test and compare them. The experiments have provided a view in how the system's stages affect each other and which techniques perform better. The Mean Shift algorithm has been individually studied and then compared with other segmentation techniques to obtain a comprehensive view in its performance resulting in a segmentation algorithm that is superior in many aspects to other existing techniques.

In this work, the acquisition of a database formed of non-melanoma images has been achieved under controlled environment. This has permitted the study of the impact that a dataset with specific characteristics introduces in a classification problem.

As a final contribution, this thesis compiles all the developed work into a GUI tool that allows the use of a full detection system with user-selected datasets. The tool is highly dynamic and modular since the technique used in each stage can be changed individually. The tool is developed in MATLAB meaning that it can be used in Microsoft Windows, Mac OS and GNU/Linux.

Future Work

The work of this thesis opens various possibilities where future research can be focused on. By studying the development process and the results obtained, the following future work is proposed with the aim of improving the existing work, work on existing challenges and find new approaches to the problem.

Regarding datasets, a big melanoma database can be acquired in order to increase the number of observations and get more complete results. A protocol for image acquisition should be established in order to access other important features like lesion size. Having a fixed distance between the camera and the skin, a ratio of pixel per mm could be calculated and the size predicted. This characteristic will, thus, generate potential studies on systems that calculate variation of lesion size over time.

The Mean Shift based segmentation algorithm presented in this thesis, while achieving a good performance, requires a large amount of time to be processed. Researches focused on improving this algorithm's performance related to computational cost could make it viable on other applications where time constrains are very strict like real time systems.

Concerning the GUI tool, its development in MATLAB makes its use restricted to individuals with the software and the computational time required to perform calculations is high due to the entire framework running under the program. A tool developed in another language such as C++ would take advantage of a much faster processing time with computer vision libraries such as OpenCV. The tool would also benefit from being a standalone program without additional software required to run it.

Most of the studies including this thesis focus on the detection of a single lesion. However, a system that is able to detect and process multiple lesions in a single image is interesting since more than one lesion sometimes appear in individual images. A system focused on parallelism and pipelines would be capable of achieving this challenge.

Real time detection is a very important topic, since the ability of obtaining results while performing the acquisition saves the time required for storing the data and processing it. Devices with video cameras would take advantage of these systems and with enough processing power; the system presented in this thesis could be easily used for real time detection.

Finally, the subject of real time detection inevitably passes through the smart phone industry. A system embedded in a smart phone not only takes advantage of its video camera and portable potential but also from its connectivity enabling a system to make use of the network to provide services such as telemedicine or cloud-connected databases.

REFERENCES

REFERENCES

[1] R.M. Pluta, A.E. Burke and R.M. Golub, "Melanoma", *JAMA*, vol. 305, p. 2368, Jun. 2011.

[2] T. Ades et al., "Cancer Facts and Figures 2014," Atlanta, GA: American Cancer Society, 2014.

[3] National Cancer Institute (2014, Nov 7). *PDQ® Melanoma Treatment* [Online]. Available: http://www.cancer.gov/cancertopics/pdq/treatment/melanoma/HealthProfessional.

[4] J.L. Bolognia, J.L. Jorizzo and J. V. Schaffer, *Dermatology*, 3rd ed., 2 vols., Philadelphia, PA: Saunders, 2012.

[5] T.M. Rünger, B. Farahvash, Z. Hatvani and A. Rees, "Comparison of DNA damage responses following equimutagenic doses of UVA and UVB: A less effective cell cycle arrest with UVA may render UVA-induced pyrimidine dimers more mutagenic than UVB-induced ones," *Photochem. Photobiol. Sci.*, vol. 11, pp. 207-215, 2012.

[6] S. Halachmi and B.A. Gilchrest, "Update on genetic events in the pathogenesis of melanoma," *Current Opinion in Oncology*, vol. 13, pp. 129-136, 2001.

[7] 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," *C.A. Cancer. J. Clin.*, vol. 35, pp. 130-151, Jun. 1985.

[8] L.A. Goldsmith et al., "NIH Consensus conference. Diagnosis and treatment of early melanoma," *JAMA*, vol. 268, p. 1314, Sep. 1992.

[9] T. Schindewolf et al., "Evaluation of different image acquisition techniques for a computer vision system in the diagnosis of malignant melanoma," *J. Am. Acad. Dermatol.*, vol. 31, pp. 33-41, Jul. 1994.

[10] M. Mogensen, L. Thrane, T.M. Joergensen, P.E. Andersen and G.B. Jemec, "Optical coherence tomography for imaging of skin and skin diseases," *Semin. Cutan. Med. Surg.*, vol. 28, pp. 196-202, Sep. 2009.

[11] J. Glaister, R. Amelard, A. Wong and D.A. Clausi, "MSIM: Multistage illumination modeling of dermatological photographs for illumination-corrected skin lesion analysis," *IEEE Trans. Biomed. Eng*., vol. 60, pp. 1873-1883, Jul. 2013.

[12] E.H. Land, "The retinex theory of color vision," *Scientific American*, vol. 237, pp. 108-128, Dec. 1977.

[13] B. Jahne, H. Haussecker and P. Geissler, *Handbook of computer vision and applications*, London, UK: Academic Press, 1999.

[14] H. J.A.M. Heijmans, "Connected morphological operators for binary images," *Computer Vision and Image Understanding*, vol. 73, pp. 99-120, Jan. 1999.

[15] E.H. Land and J.J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, pp. 1-11, 1971.

[16] M. Silveira et al., "Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," *IEEE Journal of Selected Topics in Signal Processing*, vol.3, no.1, pp.35-45, Feb. 2009.

[17] J.A. Hartigan and M.A. Wong, "Algorithm AS 136: A K-means clustering algorithm," *Applied Statistics*, vol. 28, no. 1, pp. 100-108, 1979.

[18] K. Fukunaga and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Inf. Theory*, vol. 21, pp. 32-40 ,1975.

[19] D. Ziou and S. Tabbone "Edge detection techniques: An overview," *International Journal of Pattern Recognition and Image Analysis*, vol. 8, pp. 537-559, 1998

[20] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no.6, pp.679-698, Nov. 1986.

[21] C. Xu and J.L. Prince, "Gradient vector flow: A new external force for snakes," *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Jun. 1997, pp. 66-71.

[22] I. Taranu and I. Iacovici, "Applying ABCD rule of dermatoscopy using cognitive systems," *The 2014 International Conference on Systems, Control, Signal Processing and Informatics II*, Prague, Czech Republic, Apr. 2014, pp. 111-114.

[23] D.S. Rigel et al., *Cancer of the Skin*, 2nd ed., Philadelphia, PA: Saunders, Jul. 2011.

[24] H. G. Feichtinger and T. Strohmer: *Gabor Analysis and Algorithms*, Berlin, Germany: Birkhäuser, 1998.

[25] R. Garnavi, M. Aldeen and J. Bailey, "Computer-aided diagnosis of melanoma using border- and wavelet-based texture analysis," *IEEE Trans. Technol- Biomed*., vol. 16, pp. 1239-1252, Aug. 2012.

[26] 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*, vol.8, no.3, pp.965-979, Sept. 2014.

[27] R.A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179-188, Sep. 1936.

[28] O. Ivanciuc, "Applications of support vector, machines in chemistry," *Reviews in Computational Chemistry*, ch. 6, Vch Pub, 2007.

[29] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge, UK: Cambridge University Press, 2000.

[30] J. R. Quinlan, "Induction of decision trees," *Machine Learning* vol. 1, pp. 81-106, Mar. 1986.

[31] F. Rosenblatt, "The perceptron: A probalistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, pp. 386-408, Nov. 1958.

[32] G.E. Hinton, S. Osindero and Y.W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, pp. 1527–1554, Jul. 2006.

[33] J. McCann, B. Funt and F. Ciurea, "Retinex in matlab," *Proceedings of the IS&T/SID Eighth Color Imaging Conference: Color Science, Systems and Applications*, 2000, pp 112-121.

[34] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man., Cyber.*, vol. 9, pp. 62–66, Jan. 1979.

[35] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, pp. 359-369, Mar. 1998.

[36] B. Erkol, R.H. Moss, R.J. Stanley, W.V. Stoecker and E. Hvatum, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," *Skin Res. Technol*., vol. 11, pp. 17-26, Feb. 2005.

[37] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp.790-799, Aug. 1995.

[38] D. Comaniciu and P. Meer, “Mean shift: A robust approach towards feature space analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, May 2002.

[39] T. Rückstieß, C. Osendorfer and P. Smagt, “Sequential feature selection for classification,” *AI 2011: Advances in Artificial Intelligence*, Berlin, Germany: Springer, 2011, pp. 132-141.

[40] I. Maglogiannis and D.I. Kosmopoulos, “Computational vision systems for the detection of malignant melanoma,” *Oncol. Rep*., pp. 1027-1932, Feb. 2006.

[41] Galderma S.A. (2014, Nov. 14). *Dermquest Image Library* [Online]. Available: https://www.dermquest.com/image-library/.