

# Automatic segmentation of skin lesions from dermatological photographs

by

Jeffrey Glaister

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Master of Applied Science  
in  
Systems Design Engineering

Waterloo, Ontario, Canada, 2013

© Jeffrey Glaister 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

Melanoma is the deadliest form of skin cancer if left untreated. Incidence rates of melanoma have been increasing, especially among young adults, but survival rates are high if detected early. Unfortunately, the time and costs required for dermatologists to screen all patients for melanoma are prohibitively expensive. There is a need for an automated system to assess a patient's risk of melanoma using photographs of their skin lesions. Dermatologists could use the system to aid their diagnosis without the need for special or expensive equipment.

One challenge in implementing such a system is locating the skin lesion in the digital image. Most existing skin lesion segmentation algorithms are designed for images taken using a special instrument called the dermatoscope. The presence of illumination variation in digital images such as shadows complicates the task of finding the lesion. The goal of this research is to develop a framework to automatically correct and segment the skin lesion from an input photograph. The first part of the research is to model illumination variation using a proposed multi-stage illumination modeling algorithm and then using that model to correct the original photograph. Second, a set of representative texture distributions are learned from the corrected photograph and a texture distinctiveness metric is calculated for each distribution. Finally, a texture-based segmentation algorithm classifies regions in the photograph as normal skin or lesion based on the occurrence of representative texture distributions. The resulting segmentation can be used as an input to separate feature extraction and melanoma classification algorithms.

The proposed segmentation framework is tested by comparing lesion segmentation results and melanoma classification results to results using other state-of-the-art algorithms. The proposed framework has better segmentation accuracy compared to all other tested algorithms. The segmentation results produced by the tested algorithms are used to train an existing classification algorithm to identify lesions as melanoma or non-melanoma. Using the proposed framework produces the highest classification accuracy and is tied for the highest sensitivity and specificity.

## **Acknowledgements**

I would like to thank Prof. David Clausi and Prof. Alexander Wong for supporting and guiding me, helping me learn and grow and become a better researcher over the duration of my Masters program. I would also like to thank Prof. Andrea Scott and Prof. Jonathan Histon for reviewing my thesis and all the members of the Vision and Image Processing Lab for their support.

I would like to thank the Natural Sciences and Engineering Research Council (NSERC) of Canada, Agfa Healthcare Inc., and the Ontario Centres of Excellence (OCE) for funding my work on this project.



## **Dedication**

This is dedicated to my family, friends, teachers and professors, who have helped me throughout the years.

# Table of Contents

List of Tables	ix
List of Figures	x
Nomenclature	xii
<b>1 Introduction</b>	<b>1</b>
1.1 Melanoma Screening . . . . .	1
1.2 Visual melanoma screening limitations . . . . .	3
1.3 Automated melanoma screening systems . . . . .	3
1.4 Thesis outline and contributions . . . . .	5
<b>2 Background</b>	<b>6</b>
2.1 Existing illumination correction algorithms . . . . .	6
2.1.1 General illumination correction algorithms . . . . .	6
2.1.2 Illumination correction algorithms for digital images of skin lesions	7
2.2 Existing segmentation algorithms . . . . .	8
2.2.1 Texture-based segmentation algorithms . . . . .	8
2.2.2 Lesion segmentation algorithms . . . . .	8

<b>3</b>	<b>Illumination Correction of Skin Lesion Images</b>	<b>10</b>
3.1	Problem formulation . . . . .	11
3.2	Initial non-parametric illumination modeling . . . . .	13
3.2.1	Final parametric modeling of illumination . . . . .	15
3.2.2	Reflectance map estimation . . . . .	17
3.3	Implementation . . . . .	17
3.3.1	Optimization . . . . .	17
3.3.2	Statistical region merging . . . . .	18
3.4	Summary of the MSIM algorithm . . . . .	18
3.5	Illumination correction results . . . . .	19
3.5.1	Experimental setup and results . . . . .	19
3.5.2	Visual comparison results . . . . .	20
3.5.3	Coefficient of variation results . . . . .	24
3.5.4	Segmentation results . . . . .	25
3.5.5	Classification results . . . . .	26
3.5.6	Comparison of multi-stage versus parametric illumination modeling . . . . .	28
3.6	Chapter summary . . . . .	30
<b>4</b>	<b>Normal Skin and Lesion Sparse Texture Model</b>	<b>32</b>
4.1	Sparse texture model . . . . .	33
4.2	Representative texture distributions . . . . .	33
4.3	Texture distinctiveness metric . . . . .	36
4.4	Implementation details . . . . .	38
4.4.1	Colour space . . . . .	38
4.4.2	Learning representative texture distributions . . . . .	38
4.5	Summary of the JPTD texture distinctiveness algorithm . . . . .	39
4.6	Texture distinctiveness map results . . . . .	40
4.7	Chapter summary . . . . .	43

<b>5</b>	<b>Texture-based Skin Lesion Segmentation</b>	<b>44</b>
5.1	Initial regions . . . . .	44
5.2	Distinctiveness-based segment classification . . . . .	46
5.3	Segmentation refinement . . . . .	47
5.4	Summary of the TDLS segmentation algorithm . . . . .	48
5.5	Experimental results . . . . .	48
5.5.1	Segmentation results . . . . .	49
5.5.2	Melanoma classification results using low level features . . . . .	53
5.5.3	Melanoma classification results using high-level intuitive features . .	54
5.6	Chapter summary . . . . .	55
<b>6</b>	<b>Conclusions</b>	<b>56</b>
6.1	Summary and conclusions . . . . .	56
6.2	Future work . . . . .	57
	<b>APPENDICES</b>	<b>59</b>
<b>A</b>	<b>Illumination Correction and Segmentation Results for Melanoma Images</b>	<b>60</b>
<b>B</b>	<b>Illumination Correction and Segmentation Results for Non-Melanoma Images</b>	<b>72</b>
	<b>References</b>	<b>83</b>

# List of Tables

3.1	Coefficient of Variation of Normal Skin Pixels for Uncorrected and Corrected Lesion Images . . . . .	25
3.2	Area Under ROC Curves (AUC) for Uncorrected and Corrected Skin Lesion Images . . . . .	26
3.3	Classification Results using Uncorrected and Corrected Lesion Images . . .	28
3.4	CV and AUC Results comparing Parametric and Multi-stage Modeling . .	29
5.1	Classification Results of Lesion and Normal Skin Pixels for All Images . . .	52
5.2	Classification Results of Lesion and Normal Skin Pixels for Melanoma Images	52
5.3	Classification Results of Lesion and Normal Skin Pixels for Non-melanoma Images . . . . .	52
5.4	Classification Results of Melanoma and Non-melanoma using Low-level Features for Corrected Images comparing Segmentation Algorithms . . . . .	53
5.5	Classification Results of Melanoma and Non-melanoma using High- and Low-level Features for Corrected Images comparing Segmentation Algorithms	54

# List of Figures

1.1	Examples of digital images of malignant melanoma . . . . .	2
1.2	Algorithm flow chart for automated skin lesion diagnosis . . . . .	4
3.1	Algorithm flow chart highlighting steps in the MSIM algorithm . . . . .	10
3.2	Methodology to calculate illumination map . . . . .	14
3.3	Sample posterior distribution $\hat{p}(i_{log} v_{log})$ . . . . .	15
3.4	Uncorrected and illumination-corrected skin lesion images . . . . .	21
3.5	Example of the limitations of the tested illumination correction algorithms	24
3.6	Receiver operator characteristic curves for corrected skin lesion images . .	27
3.7	Subset of digital images with significant artefacts . . . . .	29
4.1	Algorithm flow chart highlighting steps in the JPTD algorithm . . . . .	32
4.2	Texture vector . . . . .	34
4.3	Map of representative texture distributions . . . . .	35
4.4	Map of the texture distinctive metric . . . . .	37
4.5	Corrected skin lesion images and their corresponding textural distinctiveness maps . . . . .	41
5.1	Algorithm flow chart highlighting steps in the TDLS algorithm . . . . .	45
5.2	Segmentation of skin lesions using various algorithms. . . . .	50
A.1	Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm. . . . .	60

B.1	Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm. . . . .	72
-----	---	----

# Nomenclature

$\alpha(s_k s_0)$	Acceptance probability
$\mathbf{h}_s$	Local neighbourhood around pixel $s$
$\mathbf{t}_j^r$	Mean of texture distribution $T_j^r$
$\mathbf{t}_s$	Texture vector at pixel $s$
$\lambda$	Normalization term
$\mathcal{D}_R$	Regional textural distinctiveness for region $R$
$\mathcal{T}$	Set of texture vectors
$\mathcal{T}^r$	Set of representative texture distributions
$\Omega$	Set of representative samples
$\sigma$	Standard deviation or shape parameter for Gaussian function
$\Sigma_j$	Covariance of of texture distribution $T_j^r$
$\tau$	Otsu threshold
$\theta_j$	Set of parameters defining texture distribution $T_j^r$
$c$	Tukey biweight tuning constant
$D_j$	Textural distinctiveness metric associated with the $j^{th}$ texture distribution
$d_{j,k}$	Dissimilarity of the $j^{th}$ and $k^{th}$ texture distributions
$i(s)$	Illumination component at pixel location $s$



$l_{j,k}$	Similarity of the $j^{th}$ and $k^{th}$ texture distributions
$Q(s_k s_0)$	Uniform instrumental distribution
$r(s)$	Reflectance component at pixel location $s$
$S$	Set of skin pixels
$s$	Pixel location $(x, y)$
$T^r$	Representative texture distribution
$V(s)$	Value channel pixel intensity at location $s$
$y(s)$	Classification at location $s$

# Chapter 1

## Introduction

Melanoma is a cancer of the melanocytes, the cell found in the skin's epidermis that produces melanin [1]. Melanoma most commonly occurs on the trunk or lower extremities [1]. While malignant melanoma is less common than non-melanoma skin cancer [2], it is considered the most deadly form of skin cancer [1]. This is because melanoma accounts for approximately 75% of deaths associated with skin cancer [3]. In 2013, it is estimated that 76,690 people will be diagnosed with melanoma and 9,480 people will die of melanoma in the United States [4]. In Canada, 1 in 74 men and 1 in 90 women will develop melanoma in their lifetime [1].

The recent trends in melanoma incidence rates are more alarming. A study of the melanoma trends from 1992-2006 found that incidence rates for non-Hispanic white males and females were increasing at an annual rate of approximately 3% [5]. For young adults ages 15-30, melanoma is one of the most commonly diagnosed forms of cancer [6]. If melanoma is detected early, while it is classified at Stage I (less than 0.76mm thick), the 5-year survival rate is 96% [7]. However, the 5-year survival rate decreases to 5% if the melanoma is in Stage IV [7]. The cost of treatment of Stage IV melanoma is also 30 times the cost of treatment for Stage I melanoma [8]. With the rising incidence rates in certain subsets of the general population, early melanoma screening is beneficial.

### 1.1 Melanoma Screening

Skin lesions are typically visually screened for melanoma. Visual algorithms that doctors use as a guide to assess skin lesions include the ABCD [10] scale, the ABCDE scale [11] or



Figure 1.1: Examples of digital images of malignant melanoma acquired using a digital camera, from the Dermquest database [9]. In both images, the illumination variation changes horizontally.

the Glasgow 7-point checklist [12]. For example, the ABCD scale is an acronym for asymmetry, border irregularity, colour variegation, and diameter [10], and has been proposed to be expanded to include evolving [11]. Using this scale, the lesion is examined to see if it is asymmetrical, has an irregularly-shaped border, consists of multiple colours, and has a diameter greater than 6 mm. The skin lesion is also observed at multiple instances in time to see if it has changed, or evolved. After analyzing each feature, a total score is calculated to assess the risk that the lesion is malignant melanoma. Two examples of malignant melanoma are shown in Fig. 1.1. Fig. 1.1a has an irregular border and Fig. 1.1b has high colour variegation. Other visual scales consist of similar features.

Expert dermatologists may have access to a tool known as the dermatoscope to aid with diagnosis. A dermatoscope is a handheld device that optically magnifies, illuminates and enhances skin lesions, allowing the dermatologist to better view the lesion features. Use of the dermatoscope has been found to improve diagnosis, compared to the naked eye [13]. With a dermatoscope and following the ABCD scale, diagnosis of skin lesions as malignant melanoma has a reported sensitivity of 76.0-87.7% and specificity of 61.0-77.8% [14]. However, only 48% of dermatologists in the United States reported using a dermatoscope [15]. The main reasons against using the dermatoscope include a lack of training or interest [15].

## 1.2 Visual melanoma screening limitations

There are some limitations with relying on visual screening for melanoma. Analysis of inter-observer agreement by Argenziano et al. found that experienced dermatologists diagnosing melanoma using a dermatoscope had moderate agreement [14]. When inter-observer agreement of dermatologists on the assessment of individual ABCD features were analyzed, it was found that they only had fair agreement [14]. In terms of cost-effectiveness, despite increasing incidence rates, screening the general population for melanoma is not feasible. Losina et al. looked at the cost-effectiveness of one-time, annual and biennial screening of the general population with melanoma, compared to the treatment costs of early and late stage melanoma. They found that it is cost-effective to perform visual screenings of high-risk patients, such as a one-time screening of the general population over 50 and bi-annual screenings of siblings of patients with melanoma [16]. However, screenings of the general population are not cost-effective.

The limitations with visual melanoma screening can be overcome through the use of computer-aided diagnosis of melanoma. These computer algorithms take an image of the skin lesion as an input and extract a set of useful features. The features are used to identify the skin lesion as malignant melanoma. Initial feasibility studies of computer-aided diagnosis of melanoma concluded that such a system could be a low-cost diagnostic tool [17]. A study by Burroni et al. found that even using a simple computer vision analysis algorithm produced reliable diagnostic results [18].

Computer-aided melanoma diagnosis algorithms allow the general population to be screened for melanoma, especially in developing countries where it would not be possible otherwise. This new emerging field is called teledermatology [19]. Teledermatology involves the patient acquiring an image of the skin lesion using a cellular phone or personal digital assistants. The images are sent to a database where they are analyzed and diagnosed and the results are sent back to the patient. Initial teledermatology systems used remote teleconsultants to analyze the images of various skin diseases in order to assess feasibility of using such systems [20]. Recent research involves developing automated computer algorithms to analyze the image of the skin lesion and assess the risk of melanoma.

## 1.3 Automated melanoma screening systems

Early work on automated systems to assess the risk of melanoma used images acquired via a digital dermatoscope [21, 22, 23, 24]. These types of images are called dermoscopy

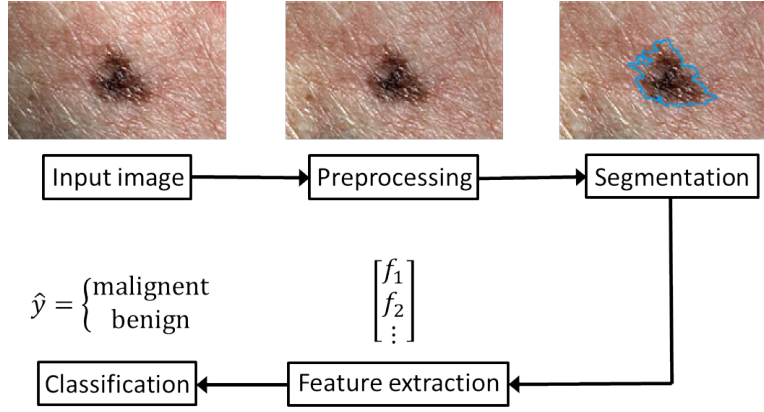


Figure 1.2: Algorithm flow chart for automated skin lesion diagnosis

images. Because a dermatoscope is used as a filter when acquiring an image of the lesion, the images have relatively low levels of noise and consistent background illumination and a pre-processing step is optional. Pre-processing for dermoscopy images includes enhancing image colours to allow the algorithm to more easily segment, or identify, the lesion from the background. However, the requirement of a dermatoscope does not allow these systems to be easily implemented as less than half of practicing dermatologists use dermatoscopes [15]. Furthermore, these automated systems require the purchase of a dermatoscope that can be attached to a mobile device to be used for teledermatology [19].

Recent automated melanoma screening algorithms allow images taken by a standard digital camera, such as those in Fig. 1.1, to be used [25, 26]. Since most of the population has access to a digital camera, these algorithms can be used by most dermatologists and the general population. A typical algorithm pipeline, as illustrated in Fig. 1.2, is to segment the lesion area, extract features from the lesion and classify the lesion in terms of the risk of melanoma.

Using a standard digital camera to acquire the image introduces some challenges. One problem is that illumination variation from skin surface reflectance can affect the algorithm. For example, healthy skin areas that are covered by shadows appear similar in colour as the lesion and as a result, might be identified as part of the lesion. This can be seen in Fig. 1.1, where both images illustrate horizontal illumination variation.

Another challenge is in the segmentation step, which is the step where the border of the lesion is identified. While images acquired using the dermatoscope have been magnified and enhanced to better identify the lesion, standard digital images taken by a digital camera do not have those advantages. Illumination variation, textural patterns of skin, and noise

can make separating normal skin and lesion difficult in digital images.

Based on these two challenges, there is a need for a robust segmentation algorithm that is designed specifically for images of lesions taken with a standard digital camera. This algorithm should incorporate correcting for illumination variation as a pre-processing step. This thesis proposes a novel skin lesion segmentation algorithm to solve these two challenges.

## 1.4 Thesis outline and contributions

The purpose of this thesis is to propose a novel skin lesion segmentation algorithm, which is part of an automated melanoma screening system. In this thesis, there are three main contributions:

1. A multi-stage illumination modeling algorithm to correct for illumination variation in a digital image of a skin lesion is designed, implemented and tested in Chapter 3,
2. An unsupervised texture learning algorithm which learns unique skin and lesion texture distributions from the digital image is presented in Chapter 4, and
3. A textural segmentation algorithm using the learned texture models to identify the lesion area is formulated and tested in Chapter 5.

The multi-stage illumination modeling algorithm uses two stages to better correct skin lesion images. The first stage is a non-parametric Monte Carlo estimate, which is followed by the fitting of a quadratic surface model. An unsupervised texture learning algorithm is proposed to learn representative texture distributions using a probabilistic framework and measuring the distinctiveness of each texture distribution. Finally, a novel textural segmentation algorithm specific to skin lesion image is proposed that uses the learned texture models to classify regions as normal skin or lesion. In addition, a review of relevant existing illumination correction and segmentation algorithms related to skin lesions is presented in Chapter 2. In Chapter 6, conclusions and future avenues to continue improving the proposed algorithm are presented.

# Chapter 2

## Background

In this chapter, existing literature relevant to solving the contributions identified in Sec. 1.4 are reviewed. These are divided into two parts. Illumination correction algorithms are discussed in Sec. 2.1 and skin lesion image segmentation algorithms are discussed in Sec. 2.2.

### 2.1 Existing illumination correction algorithms

Illumination correction algorithms adjust pixel intensities in an image based on an estimated illumination map. The goal of these algorithms is to remove any external illumination, so that the resulting image is independent of any illumination effects. Some of these effects that should be removed include shadows and bright areas caused by illumination variation. The motivation of this preprocessing step is to improve the performance of subsequent steps, including lesion segmentation and classification.

In Sec. 2.1.1, general illumination correction algorithms, which can be applied to any type of image, are discussed. In Sec. 2.1.2, correction algorithms that are designed specifically for skin lesion images are discussed.

#### 2.1.1 General illumination correction algorithms

Existing general illumination correction algorithms focus on correcting for illumination variation in standard digital images. These algorithms are general and can be applied to any image. The bright and dark pixels caused by illumination can be minimized by

adjusting the contrast of an image, using algorithms such as histogram equalization [27]. However, contrast adjustment is typically performed on a global scale, without taking into account local illumination variation. For example, histogram equalization adjusts the gray level distribution of pixels in an image without taking into account spatial location of the pixels.

The illumination-reflectance model assumes that the measured intensity of each pixel in an image is written as a multiplicative relationship between the underlying illumination and object reflectance [28, 29]. In general, illumination varies smoothly and does not change in intensity suddenly. Therefore, when assuming this model, algorithms first estimate the low-frequency illumination component and then find the reflectance component.

The first step, illumination estimation, has been widely researched. One of the earliest algorithms that use this model is called the Retinex algorithm [28]. It applies a set of Gaussian filters to the image to estimate the illumination component [28, 30]. Other algorithms estimate illumination using morphological operators [31], bilateral filters [32], Monte Carlo sampling [33] or total variation [34].

### **2.1.2 Illumination correction algorithms for digital images of skin lesions**

Unfortunately, existing general illumination correction algorithms have shortcomings when applied to digital images of skin lesions. This is caused by the dark appearance of skin lesions, which is mistaken for areas covered by shadows. General algorithms will over-brighten those areas [35]. In addition, most recent correction algorithms proposed specifically for skin lesion images can only be applied to dermoscopy images, which already have some filtering due to the dermatoscope. Rather than correcting illumination variation, the goals of these algorithms include colour calibration [36] and normalization [37] for improved lesion classification or contrast enhancement [38, 39] for improved lesion segmentation.

Recent work by Cavalcanti et al. [25, 35] proposes a correction algorithm specific for skin lesion images. The algorithm fits pixel intensities from the four corners of the photograph to a parametric surface. The disadvantage with this algorithm is that only a very small subset of pixels are used to fit the parametric surface. This results in the estimated illumination map being over- or underestimated.



## 2.2 Existing segmentation algorithms

The purpose of image segmentation algorithms is to find and outline distinct objects of importance in an image. For example, for images of skin lesions, the border of the skin lesion should be identified. Segmentation in general is a very well-researched area and many different algorithms have been proposed. Two types of segmentation algorithms relevant to this thesis are discussed in this section. First, texture segmentation algorithms can be applied to any image to find areas of importance based on their textural characteristics. Second, lesion segmentation algorithms have been designed specifically to be applied to images of skin lesions.

### 2.2.1 Texture-based segmentation algorithms

Textures include smoothness, roughness, or the presence of ridges, bumps or other deformations and are visible by variation in pixel intensities in an area [40]. Texture analysis extracts features and measurements of a texture, allowing textures from different regions to be compared. Texture analysis is useful for image segmentation because different parts of the same object will usually match in texture.

Types of texture segmentation algorithms include statistical, structural, and model-based [41]. An in-depth review of the popular texture segmentation algorithms in general has been compiled by Zhang and Tan [41]. Statistical texture segmentation algorithms describe a texture in a region as a set of statistics. Algorithms include using first level statistics, gray-level co-occurrence matrix or Haralick statistics. Model-based algorithms use probability models, such as the autoregressive model or Markov random field model, to characterize textures. Structural algorithms deconstruct and characterize the texture as a number of texture elements.

### 2.2.2 Lesion segmentation algorithms

Many algorithms have been proposed to automatically segment skin lesion images. However, like state-of-art illumination correction algorithms described in Sec. 2.1.2, the majority of proposed segmentation algorithms are only applicable to dermoscopy images, where there is better contrast between certain types of lesions and surrounding skin area [42]. A recent summary of the popular existing segmentation algorithms for dermoscopy images is compiled by Celebi et al. [42]. Algorithms compared in the summary [42] include using

simple thresholding, active contours [43], and region merging [44]. The majority of algorithms in [42] only use features derived from pixel colour to drive the segmentation. This includes the blue channel from the RGB colour space, the luminance channel from the CIE 1976  $L^*u^*v^*$  (CIELUV) or CIE 1976  $L^*a^*b^*$  (CIELAB) colour spaces, or an orthogonal transform applied to the colour channels. However, it is difficult to accurately segment lesions with fuzzy edges when relying solely on colour features. These algorithms are also have not been designed to work for digital images of skin lesions.

Some texture-based segmentation algorithms described in Sec. 2.2.1 have been applied to dermoscopy images. Stoecker et al. [45] analyzed texture in skin images using basic statistical approaches, such as the gray-level co-occurrence matrix. They found that texture analysis could accurately find regions with a smooth texture and that texture analysis is applicable to segmentation and classification of dermoscopy images. Proposed textural lesion segmentation algorithms include using gray-level co-occurrence matrix [40, 46], first-order region statistics [47], and Markov random field models [48]. Xu et al. [49] learns a model of the normal skin texture using pixels in the four corners of the image, which is later used to find the lesion.

Segmenting skin lesion images is more difficult because of the presence of shadows and bright areas caused by illumination variation. Hance et al. [50] explored a few different algorithms, including thresholding, active contours and split-and-merge, and modified them to be specific to lesion images. For example, the thresholding algorithm has to be modified to account for bright areas where there is reflection of the camera’s flash. Four separate algorithms by Cavalcanti et al. include a preprocessing step which corrects for illumination variation before applying a thresholding [25, 35, 51] or level-set segmentation algorithm [52]. Thresholding is performed on single colour channels [51], multiple colour channels [35] and a set of channels derived using principal component analysis and other processing steps [25]. Including the pre-processing step allowed the lesion segmentation algorithms proposed by Cavalcanti et al. to perform well. Therefore, the algorithm proposed in this thesis adopts the approach of applying a pre-processing step followed by the segmentation algorithm.

## Chapter 3

# Illumination Correction of Skin Lesion Images

In this chapter, the illumination correction algorithm which functions as a pre-processing step is outlined. A multi-stage illumination modeling (MSIM) algorithm is proposed to normalize illumination variation in an image of a skin lesion [53, 54, 55]. The proposed algorithm introduces multiple stages to correct illumination variation allowing it to be robust towards complicated illumination patterns and image artifacts, such as hair. A flow chart showing the stages of the MSIM algorithm is shown in Fig. 3.1. The proposed algorithm is applied to digital images of skin lesion, that is images of lesions that have been taken with a digital camera.

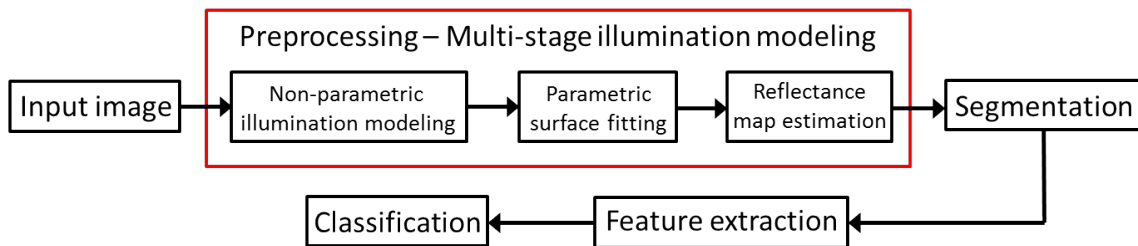


Figure 3.1: Algorithm flow chart highlighting steps in the MSIM algorithm

### 3.1 Problem formulation

In this section, the underlying illumination model for digital images of a skin lesion is first derived. The following assumptions are made about the digital images in order to derive the model. The image is assumed to be taken in a controlled environment, such as a doctor's office, with a standard digital camera. The region being photographed is assumed to be illuminated with a single source of white light. Finally, the skin lesion of interest is assumed to be in the centre of the image. Based on these assumptions, only the V (value) channel from the HSV (hue-saturation-value) colour space [56] is modified in the preprocessing step. The hue and saturation channels are not affected by illumination variation because those channels are responsible for color in an image and it is assumed that the light source is white light. These are similar assumptions as those made by Cavalcanti et al. [35].

To correct illumination variation in an image, the multiplicative illumination-reflectance model is commonly assumed [28]. Using this model, the measured V channel pixel intensity is the entrywise multiplication of the illumination component and reflectance component. This is shown in Eqn. 3.1, where  $s$  is a pixel location  $(x, y)$ ,  $v(s)$  is the V channel pixel intensity at  $s$ ,  $i(s)$  is the illumination component and  $r(s)$  is the reflectance component. The goal of the illumination correction algorithm is to first estimate the underlying illumination component and then calculate the reflectance component.

$$v(s) = i(s) \cdot r(s) \quad (3.1)$$

The logarithm transform of both sides in Eqn. 3.2 is taken so that the multiplicative model becomes additive (3.2). For brevity,  $v_{log}$ ,  $i_{log}$ , and  $r_{log}$  are still associated with a pixel of interest located at  $s$ .

$$\begin{aligned} \log(v(s)) &= \log(i(s)) + \log(r(s)) \\ v_{log} &= i_{log} + r_{log} \end{aligned} \quad (3.2)$$

Estimating  $i_{log}$  can be seen as an inverse problem. It is formulated as Bayesian least squares, where  $P(i_{log}|v_{log})$  is the posterior (Eqn. 3.3) [57].

$$\begin{aligned}
\hat{i}_{log} &= \arg \min_{i_{log}} \{E((i_{log} - \hat{i}_{log})^2 | v_{log})\} \\
&= \arg \min_{i_{log}} \left( \int (i_{log} - \hat{i}_{log})^2 P(i_{log} | v_{log}) di_{log} \right)
\end{aligned} \tag{3.3}$$

To minimize the estimate of  $i_{log}$ , take the derivative of the arg min argument from Eqn. 3.3,

$$\begin{aligned}
&\frac{\partial}{\partial \hat{i}_{log}} \int (i_{log} - \hat{i}_{log})^2 P(i_{log} | v_{log}) di_{log} \\
&= \int 2(i_{log} - \hat{i}_{log}) P(i_{log} | v_{log}) di_{log},
\end{aligned} \tag{3.4}$$

Set to zero,

$$\int i_{log} P(i_{log} | v_{log}) di_{log} = \int \hat{i}_{log} P(i_{log} | v_{log}) di_{log} \tag{3.5}$$

and simplify the right hand side,

$$\begin{aligned}
\int \hat{i}_{log} P(i_{log} | v_{log}) di_{log} &= \hat{i}_{log} \int P(i_{log} | v_{log}) di_{log} \\
&= \hat{i}_{log},
\end{aligned} \tag{3.6}$$

leading to,

$$\begin{aligned}
\hat{i}_{log} &= \int i_{log} P(i_{log} | v_{log}) di_{log} \\
\hat{i}_{log} &= E(i_{log} | v_{log}).
\end{aligned} \tag{3.7}$$

From Eqn. 3.7, the optimal estimate of the illumination component of an image is the conditional mean of  $i_{log}$  given  $v_{log}$ . Since estimating a conditional mean is complex and intensive, a Monte Carlo posterior estimation algorithm is used to estimate  $P(i_{log} | v_{log})$ . Furthermore, Monte Carlo sampling is a non-parametric approach and a parametric model for the posterior distribution  $P(i_{log} | v_{log})$  is not assumed.

### 3.2 Initial non-parametric illumination modeling

In this section, the approach taken for the importance-weighted Monte Carlo sampling [58] for illumination estimation is outlined. At the end of this stage, a grayscale map is generated, where each pixel in the map gives the initial estimate of illumination at that location. Consistent with the previous section, let  $s$  be a pixel location in the digital image. The goal of this implementation of Monte Carlo sampling is to determine a set of representative samples  $\Omega$  and their associated importance weights from the pixels  $s_k$  for a search space surrounding the pixel of interest  $s_0$ .

A subset of pixels  $s_k$  in the search space is selected randomly using a uniform instrumental distribution  $Q(s_k|s_0)$ . A uniform distribution is used to let all pixels in the search space have an equal probability of being selected. An acceptance probability  $\alpha(s_k|s_0)$  is calculated for each selected pixel by comparing the neighbourhoods around the selected pixel  $s_k$  and pixel of interest  $s_0$ , respectively. The local neighbourhoods around  $s_k$  and  $s_0$  are represented by the vectors  $\mathbf{h}_k$  and  $\mathbf{h}_0$  respectively. The acceptance probability determines if  $s_k$  is a realization of  $P(i_{log}|v_{log})$ . A Gaussian error statistic is used in this implementation, as shown in Eqn. 3.8 where  $h_k[j]$  and  $h_0[j]$  are the  $j^{th}$  sites in the neighbourhoods around  $s_k$  and  $s_0$ , respectively.

$$\alpha(s_k|s_0) = \prod_j \frac{\frac{1}{2\pi\sigma} \exp \left[ -\frac{(h_k[j] - h_0[j])^2}{2\sigma^2} \right]}{\lambda} \quad (3.8)$$

A normalization term  $\lambda$  is included in the denominator, such that  $\alpha(s_k|s_0)$  is 1 if the neighbourhoods around  $s_k$  and  $s_0$  are identical. The acceptance probability is the product of the probabilities from each site  $j$  because all elements in the neighbourhoods are assumed to be independent. The parameter  $\sigma$  is based on local variance of the differences of the Value channel intensities and controls the shape and extent of the Gaussian function. The value of  $\alpha(s_k|s_0)$  gives the probability that the pixel  $s_k$  is accepted into the set  $\Omega$  and is used to estimate the posterior distribution  $P(i_{log}|v_{log})$ . Furthermore,  $\alpha(s_k|s_0)$  is also the associated importance weight for the accepted pixel. A random value  $u$  is drawn from a uniform distribution  $U(0, 1)$  to determine whether to accept pixel  $s_k$ . If  $u \leq \alpha(s_k|s_0)$ , then  $s_k$  is accepted into  $\Omega$ . Otherwise,  $s_k$  is discarded.

The selection and acceptance process is repeated until the desired number of samples has been selected from the search space. Then, the posterior distribution is calculated as a weighted histogram. An example of the posterior distribution as a weighted histogram is shown in Fig. 3.3, where each element of the stacked bar chart corresponds to a pixel

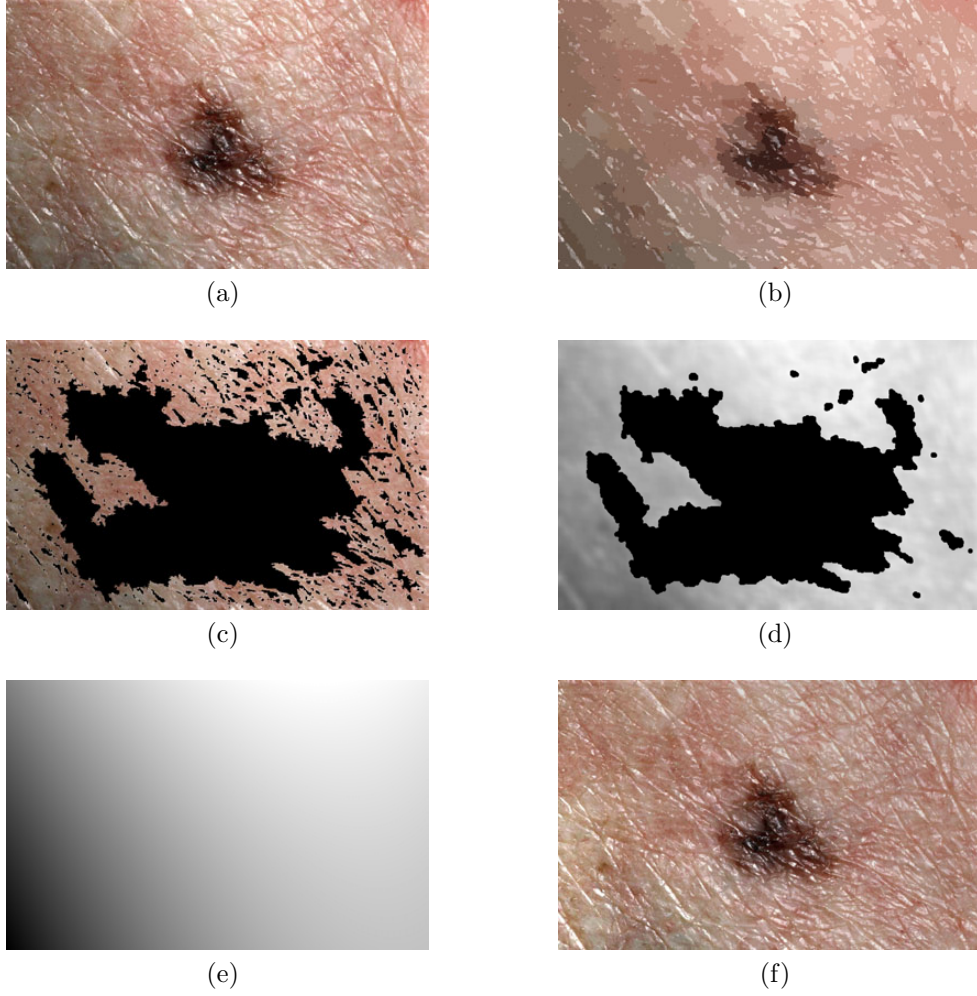


Figure 3.2: Methodology to calculate illumination map: (a) original digital image of a skin lesion with illumination variation; (b) segmentation map; (c) regions included in the subset of skin pixels, where pixels black in colour are not classified as normal skin; (d) initial illumination map estimated via non-parametric modeling using Monte Carlo sampling; (e) final illumination map determined by using (d) as a prior to the parametric surface model; (f) image corrected for illumination variation using the MSIM algorithm.

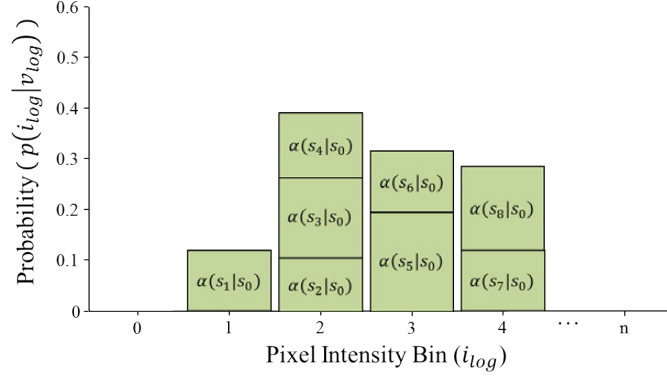


Figure 3.3: Sample posterior distribution  $\hat{p}(i_{log}|v_{log})$ , built from pixels accepted in the set  $\Omega$ . Each stacked element corresponds to a pixel  $s_k$  in  $\Omega$ , where the height is  $\alpha(s_k|s_0)$  and bin location is  $s_k$ . The histogram is normalized so that  $\sum_k \hat{p}(i_{log_k}|v_{log_k}) = 1$ .

$s_k$  from the set  $\Omega$ . The height of the element is equal to the importance weight  $\alpha(s_k|s_0)$  and each element is added to the histogram bin that corresponds to the value of the pixel intensity  $s_k$ . In the example histogram in Fig. 3.3, the set  $\Omega$  contains eight pixels ( $s_1$  to  $s_8$ ) and the range of the pixel intensities depends on the dynamic range of the input image. The histogram is normalized, so that  $\sum_k \hat{p}(i_{log_k}|v_{log_k}) = 1$ . Using the estimated posterior distribution, the conditional mean from Eqn. 3.7 can be calculated to find the estimate of the illumination component in log-space at pixel  $s_0$ . The initial illumination estimate  $\hat{i}$  is found by taking the exponential of  $\hat{i}_{log}$ . During testing, a  $21 \times 21$  search space and  $7 \times 7$  neighbourhood was used.

### 3.2.1 Final parametric modeling of illumination

The initial non-parametric modeling stage results in an illumination map  $\hat{i}$  where the lesion appears as shading. This is because the lesion is darker than its surroundings, which leads to an inaccurate estimate of illumination within the lesion area. If the initial illumination map is used to correct for illumination variation, the lesion would be greatly brightened and the contrast between the lesion and surrounding skin would be reduced. Instead, parametric modeling of illumination via a quadratic surface model is performed using the initial non-parametric illumination estimate as a prior to obtain the final illumination map. This parametric model uses assumptions made about the lighting conditions to learn a surface model based on the initial illumination estimates corresponding to the healthy skin pixels and find the correct illumination map for the lesion area.



Only normal skin pixels are used to learn the parametric illumination model. Therefore, the first step is to obtain a rough estimate of the normal skin pixels and lesion pixels. A segmentation algorithm is applied to the original digital image, which divides the image into a large number of regions. The segmentation algorithm used in the implementation is Statistical Region Merging and is described in Section 3.3.2. Because the lesion is assumed to be located in the center of the image, any regions that touch the border are classified as normal skin. This heuristic provides an estimate of which pixels are skin. Fig. 3.2b shows an example of the segmentation results. Fig. 3.2c shows the areas classified as skin using the heuristic. As seen in Fig. 3.2c, the segmentation algorithm does not accurately identify all the normal skin regions. However, it is not necessary to have a perfect segmentation algorithm because these pixels are only used to fit the surface model.

The second step is to learn a quadratic surface model, expressed in Eqn. 3.9, which is fit based on the initial non-parametric estimates corresponding to healthy skin pixels [35]. A surface model is used because the final illumination is assumed to be smooth. Fitting a smooth surface model enforces that assumption. Another assumption is that the region is illuminated by an overhead white light source when being photographed. This means that the illumination variation is caused by the curvature of the skin surface in the image. For example, illumination variation in an image of a lesion on the arm is caused by the curvature of the arm. Therefore, we expect a quadratic surface to best fit the illumination variation without being too complex because a quadratic surface can approximate most skin surfaces.

$$i'(x, y) = P_1x^2 + P_2xy + P_3y^2 + P_4x + P_5y + P_6 \quad (3.9)$$

Using this model, the parameters of the final illumination map can be estimated using a maximum likelihood estimation based on the initial estimate  $\hat{i}$  and the set of skin pixels  $S$ . Outliers may exist in the illumination map because of noise or the reflection of light directly into the camera. To minimize their effects, a robust fitting algorithm using the Tukey biweight is implemented, shown in Eqn. 3.10 where  $c$  is the tuning constant [59].

$$\hat{i}' = \arg \min_{\hat{i}'} \sum_{s \in S} \rho(\hat{i}(s) - \hat{i}'(s)) \quad (3.10)$$

$$\text{where } \rho(z) = \begin{cases} \frac{c^2}{6} (1 - [1 - (z/c)^2]^3), & \text{for } |z| \leq c \\ \frac{c^2}{6}, & \text{for } |z| > c \end{cases}$$

The value for  $c$  was set to 4.6851, which is the standard Tukey biweight tuning constant [59]. Fig. 3.2d shows the pixel intensities from the initial illumination map that are used to fit the surface and Fig. 3.2e shows the resulting surface after being fit.

### 3.2.2 Reflectance map estimation

The illumination-reflectance model assumes that the V channel is the entrywise multiplication of the illumination and reflectance components. Therefore, the reflectance map can be estimated by dividing the original V channel pixel intensity  $v(s)$  by the final estimated illumination map  $\hat{i}'(s)$ . The reflectance channel is combined with the original hue and saturation channels to correct the digital image for illumination variation. Fig. 3.2f shows an example of a corrected image. In Fig. 3.2f, the illumination variation seen in the uncorrected image (Fig. 3.2a) has been corrected.

## 3.3 Implementation

In this section, the implementation details of the MSIM algorithm are described, including additional steps added for optimization and the algorithm used to estimate normal skin regions.

### 3.3.1 Optimization

The Monte Carlo algorithm is computationally complex and intensive, so additional steps are included to increase the computation speed of the MSIM algorithm. The MSIM algorithm is implemented in MATLAB and run on a computer with an Intel Core i5-2400S CPU (2.5 GHz, 6 GB RAM). Without any optimization, the MSIM algorithm takes on average 118.61s to correct illumination variation of a  $1640 \times 1043$  image. After the modifications outlined in this section, it takes, on average, 59.57s to correct the same image. Because the illumination is assumed to be fairly smooth, the increase in computation speed does not result in a loss of accuracy.

Two steps are added prior to the first stage of illumination estimation. First, the digital image is downsampled by a factor of two in both the horizontal and vertical directions, and the Monte Carlo illumination estimation is applied to the smaller image. Since the image is acquired in a controlled environment with overhead lighting, the illumination is assumed to

be low frequency. This means that the loss of details due to downscaling does not impact the final illumination estimation. The initial nonparametric illumination estimation is rescaled to the original image’s size before obtaining the parametric illumination estimation.

The second optimization step is to estimate the pixels that belong to healthy skin and only apply the Monte Carlo sampling algorithm to those pixels. In the final parametric illumination map, only the initial illumination estimates associated with healthy skin pixels are considered in the robust fitting algorithm. Therefore, there is no need to calculate the initial estimate for lesion pixels. The method to find the healthy skin pixels is described in the next section. A morphological closing operator with a small disc structuring element is applied to the skin pixel map to remove the majority of gaps and holes.

### 3.3.2 Statistical region merging

To find an estimate of regions classified as normal skin, a segmentation algorithm is applied to the original colour digital image. The segmentation algorithm implemented is the statistical region merging (SRM), which tends to over-segment an image [60]. SRM groups nearby pixels that are similar in colour together into segments. Any segments that touch the border are classified as normal skin and are added to the set  $S$  of normal skin pixels. Since the lesion is assumed to be found in the centre of the image, only segments corresponding to normal skin are found near the borders of the image.

SRM uses two important steps to segment a colour image. First, the image is set up as a 4-connected graph. This motivates the sorting step, where pairs of pixels are organized based on the maximum difference in the RGB values. Pixels that are closer in colour are merged first. As pixels are merged, they are treated as regions. Second, a merging predicate is introduced, which will merge two regions based on similarity between the average RGB intensities of the two regions of interest. The SRM algorithm contains a single tunable parameter,  $Q$ , which was set to 256 based on testing.

## 3.4 Summary of the MSIM algorithm

1. Apply Statistical Region Merging to the original RGB digital image (Section 3.3.2) to classify normal skin pixels.
2. Downscale V channel of the original HSV image by factor 2 (Section 3.3.1).

3. Iterating through each pixel of interest  $s_0$  that is part of the set of normal skin pixels  $S$  in the downscaled image, randomly draw a sample  $s_k$  from a search space around  $s_0$  using a uniform instrumental distribution.
4. Compute acceptance probability  $\alpha(s_k|s_0)$  using Eqn. 3.8.
5. Generate a random value  $u$  from a uniform distribution  $U(0, 1)$ . If  $u \leq \alpha(m_k|m_0)$ , include  $s_k$  in the set of accepted samples  $\Omega$  with an importance weight equal to  $\alpha(s_k|s_0)$ . Otherwise,  $s_k$  is discarded.
6. Repeat steps 3-5 until the desired number of samples  $N$  are selected from the search space.
7. Estimate the posterior distribution  $\hat{p}(i_{log}|v_{log})$  based on the samples in  $\Omega$  and their importance weights.
8. Compute the initial illumination map estimate in log-space  $\hat{i}_{log}$  for pixel  $s_0$  as the conditional mean in Eqn. 3.6. Rescale the initial illumination map to the dimensions of the original image.
9. Using the initial illumination map corresponding to healthy skin pixels, calculate the parametric surface using a robust fitting algorithm to obtain the final estimated illumination map.
10. Estimate the reflectance map (Sec. 3.2.2) and replace the original V channel with the reflectance map to obtain the corrected image.

## 3.5 Illumination correction results

In this section, the experimental setup to compare the MSIM algorithm to other illumination correction algorithms is described. The algorithms are compared visually and quantitatively using the coefficient of variation, segmentation accuracy and classification accuracy.

### 3.5.1 Experimental setup and results

Three state-of-art illumination correction algorithms are used for comparison. The algorithms are the Frankle-McCann Retinex algorithm (as implemented by Funt et al. [61]),

morphological illumination estimation [31], and the Cavalcanti et al. algorithm [35]. The Cavalcanti et al. algorithm is the only compared algorithm that is designed specifically for digital images of skin lesions. It uses the V channel values from pixels in  $20 \times 20$  patches in the four corners of the image to fit a parametric surface.

To allow for a fair comparison, the corrected images are normalized by adjusting the average V channel values of the healthy skin class to be equal. This normalization process is done by computing the average V channel intensity for healthy skin pixels in the uncorrected image and corrected image. Then, all pixels in the corrected image are multiplied by the ratio of the average healthy skin V channel intensity in the uncorrected and corrected image. Normalization allows a fair comparison as the images are adjusted to have a similar dynamic range of skin pixel intensities. A manual ground truth segmentation is used to obtain the true classification of the pixels.

A set of 190 digital images from the Dermatology Information System (DermIS) database [62] and the DermQuest database [9] are used for testing. Sixty-four images tested are from the DermIS database and 126 are from the DermQuest database. 108 images tested had melanomas, while 82 images had other types of skin lesions. Some examples of the images tested are shown in the first column in Fig. 3.4. The DermQuest images are also used to test the impact of the illumination correction algorithm on classification of melanoma and non-melanoma. The entire set of DermQuest images used to test the illumination correction and segmentation algorithms is provided in Appendices A and B.

### 3.5.2 Visual comparison results

Images in Fig. 3.4 show the uncorrected and corrected images after applying each algorithm. The Retinex and morphological algorithms can noticeably alter the lesion colour. For example, in Fig. 3.4a, 3.4b, 3.4d and 3.4f, areas of the lesion are visibly brighter in the morphological corrected image than in the original image. This is because the Retinex and morphological algorithms do not include a parametric modeling step. The lesion is mistaken as a shadow and the illumination correction algorithm brightens that area. This reduces the contrast between normal skin and lesion, which can affect the results of the segmentation step. Furthermore, the colour of the lesion is affected, which can diminish the ability to classify the images as benign melanoma. In Fig. 3.4c, 3.4g, and 3.4i, the Retinex algorithm is unable to correct illumination variation on the left side of the image.

In cases where there is complex skin texture, such as Fig. 3.4d and 3.4h, the Cavalcanti et al. algorithm has difficulties removing the illumination variation, while it is removed by the MSIM algorithm. In Fig. 3.4d, the top left corner is not corrected by the Cavalcanti

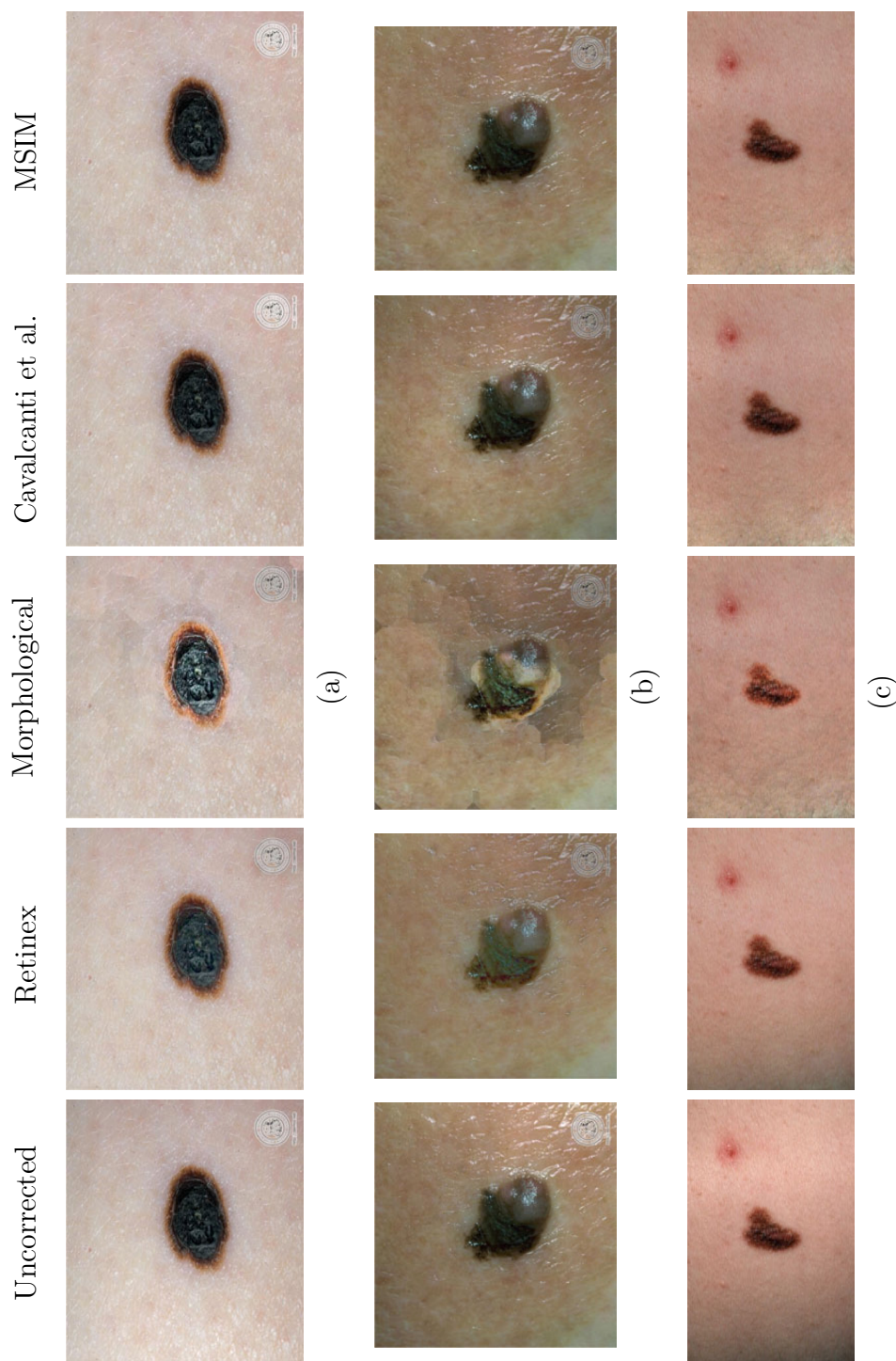


Figure 3.4: Uncorrected and corrected skin lesion images. In images (a), (b), (e) and (g), the Retinex and Morphological algorithms cause a change in the lesion colour and a decrease in contrast between lesion and skin. In images (d) and (h), the Cavalcanti et al. algorithm does not adequately correct for illumination variation, while the MSIM algorithm does.



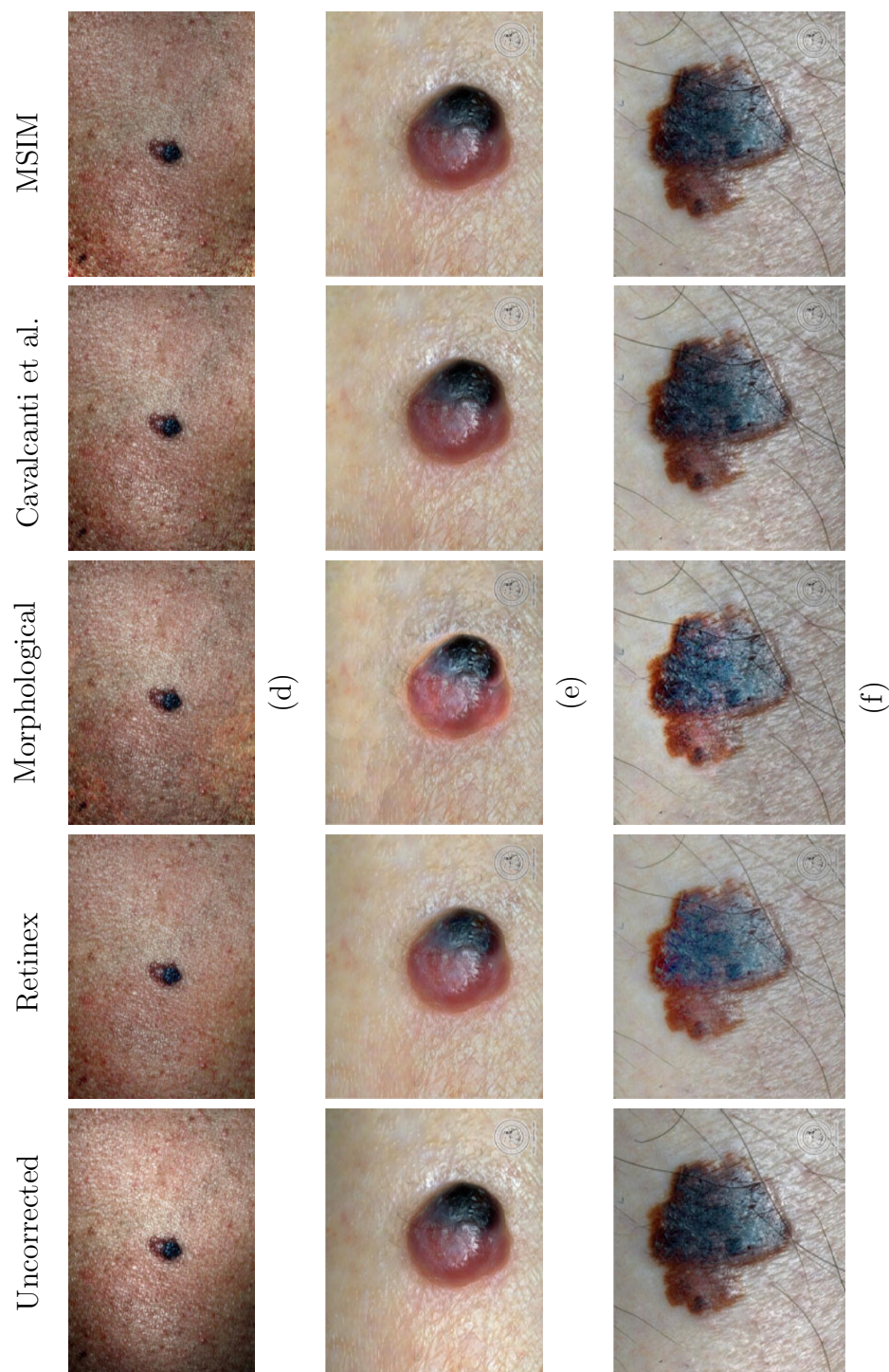


Figure 3.4: Uncorrected and corrected skin lesion images, cont.

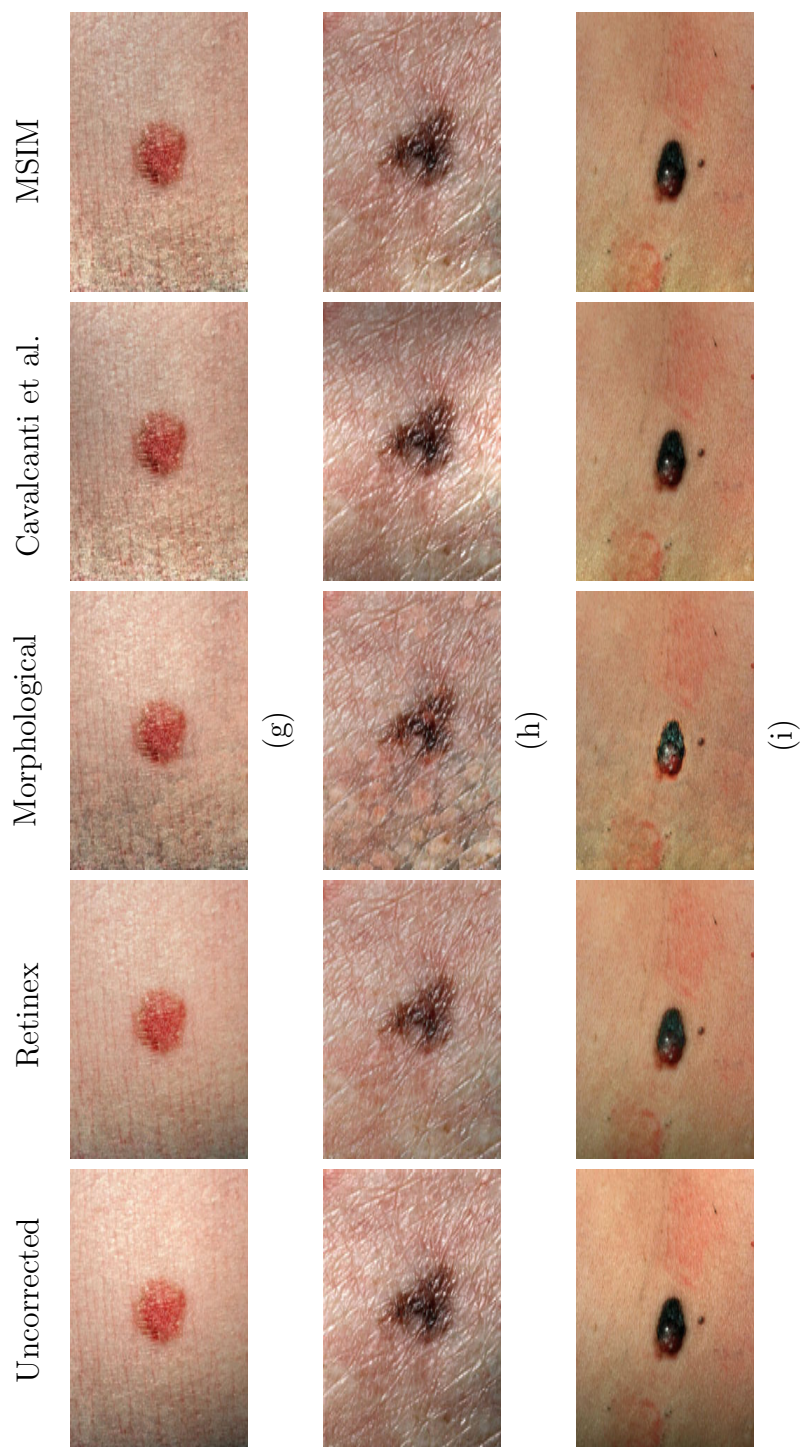


Figure 3.4: Uncorrected and corrected skin lesion images, cont.





Figure 3.5: Example of the limitations of the tested illumination correction algorithms. None of the tested illumination correction algorithms are able to model and correct the light reflected off of the center of skin lesion which oversaturates the image in that area. Note that the proposed MSIM algorithm is able to correct some of the illumination variation excluding the aforementioned oversaturated region.

et al. algorithm but is corrected by the MSIM algorithm. Similarly, this occurs in the left side of Fig. 3.4h. This is because the MSIM algorithm uses a greater number of pixels to fit the parametric model, compared to Cavalcanti et al. Using more pixels results in a more accurate illumination model.

Fig. 3.5 shows a digital image with complicated, highly oversaturated illumination variation that is not corrected satisfactorily by all of the tested methods. This is caused by the reflection of light off the center of the lesion, which saturates the image in that area and makes it challenging for all tested methods to model illumination in that region. The proposed MSIM algorithm is able to correct some of the illumination variation in the image, excluding the oversaturated region.

### 3.5.3 Coefficient of variation results

Coefficient of variation is used to measure the variance of background pixel intensities before and after correcting for illumination variance. The set of background pixels is assumed to be all healthy skin pixels. This metric has been used to quantify the performance of other illumination correction algorithms [63]. It compares the mean skin pixel intensity to the variance of the skin pixel intensities and is given in Eqn. 3.12 [63]. Coefficient of variation is useful to compare a image before and after correcting for illumination. A lower coefficient of variation after correction is desirable, as the only variation in pixel intensities should be from texture characteristics. The pixel intensities of healthy skin in the images were normalized so that the dominant cause of a change in coefficient of variation is due to a change in variance.

Table 3.1: Coefficient of Variation [63] of Normal Skin Pixels for Uncorrected and Corrected Lesion Images

Type	Uncorrected	Retinex	Morphological	Cavalcanti et al.	MSIM
Average <i>cv</i>	0.1473	0.0900	0.0888	0.1032	<b>0.0851</b>
Average melanoma <i>cv</i>	0.1474	0.0902	<b>0.0842</b>	0.1016	0.0849
Average non-melanoma <i>cv</i>	0.1472	0.0896	0.0951	0.1053	<b>0.0853</b>

$$cv = \frac{\sigma}{\mu} \quad (3.12)$$

Table 3.1 shows the average coefficient of variation using the entire dataset and after partitioning the dataset into ‘Melanoma’ and ‘Non-melanoma’ categories. Using the entire dataset or only the ‘Non-melanoma’ images, the average coefficient of variation for images corrected using the MSIM algorithm is the lowest. Looking at the ‘Melanoma’ images, the morphological algorithm produces images with the lowest average coefficient of variation, but the MSIM algorithm has a close average coefficient of variation.

### 3.5.4 Segmentation results

The uncorrected and corrected images were compared based on the ability to segment lesion and normal skin pixels. While a segmentation algorithm for skin lesion images is proposed and tested in this thesis, we were interested in analyzing the affects of illumination correction using an existing segmentation algorithm. A segmentation algorithm described by Cavalcanti et al. was used [35]. Separate thresholds are found for the R, G, and B colour channels to separate normal skin and lesion pixel intensities. If a pixel’s intensity is below those thresholds in at least two channels, then that pixel is classified as lesion. Finding the optimal thresholds for each image greatly depends on the dynamic range and contrast of the images. While Cavalcanti et al. [35] proposes finding Otsu’s threshold [64] for each channel, we perform a more comprehensive test by using a subset of all possible values for the thresholds and plotting the receiver operator characteristic (ROC) curves.

ROC curves plot the true positive rate (*TPR*) versus false positive rate (*FPR*) found using different thresholds [65]. The *TPR* is the number of true positives divided by the

Table 3.2: Area Under ROC Curves (AUC) for Uncorrected and Corrected Skin Lesion Images

Type	Uncorrected	Retinex	Morphological	Cavalcanti et al.	MSIM
Average <i>AUC</i>	0.9225	0.9435	0.9114	0.9788	<b>0.9830</b>
Average melanoma <i>AUC</i>	0.9368	0.9581	0.9278	0.9850	<b>0.9871</b>
Average non-melanoma <i>AUC</i>	0.9058	0.9239	0.8934	0.9655	<b>0.9750</b>

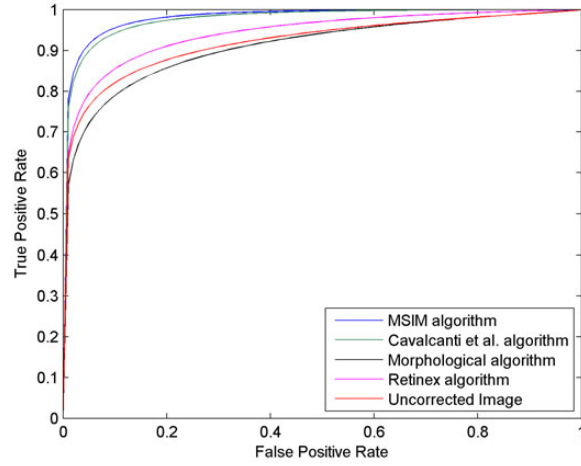
number of true and false positives and the  $FPR$  is the number of false positives divided by the sum of false positives and true negatives. An ideal ROC curve passes through the point where  $TPR = 1$  and  $FPR = 0$ , meaning that there is perfect segmentation for some threshold value. Otherwise, curves that are closer to the top left of the graph correspond to a more accurate segmentation. The ROC curves are shown in Fig. 3.6.

The area under the curve ( $AUC$ ) is calculated as a metric for the ROC curve [66]. A higher  $AUC$  corresponds to more accurate segmentation results. The  $AUC$ s for the full dataset and ‘Melanoma’ and ‘Non-Melanoma’ subsets are shown in Table 3.2.

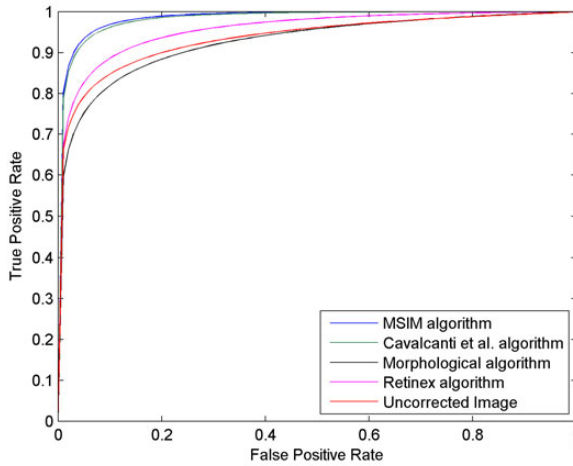
From Fig. 3.6 and Table 3.2, the MSIM algorithm performs the best for image segmentation, followed by the Cavalcanti et al. algorithm. This is because the MSIM algorithm uses a larger set of normal skin pixels to fit the parametric surface compared to the Cavalcanti et al. algorithm. The other two algorithms do not fit any surface for illumination correction. Therefore, the lesion in the images is more difficult to separate using segmentation algorithms.

### 3.5.5 Classification results

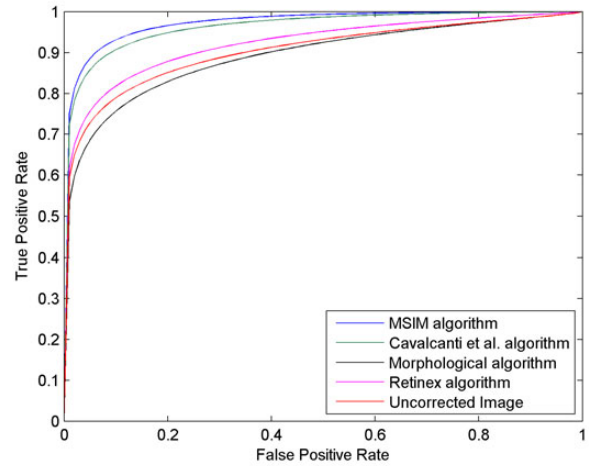
Using images corrected by each illumination correction algorithm, a separate classifier was trained to segment the image, extract features based on the ABCD scale of dermatoscopy [10], and assess the lesion for risk of melanoma. The algorithm and features used are implemented by Cavalcanti et al. [25]. Of the set of 52 features, 11 characterize asymmetry, 12 characterize border irregularity, 25 characterize colour and 4 characterize differential



(a) All



(b) Melanoma



(c) Non-melanoma

Figure 3.6: Receiver operator characteristic (ROC) curves corresponding to segmentation results using a set of different thresholds for melanoma and non-melanoma skin lesion images. Graph (a) uses the result from the entire dataset, while graphs (b) and (c) use only the melanoma or non-melanoma images, respectively.

Table 3.3: Classification Results using Uncorrected and Corrected Lesion Images

Correction Algorithm	Sensitivity	Specificity	Accuracy
Uncorrected	64.3%	65.7%	65.1%
Retinex	69.7%	75.0%	72.2%
Morphological	74.5%	70.7%	72.2%
Cavalcanti et al.	74.2%	77.9%	76.0%
MSIM	<b>82.8%</b>	<b>82.4%</b>	<b>82.5%</b>

structure.

A support vector machine (SVM) classifier was used for training and testing. The SVM algorithm was implemented using LIBSVM [67], with a linear kernel. The optimal soft-margin cost and benign class weight parameters were found using a grid-search algorithm to maximize the harmonic mean of the precision and recall metrics. The feature scores were normalized to have 0-mean and unit variance. To compare results from the different correction methods, sensitivity, specificity, and accuracy are used as metrics using a leave-one-out cross-validation scheme. Only the 126 images from the DermQuest database were used for classification, because they had appropriate image quality and resolution.

Table 3.3 shows the classification sensitivity, specificity and accuracy. The MSIM algorithm has the highest sensitivity, while MSIM has the highest specificity and accuracy. This is partly due to the illumination correction algorithm increasing the separability of the lesion and normal skin, allowing for more accurate segmentation. Eleven of the features used depend on the morphology of the segmentation border. Furthermore, the illumination correction algorithm can increase class separability of features based on pixel intensities.

### 3.5.6 Comparison of multi-stage versus parametric illumination modeling

An investigation was conducted to determine the effect that the non-parametric modeling stage has in the proposed MSIM algorithm. The investigation involves comparing the results obtained: a) using solely the parametric modeling stage, and b) using both the non-parametric and parametric modeling stages. It is found that a key benefit to using an initial non-parametric modeling stage is that this type of modeling strategy is more robust to the presence of outlier detail in the image compared to a purely parametric modeling strategy. This robustness to outlier detail is particularly important in images

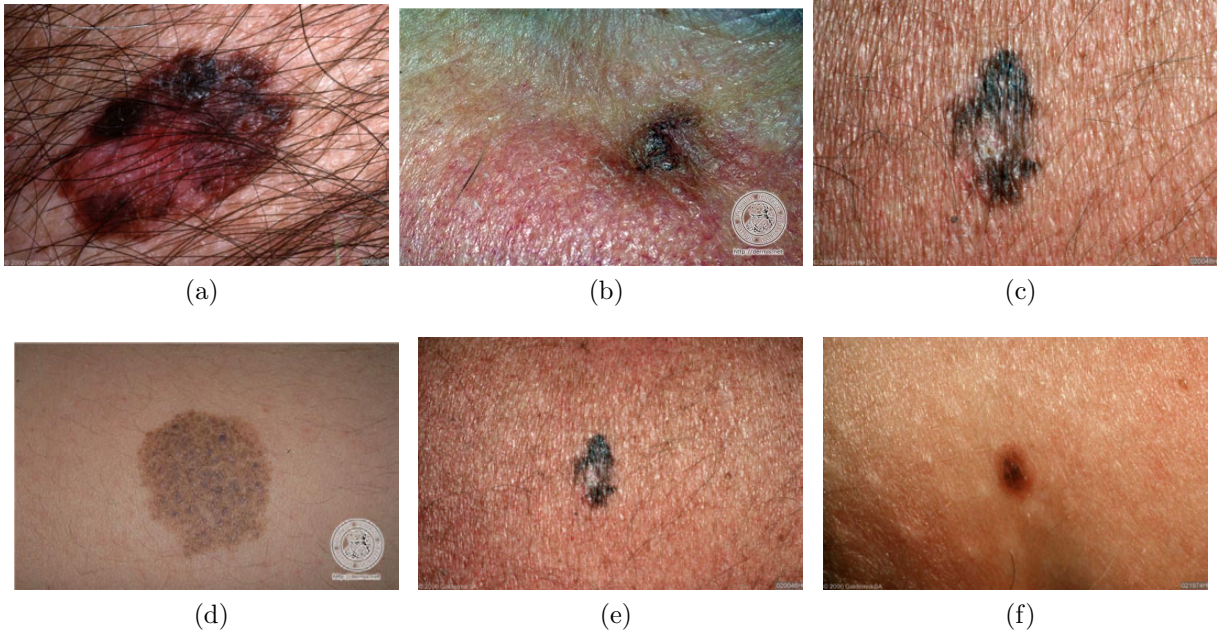


Figure 3.7: The subset of digital images with significant artefacts. These images were used to compare the results of using both stages of the MSIM algorithm to correct for illumination variation compared to using solely the parametric modeling stage.

Table 3.4: CV and AUC Results comparing Parametric and Multi-stage Modeling

Metric	Parametric Modeling Only	Multi-stage Modeling
Average $cv^\dagger$	0.1501	<b>0.1442</b>
Average $AUC^\dagger$	0.8966	<b>0.9085</b>

<sup>†</sup> A lower  $cv$  and higher  $AUC$  indicates better illumination correction.



with significant amounts of hair, which can lead to very poor illumination estimates if a purely parametric modeling approach is employed. The use of an initial outlier-robust non-parametric estimate as a prior to a parametric modeling stage allows for significantly improved illumination estimates.

A subset of six digital images with significant artefacts were corrected using both stages of the MSIM algorithm and using solely the parametric modeling stage. The subset included images from both the Dermatology Information System and Dermquest databases and images with melanoma and with non-melanoma lesions. These six images are chosen to demonstrate the need for the non-parametric modeling step when correcting images with artefacts. The images in this subset are shown in Fig. 3.7. Fig. 3.7a has a significant amount of hair. Fig. 3.7b and 3.7d have complicated illumination variation patterns and large watermarks. Fig. 3.7c and 3.7e have texture patterns in the healthy skin areas and Fig. 3.7f has a difficult illumination component to model. Coefficient of variation, as described in Sec. 3.5.3 is used to quantify the performance of using only one or both stages of the MSIM algorithm. ROC curve analysis is also performed on the subset, as described in Sec. 3.5.4.

The results from comparing parametric and multi-stage illumination modeling using the subset of six images are presented in Table 3.4. The average coefficient of variation of the set of images corrected using both stages was 0.1442 and the average coefficient of variation of the images corrected using only the parametric modeling stage was 0.1501. Therefore, using both stages resulted in a lower coefficient of variation than if only the parametric modeling stage is performed. The area under the ROC curve (AUC) for the images corrected using both stages was 0.9085, while the AUC for the images corrected using only the parametric modeling stage was 0.8966. Since the ROC AUC is higher for the subset of images corrected using both stages, it indicates that the skin pixels and lesion pixels are more separable in those corrected images. From the coefficient of variation and ROC results, the combination of a non-parametric modeling stage with a parametric modeling stage results in a final image that is noticeably better corrected for illumination variation.

## 3.6 Chapter summary

A multi-stage illumination modeling (MSIM) algorithm is proposed to remove illumination variation in digital images of skin lesions. Illumination variation is a prominent artifact in digital images and because of the similar appearance of the lesion and areas in shadow, it can affect segmentation results. The proposed algorithm is tested on a set of 190 digital

images from the DermIS [62] and DermQuest [9] databases. The corrected images are compared to images correcting using three existing illumination correction algorithms. Performance of the algorithms is quantified using the coefficient of variation, segmentation accuracy and classification accuracy. Coefficient of variation is a measure of the ratio of variance and mean intensities of background pixels, where a lower coefficient of variance indicates less illumination variation. Results find that the proposed algorithm had the lowest average coefficient of variation across the entire dataset. An existing segmentation algorithm is applied to the corrected images and ROC curves were plotted. The proposed algorithm has the highest area under the receiver operator characteristic (ROC) curve, indicating better segmentation performance. Finally, an existing melanoma classification algorithm is trained using the corrected images to identify lesions as malignant melanoma based on a set of features. The proposed algorithm has the highest classification accuracy using the classifier. With the skin lesion image corrected for illumination variation, a set of representative texture distributions is learned from the corrected image in Chapter 4.



## Chapter 4

# Normal Skin and Lesion Sparse Texture Model

In this section, the sparse texture modeling approach used to learn skin and lesion texture distributions and measure their distinctiveness is outlined. The representative texture distributions used to identify pixels that belong to the lesion and skin classes and to find the border of the skin lesion. This sparse texture modeling algorithm is referred to as the Joint Probability Texture Distinctiveness (JPTD) algorithm. Sec. 4.1 describes existing sparse texture models. In Sec. 4.2, sparse texture distributions are introduced, derived from the local texture patches. In Sec. 4.3, a metric is introduced to measure distinctiveness of the texture distributions from each other. Implementation details are given in Sec. 4.4. A flow chart highlighting steps in the JPTD algorithm is shown in Fig. 4.1.

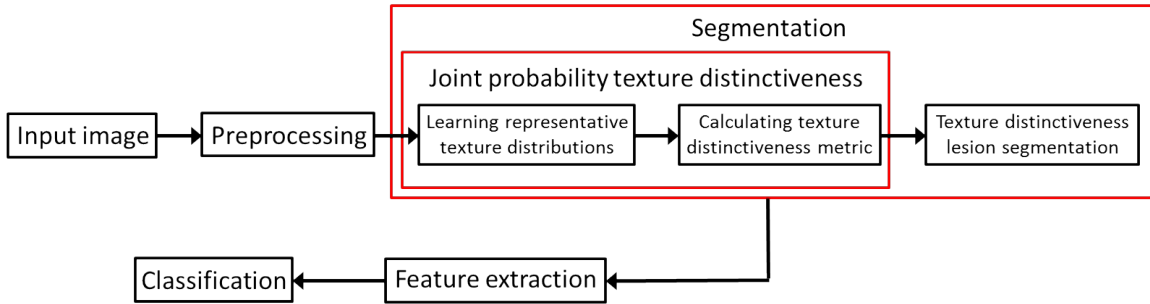


Figure 4.1: Algorithm flow chart highlighting steps in the JPTD algorithm

## 4.1 Sparse texture model

Sparse texture models find a small number of texture representations, such as texture patches, to characterize an entire image [68]. Sparse texture models learn important local texture details present in an image. Using a sparse texture model allows the image to be stored efficiently and allows for efficient computation of algorithms that involve textures from the image. There are many ways to learn the model, including clustering or by formulating the problem as an optimization problem [69]. A common method to learn a sparse texture model is by employing a dictionary learning algorithm [68], where a set of texture patches that can best match details in the original image is learned.

## 4.2 Representative texture distributions

In the JPTD algorithm, a sparse texture model algorithm [70] is modified to find representative sparse texture distributions from an input image, which has been corrected for illumination variation. Then, in Chapter 5, the texture distributions are classified as representing normal skin or lesion textures.

To learn the sparse texture distributions, a local texture vector is obtained for each pixel in the image. The corrected input image contains  $N \times M$  pixels and each pixel has  $a$  channels. The texture vector contains pixels in the neighbourhood of size  $n$  centered on the pixel of interest. Each row in the neighbourhood is concatenated sequentially. Let  $s$  be a pixel location in the image. Then, the vector  $\mathbf{t}_s$  represents the  $n \times n \times a$  texture patch centered at pixel  $s$ . The process of extracting the texture vector for a pixel in a single channel is illustrated in Fig. 4.2. In the case of  $d$  multiple channels, the texture vectors for each single channel are concatenated sequentially into a single vector. A texture vector is extracted for each pixel in the image to obtain  $\mathcal{T}$ , which is a set of  $N \times M$  texture vectors, each of size  $n \times n \times a$ .

$$\mathcal{T} = \{\mathbf{t}_{s_j} | 1 \leq j \leq N \times M\} \quad (4.1)$$

Using the set of all texture vectors extracted from an image, we find a set of  $K$  representative texture distributions, defined as  $\mathcal{T}^r$ . By characterizing the sparse model as a set of distributions, we can capture both local and global characteristics in the image. The representative texture distributions are defined as  $\mathcal{T}^r$  and are able to capture the commonly occurring texture patterns found in lesion and normal skin regions. By using a small set

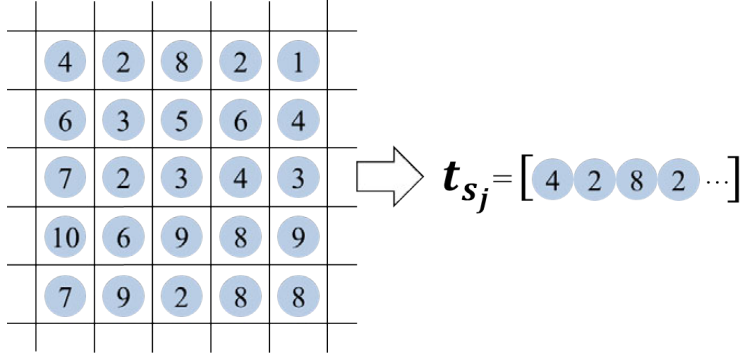


Figure 4.2: Extracting a texture vector. For images with multiple channels, a separate vector is obtained for each channel and concatenated sequentially. In this example,  $s$  is the center pixel with an intensity of 3.

of representative texture distributions instead of using all the local texture vectors, the computational complexity and memory requirements are reduced.

$$\mathcal{T}^r = \{T_j^r | 1 \leq j \leq K\} \quad (4.2)$$

Each texture vector belongs to a single representative texture distribution, which best corresponds with that texture vector. All parameters needed to characterize the  $j^{th}$  texture distribution are contained in  $\theta_j$ . Each distribution has its own distinct set of parameters.

A mixture model is used to represent the set of texture distributions associated with the input image. The parameters that characterize the texture distributions are chosen to maximize the log-likelihood of the mixture model [71]:

$$\hat{\mathcal{T}}^r = \arg \max_{\mathcal{T}^r} \sum_{j=1}^K \sum_{\mathbf{t}_{s_k} \in C_j} \log (P(\mathbf{t}_{s_k} | T_j^r)) \quad (4.3)$$

To find the representative texture distributions and the sets of texture vectors corresponding to each representative distribution, an unsupervised clustering algorithm is used. The set  $C_j$  is comprised of the texture vectors corresponding to texture distribution  $T_j^r$ . The algorithm used in this implementation is described in Sec. 4.4.2. A Gaussian distribution is assumed, and  $\theta_j$  contains the two required parameters to define a multivariate Gaussian distribution. The mean and covariance of the  $j^{th}$  texture distribution are represented by  $\mathbf{t}_j^r$  and  $\Sigma_j$ , respectively.

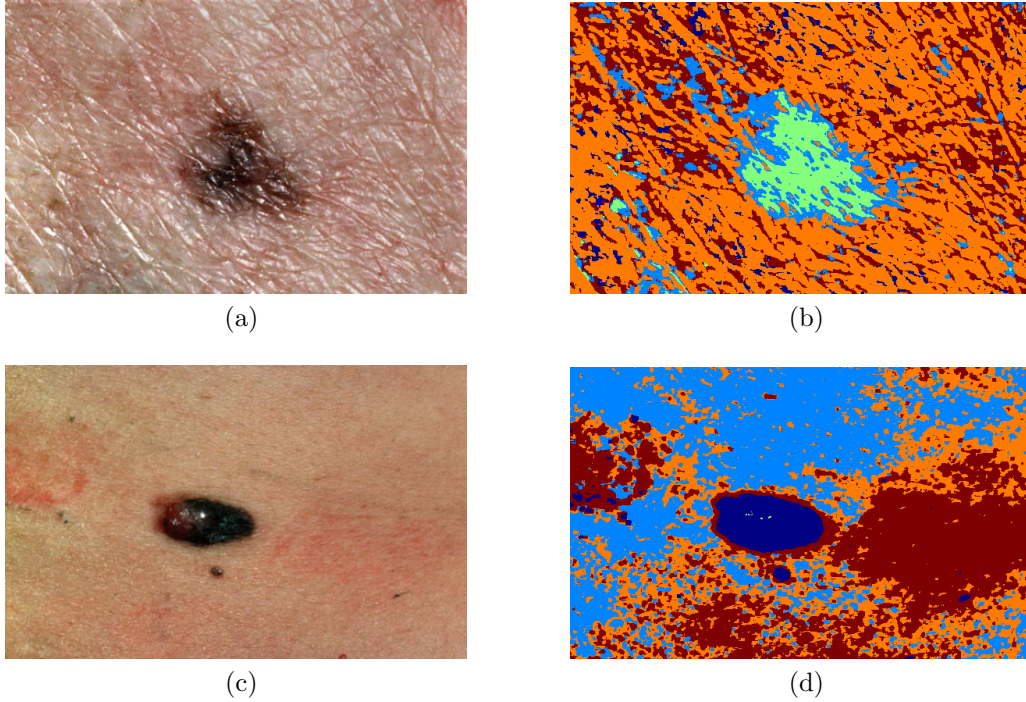


Figure 4.3: Map of representative texture distributions. In (a) and (c), the original images are shown. In (b) and (d), five representative texture distributions have been learned and each pixel in the image is replaced by one of five colours, depending on which texture distribution that pixel is associated with.

Examples of maps illustrating the results of finding the representative texture distributions are shown in Fig. 4.3. In the two examples, five representative texture distributions have been learned. As described above, each pixel in the image have also been associated with one of the five representative texture distributions. In Fig. 4.3a and 4.3c, the original image is shown. In Fig. 4.3b and 4.3d, each pixel from the original image has been replaced with one of five colours to represent which representative texture distribution they are associated with. This example shows that different textures are learned for healthy skin pixels and lesion pixels.

### 4.3 Texture distinctiveness metric

Using the learned sparse texture model, the last step of the JPTD algorithm is to find a texture distinctiveness metric. Since we are only interested in two classes, normal skin and lesion, but have learned many texture distributions, multiple texture distributions must represent the same class. To measure similarity of two texture distributions  $l_{j,k}$ , we measure the probability that the mean of one texture distribution is a realization of the mean of the other texture distribution, as defined in Eqn. 4.4. Because we assume that the texture distributions are Gaussian,  $\mathbf{t}_j^r$  and  $\Sigma_j$  are the mean and covariance of distribution  $T_j^r$ .

$$l_{j,k} = \frac{1}{\sqrt{(2\pi)^{n*n*a}|\Sigma_k|}} \exp\left(-\frac{1}{2}(\mathbf{t}_j^r - \mathbf{t}_k^r)^T \Sigma_k^{-1} (\mathbf{t}_j^r - \mathbf{t}_k^r)\right) \quad (4.4)$$

However, we are interested in finding distinct texture distributions. For example, lesion texture distributions are both dissimilar from the normal skin texture atoms and also from other texture distributions, due to colour variegation and textural patterns found in skin lesions. The probability that a texture distribution is distinct from another texture distribution is given by  $d_{j,k}$ :

$$d_{j,k} = 1 - l_{j,k} \quad (4.5)$$

Using the texture distributions and probabilities of distinctiveness, a weighted graphical model can be constructed to characterize all pair-wise relationships. The graphical model is defined as  $G = \{V, E\}$ .  $V$  represents the set of vertices for the graphical model, which are the texture distributions associated with each pixel in the image.  $E$  represents the set of edges between every pair of texture distributions, which are given a weight based on the probability of distinctiveness,  $d_{j,k}$ .

A textural distinctiveness metric  $D_j$  is used to capture the dissimilarity of texture distribution  $T_j^r$  from other texture distributions. The metric is defined in 4.6 and measures the expected distinctiveness of  $T_j^r$  given the image  $I$ , where  $P(T_k^r|I)$  is the probability of occurrence of a pixel being associated with a texture distribution  $T_k^r$ .

$$D_j = \sum_{k=1}^K d_{j,k} P(T_k^r|I) \quad (4.6)$$

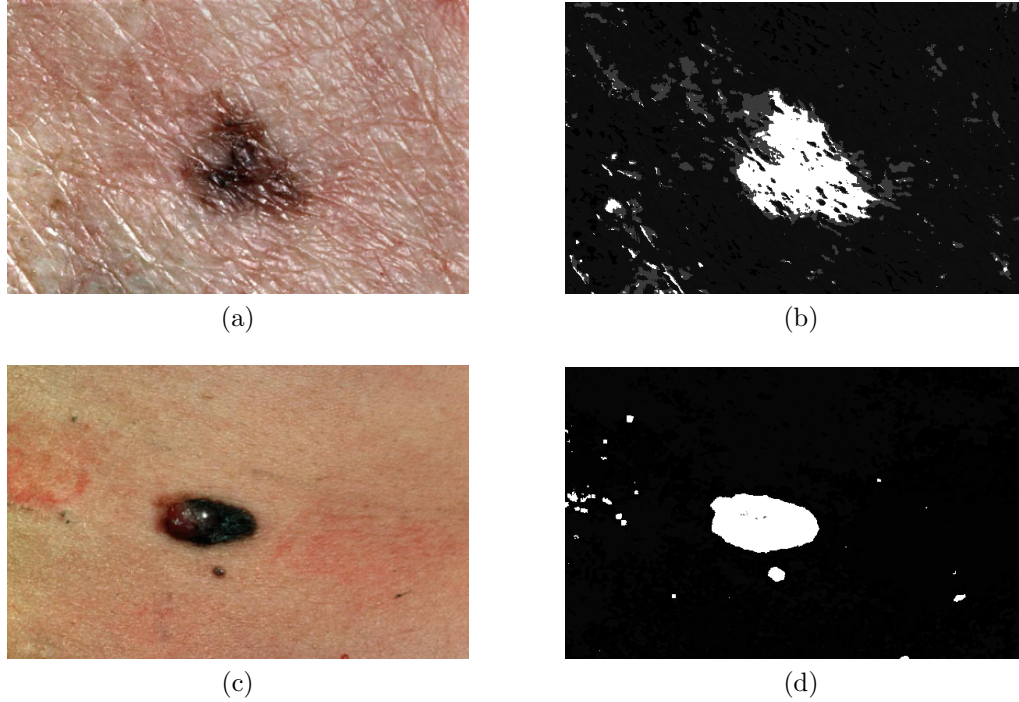


Figure 4.4: Map of the texture distinctive metric. In (a) and (c), the original images are shown. In (b) and (d), maps of the texture distinctive metric are constructed. Five representative texture distributions are learned and the texture distinctiveness metric is calculated for each distribution. The pixel intensities in (b) and (d) depend on the texture distinctiveness of the texture distribution associated with each pixel.

In the case of normal skin texture distributions, the dissimilarity of one skin texture distribution from other skin texture distributions is very small. The textural distinctiveness metric is small overall, despite being dissimilar from lesion texture distributions. Lesion texture distributions are dissimilar from other skin and lesion texture distributions, so the textural distinctiveness metric is large. Fig. 4.4b and 4.4d illustrate this effect by displaying the textural distinctiveness metric for each pixel in the image. In both figures, the lesion is predominately white, which corresponds to a highest textural distinctiveness metric.

## 4.4 Implementation details

### 4.4.1 Colour space

In our implementation of the JPTD algorithm, the image is in the RGB domain and has three channels ( $a=3$ ). However, the algorithm can be expanded to take into account multi- or hyper-spectral images of a skin lesion, where  $a$  is much greater. Furthermore, the image is converted to the XYZ colour space. Work by Terrillon et al. found that the XYZ colour space proved to be an efficient colour space in which to segment the skin region of human faces [72]. This colour space is designed to better model colour perception and reduce correlation between the XYZ channels, compared to the standard RGB colour space.

### 4.4.2 Learning representative texture distributions

In our implementation, a two-step clustering algorithm is used to learn a set of representative texture distributions that satisfies Eqn. 4.3. First, a k-means clustering algorithm is run, which is followed by learning a finite mixture model. K-means clustering is used as an the initial step to increase the robustness and to speed up the number of iterations required for the finite mixture model to converge. K-means clustering finds  $K$  clusters of texture data points that minimizes the sum of squared error between cluster members and the cluster mean. The optimization function for k-means clustering is shown in Eqn. 4.7, where  $C_k$  is the  $k^{th}$  set of texture vectors, and  $\boldsymbol{\mu}_k$  is defined as the mean vector for the  $k^{th}$  set. Implementation details for k-means clustering is described by Jain et al. [73]. The initial cluster means are randomly assigned.

$$\hat{C} = \arg \min_C \sum_{k=1}^K \sum_{\mathbf{t}_{s_j} \in C_k} \|\mathbf{t}_{s_j} - \boldsymbol{\mu}_k\|^2 \quad (4.7)$$

One limitation with k-means clustering is that it does not take into account any probabilistic information. Therefore, the second step is to apply finite mixture model clustering. To fit the finite mixture model, the model parameters in the set  $\Theta$  are found to maximize the log-likelihood function shown in Eqn. 4.8. In this implementation, a Gaussian distribution is assumed for all clusters and the model parameters are the distribution mean  $\mu$  and distribution covariance  $\Sigma$ .  $\Theta$  also contains the parameter  $\alpha$ , which is the mixing proportion. No closed form solution exists for Eqn. 4.8 in general, so an expectation-maximization iterative algorithm is used, as described by Figueiredo et al. [71]. The expectation-maximization

algorithm is initialized using cluster means, covariances and mixing proportions based on the results of the k-means clustering.

$$\hat{\Theta} = \arg \max_{\Theta} \sum_{j=1}^n \sum_{k=1}^K \log (\alpha_k P(\mathbf{t}_{s_j} | \mu_k, \Sigma_k)) \quad (4.8)$$

where  $\sum_{k=1}^K \alpha_k = 1$  and

$$\Theta = \{\mu_1, \mu_2, \dots, \mu_K, \Sigma_1, \Sigma_2, \dots, \Sigma_K, \alpha_1, \alpha_2, \dots, \alpha_K\}$$

Expectation-maximization is an iterative algorithm. The initial parameters for the Gaussian mixture model are obtained from the results of the k-means clustering. That is, the initial Gaussian means are equal to the k-means cluster means:

$$\mu_k = \mu_{C_k} \quad (4.9)$$

and the distribution covariances and mixing proportions are also dependent on the cluster results:

$$\Sigma_k = \Sigma_{C_k} \quad (4.10)$$

$$\alpha_k = P(\mathbf{t}_{s_j} \in C_k). \quad (4.11)$$

The parameters defining the  $K$  representative texture distributions are taken to be the mean and covariances for the  $K$  estimated Gaussian distributions ( $\hat{\mathbf{t}}_{\mathbf{k}}^r = \boldsymbol{\mu}_{\mathbf{k}}$ ). Furthermore, each texture vector is assigned to belong to the distribution which maximizes the weighted probability  $\alpha_k P(\mathbf{t}_{s_j} | \mu_k, \Sigma_k)$ .

## 4.5 Summary of the JPTD texture distinctiveness algorithm

1. Convert the corrected image to the XYZ colour space.



2. For each pixel  $s$  in image  $I$ , extract the texture vector  $\mathbf{t}_s$  to obtain the set of texture vectors  $T$  (Eqn. 4.1).
3. Cluster the texture vectors in  $T$ , as described in Sec. 4.4.2, to obtain the representative texture distributions.
4. Calculate probability that two texture distributions are distinct  $d_{j,k}$  using Eqn. 4.5 for all possible pairs of texture distributions.
5. Calculate the textural distinctiveness metric  $D_j$  (Eqn. 4.6) for each texture distribution.

## 4.6 Texture distinctiveness map results

This thesis introduces the use of probabilistic information to determine representative texture distributions and to measure texture distinctiveness. To determine if incorporating this information is useful, the resulting texture distinctiveness maps produced using the JPTD algorithm were compared to distinctiveness maps produced using an algorithm adapted from the Texture Distinctiveness (TD) algorithm [70], which only uses the k-means clustering algorithm to find the representative texture distributions. Furthermore, the TD algorithm does not take into account the covariance corresponding to each cluster when calculating the distinctiveness metric (Eqn. 4.5). Finally, because the TD algorithm is designed to compute saliency maps, the distinctiveness metric includes an additional term based on the distance between a pixel and the center of the image. Since we are interested in understanding the effect of the additional probabilistic information, this term was omitted in the comparisons.

Fig. 4.5 shows examples of the corrected lesion image and their corresponding distinctiveness map, produced by the TD and JPTD algorithms. Examples shown are those where there are significant differences when using the two algorithms. Also, the dynamic range of pixels is scaled to the maximum pixel intensity and minimum pixel intensity, resulting in a different dynamic range for each texture distinctiveness map.

Some interesting observations can be made from the examples. First, lesions in Fig. 4.5a, 4.5c and 4.5f are comprised of two distinct textures. For example, in Fig. 4.5a, there are pronounced dark areas and lighter red areas. However, when using the TD algorithm, only the first texture is highlighted in the texture distinctiveness map. Using the JPTD algorithm, both textures are highlighted. This is also seen in Fig. 4.5c and 4.5f.

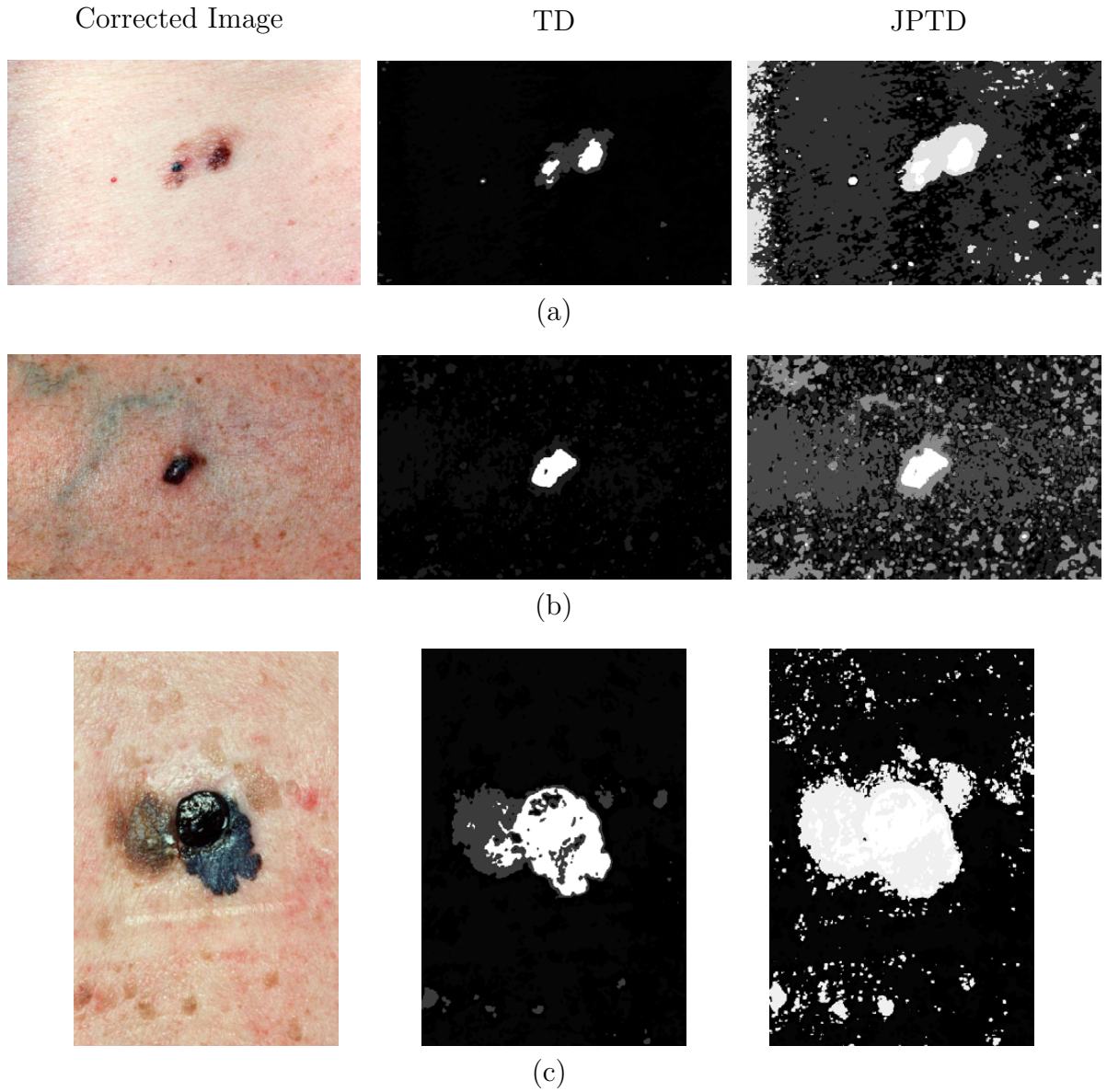


Figure 4.5: Corrected skin lesion images and their corresponding textural distinctiveness maps. The textural distinctiveness maps are produced using the TD algorithm [70] and the JPTD algorithm. The pixel intensity corresponds to the texture distinctiveness of the pixel's associated texture distribution. The JPTD algorithm is able to better highlight the lesion area, compared to the TD algorithm. However, in (a), (c) and (f), non-lesion areas are also highlighted by the JPTD algorithm.

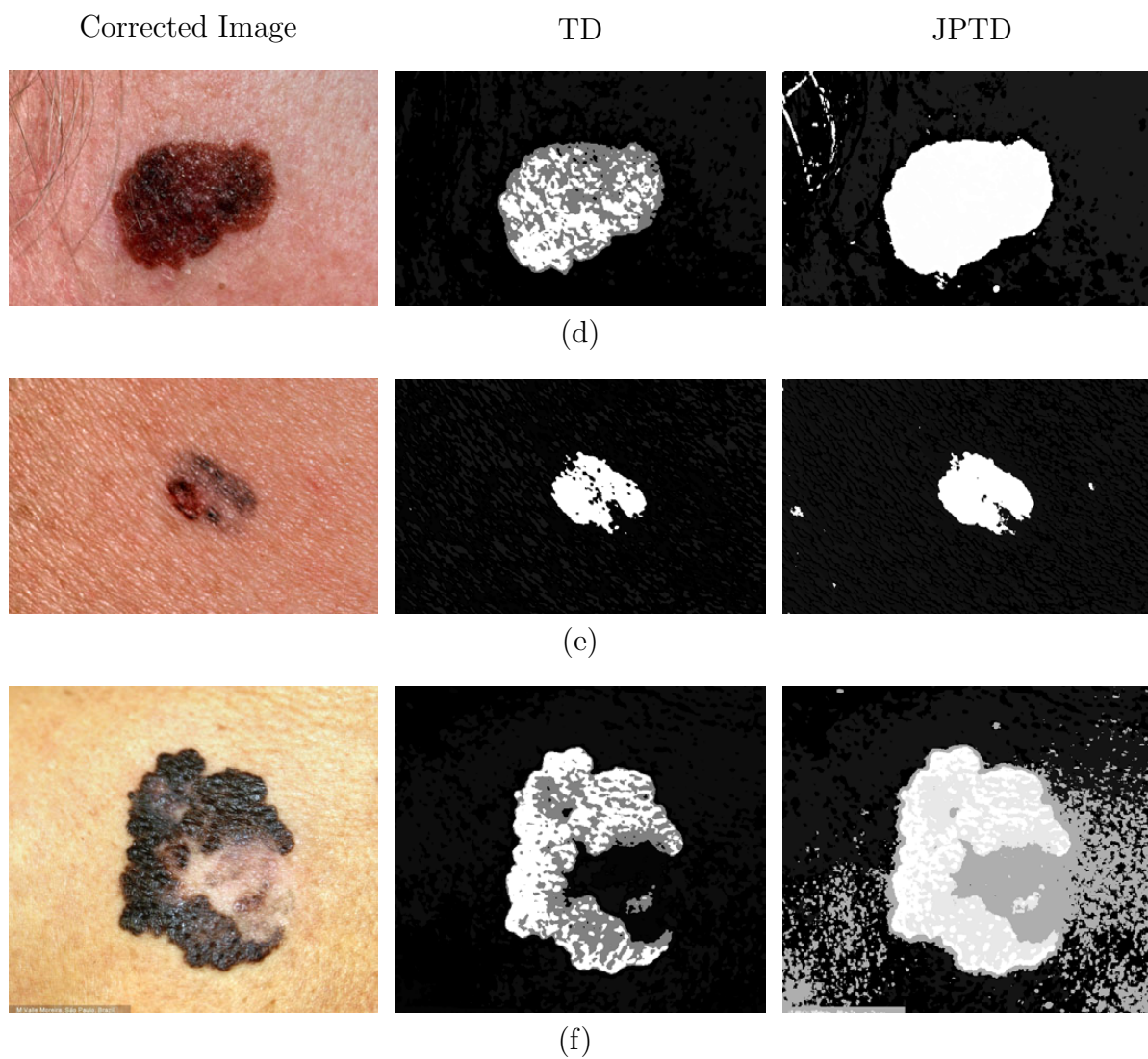


Figure 4.5: Corrected skin lesion images and their corresponding textural distinctiveness maps, cont.

As a trade-off, non-lesion areas that occur due to natural pigmentation and texture characteristics of the skin are also highlighted when using the JPTD algorithm. For example, in Fig. 4.5a, the presence of shading on the left side of the image is highlighted when using the JPTD algorithm, but not when using the TD algorithm. This motivates use of the texture-based segmentation algorithm proposed in Sec. 5 to find the lesion, rather than using the textural distinctiveness maps directly.

## 4.7 Chapter summary

In this section, we propose a novel sparse texture modeling approach called the Joint Probability Texture Distinctiveness (JPTD) algorithm to learn skin and lesion texture distributions. A texture distinctiveness metric is introduced to measure the dissimilarity of a learned texture distribution from the other texture distributions. The JPTD algorithm is compared to a similar sparse texture modeling algorithm. Visually, the introduction of probability information allows the JPTD algorithm to better highlight multiple texture distributions that belong to the same lesion. As a consequence, the JPTD algorithm can also highlight non-lesion areas. To avoid misclassifying pixels, the texture distinctiveness metric is used in conjunction with a region classification algorithm in Chapter 5 to classify regions as being in the normal skin or lesion class.

# Chapter 5

## Texture-based Skin Lesion Segmentation

In this chapter, a texture-based skin lesion segmentation algorithm is proposed. This novel skin lesion segmentation algorithm is designed to be used for images taken by a digital camera, after being corrected for illumination variation. The segmentation algorithm uses a set of learned texture distributions and their texture distinctiveness metric, as defined in Chapter 4, to find pixels that belong to the skin lesion. The proposed segmentation algorithm is referred to as the Texture Distinctiveness Lesion Segmentation (TDLS). First, the input image, which has been corrected for illumination variation, is over-segmented. This results in the image being divided into a large number of regions. Next, each region is independently classified as representing normal skin or lesion based on the textural contents of that region. Finally, post-processing steps refine the lesion segmentation. A flow chart showing the major steps in the TDLS algorithm is shown in Fig. 5.1.

The TDLS algorithm assumes that there is a single lesion in the digital image and that the lesion is roughly in the center of the image. Since the algorithm uses learned textures that represent the skin and lesion classes, it is assumed that the input image has been corrected for illumination variation. Otherwise, textures learned from areas of the skin in shadows may be misclassified as part of the lesion class.

### 5.1 Initial regions

The first step in the TDLS algorithm is to over-segment the corrected lesion image, dividing the image into a number of regions. This initial over-segmentation step is incorporated to

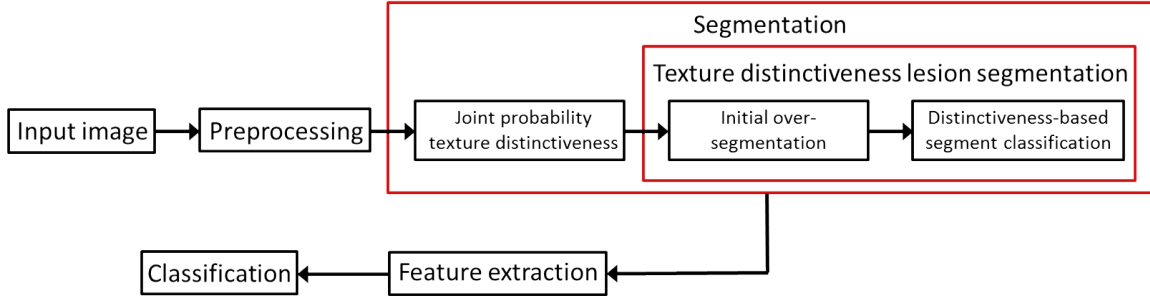


Figure 5.1: Algorithm flow chart highlighting steps in the TDLS algorithm

increase the TDLS algorithm’s robustness to noise. Furthermore, it allows for the use of an efficient and fast classification algorithm, described in Sec. 5.2, to find the areas belonging to the skin lesion. The initial over-segmentation algorithm, adapted from the statistical region merging (SRM) [60] algorithm, is described below. The main difference is that the SRM algorithm uses the image in the RGB colour space, while the TDLS algorithm converts the image to the XYZ colour space, as mentioned in Sec. 4.4.1. The advantages of using the SRM algorithm as the initial over-segmentation algorithm is that it directly takes into account pixel location, is simple and computationally efficient.

SRM contains two main steps: a sorting step and a merging step. SRM sorts pixels in an image to determine the order in which pixels are compared, and then merges pairs of pixels into regions based on their similarity. A 4-connected graph is constructed so that each pixel in the image is connected with its neighbours. The pixels are sorted based on the similarity function defined in Eqn. 5.1, where  $s$  is defined as a pixel location  $(x, y)$ ,  $s'$  is a neighbouring pixel and  $\rho(s)$  and  $\rho(s')$  are the XYZ pixel intensities at those locations. Both horizontal and vertical neighbouring pixels are considered when sorting the pixels.

$$f(s, s') = \|\rho(s) - \rho(s')\|_\infty \quad (5.1)$$

Sorting pixels determines the order in which pixels are merged into regions. The merging predicate determines whether two regions are merged together, based on pixel intensities. The predicate depends on the difference between average pixel intensity for each channel for the two regions. Furthermore, it depends on the number of pixels in the regions. The predicate used by SRM is defined in Eqn. 5.2 where  $\bar{R}$  is the average pixel intensity across a region and  $|R|$  is the number of elements in the set [60]. The parameter  $g$  is the maximum pixel intensity,  $\mathcal{R}_{|R|}$  is the set of regions with the same number of pixels as the region  $R$  and  $Q$  is a tunable parameter to change the likelihood that two regions are merged.



$$\mathcal{P}(R, R') = \begin{cases} true & \text{if } \|\bar{R}'_d - \bar{R}_d\| \leq \sqrt{b^2(R) + b^2(R')} \forall \text{ channel } d \\ false & \text{otherwise} \end{cases} \quad (5.2)$$

$$\text{where } b(R) = g \sqrt{\frac{1}{2Q|R|} \ln(6|I|^2|\mathcal{R}_{|R|}|)}$$

Here, the parameter  $Q$  is set to 128 following experimental testing. Additional implementation details can be found in the paper by Nock and Nielsen [60]. The result of the initial over-segmentation step is a map of several regions which correspond to the normal skin or lesion classes. To reduce the number of regions, all segments that touch the edges of the image are merged into a single region. This is because we assume that the lesion is in the center of the image, so regions touching the edges are all likely to be part of the normal skin class.

## 5.2 Distinctiveness-based segment classification

Following the initial oversegmentation step, each region must be classified as belonging to the normal skin class or lesion class based on a criterion. The classification step is illustrated in Eqn. 5.3, where  $y$  is the resulting segmentation map. Each element in  $y$  is either 1 (lesion) or 0 (normal skin), depending on the classification results for that element's corresponding region. The threshold is denoted by  $\tau$  and it represents the decision boundary between the normal skin and lesion class. The feature used to discriminate between the two classes is the regional textural distinctiveness metric  $\mathcal{D}_R$ . This metric is based on the texture distinctiveness across a region.

$$y(R) = \begin{cases} 1 & \text{if } \mathcal{D}_R \geq \tau \text{ (lesion)} \\ 0 & \text{otherwise (normal skin)} \end{cases} \quad (5.3)$$

From Chapter 4, each pixel in the input image is associated with a texture distribution. A texture distinctiveness metric  $D$  is calculated for each texture distribution based on the probability of it being similar to other texture distributions, as described in Chapter 4. This information is combined with the contents of each region to determine a regional texture distinctiveness metric,  $\mathcal{D}_R$ . This metric represents the average texture distinctiveness across a region (Eqn. 5.4), where  $P(T_j^r|R)$  is the probability of a pixel being associated with the  $j^{th}$  texture distribution in the region  $R$ .

$$\mathcal{D}_R = \sum_{j=1}^K D_j P(T_j^r | R) \quad (5.4)$$

Finally, a threshold  $\tau$  is defined to divide the set of representative texture distributions into two classes, normal skin and lesion, and is also based on the texture distinctiveness metrics derived in Chapter 4. There are many ways to find two classes from a one-dimensional set of features. In the TDLS algorithm, the threshold is found that divides the set of texture distributions into two classes such that the total intra-class variance of the texture distinctiveness metric for each class is minimized, as shown in Eqn. 5.5.

$$\tau = \arg \min_{\tau} (\sigma_{C_1(\tau)}^2 P(T^r | C_1(\tau)) + \sigma_{C_2(\tau)}^2 P(T^r | C_2(\tau))) \quad (5.5)$$

The threshold  $\tau$  is used to divide the set of texture distributions into two classes  $C_1(\tau)$  and  $C_2(\tau)$ . The classes depend directly on  $\tau$  because if the distinctiveness metric of the associated texture distribution is above  $\tau$ , that texture distribution is in class  $C_1(\tau)$ . Likewise, if it is below  $\tau$ , it is in class  $C_2(\tau)$ . The probability that a texture distribution is in the class  $C$  for a given  $\tau$  is  $P(T^r | C(\tau))$  and the variance of the texture distinctiveness based on the elements in the class is  $\sigma_{C(\tau)}$ . This threshold is also known as the Otsu's threshold [64].

### 5.3 Segmentation refinement

After the regions are classified as being normal skin or lesion, the following post-processing steps are applied to refine the lesion border: morphological dilation and region selection. First, the morphological dilation operator is applied to fill holes and smooth the border. Morphological dilation is a process that expands binary masks to fill small holes [74]. The shape and amount that the binary mask is expanded is controlled by a structuring element, which is a disc with a radius of 5 pixels in the TDLS algorithm.

Next, since multiple non-contiguous regions may have been identified as part of the lesion class, the number of regions is reduced to one. To eliminate the small regions, the number of pixels in each contiguous region is counted. The contiguous region with the largest number of pixels is assumed to correspond to the lesion class and any other regions are converted to the normal skin class. In practice, there could be multiple lesions in a single image, but the classification algorithms used for testing the segmentation results can only accept a single lesion. This gives the final lesion segmentation.



## 5.4 Summary of the TDLS segmentation algorithm

1. Apply Statistical Region Merging to the original XYZ image (Sec. 5.1) to oversegment the image.
2. Calculate the region distinctiveness metric for each region found in step 1 using Eqn. 5.4.
3. Using the texture distinctiveness metrics found in the previous chapter, calculate the threshold  $\tau$  between the normal skin and lesion classes (Eqn. 5.5).
4. Classify each region as normal skin or lesion based on the results of steps 2 and 3 (Eqn. 5.3).
5. Apply a morphological dilation operator to the initial lesion classification.
6. For each contiguous region in the initial segmentation, count the number of pixels in the region.
7. As the final lesion segmentation, return the contiguous region consisting of the most pixels.

## 5.5 Experimental results

The TDLS algorithm is compared to four state-of-art lesion segmentation algorithms. The first algorithm (L-SRM) is designed for dermoscopy images, but can be applied to lesion images as well. It applies the SRM algorithm outlined in Sec. 5.1 and uses the normal skin colour to find the regions corresponding to the lesion. The three other algorithms are proposed by Cavalcanti et al. and are designed specifically for lesion images. One algorithm (Otsu-R) finds the Otsu threshold using the red colour channel [51]. The second (Otsu-RGB) uses all three RGB colour channels and finds Otsu thresholds for each channel [35]. The final algorithm (Otsu-PCA) processes the RGB colour channels to find three more efficient channels to threshold. A texture channel is obtained using Gaussian filtering, a colour channel is obtained using the inverse of the red colour channel and the third channel is found using principal component analysis (PCA) [25]. For simplicity, this algorithm is referred to as Otsu-PCA. All algorithms have additional post-processing steps to clean up the contour, and these steps have been implemented as described in their publication.

The set of 126 images from the Dermquest database are used to test the segmentation algorithms. There are 66 digital images with lesions diagnosed as melanoma and 60 digital images with lesions diagnosed as non-melanoma. All tested images were first corrected using the illumination correction algorithm outlined in Sec. 3. The segmentation algorithms are compared to the ground truth, found by manual segmentation, and by using a melanoma classification algorithm. The entire set of DermQuest images with segmentation results using the TDLS algorithm is provided in Appendices A and B.

The JPTD and TDLS algorithms are implemented in MATLAB on a computer with an Intel Core i5-2400S CPU (2.5 GHz, 6 GB RAM). To segment a skin lesion in a  $1640 \times 1043$  image, the two algorithms have a combined runtime of 62.45s.

### 5.5.1 Segmentation results

Each segmentation algorithm is applied to the corrected images and the resulting segmentation is compared to the ground truth. This can be seen as a classification problem, with each pixel being classified as lesion or normal skin. To quantify the performance of each algorithm at classifying pixel correctly, sensitivity, specificity and accuracy are calculated. Their formulas are given in Eqn. 5.6, 5.7 and 5.8, where TP is the number of true positive pixels, FP is the number of false positive pixels, TF is the number of true negative pixels and FN is the number of false negative pixels.

$$Sensitivity = \frac{TP}{TP + FN} \quad (5.6)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5.7)$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (5.8)$$

Tables 5.1, 5.2 and 5.3 show the average sensitivity, specificity and accuracy across the entire set of images or for just the melanoma or non-melanoma images. Interesting examples of segmentation results are shown in Fig. 5.2, along with the ground truth.

Table 5.1 shows that the TDLS algorithm has the highest accuracy across all tested images. The TDLS algorithm also has the second highest sensitivity and specificity. This trend also follows when looking at subsets of melanoma or non-melanoma images, as seen in Tables 5.2 and 5.3. The Otsu-PCA algorithm has similar specificity and accuracy results,

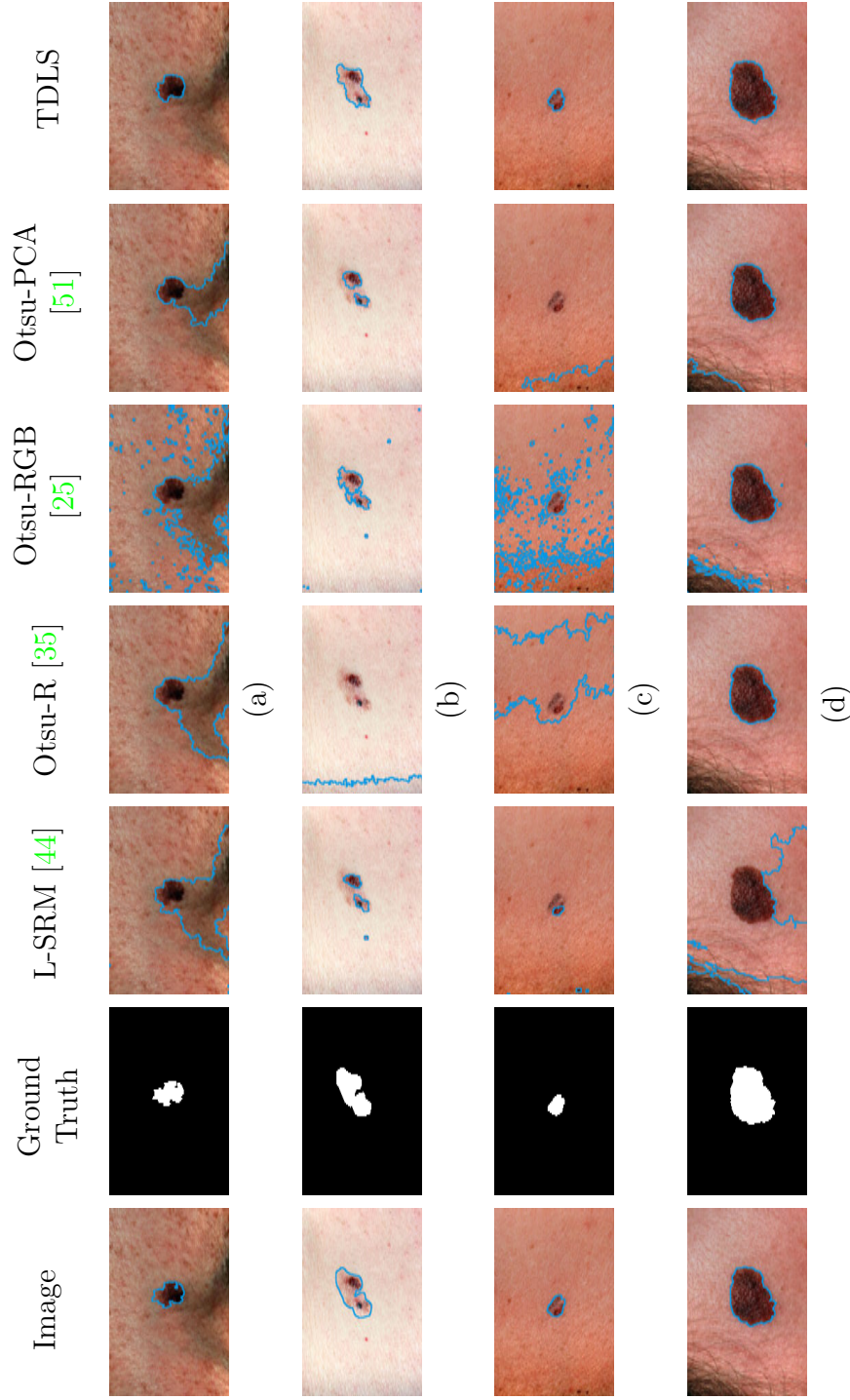


Figure 5.2: Segmentation of skin lesions using various algorithms.

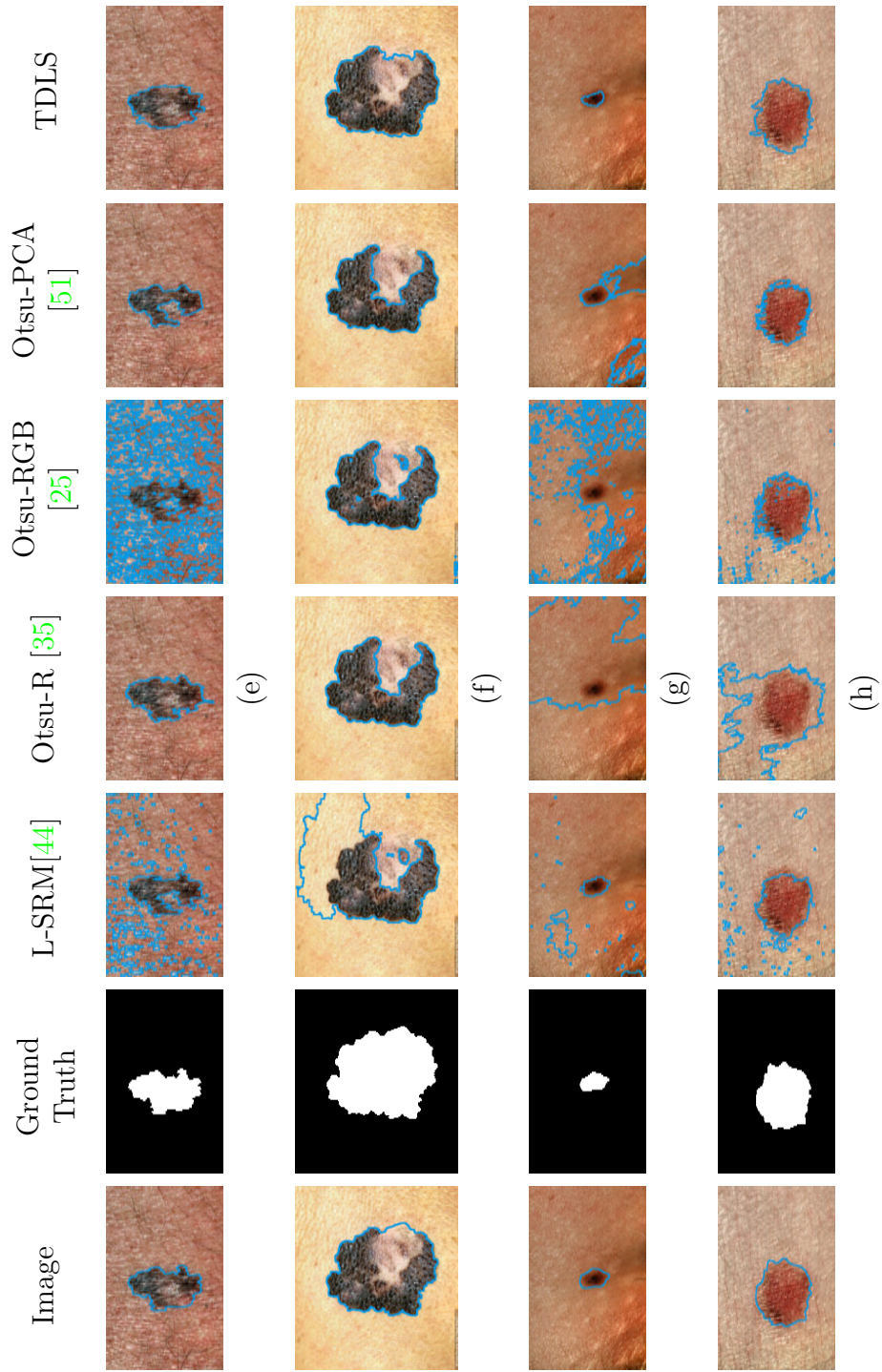


Figure 5.2: Corrected skin lesion images and their corresponding textural distinctiveness maps, cont.

Table 5.1: Classification Results of Lesion and Normal Skin Pixels for All Images

Segmentation Algorithm	Sensitivity	Specificity	Accuracy
L-SRM [44]	89.4%	92.7%	92.3%
Otsu-R [35]	87.3%	85.4%	84.9%
Otsu-RGB [25]	<b>93.6%</b>	80.3%	80.2%
Otsu-PCA [51]	79.6%	<b>99.6%</b>	98.1%
TDLS	91.6%	98.9%	<b>98.3%</b>

Table 5.2: Classification Results of Lesion and Normal Skin Pixels for Melanoma Images

Segmentation Algorithm	Sensitivity	Specificity	Accuracy
L-SRM [44]	90.0%	92.5%	92.1%
Otsu-R [35]	87.4%	91.5%	90.3%
Otsu-RGB [25]	92.2%	85.5%	85.0%
Otsu-PCA [51]	81.2%	<b>99.5%</b>	97.6%
TDLS	<b>92.7%</b>	98.6%	<b>97.9%</b>

Table 5.3: Classification Results of Lesion and Normal Skin Pixels for Non-melanoma Images

Segmentation Algorithm	Sensitivity	Specificity	Accuracy
L-SRM [44]	88.7%	93.0%	92.6%
Otsu-R [35]	87.3%	78.7%	78.9%
Otsu-RGB [25]	<b>95.2%</b>	74.6%	75.0%
Otsu-PCA [51]	77.8%	<b>99.7%</b>	98.6%
TDLS	90.47%	99.3%	<b>98.8%</b>

Table 5.4: Classification Results of Melanoma and Non-melanoma using Low-level Features for Corrected Images comparing Segmentation Algorithms

Segmentation Algorithm	Sensitivity	Specificity	Accuracy
L-SRM [44]	76.9%	<b>83.6%</b>	80.2%
Otsu-R [35]	73.1%	81.4%	77.0%
Otsu-RGB [25]	65.2%	65.0%	65.1%
Otsu-PCA [51]	77.4%	81.3%	79.4%
TDLS	<b>83.1%</b>	<b>83.6%</b>	<b>83.3%</b>

while the Otsu-RGB algorithm has better sensitivity. However, the TDLS algorithm is able to perform well in all three metrics.

Fig. 5.2 shows examples where the TDLS algorithm is able to find the lesion and other algorithms are unable to. Fig. 5.2a-f are images of melanoma lesions and Fig. 5.2g and 5.2h are non-melanoma lesions. In the compared algorithms, areas where illumination variation has not been fully corrected are often misclassified as part of the lesion, as seen in Fig. 5.2a and 5.2g. Complicated texture patterns, as seen in Fig. 5.2e, and hair, as seen in Fig. 5.2d, are also misclassified as part of the lesion. However, the TDLS algorithm is able to reasonably segment the lesion in those images.

Lesions may also be comprised of different colours and textures, such as in Fig. 5.2b and 5.2f. In fact, colour variegation across a lesion is a feature that is used to classify lesions as melanoma. It is critical that segmentation algorithms can account for the colour and texture variation when locating the skin lesion. The compared algorithms only find the most prominent colour or texture and fail to include the subtler regions as part of the lesion. However, because the TDLS algorithm learns the lesion textures and normal skin textures, it is able to locate the entire lesion.

### 5.5.2 Melanoma classification results using low level features

To quantify the effects that each algorithm has on the performance of classification of the lesion as melanoma or non-melanoma, a set of features proposed by Cavalcanti et al. [25] are extracted and used to train an SVM classifier. This is the same approach used in Sec. 3.5.5. A set of 52 features are obtained from 126 corrected DermQuest images and are based on the segmentation results produced by the compared algorithms. There are 11 that characterize asymmetry, 12 that characterize border irregularity, 25 that characterize colour and 4 that characterize differential structure. The same normalization process and



Table 5.5: Classification Results of Melanoma and Non-melanoma using High- and Low-level Features for Corrected Images comparing Segmentation Algorithms

Segmentation Algorithm	Sensitivity	Specificity	Accuracy
L-SRM [44]	79.0%	82.4%	81.0%
Otsu-PCA [51]	82.5%	<b>87.3%</b>	<b>84.9%</b>
TDLS	<b>85.7%</b>	82.9%	84.1%

parameter optimization are used as described in Sec. 3.5.5. Sensitivity, specificity, and accuracy are used as metrics using a leave-one-out cross-validation scheme.

An additional post-processing step is added to the compared algorithms to ensure that only a single contiguous region is identified as the lesion area. This is required because the 52 features assume that there is only a single lesion area. If an algorithm produces multiple possible lesion regions, the region with its centroid closest to the center of the image is taken as the final lesion segmentation.

Table 5.4 show the classification results using low-level features for each segmentation algorithm. The TDLS algorithm has the highest sensitivity and is tied with L-SRM for the highest specificity. The TDLS algorithm has the highest accuracy. The improvements in the classification results are due to improved segmentation, since eleven of the features used are based on the morphology of the lesion border.

### 5.5.3 Melanoma classification results using high-level intuitive features

The TDLS segmentation algorithm is also used with a set of high-level intuitive features (HLIFs) for melanoma classification. HLIFs more closely model the human-observable phenomenon and can be interpreted more intuitively [75, 76]. The HLIFs have been tested using images corrected by the MSIM algorithm, but in that test, manual segmentation was performed to find the lesion. In this test, the three segmentation algorithms with high accuracy, L-SRM, Otsu-PCA and the TDLS algorithm, are used with the HLIFs for lesion classification. Only three algorithms are tested because there are difficulties extracting HLIFs from the other algorithms which produced inaccurate segmentation results.

Nine HLIFs have been proposed, including six features that measure lesion asymmetry and three features that measure border irregularity . To get a complete set of features, the low-level features used in Sec. 5.5.2 are combined with the HLIFs. An SVM classifier is also

used with this set of features. The same normalization process, parameter optimization and leave-one-out cross validation scheme are used as described in Sec. 5.5.2.

Table 5.5 show the classification results using both HLIFS and low-level features for each segmentation algorithm. HLIFs improves the classification accuracy for all three algorithms that are compared. Otsu-PCA improves the most and has the highest accuracy and specificity. The TDLS algorithm has the second highest accuracy and specificity. Its accuracy is close to the performance of the Otsu-PCA algorithm.

Another important observation is that the TDLS algorithm has the highest sensitivity. This is of interest because a low sensitivity would correspond to many malignant melanoma cases being misclassified as benign, which could lead to patient deaths. Therefore, improving sensitivity using the HLIFs and the TDLS algorithm is crucial for the melanoma detection framework.

## 5.6 Chapter summary

In this chapter, we propose a novel texture-based skin lesion segmentation algorithm to find the skin lesion in digital images, referred to as the Texture Distinctiveness Lesion Segmentation (TDLS) algorithm. The segmentation algorithm incorporates learned representative texture distributions and the texture distinctiveness metric to classify regions as belonging to the skin class or lesion class. The segmentation algorithm is compared to four other state-of-the-art algorithms in terms of segmentation and classification accuracy. The proposed algorithm is found to have the highest segmentation accuracy. For classification accuracy, an existing melanoma classification algorithm is used to extract features based on the segmentation results and identify if a lesion is melanoma. When using a set of existing low-level features for melanoma classification, the TDLS algorithm has the highest accuracy.



# Chapter 6

## Conclusions

### 6.1 Summary and conclusions

In summary, the objective of this thesis is to develop a framework to segment images of skin lesions taken using a standard digital camera. To do this, a novel automatic lesion segmentation framework is proposed which includes a multi-stage illumination modeling algorithm as a pre-processing step and a proposed texture-based segmentation algorithm. The framework is to be incorporated to a larger framework to classify and detect of melanoma lesions.

In Chapter 3, the multi-stage illumination modeling algorithm is introduced. The MSIM algorithm uses Monte Carlo sampling to determine an initial estimate of illumination variation, followed by a parametric illumination model. The algorithm is compared to other state-of-art correction algorithms, including an algorithm designed specifically for digital images of skin lesions [35]. The MSIM algorithm produces images with a low coefficient-of-variation (Table 3.1), high segmentation accuracy (Table 3.2) and lesion classification accuracy (Table 3.3). Due to the multiple stages used to model illumination, MSIM is able to correct images with complex texture patterns or hair without drastically altering the colour of the lesion itself.

In Chapter 4, a probabilistic texture distinctiveness metric is introduced based on a learned model of normal skin and lesion textures. Representative texture distributions are learned from the image itself and the texture distinctiveness metric captures the dissimilarity between pairs of texture distributions. The proposed algorithm is visually compared to the results after using an existing texture distinctiveness algorithm [70] that does not

take into account the probabilistic information for representative textures. Incorporating probabilistic information in the proposed algorithm allows skin lesions made up of multiple colours or textures to be fully highlighted in the distinctiveness map (Fig. 4.5). However, this causes some areas that do not correspond to the lesion to also be highlighted. This motivates a proposed texture-based segmentation algorithm which uses the resulting texture distinctiveness map.

Chapter 5 proposes a texture-based segmentation algorithm which divides the image into numerous smaller regions and classifies those regions as lesion or skin based on the texture distinctiveness map. The entire proposed framework is tested by using the illumination corrected images as the input to the texture-based segmentation algorithm. It is compared to state-of-art lesion segmentation algorithms, including three algorithms designed for lesion images [25, 35, 51]. The proposed framework produces the highest segmentation accuracy (Table 5.1), using manually segmented images as ground truth. A lesion classifier is also trained using a set of low-level features obtained from images segmented using the existing algorithms. The proposed algorithm also has the highest classification accuracy and is tied for the highest sensitivity and specificity, when using that feature set. Overall, the TDLS segmentation framework is able to locate the skin lesion in digital images accurately, compared to other state-of-the-art algorithms.

## 6.2 Future work

There are a few avenues to extend the work proposed in this thesis. Future research topics include the following:

- **Incorporate additional high-level intuitive lesion features (HLIFs) to improve lesion classification.** The framework was tested in Sec. 5.5.3 using HLIFs combined with the low-level features to give a comprehensive set of features to capture criteria from the ABCD scale. In the future, HLIFs can replace the 52 low-level features used by Cavalcanti et al. [25] with a smaller number of features, greatly reducing the dimensionality of the feature space.
- **Investigate if the probabilistic texture distinctiveness model introduced in Chapter 4 can improve saliency detection in general images.** The probabilistic model can be used with the saliency detection framework proposed by Scharfenberger et al. [70] to formulate a novel saliency detection strategy.

- **Adapt the lesion segmentation framework for use with multispectral cameras.** Multispectral cameras are sensitive to light frequencies outside of the visible wavelengths. Furthermore, while traditional cameras capture red, green and blue channels, multispectral cameras can capture more channels comprised of different frequencies of light. Studies have found that multispectral images could be processed to find biological structures beneath the outermost layer of skin [77], which can be used to improve the detection of melanoma. The lesion segmentation framework has been formulated in such a way that it is possible to incorporate information from the additional channels from multispectral images.

# APPENDICES

# Appendix A

## Illumination Correction and Segmentation Results for Melanoma Images

This appendix contains the 66 digital images of melanoma lesions from the DermQuest database [9] that are used to test both the proposed illumination correction algorithm and segmentation algorithm.



Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm.

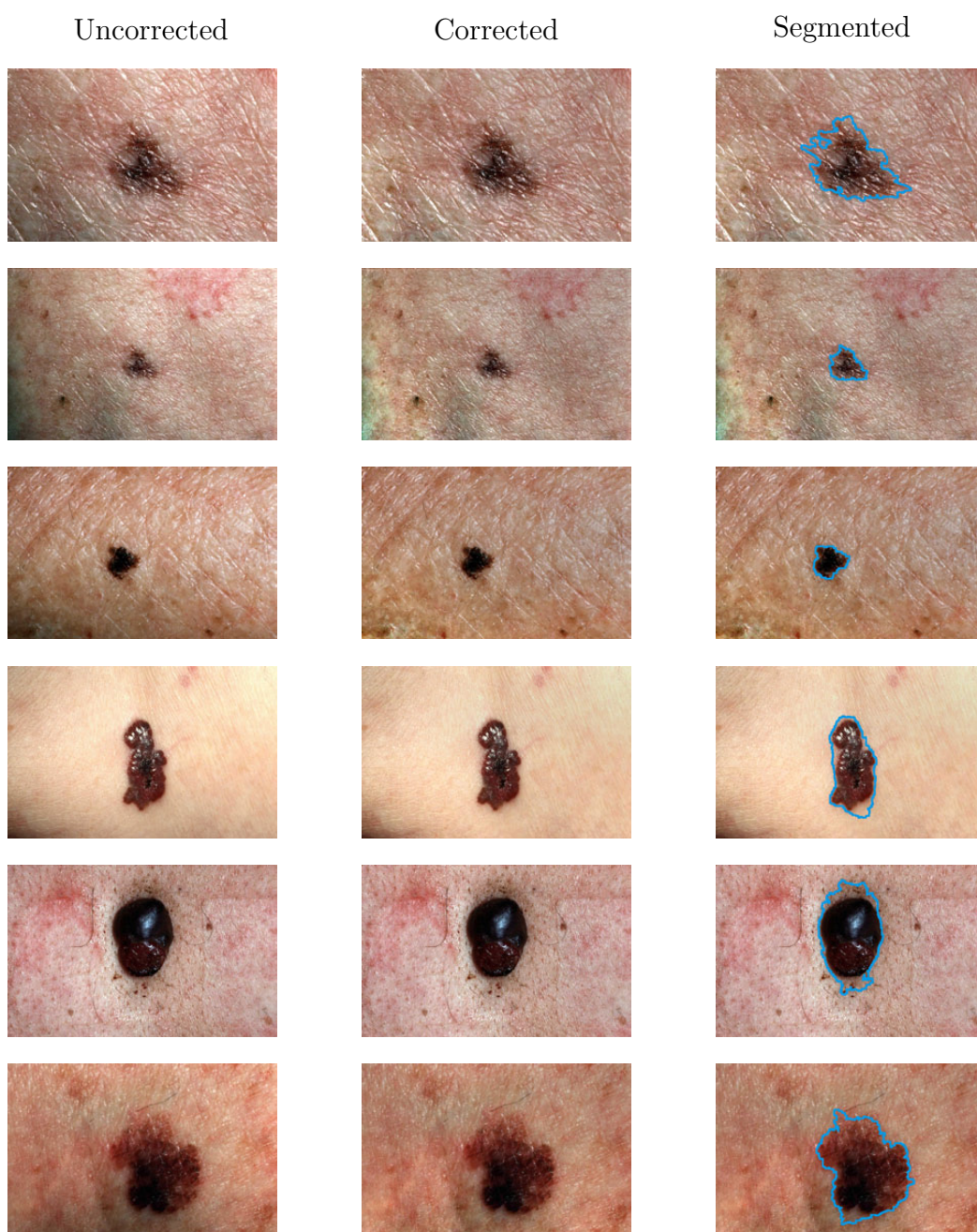


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.



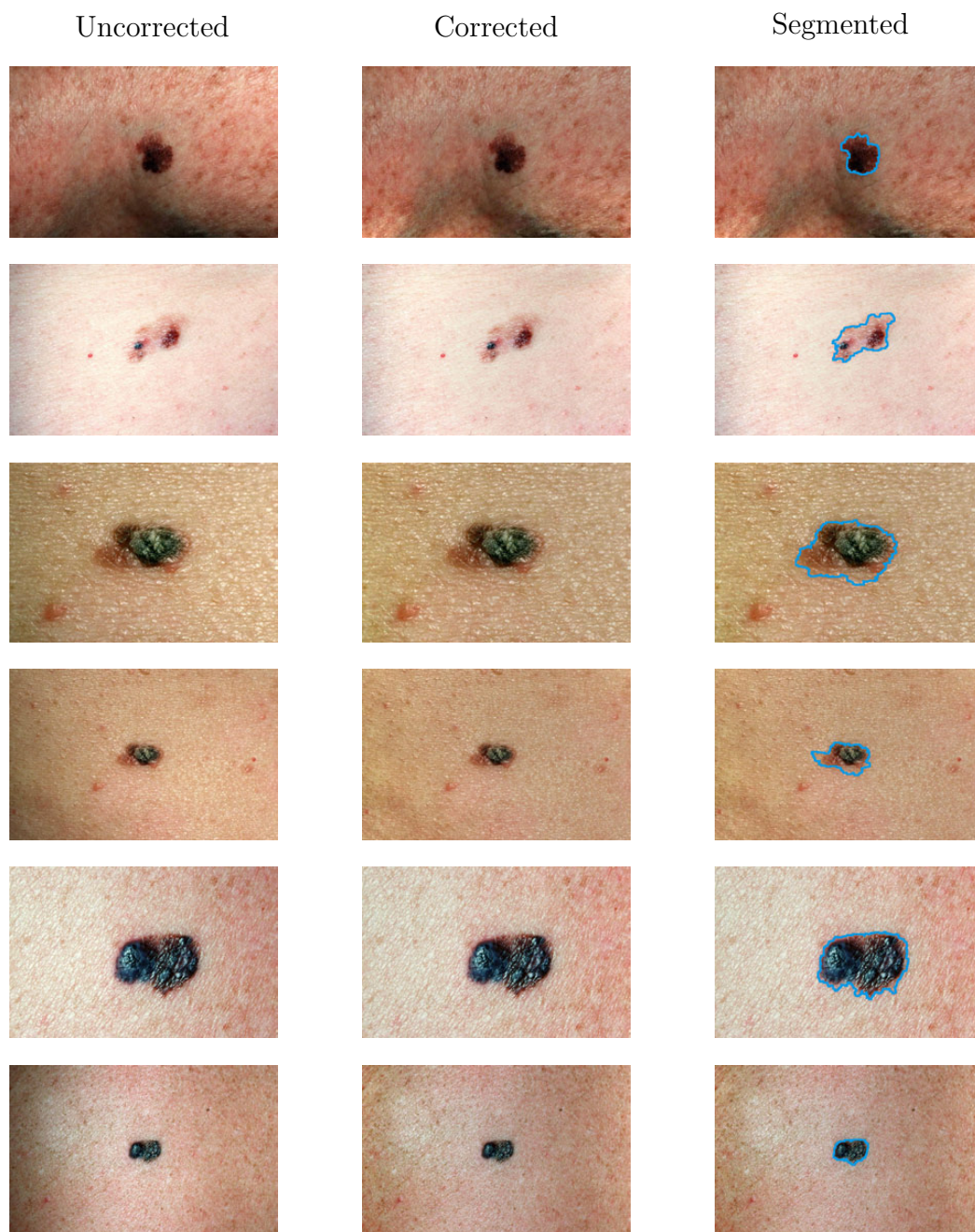


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.

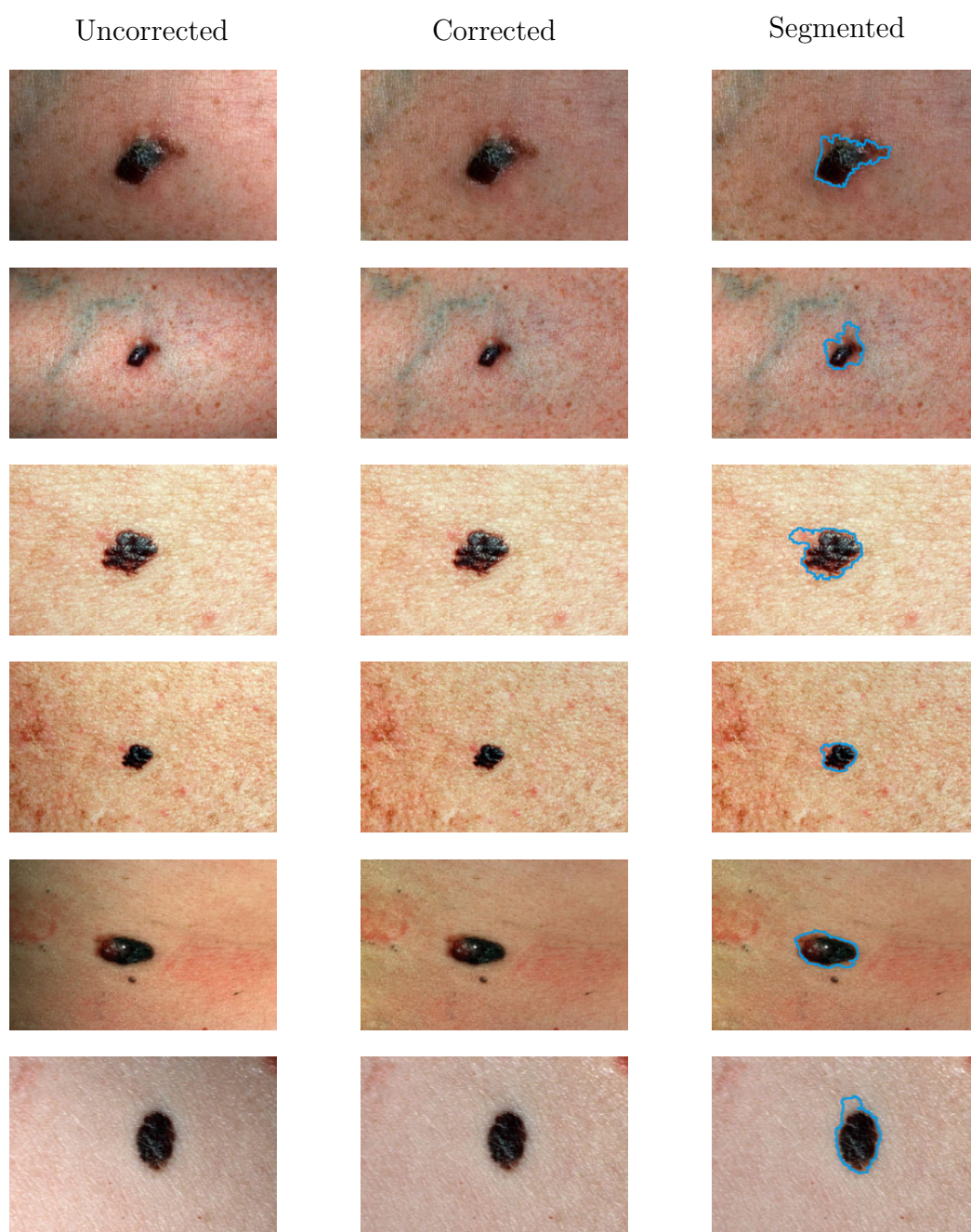


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.



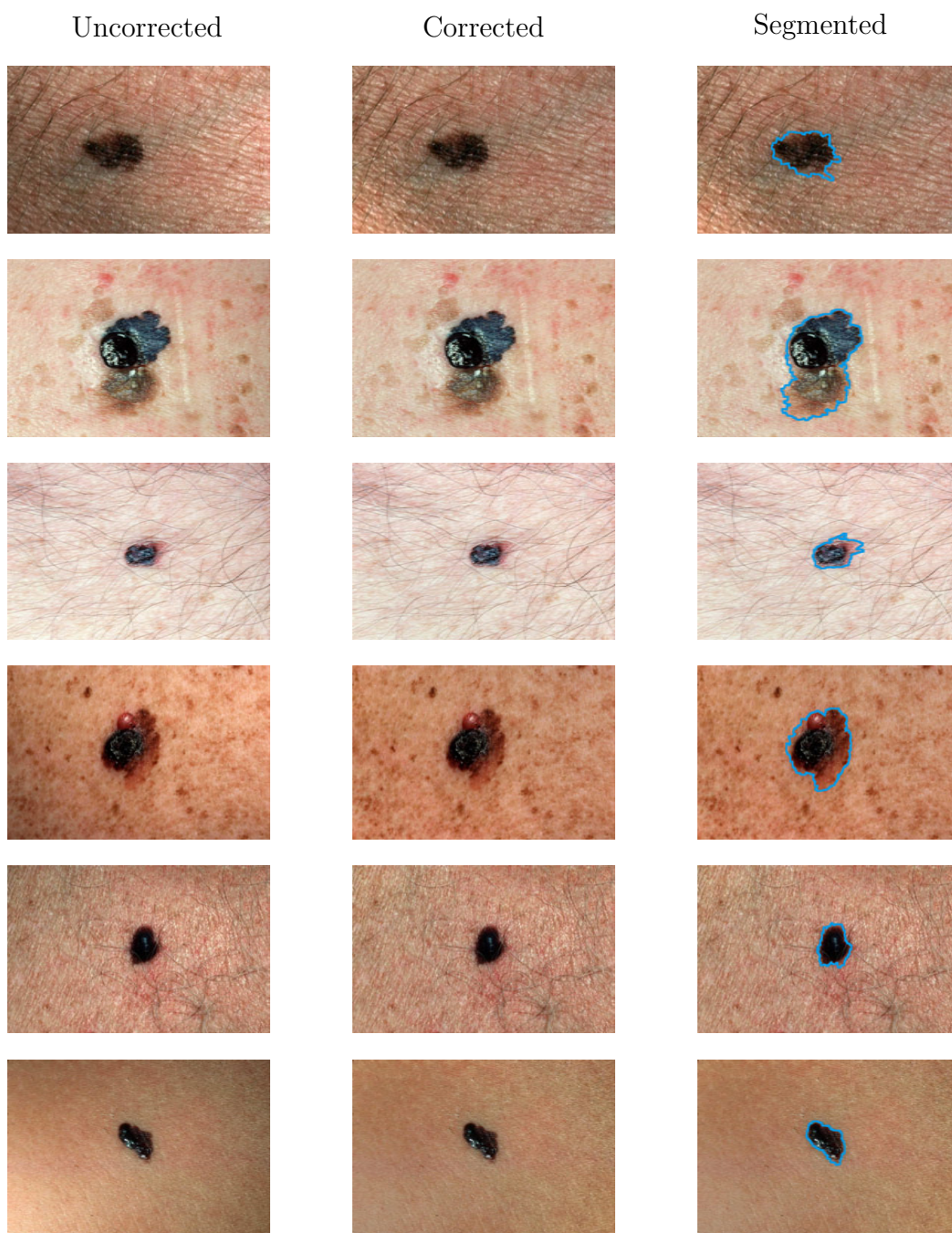


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.

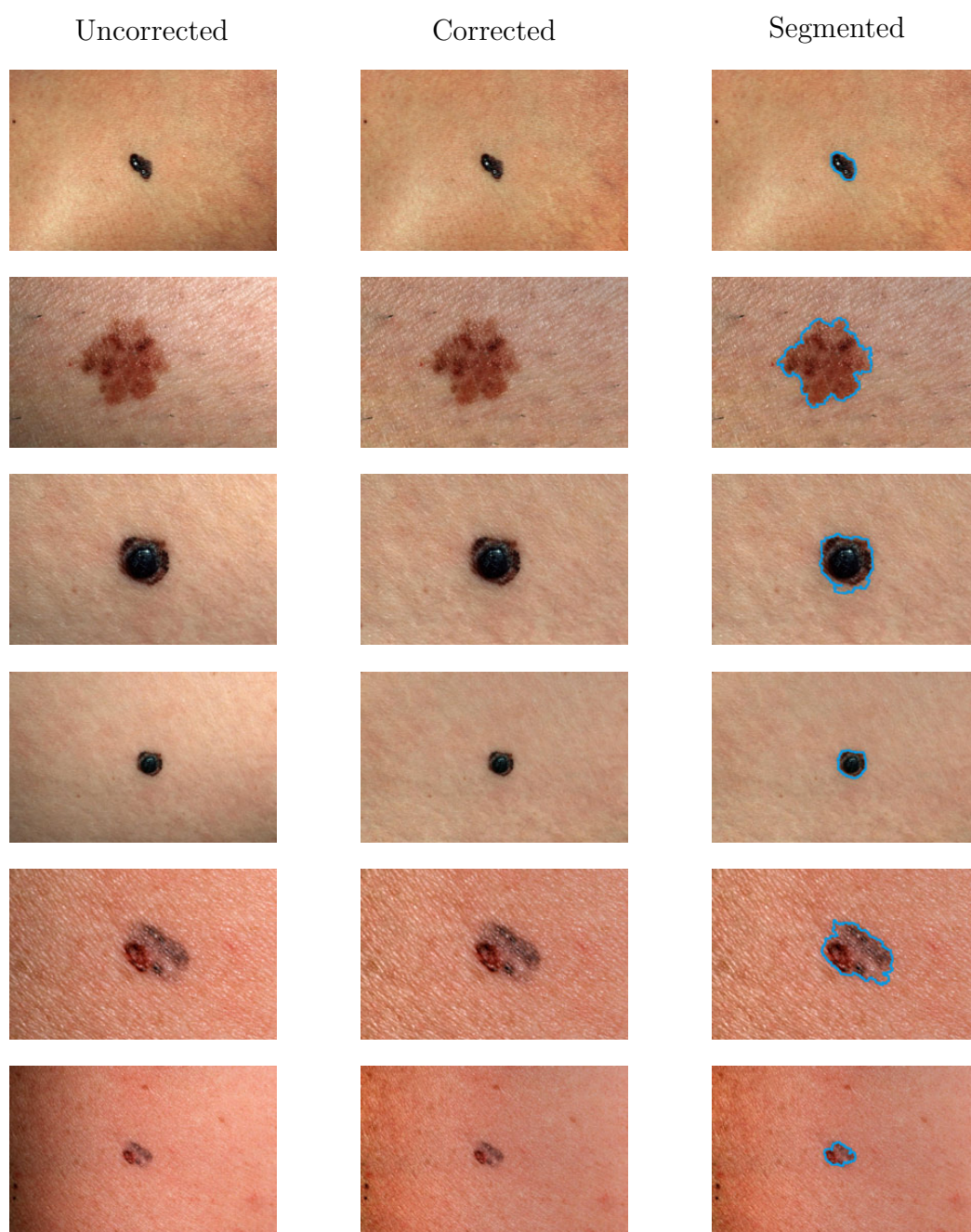


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.



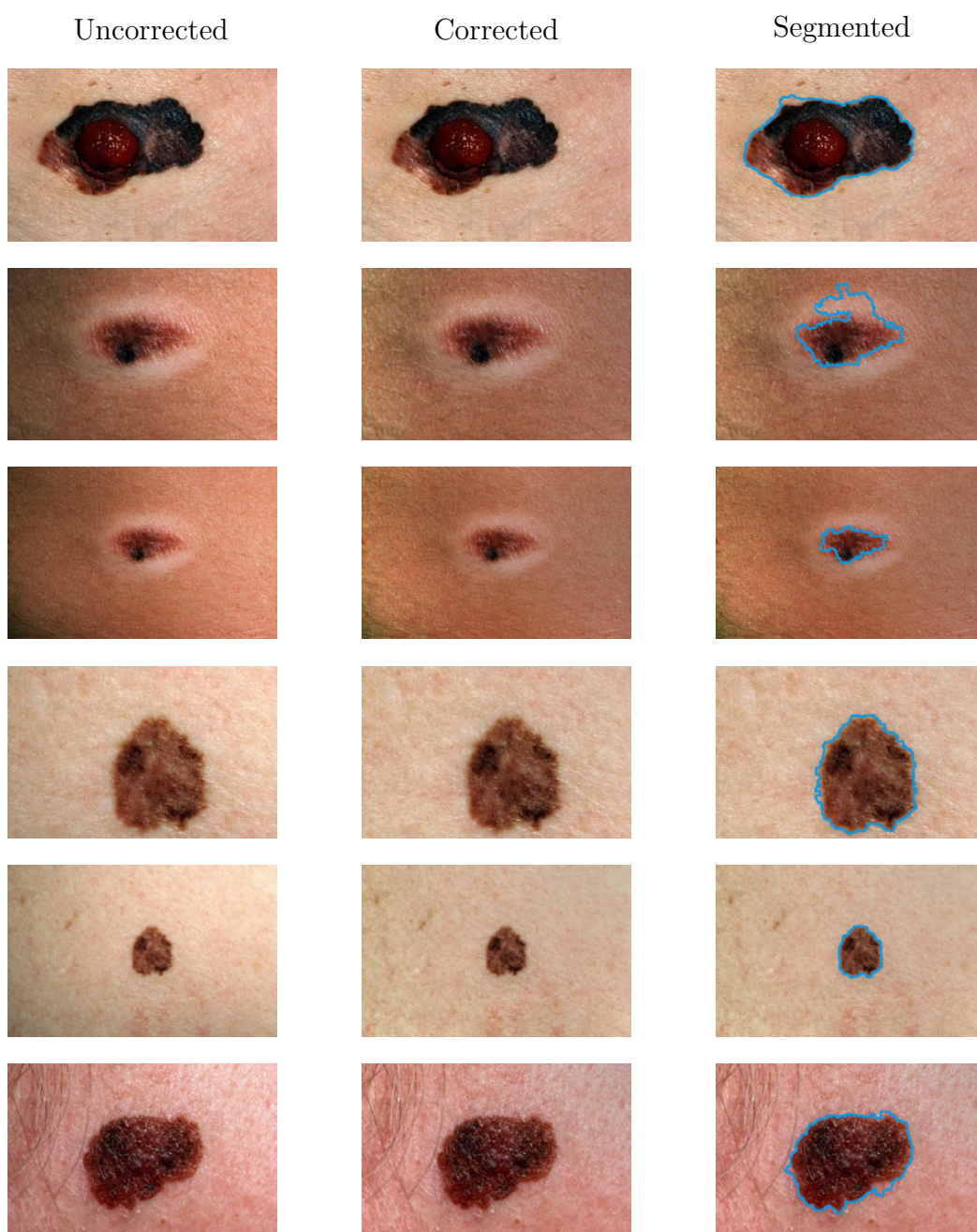


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.

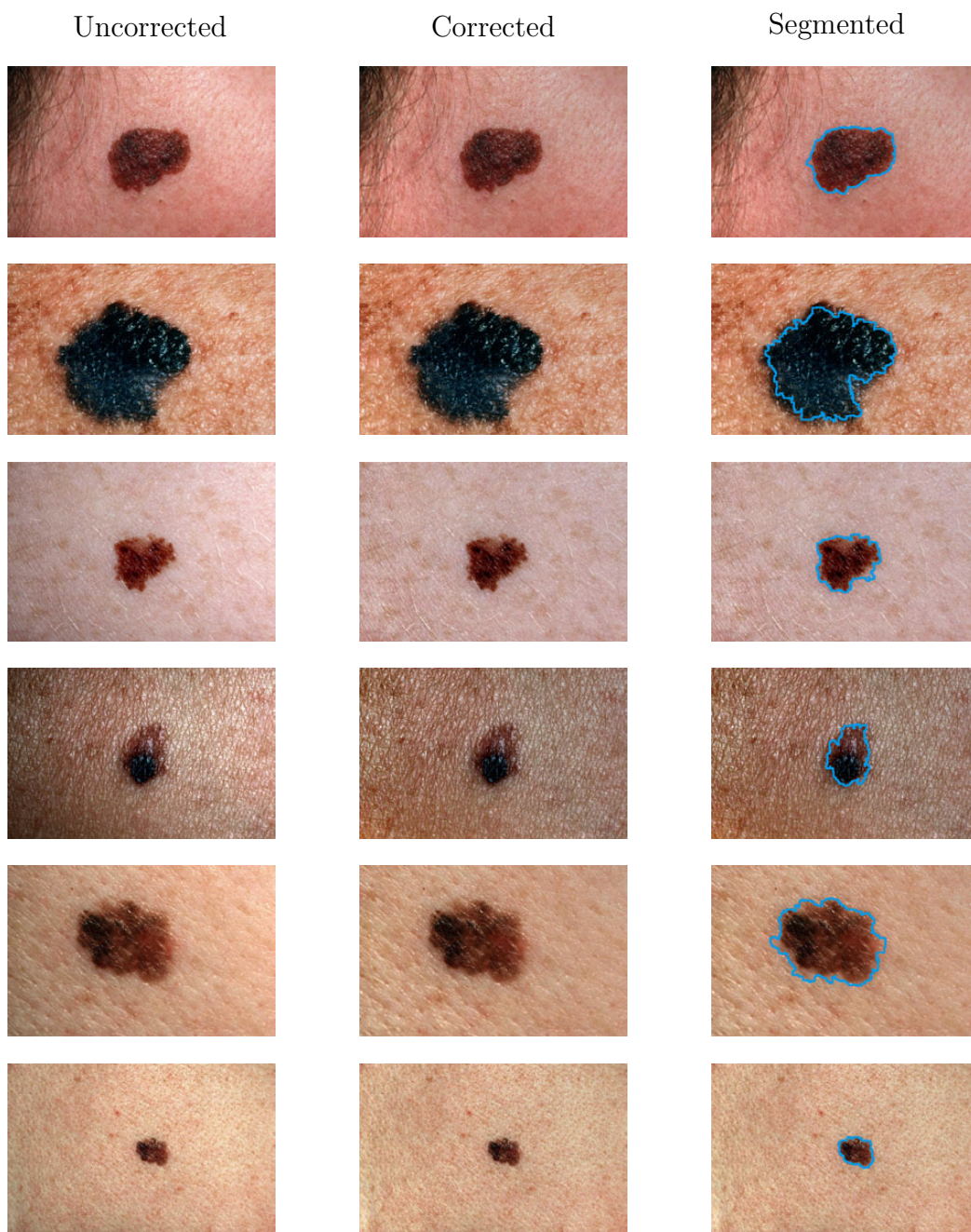


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.



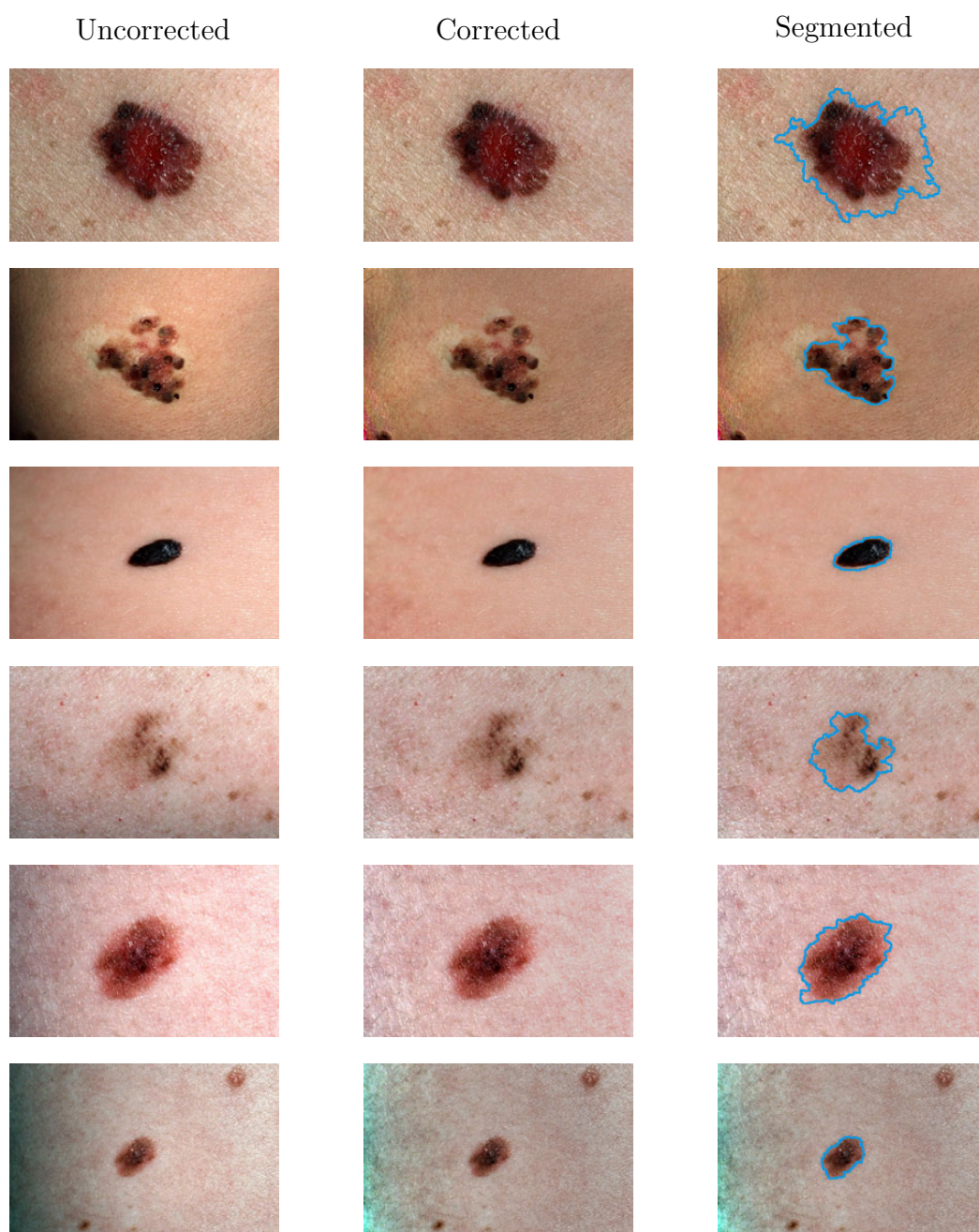


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.

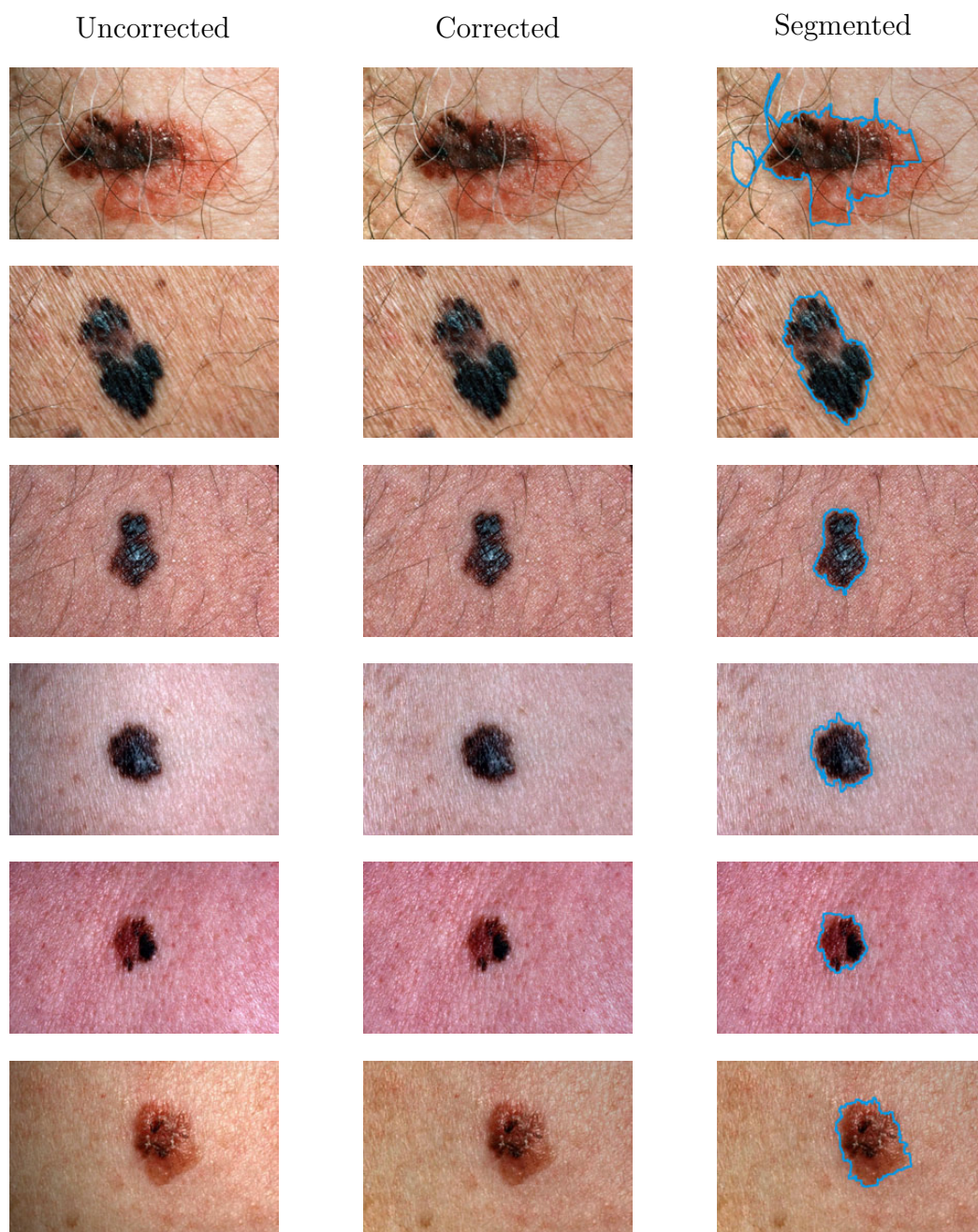


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.



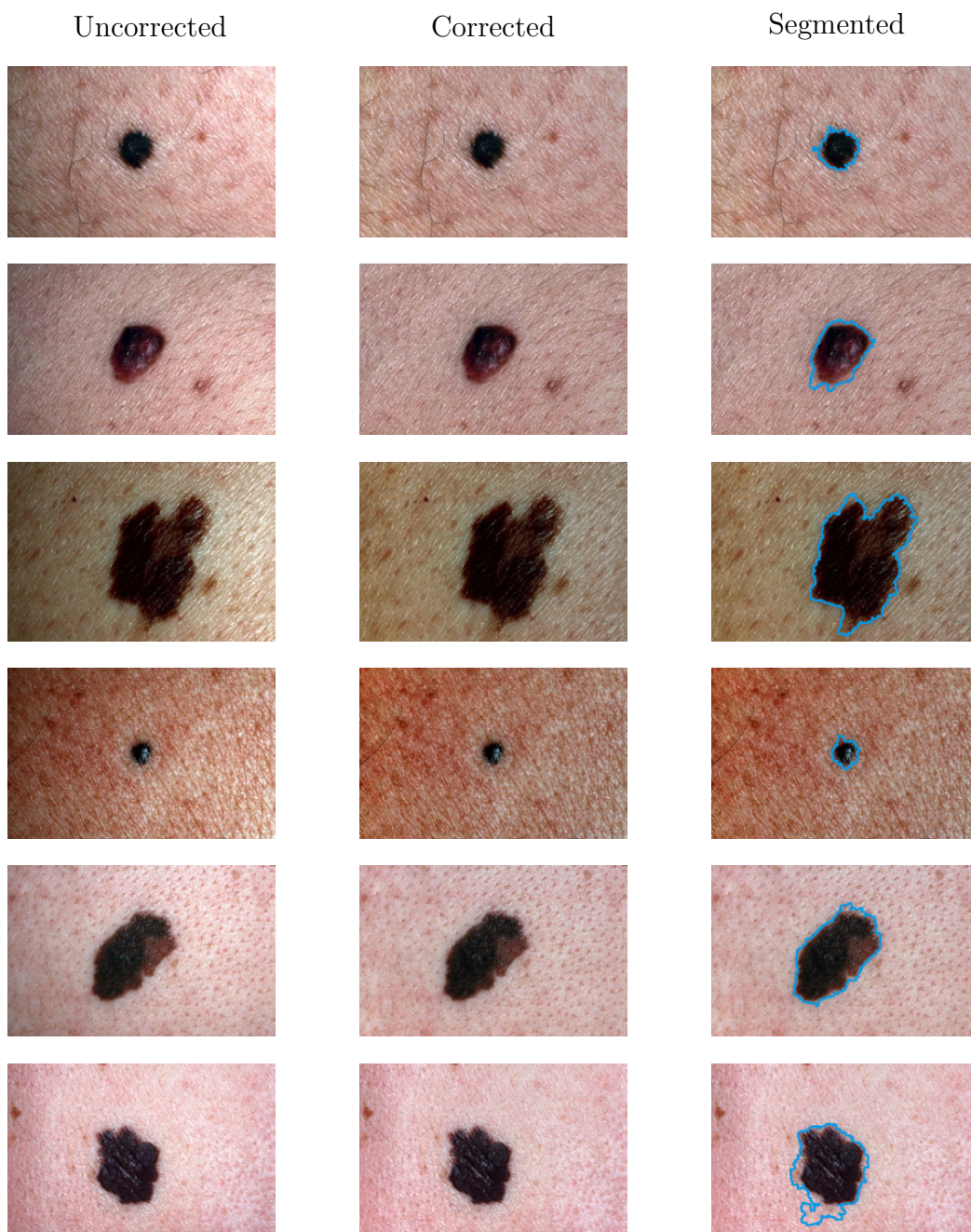


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.

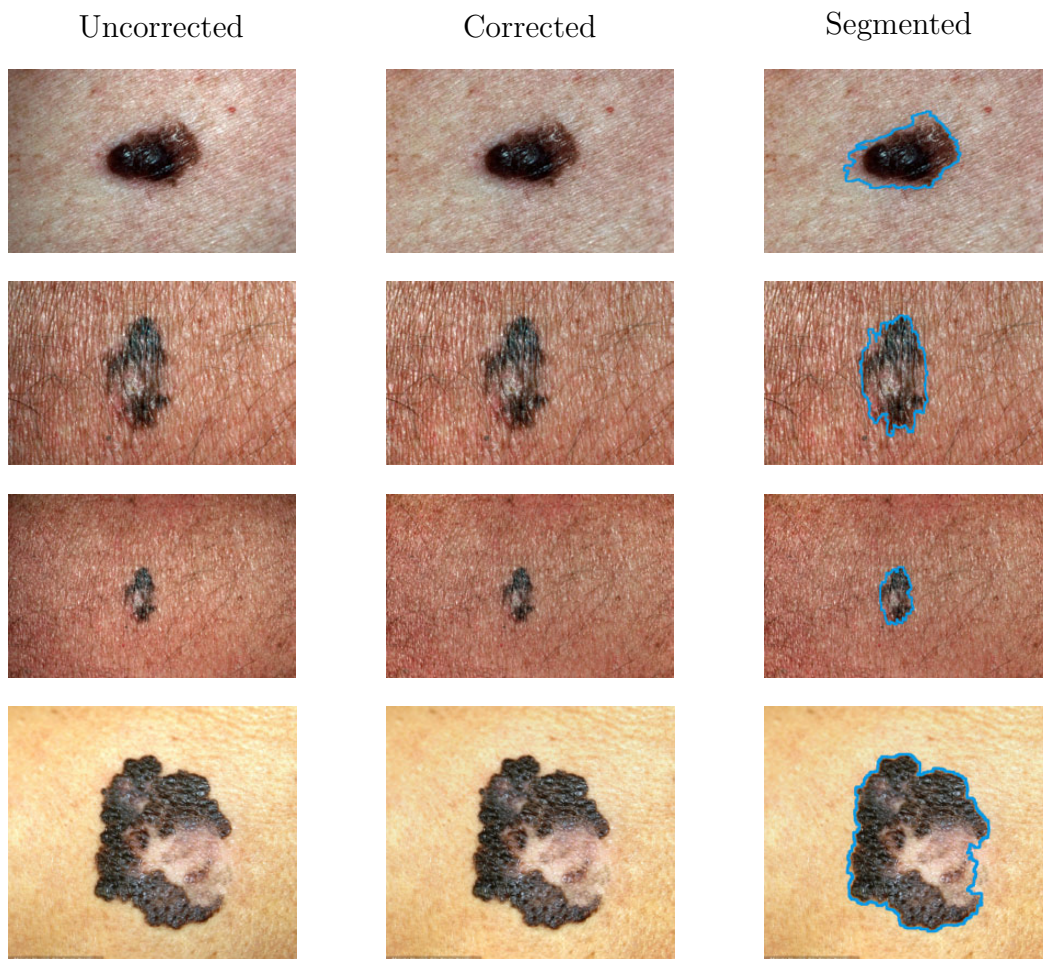


Figure A.1: Uncorrected, illumination-corrected and segmented images of melanoma skin lesions using the proposed algorithm, cont.



## Appendix B

### Illumination Correction and Segmentation Results for Non-Melanoma Images

This appendix contains the 60 digital images of non-melanoma lesions from the DermQuest database [9] that are used to test both the proposed illumination correction algorithm and segmentation algorithm.



Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm.

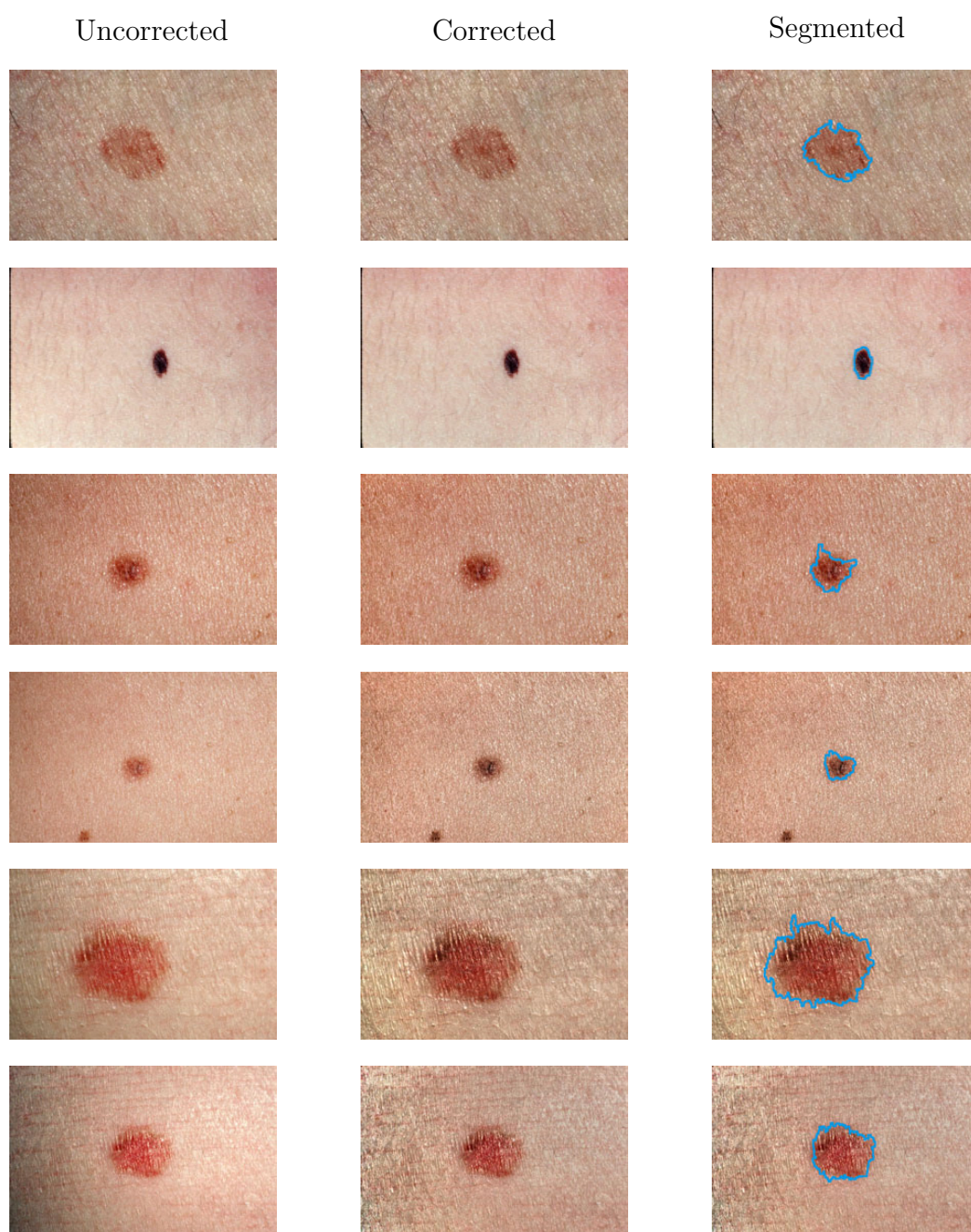


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.

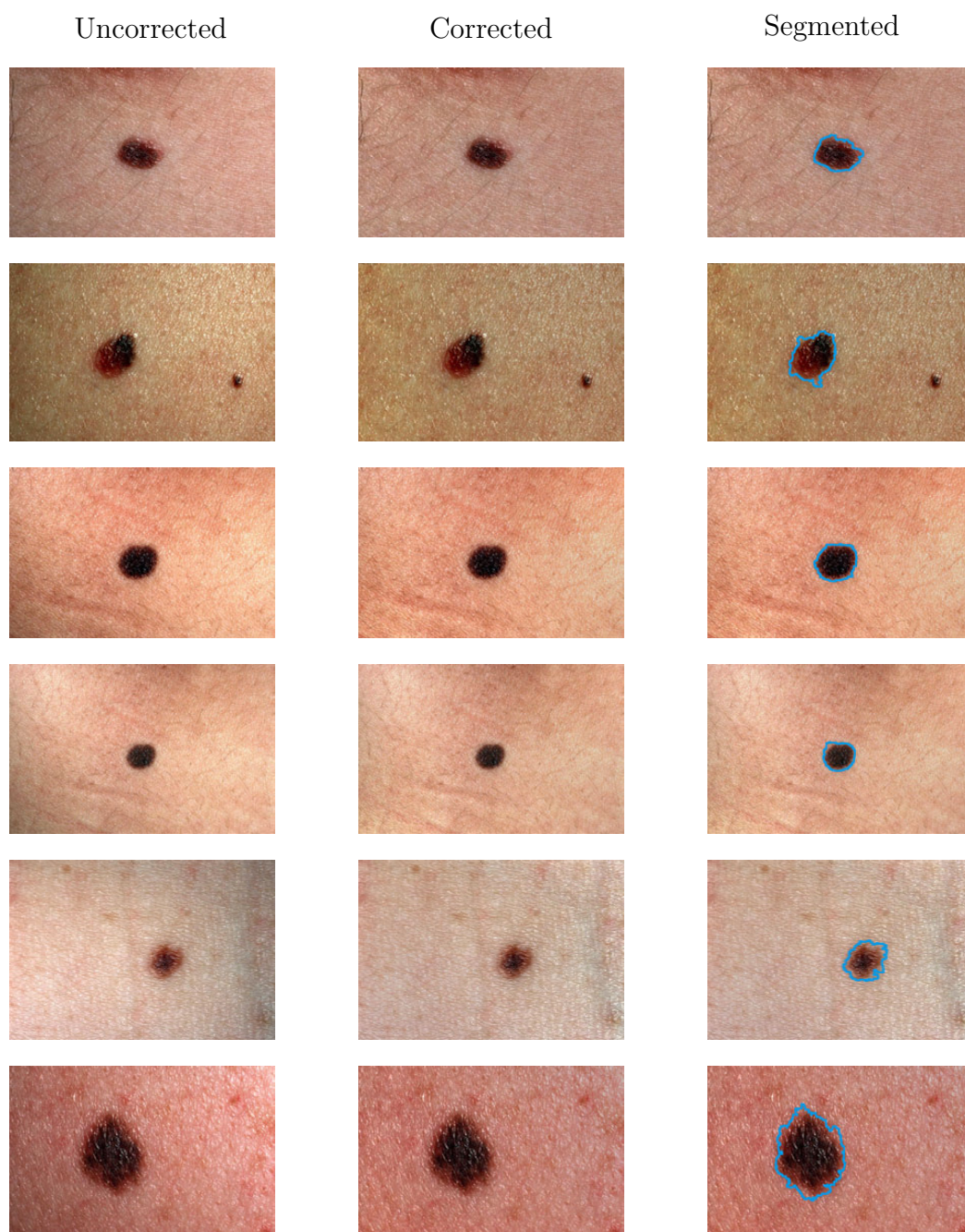


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.





Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.

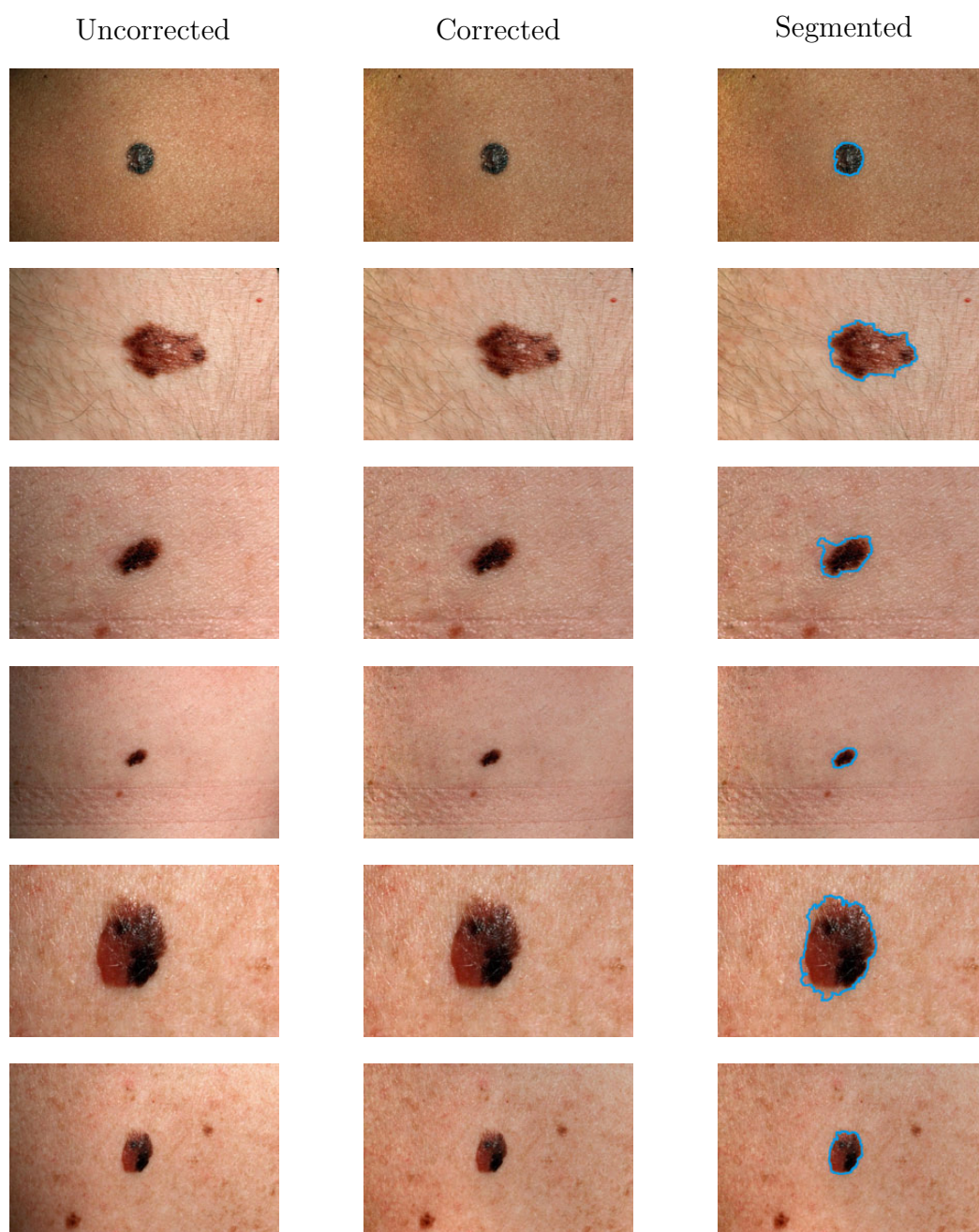


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.



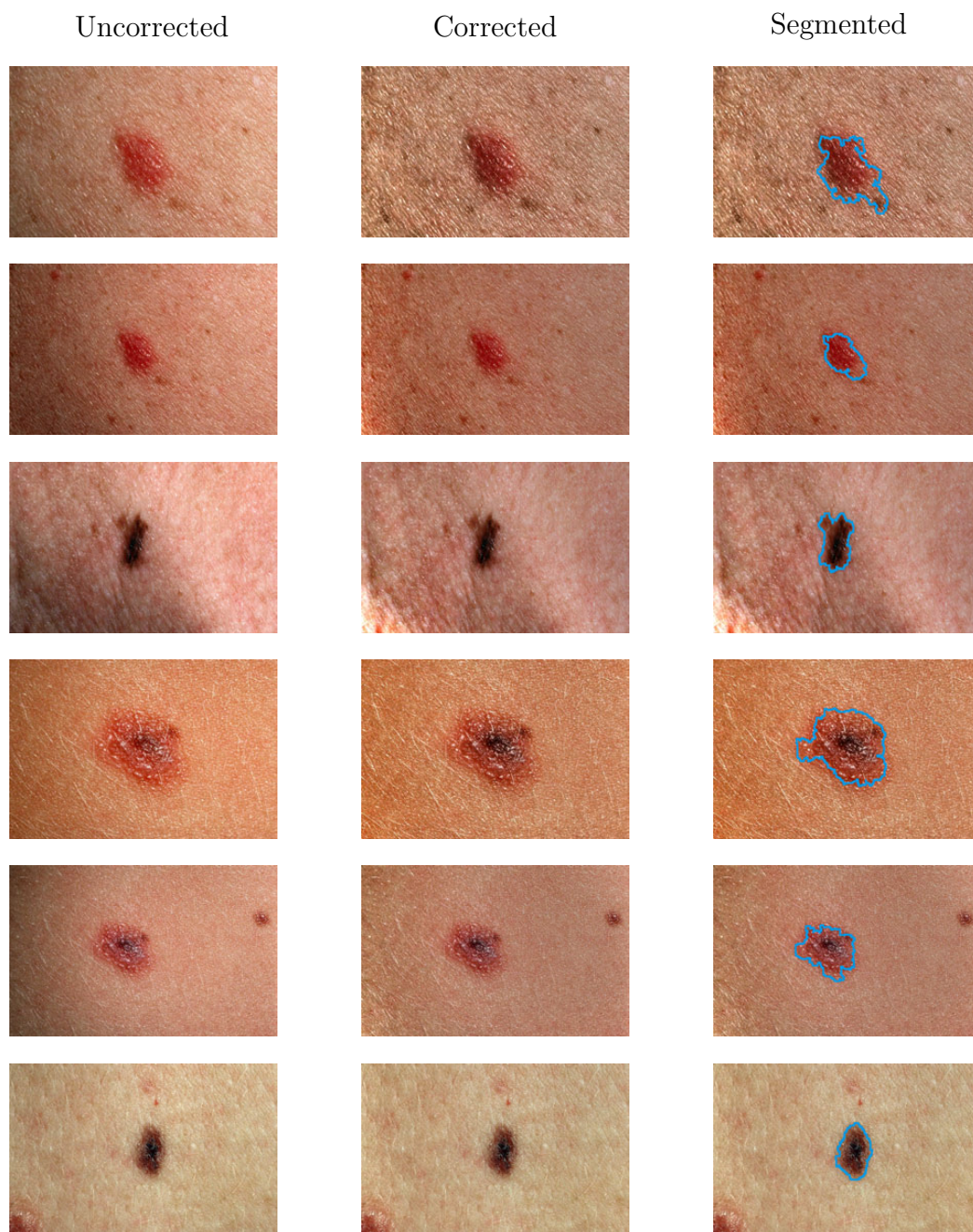


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.

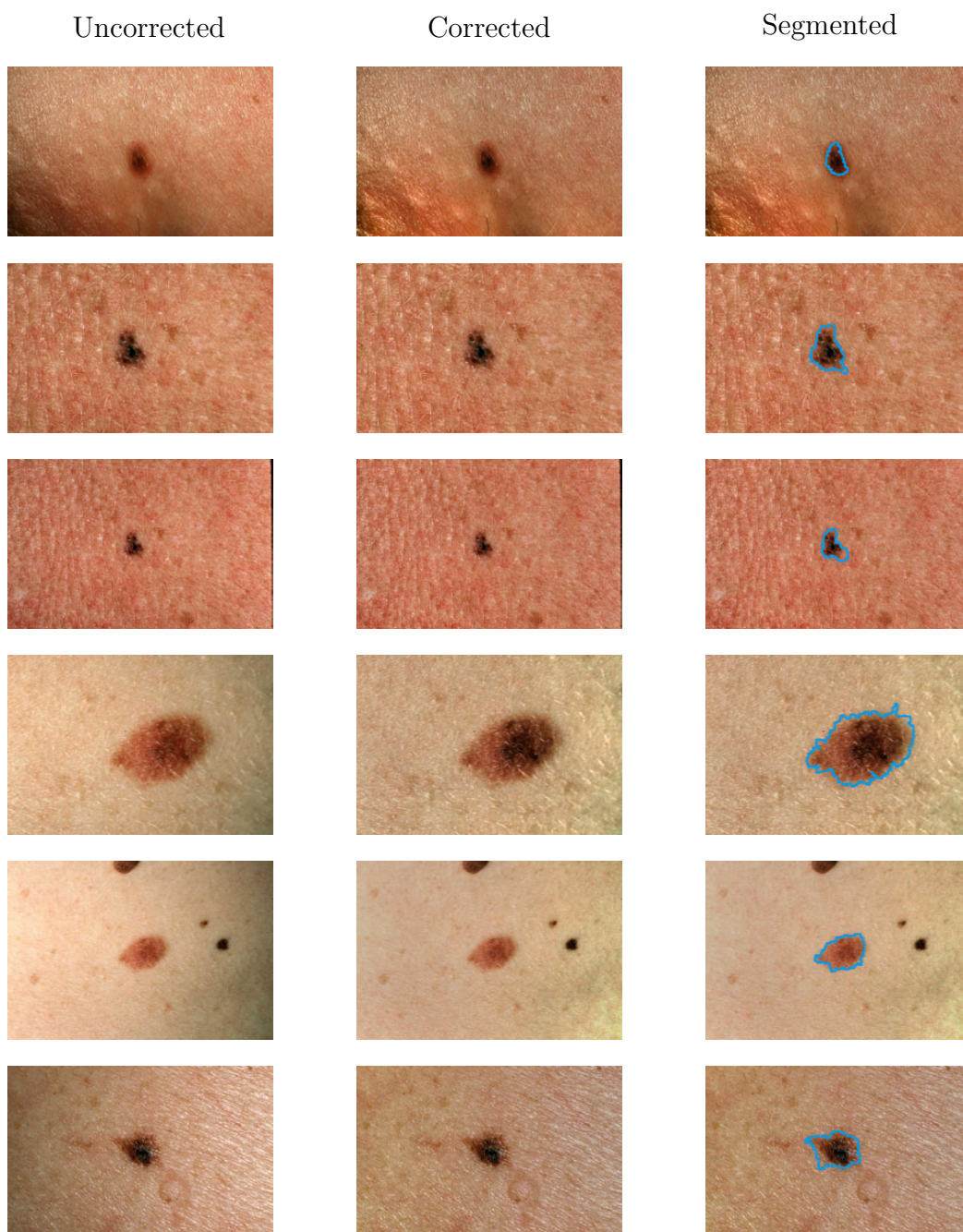


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.





Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.



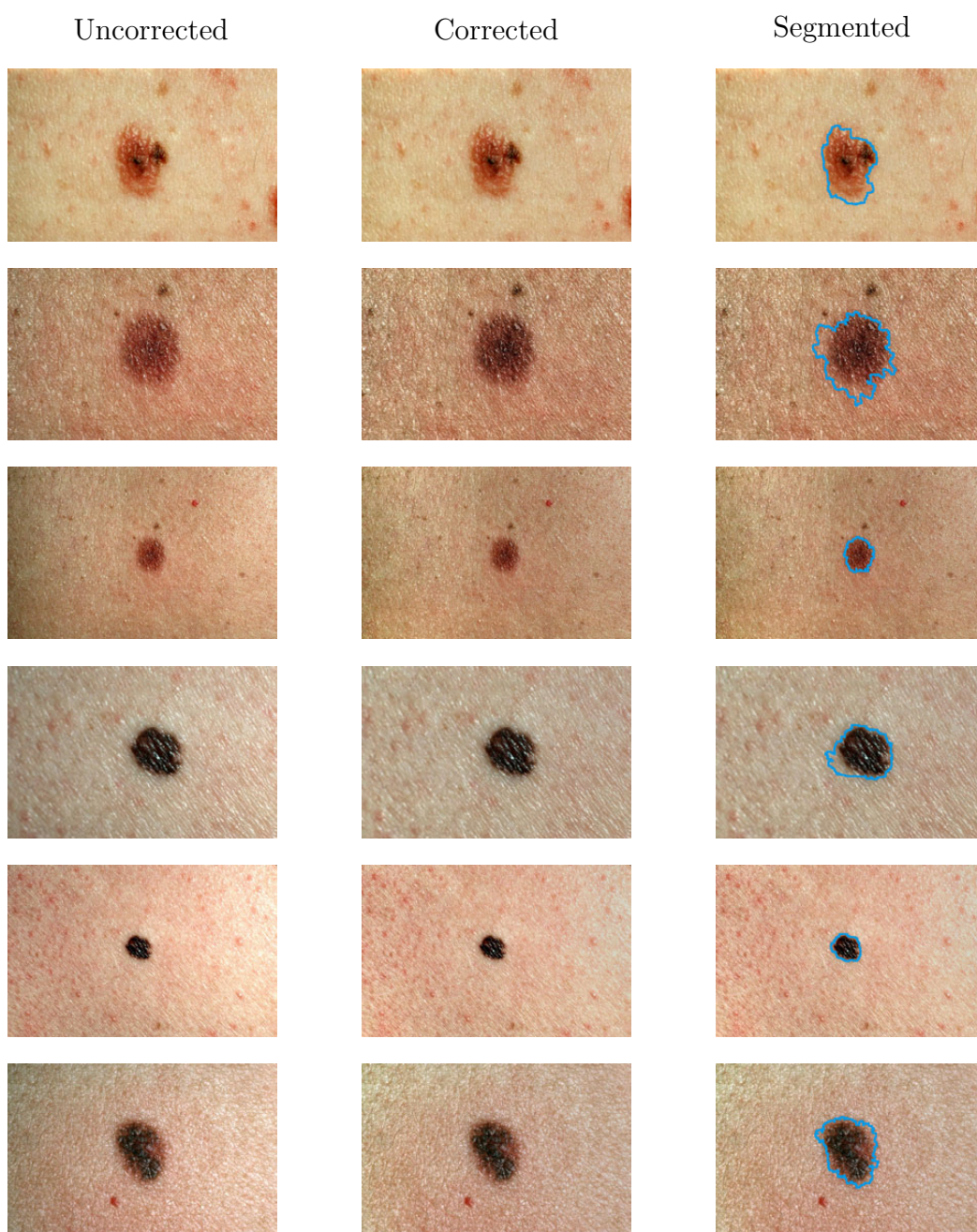


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.

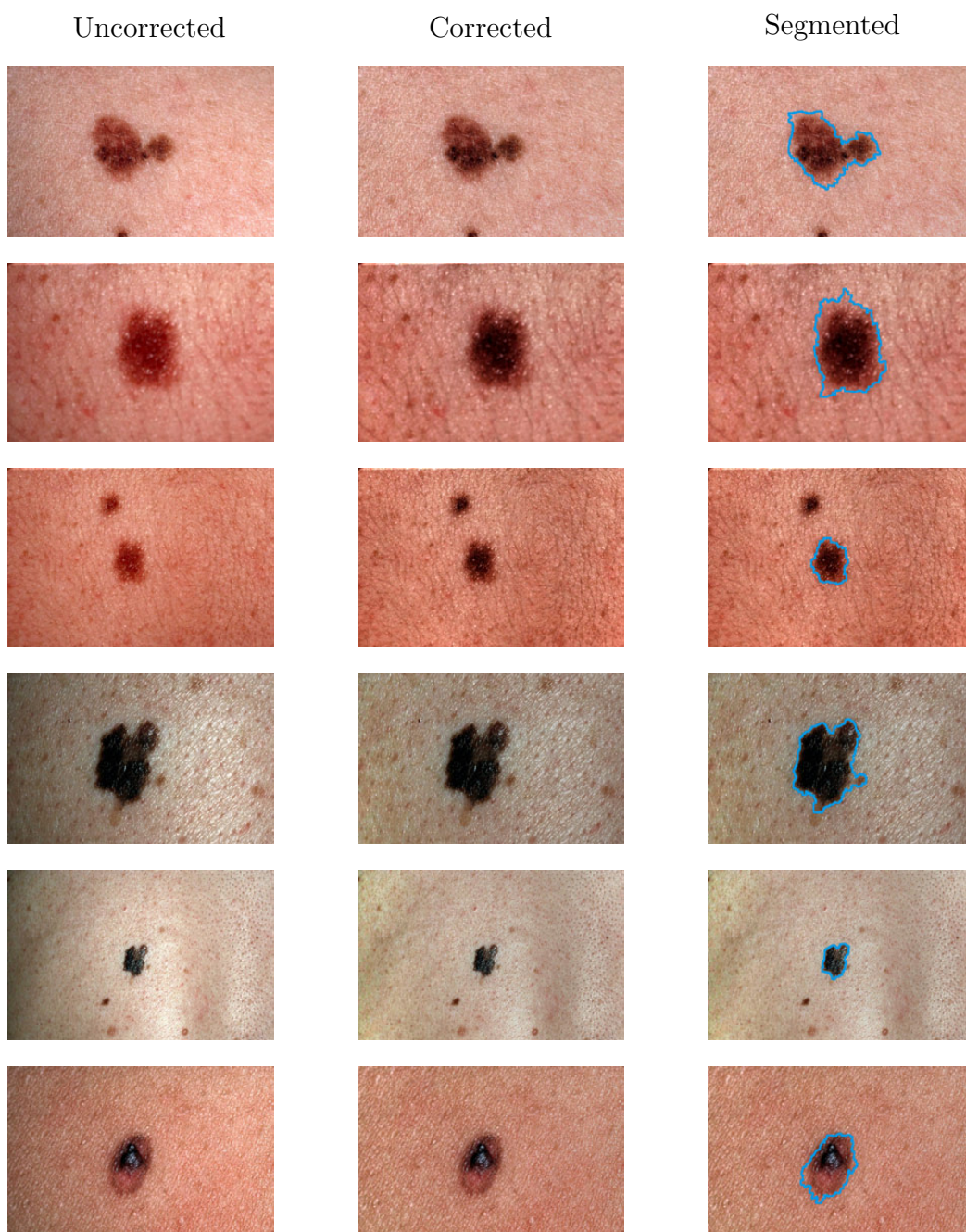


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.



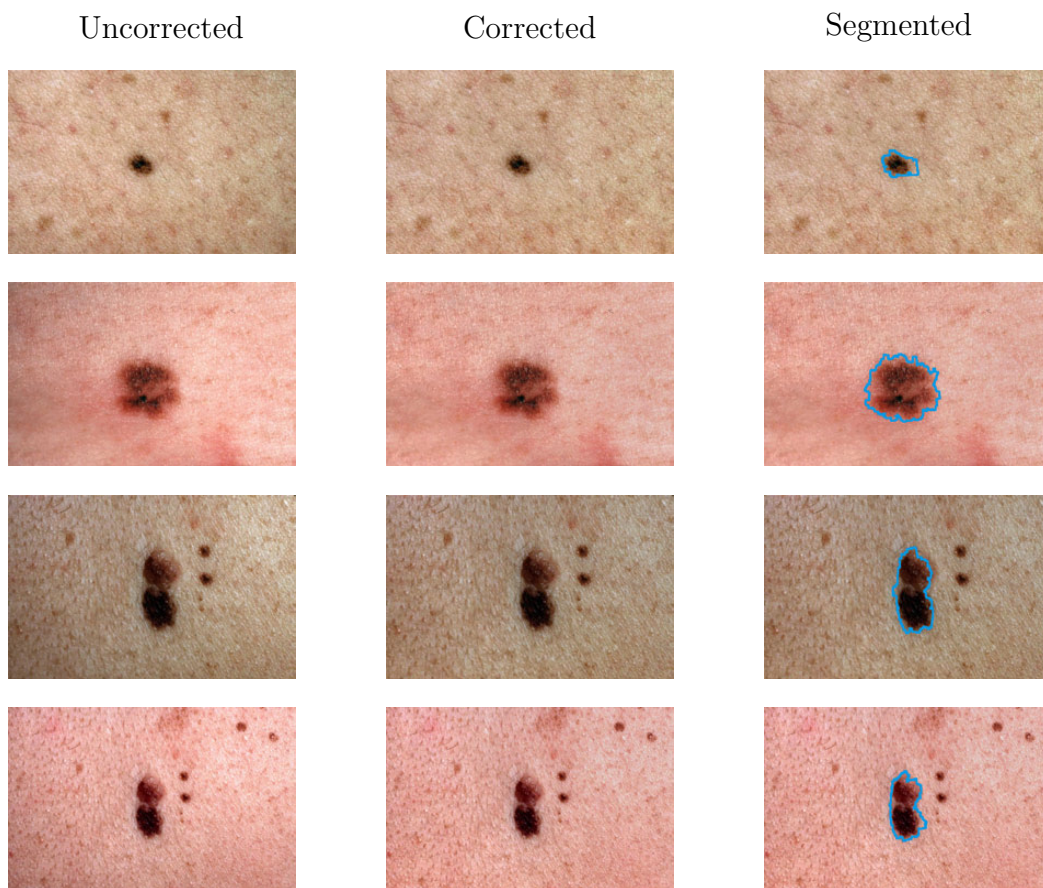


Figure B.1: Uncorrected, illumination-corrected and segmented images of non-melanoma skin lesions using the proposed algorithm, cont.

# References

- [1] Public Health Agency of Canada, “Melanoma skin cancer,” [http://www.phac-aspc.gc.ca/cd-mc/cancer/melanoma\\_skin\\_cancer-cancer\\_peau\\_melanome-eng.php](http://www.phac-aspc.gc.ca/cd-mc/cancer/melanoma_skin_cancer-cancer_peau_melanome-eng.php), 2013, Accessed: 29 Apr 2013.
- [2] T. L. Diepgen and V. Mahley, “The epidemiology of skin cancer,” *British Journal of Dermatology*, vol. 146, no. Suppl. 61, pp. 1–6, 2002.
- [3] A. F. Jerants, J. T. Johnson, C. D. Sheridan, and T. J. Caffrey, “Early detection and treatment of skin cancer,” *American Family Physician*, vol. 62, no. 2, pp. 1–6, July 2000.
- [4] N. Howlader, A. M. Noone, M. Krapcho, J. Garshell, N. Neyman, S. F. Altekruse, C. L. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, H. Cho, A. Mariotto, D. R. Lewis, H. S. Chen, E. J. Feuer, and K. A. Cronin, “SEER cancer statistics review, 1975-2010,” Bethesda, MD, Tech. Rep., 2013.
- [5] A. Jemal, M. Saraiya, P. Patel, S. S. Cherala, J. Barnholtz-Sloan, J. Kim, C. L. Wiggins, and P. A. Wingo, “Recent trends in cutaneous melanoma incidence and death rates in the united states, 1992-2006,” *Journal of the American Academy of Dermatology*, vol. 65, no. 5, pp. S17.e1–S17.e11, Nov 2011.
- [6] C. Herzog, A. Pappo, M. Bondy, A. Bleyer, and J. Kirkwood, *Cancer Epidemiology in Older Adolescents and Young Adults 15 to 29 Years of Age*. Bethesda, MD: National Cancer Institute, 2006, ch. Malignant Melanoma, pp. 53–64.
- [7] K. A. Freedberg, A. C. Geller, D. R. Miller, R. A. Lew, and H. K. Koh, “Screening for malignant melanoma: A cost-effectiveness analysis,” *Journal of the American Academy of Dermatology*, vol. 41, no. 5, pt. 1, pp. 738–745, Nov 1999.

- [8] H. Tsao, G. S. Rogers, and A. J. Sober, "An estimate of the annual direct cost of treating cutaneous melanoma," *Journal of the American Academy of Dermatology*, vol. 38, pp. 669–680, May 1998.
- [9] DermQuest, <http://www.dermquest.com>, 2012, Accessed: 08 Nov 2012.
- [10] R. J. Friedman, D. S. Rigel, and A. W. Kopf, "Early diagnosis of cutaneous melanoma: Revisiting the ABCD criteria," *CA: A Cancer Journal for Clinicians*, vol. 35, no. 3, pp. 130–151, May 1985.
- [11] N. R. Abbasi, H. M. Shaw, D. S. Rigel, R. J. Friedman, W. H. McCarthy, I. Osman, A. W. Kopf, and D. Polsky, "Early diagnosis of cutaneous melanoma: Revisiting the ABCD criteria," *The Journal of the American Medical Association*, vol. 292, no. 22, pp. 2771–2776, Dec 2004.
- [12] R. M. MacKie, "Clinical recognition of early invasive malignant melanoma," *British Medical Journal*, vol. 301, no. 6759, pp. 1005–1006, Nov 1990.
- [13] M. E. Vestergaard, P. Macaskill, P. E. Holt, and S. W. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: A meta-analysis of studies performed in a clinical setting," *British Journal of Dermatology*, vol. 159, no. 3, pp. 669–676, September 2008.
- [14] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, and G. Ferrara, "Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the internet," *Journal of the American Academy of Dermatology*, vol. 48, no. 5, pp. 679–693, 2003.
- [15] H. C. Engasser and E. M. Warshaw, "Dermatoscopy use by US dermatologists: A cross-sectional survey," *Journal of the American Academy of Dermatology*, vol. 63, no. 3, pp. 412–419, 2010.
- [16] E. Losina, R. P. Walensky, A. Geller, F. C. Beddingfield III, L. L. Wolf, B. A. Gilchrest, and K. A. Freedberg, "Visual screening for malignant melanoma: A cost-effectiveness analysis," *Archives of Dermatology*, vol. 143, no. 1, pp. 21–28, Jan 2007.
- [17] N. Cascinelli, M. Ferrario, T. Tonelli, and E. Leo, "A possible new tool for clinical diagnosis of melanoma: The computer," *Journal of the American Academy of Dermatology*, vol. 16, no. 2.1, pp. 361–367, Feb 1987.

- [18] M. Burroni, R. Corona, G. DellEva, F. Sera, R. Bono, P. Puddu, R. Perotti, F. Nobile, L. Andreassi, and P. Rubegni, “Melanoma computer-aided diagnosis: Reliability and feasibility study,” *Clinical Cancer Research*, vol. 10, pp. 1881–1886, March 2004.
- [19] C. Massone, R. Hofmann-Wellenhof, V. Ahlgrimm-Siess, G. Gabler, C. Ebner, and H. P. Soyer, “Melanoma screening with cellular phones,” *PLOS One*, vol. 2, no. 5, p. e483, 2007.
- [20] K. Tran, M. Ayad, J. Weinberg, A. Cherng, M. Chowdhury, S. Monir, M. E. Hariri, and C. Kovarik, “Mobile teledermatology in the developing world: Implications of a feasibility study on 30 egyptian patients with common skin diseases,” *Journal of the American Academy of Dermatology*, vol. 64, no. 2, pp. 302 – 309, 2011.
- [21] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, and R. H. Moss, “A methodological approach to the classification of dermoscopy images,” *Computerized Medical Imaging and Graphics*, vol. 31, no. 6, pp. 362–373, Sept 2007.
- [22] S. W. Menzies, L. Bischof, H. Talbot *et al.*, “The performance of solarscan: An automated dermoscopy image analysis instrument for the diagnosis of primary melanoma,” *Archives of Dermatology*, vol. 141, no. 11, pp. 1388–1396, Nov 2005.
- [23] H. Iyatomi, H. Oka, M. E. Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka, and K. Ogawa, “An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm,” *Computerized Medical Imaging and Graphics*, vol. 32, no. 7, pp. 566 – 579, Oct 2008.
- [24] H. Ganster, A. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler, “Automated melanoma recognition,” *IEEE Transactions on Medical Imaging*, vol. 20, no. 3, pp. 233–239, Mar 2001.
- [25] P. G. Cavalcanti and J. Scharcanski, “Automated prescreening of pigmented skin lesions using standard cameras,” *Computerized Medical Imaging and Graphics*, vol. 35, no. 6, pp. 481–491, Sept 2011.
- [26] J. Alcon, C. Ciuhu, W. ten Kate, A. Heinrich, N. Uzunbajakava, G. Krekels, D. Siem, and G. De Haan, “Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 14–25, Feb 2009.

- [27] S. Shan, W. Gao, B. Cao, and D. Zhao, “Illumination normalization for robust face recognition against varying lighting conditions,” in *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*, 2003, pp. 157–164.
- [28] E. H. Land and J. J. McCann, “Lightness and retinex theory,” *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, Jan 1971.
- [29] B. K. P. Horn, “Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis,” *Computer Graphics and Image Processing*, vol. 3, no. 1, pp. 277–299, Dec 1974.
- [30] J. Frankle and J. J. McCann, “Method and apparatus for lightness imaging,” Patent, 1983.
- [31] P. Soille, “Morphological operators,” B. Jaehne, H. Haussecker, and P. Geissler, Eds., vol. 2.
- [32] M. Elad, “Retinex by two bilateral filters,” in *Scale Space and PDE Methods in Computer Vision*, ser. Lecture Notes in Computer Science, R. Kimmel, N. A. Sochen, and J. Weickert, Eds. Springer Heidelberg, 2005, vol. 3459, pp. 217–229.
- [33] A. Wong, D. A. Clausi, and P. Fieguth, “Adaptive monte carlo retinex method for illumination and reflectance separation and color image enhancement,” in *CRV’09: Canadian Conference on Computer and Robot Vision*, Kelowna, may 2009, pp. 108–115.
- [34] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, “Total variation models for variable lighting face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1519–1524, 2006.
- [35] P. G. Cavalcanti, J. Scharcanski, and C. B. O. Lopes, “Shading attenuation in human skin color images,” in *Advances in Visual Computing*, ser. Lecture Notes in Computer Science, G. Bebis, R. Boyle, B. Parvin, D. Koracin, R. Chung, R. Hammoud, M. Hussain, T. Kar-Han, R. Crawfis, D. Thalmann, D. Kao, and L. Avila, Eds. Springer Heidelberg, 2010, vol. 6453, pp. 190–198.
- [36] Y. V. Haeghen, J. M. A. D. Naeyaert, I. Lemahieu, and W. Philips, “An imaging system with calibrated color image acquisition for use in dermatology,” *IEEE Transactions on Medical Imaging*, vol. 19, no. 7, pp. 722–730, july 2000.

- [37] H. Iyatomi, M. E. Celebi, G. Schaefer, and M. Tanaka, "Automated color calibration method for dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 35, no. 2, pp. 89–98, 2011.
- [38] M. E. Celebi, H. Iyatomi, and G. Schaefer, "Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure," in *ICIP'09: 16th IEEE International Conference on Image Processing*, Cairo, November 2009, pp. 2601–2604.
- [39] G. Schaefer, M. I. Rajab, M. Emre Celebi, and H. Iyatomi, "Colour and contrast enhancement for improved skin lesion segmentation," *Computerized Medical Imaging and Graphics*, vol. 35, no. 2, pp. 99–104, 2011.
- [40] M. Anantha, R. H. Moss, and W. V. Stoecker, "Detection of pigment network in dermoscopy images using texture analysis," *Computerized Medical Imaging and Graphics*, vol. 28, no. 5, pp. 225–234, 2004.
- [41] J. Zhang and T. Tan, "Brief review of invariant texture analysis methods," *Pattern Recognition*, vol. 35, no. 3, pp. 735–747, 2002.
- [42] M. Celebi, H. Iyatomi, G. Schaefer, and W. V. Stoecker, "Lesion border detection in dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 33, no. 2, pp. 148–153, 2009.
- [43] B. Erkol, R. H. Moss, R. Joe Stanley, W. V. Stoecker, and E. Hvatum, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," *Skin Research and Technology*, vol. 11, no. 1, pp. 17–26, 2005.
- [44] M. E. Celebi, H. A. Kingravi, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, R. H. Moss, J. M. Maltes, J. M. Grichnik, A. A. Marghoob, H. S. Rabinovitz, and S. W. Menzies, "Border detection in dermoscopy images using statistical region merging," *Skin Research and Technology*, vol. 14, no. 3, pp. 347–353, 2008.
- [45] W. V. Stoecker, C.-S. Chiang, and R. H. Moss, "Texture in skin images: Comparison of three methods to determine smoothness," *Computerized Medical Imaging and Graphics*, vol. 16, no. 3, pp. 179–190, 1992.
- [46] A. P. Dhawan and A. Sim, "Segmentation of images of skin lesions using color and texture information of surface pigmentation," *Computerized Medical Imaging and Graphics*, vol. 16, no. 3, pp. 163–177, 1992.



- [47] M. Silveira, J. Nascimento, J. Marques, A. R. S. Marcal, T. Mendonca, S. Yamauchi, J. Maeda, and J. Rozeira, "Comparison of segmentation methods for melanoma diagnosis in dermoscopy images," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 1, pp. 35–45, 2009.
- [48] C. Serrano and B. Acha, "Pattern analysis of dermoscopic images based on Markov random fields," *Pattern Recognition*, vol. 42, no. 6, pp. 1052–1057, 2009.
- [49] L. Xu, M. Jackowskia, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, and A. Huntley, "Segmentation of skin cancer images," *Image and Vision Computing*, vol. 17, pp. 65–74, 1999.
- [50] G. Hance, S. Umbaugh, R. Moss, and W. Stoecker, "Unsupervised color image segmentation: with application to skin tumor borders," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 15, no. 1, pp. 104–111, 1996.
- [51] P. Cavalcanti, Y. Yari, and J. Scharcanski, "Pigmented skin lesion segmentation on macroscopic images," in *Image and Vision Computing New Zealand (IVCNZ), 2010 25th International Conference of*, 2010, pp. 1–7.
- [52] P. Cavalcanti, J. Scharcanski, L. Di Persia, and D. Milone, "An ICA-based method for the segmentation of pigmented skin lesions in macroscopic images," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 5993–5996.
- [53] J. Glaister, A. Wong, and D. A. Clausi, "Illumination correction in dermatological photographs using multi-stage illumination modeling for skin lesion analysis," in *EMBC'12: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, August 2012, pp. 102–105.
- [54] R. Amelard, J. Glaister, A. Wong, and D. A. Clausi, "Melanoma decision support using lighting-corrected intuitive feature models," in *Computer Vision Techniques for the Diagnosis of Skin Cancer*, J. Scharcanski and M. E. Celebi, Eds. Springer, *Accepted*, 2013.
- [55] J. Glaister, R. Amelard, A. Wong, and D. A. Clausi, "MSIM: Multi-stage illumination modeling of dermatological photographs for illumination-corrected skin lesion analysis," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 7, pp. 1873–1883, July 2013.

- [56] A. R. Smith, “Color gamut transform pairs,” *SIGGRAPH Comput. Graph.*, vol. 12, no. 3, pp. 12–19, Aug 1978.
- [57] P. Fieguth, *Statistical image processing and multidimensional modeling*. New York, NY: Springer, 2010, vol. 155, p. 65.
- [58] M.-H. Chen, “Importance-weighted marginal bayesian posterior density estimation,” *Journal of the American Statistical Association*, vol. 89, no. 427, pp. 818–824, 1994.
- [59] P. W. Holland and R. E. Welsch, “Robust regression using iteratively reweighted least-squares,” *Communications in Statistics: Theory and Methods*, vol. A6, pp. 813–827, 1977.
- [60] R. Nock and F. Nielsen, “Statistical region merging,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452–1458, nov 2004.
- [61] B. Funt, F. Ciurea, and J. J. McCann, “Retinex in matlab,” *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 48–57, 2004.
- [62] Dermatology Information System, <http://www.dermis.net>, 2012, Accessed: 08 Nov 2012.
- [63] E. Ardizzone, R. Pirrone, and O. Gambino, “Illumination correction on MR images,” *Journal of Clinical Monitoring and Computing*, vol. 20, no. 6, pp. 391–398, July 2006.
- [64] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [65] T. Fawcett, “An introduction to roc analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [66] J. A. Hanley and B. J. McNeil, “The meaning and use of the area under a receiver operating characteristic (roc) curve,” *Radiology*, vol. 143, no. 1, pp. 29–36, April 1982.
- [67] C. Chang and C. Lin, “Libsvm: A library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 27:1–27:27, 2011.
- [68] G. Peyre, “Sparse modeling of textures,” *Journal of Mathematical Imaging and Vision*, vol. 34, no. 1, pp. 17–31, 2009.
- [69] J.-L. Starck, M. Elad, and D. Donoho, “Image decomposition via the combination of sparse representations and a variational approach,” *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1570–1582, 2005.

- [70] C. Scharfenberger, A. Wong, K. Fergani, J. S. Zelek, and D. A. Clausi, “Statistical textural distinctiveness for salient region detection in natural images,” in *IEEE Conference on Computer Vision and Pattern Recognition*, Oregon, June 2013, Accepted.
- [71] M. A. T. Figueiredo, J. M. N. Leitao, and A. K. Jain, “On fitting mixture models,” in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, ser. Lecture Notes in Computer Science, E. R. Hancock and M. Pelillo, Eds. Springer Berlin Heidelberg, 1999, vol. 1654, pp. 54–69.
- [72] J.-C. Terrillon, M. Shirazi, H. Fukamachi, and S. Akamatsu, “Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images,” in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, 2000, pp. 54–61.
- [73] A. K. Jain, R. P. W. Duin, and J. Mao, “Statistical pattern recognition: a review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan 2000.
- [74] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2008.
- [75] R. Amelard, A. Wong, and D. A. Clausi, “Extracting high-level intuitive features (HLIF) for classifying skin lesions using standard camera images,” in *2012 Ninth Conference on Computer and Robot Vision*, Toronto, May 2012, pp. 396–399.
- [76] —, “Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification,” in *EMBC’12: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, August 2012, pp. 4458–4461.
- [77] A. Pelagotti, P. Ferrara, L. Pescitelli, G. Gerlini, A. Piva, and L. Borgognoni, “Noninvasive inspection of skin lesions via multispectral imaging,” *Proc. SPIE 8792, Optical Methods for Inspection, Characterization, and Imaging of Biomaterials*, pp. 87 920C–1–87 920C–13, 2013.