Classification of Pigmented Skin Lesions based on Color Features

Ana Filipa Brissos Afonso

Thesis to obtain the Master of Science Degree in Biomedical Engineering

Examination Committee
Chairperson: Prof. João Pedro Estrela Rodrigues Conde
Supervisors: Prof. Maria Margarida Campos da Silveira
Prof. Teresa Maria de Gouveia Torres Feio Mendonça
Member of the Committee: Prof. João Miguel Raposo Sanches

December 2012
Anybody who has been seriously engaged in scientific work of any kind realizes that over the entrance to the gates of the temple of science are written the words: Ye must have faith.

Max Planck
Acknowledgments

First of all, I would like to mention that this thesis would not have been possible without the help, support and patience of my principal supervisor, professor Margarida Silveira. I am extremely grateful to her for all the effort and invaluable academic and personal level support.

Previously, I had the pleasure to be involved in a R&D fellowship regarding ADDI (Automatic computer-based Diagnosis system for Dermoscopy Images) project. The ideas and advices of professor Margarida Silveira inspired me to continue my research, increase my engineering knowledge and also awakened in me an amazing interest in the area of dermoscopy. Therefore, I chose this subject for my final paper.

It would not have been possible to write this thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here: professor Teresa Mendonça and Doctor Jorge Rozeira for all the information and material provided and Carlos Cabral who helped me most in the final part of the project. Of course, that during those 6 months there were also people that I would like to thank for the good advice, support and friendship they offered me during this time: my family and friends. I would like to address a special thank to Miguel for all the comprehension during the most stressed moments and for sharing with me one of the most important steps in my life!

Above all, I want to give my special thanks to my mother Teresa, my father Artur and my grandmother Luísa for teaching me all the values, made me what I am today, believe in my future success and build my self confidence.

Last and not the least, I would like to acknowledge the financial, academic and technical support of the University Instituto Superior Técnico.

This work was supported by Fundação para a Ciência e Tecnologia (FCT/MCTES) through the ADDI project (PTDC/SAU-BEB/103471/2008).
Abstract

Melanoma is a type of cutaneous cancer which has been leading, in recent years, to a significant increase of mortality rates. Nevertheless, its early diagnosis can prevent the development of the disease. Dermoscopy emerge as a noninvasive technique used to observe dermoscopic lesions which allows a magnification up to 100x. The analysis and classification is usually performed by some medical procedures, such as the ABCD rule and 7-point checklist. The subjectivity inherent to the process required the development of automatic systems for diagnosis used for medical support. One of the main features used for skin lesion classification is the color. The literature usually studies this feature along with others, such as the texture and the shape of the lesion. This thesis focuses on an automatic system for skin lesion classification in melanomas and nevi based on an extraction of color features using parametric and non-parametric methods regarding both the region of the lesion and the skin. Different color spaces are also considered. The database used contains 148 images provided by Hospital Pedro Hispano and ADDI Project. Three classifiers are used: kNN, SVM and GNB. In conclusion, the best results were achieved for the parametric method with a Sensitivity=100%, a Specificity=97.76% and an Accuracy=98.88%. The non-parametric approach also reveals successful results with a Sensitivity=94.03%, a Specificity=99.25% and an Accuracy=96.64%. These performances consider the RGB color space and an extraction of features from the entire image, including both the lesion and skin.

Keywords

Melanoma, Dermoscopy, Classification of dermoscopic lesions, Color analysis, Parametric and Non-Parametric methods
Resumo

O melanoma é um tipo de cancro cutâneo que tem conduzido, nos últimos anos, a um aumento considerável dos níveis de mortalidade. No entanto, o seu diagnóstico precoce pode prevenir o desenvolvimento da doença. A dermoscopia surge como um método não invasivo de observação de lesões dermoscópicas, o qual permite um nível de amplificação até cerca de 100x. A análise e classificação é usualmente realizada com recurso a algoritmos médicos, como a regra ABCD e o método dos 7 pontos. A subjectividade inerente ao processo levou à necessidade de desenvolver métodos automáticos de diagnóstico para auxílio médico. Uma das características mais utilizadas para a classificação de imagens é a cor. Na literatura, esta característica é usualmente abordada em conjunto com outras, como a textura e a forma da lesão. Esta tese propõe um método automático de classificação de lesões dermoscópicas em melanomas e nevos baseado apenas na extração de características de cor, recorrendo a métodos paramétricos e não-paramétricos tendo em conta as zonas da lesão e da pele. Diferentes espaços de cor são também considerados. A base de dados utilizada contém 148 imagens fornecidas pelo Hospital Pedro Hispano no âmbito do projecto ADDI. Três classificadores são utilizados: o kNN, o SVM e o GNB. Em suma, os melhores resultados foram obtidos para o método paramétrico com uma Sensibilidade=100%, Especificidade=97.76% e Precisão=98.88%. A abordagem não-paramétrica também teve bons resultados com uma Sensibilidade=94.03%, Especificidade=99.25% e Precisão=96.64%. Estes desempenhos consideram o espaço de cor RGB e uma extração das características de toda a imagem, incluindo tanto a lesão como a pele.

Palavras Chave

Melanoma, Dermoscopia, Classificação de lesões dermoscópicas, Análise da cor, Métodos Paramétricos e Não Paramétricos
Contents

1 Introduction
  1.1 Motivation ...................................................... 2
  1.2 Skin Lesion Classification ................................. 3
  1.3 Color Analysis ............................................... 4
  1.4 Objectives and Structure ................................. 5
  1.5 Contributions .................................................. 5

2 Dermoscopy Analysis ............................................. 7
  2.1 Dermoscopy ...................................................... 8
  2.2 Medical Procedures ........................................... 8
    2.2.1 ABCD rule .................................................. 9
    2.2.2 7-point checklist ....................................... 10
  2.3 Automatic Systems for Diagnosis ......................... 12
    2.3.1 Segmentation ........................................... 12
    2.3.2 Feature Extraction ..................................... 13
    2.3.3 Extraction of color features ......................... 14
    2.3.4 Feature Selection ..................................... 15
    2.3.5 Classification ......................................... 15

3 Feature Extraction and Selection ................................ 17
  3.1 Preprocessing ................................................ 18
    3.1.1 Otsu's Method ......................................... 18
  3.2 Color Spaces .................................................. 21
    3.2.1 RGB color space ....................................... 21
    3.2.2 HSV color space ....................................... 22
    3.2.3 CIELab color space ................................... 24
    3.2.4 Opponent color space ................................ 25
  3.3 Color Density Estimation using Histograms ............... 26
  3.4 Color Density Estimation using Gaussian Mixture Models .. 27
    3.4.1 Expectation Maximization Algorithm .................. 28
  3.5 Feature Selection ........................................... 31
    3.5.1 Correlation-Based Method .............................. 31
## List of Figures

1. Illustration of (a) a non-melanocytic lesion (basal cell carcinoma) and (b) a melanocytic lesion (melanoma) ................................................. 3

1.2 Classification tree of skin lesions (Adapted from [1]) .......................... 4

1.3 Different types of melanocytic lesions associated with different colors: (a) and (c) illustrate nevi and (b) refers to a melanoma [1] ................. 4

2.1 Melanoma classified by the 7-point checklist method. The presence of atypical pigment network (blue arrow), streaks (black arrow) and blotches (white arrow) led to a final score of 4 (Adapted from [2]) ........................................ 12

2.2 Scheme illustrating an automatic system for diagnosis .......................... 12

3.1 Figure illustrating (a) the original image and the three specific regions of feature extraction: (b) lesion, (c) skin and (d) the entire image except for the dark corners ........................................ 18

3.2 Grayscale image histogram marked at the selected threshold computed based on the Otsu’s method ....................................................... 20

3.3 Figure representing the original image (a) and its black corners (b) ........... 21

3.4 Representation of RGB color cube (Adapted from [3]) .......................... 22

3.5 Representation of HSV color space geometry [4] .................................. 23

3.6 Representation of CIELAB color space geometry [5] .............................. 24

3.7 An “a-b” plane defined by a specific value of luminance L [6] .................. 24

3.8 Image stored in RGB color space ......................................................... 26

3.9 Example of the color distribution of each color channel along the lesion region of an RGB color space ................................................. 26

3.10 Comparison between the histogram and estimated probability density model in each channel in the lesion region. The red color refers to the histogram and blue color to the probability density model ................................................. 31

4.1 Representation of data points being classified by k-Nearest Neighbors classifier. The parameter k equals a value of 3 and the point takes the label of the class that is in majority which is the red one [7] ........................................ 35

4.2 Definition of margin and decision boundary represented by the red line in Support Vector Machine [7] ......................................................... 36
4.3 Figure representing the decision boundary and its relation with the SVM parameters and a data point $x$. ................................................................. 37

4.4 Scheme of the values $\xi_i$ assigned to data points according to its position relative to hyperplane, in which $\xi_i \geq 0$. The circles around some points define the support vectors [7]. ................................................................. 38

4.5 Scheme of Naive Bayes classifier, assuming the conditional independence of features $(x_1, x_2, ..., x_F)$ given a class $t$ [Adapted from [8]]. ................................................................. 39

5.1 Illustration of four dermoscopic images: the lesions depicted on the right side are melanomas while the other belong to nevi class. ................................................................. 42

5.2 Original image of a melanoma and the new one obtained by random rotational transformation. ................................................................. 43

5.3 Probability density function estimated using Gaussian Mixture Model generated by 2 Gaussians compared to color histogram for the red channel of an RGB image considering different quantizations. ................................................................. 46

5.4 Color distribution of dermoscopic images stored in RGB color space, considering the three channels separately. The red color refers to melanomas and blue to nevi. ................................................................. 47

5.5 Color distribution of dermoscopic images stored in HSV color space, considering the three channels separately. The red color refers to melanomas and blue to nevi. ................................................................. 47

5.6 Color distribution of dermoscopic images stored in CIELab color space, considering the three channels separately. The red color refers to melanomas and blue to nevi. ................................................................. 47

5.7 Intensity Histograms of dermoscopic images stored in Opponent color space, considering the three channels separately. The red color refers to melanomas and blue to nevi. ................................................................. 48

5.8 Cases of nevi correctly classified (TN) by kNN classifier. ................................................................. 52

5.9 Cases of melanomas correctly classified (TP) by kNN classifier. ................................................................. 53

5.10 Nevu misclassified as a melanoma (FP) by kNN classifier. ................................................................. 53

5.11 Cases of nevi correctly classified (TN) by SVM classifier. ................................................................. 57

5.12 Melanoma misclassified as a nevus (FN) by SVM classifier. ................................................................. 57

5.13 Cases of melanomas correctly classified (TP) by SVM classifier. ................................................................. 57

5.14 Cases of nevi misclassified as melanomas (FP) by SVM classifier. ................................................................. 57

5.15 ROC curves obtained by GNB classifier for color densities estimated with histograms in the three extraction regions. The symbol (o) refers to the output result of the classifier for each color space expressed in tables 5.25 and 5.26. Color blue refers to RGB; color green to HSV; color red to CIELab and color black to Opponent. ................................................................. 58

5.16 ROC curves obtained by GNB classifier for color densities estimated using GMM in the three extraction regions. The symbol (o) refers to the output result of the classifier for each color space expressed in tables 5.25 and 5.26. Color blue refers to RGB; color green to HSV; color red to CIELab and color black to Opponent. ................................................................. 60
5.17 Cases of nevi correctly classified (TN) by GNB classifier. 61
5.18 Melanoma misclassified as a nevu (FN) by GNB classifier. 61
5.19 Cases of melanomas correctly classified (TP) by GNB classifier. 61
5.20 Cases of nevi misclassified as melanomas (FP) by GNB classifier. 62
5.21 Bar graph of the optimal performance measures achieved by each classifier used to study the color features. 62
5.22 Bar graph of the performance measures achieved by each classifier used to study the color features considering only the region of interest (Lesion). 63
5.23 Example of a nevu misclassified as a melanoma by all classifiers. 64
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of dermoscopic features present in nevi and melanomas (Adapted from [9]).</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Criteria assessed in the ABCD rule used in dermoscopy (Adapted from [1]).</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>Standard Major Criteria in 7-point checklist (Adapted from [1] [10]).</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>Standard Minor Criteria in 7-point checklist (Adapted from [1] [10]).</td>
<td>11</td>
</tr>
<tr>
<td>2.5</td>
<td>Summary of important studies based on Automatic Diagnostic Systems (Database section refers to the number of images analysed and the number of melanomas considered in parenthesis).</td>
<td>16</td>
</tr>
<tr>
<td>3.1</td>
<td>Dimension of histograms used in this work considering the color space and the number of bins.</td>
<td>27</td>
</tr>
<tr>
<td>5.1</td>
<td>Database used after a prior review.</td>
<td>42</td>
</tr>
<tr>
<td>5.2</td>
<td>Database used in this work.</td>
<td>43</td>
</tr>
<tr>
<td>5.3</td>
<td>General Confusion Matrix.</td>
<td>44</td>
</tr>
<tr>
<td>5.4</td>
<td>Dimension of histograms used in this work considering the color space and the number of bins.</td>
<td>45</td>
</tr>
<tr>
<td>5.5</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated with histograms in the region of interest (Lesion).</td>
<td>49</td>
</tr>
<tr>
<td>5.6</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated with histograms in the entire image.</td>
<td>49</td>
</tr>
<tr>
<td>5.7</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated with histograms in the lesion and in the skin extracted separately.</td>
<td>49</td>
</tr>
<tr>
<td>5.8</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated using GMM in the region of interest (Lesion).</td>
<td>50</td>
</tr>
<tr>
<td>5.9</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated using GMM in the entire image.</td>
<td>50</td>
</tr>
<tr>
<td>5.10</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated using GMM in the lesion and in the skin extracted separately.</td>
<td>50</td>
</tr>
<tr>
<td>5.11</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the covariance matrices in the region of interest (Lesion).</td>
<td>51</td>
</tr>
<tr>
<td>5.12</td>
<td>Results of SE, SP and ACC achieved by the kNN classifier considering the covariance matrices in the entire image.</td>
<td>51</td>
</tr>
</tbody>
</table>
5.13 Results of SE, SP and ACC achieved by the kNN classifier considering the covariance matrices in the lesion and in the skin extracted separately. 51

5.14 Best results obtained by the kNN classifier using color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM as color features. 52

5.15 Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated with histograms in the region of interest (Lesion). 54

5.16 Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated with histograms in the entire image. 54

5.17 Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated with histograms in the lesion and in the skin extracted separately. 54

5.18 Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated using GMM in the region of interest (Lesion). 54

5.19 Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated using GMM in the entire image. 55

5.20 Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated using GMM in the lesion and in the skin extracted separately. 55

5.21 Results of SE, SP and ACC achieved by the SVM classifier considering the covariance matrices obtained by GMM in the region of interest (Lesion). 55

5.22 Results of SE, SP and ACC achieved by the SVM classifier considering the covariance matrices obtained by GMM in the entire image. 55

5.23 Results of SE, SP and ACC achieved by the SVM classifier considering the covariance matrices obtained by GMM in the lesion and in the skin extracted separately. 56

5.24 Best results obtained by the SVM classifier using color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM as color features. 56

5.25 Results of SE and SP achieved by GNB classifier for color densities estimated with histograms. 58

5.26 Results of ACC achieved by GNB classifier for color densities estimated with histograms. 58

5.27 Maximum values of SE and SP achieved for color densities estimated with histograms. 58

5.28 Results of AUC achieved by GNB classifier for color densities estimated with histograms. 59

5.29 Results of SE and SP achieved by GNB classifier for color densities estimated using GMM. 59

5.30 Results of ACC achieved by GNB classifier for color densities estimated using GMM. 59

5.31 Maximum values of SE and SP achieved for color densities estimated using GMM. 60

5.32 Results of AUC achieved by GNB classifier for color densities estimated using GMM. 60

5.33 Best results obtained by the GNB classifier using color densities estimated with histograms and color densities estimated using GMM. 61
5.34 Optimal performance measures achieved by each set of features used by kNN, SVM and GNB. ................................................................. 62

5.35 Optimal performance measures achieved by each set of features used by kNN, SVM and GNB considering only the lesion region. ......................................................... 63
Acronyms

**CAD**  Computer-Aided Diagnostic

**ELM**  Epiluminescence Microscopy

**TDS**  Total Dermoscopy Score

**CFS**  Correlation-based Feature Selection

**kNN**  k-Nearest Neighbors

**SVM**  Support Vector Machine

**ANN**  Artificial Neural Network

**LMT**  Logistic Model Trees

**EM**  Expectation-Maximization

**GMM**  Gaussian Mixture Model

**kNN**  k-Nearest Neighbors

**SVM**  Support Vector Machine

**GNB**  Gaussian Naive Bayes

**TP**  True Positives

**FP**  False Positives

**TN**  True Negatives

**FN**  False Negatives

**SE**  Sensitivity

**SP**  Specificity

**ACC**  Accuracy

**ROC**  Receiver Operating Characteristic

**TPR**  True Positive Rate
FPR  False Positive Rate

AUC  Area Under Curve
Introduction

Contents

1.1 Motivation .......................................................... 2
1.2 Skin Lesion Classification ....................................... 3
1.3 Color Analysis ....................................................... 4
1.4 Objectives and Structure ......................................... 5
1.5 Contributions ....................................................... 5
1.1 Motivation

The related concepts of melanoma and dermoscopy play a central role throughout this thesis. Malignant melanoma is a type of skin cancer which results from a proliferation of melanocytes which have the ability to metastasize [1]. These cells are responsible for skin coloration by producing a dark pigment called melanin.

Despite being the less common cutaneous cancer, it is the most dangerous and, in recent years, it has been leading to an increase of mortality in many European countries [11]. However, the diagnosis of melanoma in early stages can prevent the development of the disease and its cure can be possible by a simple surgical excision of this melanocytic lesion. On the other hand, a late detection gets an unfavorable prognosis. Therefore, the research of effective methods that can detect prematurely these type of lesions is a major step in science and in the development of efficient treatments for skin cancer.

Over the last years, dermoscopy revealed to be a fundamental medical technique in the diagnosis of melanomas [1] [12]. It uses optical magnification to allow the observation of morphological structures located in different regions of the skin that are important to differentiate between different types of lesions and are not visible to the naked eye. The practice of dermoscopy requires a trained and experienced physician to allow an efficient analysis of skin lesions. As a noninvasive tool, it provides a way of avoiding the excisions of benign melanocytic lesions.

Nowadays, some medical procedures are used to perform dermoscopic analysis in order to simplify the classification of lesions and also detect malignant melanomas. These algorithms, among which the ABCD rule [1] and 7-point checklist [1] stand out, rely on important image features that are key components in the diagnosis of pigmented skin lesions. The accuracy of dermatologists in detecting melanoma is estimated to be in a range between 75% and 84% [13].

The development of Computer-Aided Diagnostic (CAD) tools has been rising lately in order to assist dermatologists in the detection of melanomas. Various studies have been proposed in the literature which described an extraction of features that are used to train a classifier so that it will be able to differentiate melanomas from other skin lesions. Different types of features related to the structure and border irregularity, shape, texture and color extracted from pigmented skin lesions are usually considered.

Color is probably the most important feature described by the authors due to the variety of colors that a pigmented skin lesion can exhibit. Furthermore, it provides information about the intensity variations present in the skin lesion. An analysis based on color should consider different color spaces which complement the information provided by each other. Also, its invariance with respect to scaling, translation and rotation of an image contributes to an efficient classification of skin lesions and detection of melanomas.

Despite the importance of color, most of the studies consider this feature combined with others. Therefore, a new approach for automatic classification based only on color features is required in order to understand its role in skin lesion classification.
1.2 Skin Lesion Classification

The first step of any dermoscopic analysis is the distinction between melanocytic and non-melanocytic lesions [1]. These two classes differ since melanocytic lesions arise from a proliferation of melanocytes which produce a pigment called melanin. Both classes include malignant and benign lesions. Nevertheless, as melanoma which belongs to melanocytic class is considered the most problematic lesion, it becomes crucial the differentiation between this type of lesion and the remaining.

There are four types of non-melanocytic lesions known by Basal cell carcinoma, Seborrheic keratosis, Vascular lesions and Dermatofibroma which are characterized by specific morphological features.

Basal cell carcinoma is the most incident malignant neoplasm in humans. Nonetheless, as its growth is slow, they are treated as harmless lesions. However, if appropriate treatment is not carried out, they can produce numerous structural destruction in tissues leading to death. The other three lesions are benign from which Dermatofibroma occurs with more regularly, in different parts of the body, mainly extremities. Figure 1.1 illustrates two types of skin lesions.

![Figure 1.1: Illustration of (a) a non-melanocytic lesion (basal cell carcinoma) and (b) a melanocytic lesion (melanoma).](image)

On the other hand, melanocytic lesions are also distinguished between the malignant and benign ones. The coloration that often characterizes a malignant melanoma includes brown and black, but also regions of red, white or blue can appear. The concept of melanoma in situ refers to a stage of neoplasm which is present on epidermis. As it is not in skin dermis, there is no continuity with vascular plexus and it does not have a potential for metastasis.

Nevi are benign melanocytic lesions which are classified as congenital when present at birth or acquired when genetic changes are not their cause. Estimates suggest that about 50% of melanomas arise from pre-existing melanocytic nevi. One of the most important types is Clark nevu as it is considered a relevant precursor of melanoma. It is also called dysplastic or atypical nevi as its appearance is different from the regular ones. Generally, it is large and has irregular borders. Furthermore, its color is not uniform, ranging from pink to dark brown. A summary of different types of skin lesions is represented in figure 1.2.
One of the major challenges in dermoscopy is the distinction between melanomas and other melanocytic lesions. Sometimes, the detection of malignant neoplasms by dermatologists is tough due to the similarities that many melanomas share with nevi. This fact constitutes a motivation to develop CAD systems to allow a more efficient detection of these diseases.

### 1.3 Color Analysis

Color plays an important role in classification processes as each skin lesion has its characteristic colorations that allow a better differentiation among the different types. Melanin is the most important pigment responsible for coloring the melanocytic neoplasms. The color differences observed in different types of lesions is explained by the location of melanin in the skin \[14\]. While color blue is characteristic of melanin located in-depth in the skin, a black color appears on lesions when melanin is located in the outer layers of the epidermis. A red color is related with an increase in vessel dilation or neovascularization. Thus, analysis based on this image property can classify different types of skin lesions and constitutes an important tool for the detection of melanomas in early stages. Figure 1.3 illustrates different colors present in skin lesions.

![Figure 1.3: Different types of melanocytic lesions associated with different colors: (a) and (c) illustrate nevi and (b) refers to a melanoma [1].](image-url)
Considering the importance of color, both medical procedures and automatic methods use this feature in order to classify dermoscopic lesions and accurately detect melanomas. Regarding the ABCD rule, the analysis of color is essential as it is only outweighed by the asymmetry feature. The 7-point checklist method also emphasizes the color as in its major and minor criteria the coloration of pigmentation and morphological structures are considered.

Various studies based on color features extracted from dermoscopic images have been proposed in order to create automatic algorithms for classification [15] [16] [17] [18], achieving successful results. Among the most common color descriptors are mean and standard deviation computed over one image channel [15] [17] [19], histograms [15] [20], the maximum and minimum value of vector components of each channel [13] [21] and covariance matrices [22]. Some of these features are extracted from different color spaces in order to increase the robustness of the results.

Therefore, an automatic classification of dermoscopic lesions based on color features and posterior evaluation regarding the diagnosis performed by the dermatologist is an important tool for the detection of melanomas.

1.4 Objectives and Structure

The aim of the thesis is to develop a system capable of detecting melanomas based on different color features and understand the role of color in pigmented skin lesion classification. To achieve these objectives, an approach based on three steps was followed:

- **First Step:** Extraction of features;
- **Second Step:** Selection of the most relevant features for classification;
- **Third Step:** Classification of pigmented skin lesions into one of two classes, melanoma (1) and nevu (0).

As it was mentioned, skin lesions can assume a variety of colors, such as blue, black, brown or red, hence features based on color are crucial in classification processes.

This thesis is organized as follows: Chapter 2 presents a review of the state of the art related to the analysis of dermoscopic images in order to understand the studies that have been done in the area as well as some medical procedures and computational methods for the diagnosis of cutaneous lesions. Chapter 3 describes the methods used to extract different color features and explains how the selection of the most relevant features is performed in order to reduce data dimensionality. Chapter 4 describes three well-known classifiers used to label images. Chapter 5 presents experimental results and Chapter 6 addresses the main conclusions and future work.

1.5 Contributions

This thesis provides a new approach to the classification of melanocytic lesions present in dermoscopic images. Based on a set of extracted features related to color, it is possible to distinguish melanomas from nevi. The main contributions of this work are summarized as follows:
• An approach based only on color features is important as much of the state of the art considers color information combined with other features, such as texture, shape and structure. This thesis focuses only on color information provided by the intensity values of the three image channels simultaneously and uses color distribution and covariance matrices as features.

• This work considers four color spaces obtained by the conventional RGB format. An extraction of features is performed in each of the color spaces in order to conclude which of the four achieves better results in the classification process.

• Three types of color features extracted using parametric and non-parametric methods were considered in this thesis. The goal was to compare different features and discuss its relevance to the classification of lesions as previous studies did not perform this analysis.

• The classification based on more than one classifier is also important as each one considers a different model to relate features and consequently achieves different results and conclusions.

• The most common studies only perform feature extraction in the region of interest (lesion). This thesis considers not only the lesion but also the skin in order to understand its influence in the lesion classification.
2

Dermoscopy Analysis

Contents

2.1 Dermoscopy ................................................................. 8
2.2 Medical Procedures ..................................................... 8
2.3 Automatic Systems for Diagnosis ................................. 12
2.1 Dermoscopy

Dermoscopy, also called Epiluminescence Microscopy (ELM), is a noninvasive method for the in vivo detection of pigmented skin lesions. As it provides an optical magnification of skin structures and substructures that are not observed by the naked eye, it is useful in the early diagnosis of melanomas [1] [12]. Since the research conducted by Kolhaus in 1663 on the small vessels in the nailfold using skin surface microscopy, this technique has become essential in the diagnostic over clinical visual inspection [14].

This diagnostic tool provides a way of acquiring dermoscopic images of skin lesions. The first step consists in placing various immersion liquids as mineral oil, alcohol or even water, under the region of interest. This preparation stage allows the elimination of reflection and renders the lesion more translucent, avoiding the light to be dispersed, absorbed or reflected. The following step takes into account an appropriate instrument, such as a hand-held lens, a dermatoscope, a stereomicroscope, a camera or a digital imaging system [1]. The traditional dermatoscope is the most widely used as it provides a relatively high magnification in addition to its low costs, rapidity and simple use. However, the use of digital imaging systems is also becoming quite popular as it allows an easy storage, retrieval and follow-up of dermoscopic lesions [14]. The goal is to inspect the lesion accurately as the magnification of these instruments range from 6x up to 100x. Dermatologists can analyse the region of interest and its morphological components in detail in order to correctly differentiate between the malignant and benign skin lesions and consequently achieve a successful diagnosis.

Recently, several studies proved that 65% to 80% of melanomas can be detected efficiently by clinical inspection and evaluation performed by dermatologists while others have shown that an increase of melanoma detection accuracy of 10%-27% is achieved by dermatologists in comparison with the naked eye evaluation [1]. Nevertheless, although it is demonstrated that dermoscopy increases diagnostic accuracy, it highly depends on the experience of the physician as some skin lesions can be underdiagnosed leading to a dangerous misclassification.

2.2 Medical Procedures

The dermoscopy technique allows the enhancement of different dermoscopic features and structures that can be analysed by some medical approaches: Pattern Analysis, ABCD rule, 7-point checklist, Menzies method, C.A.S.H (Color, Architecture, Symmetry, Homogeneity) and ABCDE rule (Asymmetry, Border, Color, Diameter and Evolution) which are used by dermatologists in order to efficiently distinguish between malignant and benign lesions [1] [14] [23].

The dermoscopic features can be divided into global and local [1]. The global features constitute more general morphological patterns which are predominantly distributed throughout the region of the lesion. This class includes: reticular pattern which is the most common global feature as well as globular, cobblestone, homogeneous and starburst pattern, among others. On the other hand, the local features are present in specific regions which allows a more detailed diagnosis of the lesion. Pigment network, dots, globules and hypopigmentation are some of the local patterns considered in
The most classic diagnostic approach is known by pattern analysis which was first described by Perhamberger et al. in 1987 [1]. This method is based on the assessment of specific criteria and includes two steps. In the first one, the physician needs to identify if the lesion belongs to melanocytic or non-melanocytic class. This classification is mainly based on global features. The second step consists in the distinction between malignant and benign lesions, taking into account the melanocytic lesions selected in the previous step. This analysis is based on global and local features. The following table presents some of the dermoscopic features considered in this method. A comparison between nevi and melanoma is also illustrated.

**Table 2.1:** Comparison of dermoscopic features present in nevi and melanomas (Adapted from [9]).

<table>
<thead>
<tr>
<th>Dermoscopic features</th>
<th>Atypical Nevi</th>
<th>Cutaneous Melanoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pigmented network</td>
<td>Irregular, discrete, locally proeminent, abruptly ends or thins at periphery</td>
<td>Irregular, proeminent, wide, abruptly ends or thins at periphery</td>
</tr>
<tr>
<td>Diffuse pigmentation</td>
<td>Irregular, intense, inhomogeneous, center, abruptly ends at periphery</td>
<td>Irregular, inhomogeneous, abruptly ends or thins at periphery</td>
</tr>
<tr>
<td>Depigmentation</td>
<td>Irregular and periphery</td>
<td>Irregular, bizarre, pink-and-white center and periphery</td>
</tr>
<tr>
<td>Brown globules</td>
<td>Varied in size and shape, irregularly distributed</td>
<td>Often present, varied in size and shape, irregularly distribut</td>
</tr>
<tr>
<td>Black dots</td>
<td>Rare, regularly distributed throughout lesion</td>
<td>Often present, varied in size and shape, irregularly distribu</td>
</tr>
<tr>
<td>Radial streming</td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>Pseudopods</td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>Gray-blue veil</td>
<td>Absent</td>
<td>Present</td>
</tr>
</tbody>
</table>

Despite the increase in the accuracy rate by 10% to 30% of correctly diagnosed lesions achieved by this method, some problems related to the reliability and reproducibility inherent to the subjectivity of the process required the development of new algorithms [14] [2]. ABCD rule and 7-point checklist are two common medical procedures which require a previous step to select melanocytic lesions in order to distinguish between melanomas and other benign lesions.

### 2.2.1 ABCD rule

The ABCD rule [1] is a semiquantitative method which allows the differentiation between melanoma and other types of melanocytic lesions. It was first described by Stolz et al. in 1994, becoming a simple and objective analysis technique. As the name suggests, it is based on four image features, Asymmetry (A), Border (B), Color (C) and Differential Structures (D). For each one, a score is assigned in order to compute a parameter called Total Dermoscopy Score (TDS) [1]. Furthermore, it is necessary to take into account different weight factors depending on each property as described in the following equation:
The assessment of asymmetry requires the division of the lesion into two perpendicular axes with specific orientation in order to obtain the lowest score. If the lesion is totally symmetric regarding its contour, colors and differential structures, it is assigned a score of 0. On the other hand, if the asymmetry is present in both axes, the score is 2 [24]. To analyse the border, the lesion is divided into eighths equally spaced by four axes crossed in a center point. Depending on the transition between the lesion pattern and the skin surrounding it, a score will be assigned. An abrupt transition results in a score of 1, otherwise it is assigned a score of 0. At the end, the final score for border will be the sum of each value assigned to every eighth. The colors considered in this analysis are white, red, light-brown, dark-brown, blue-gray and black. The score assigned to this feature is the number of those colors present in the lesion. Finally, the differential structures are evaluated in a similar way. Based on five structures: pigment network, structureless or homogeneous areas, streaks, dots and globules, the score is the number of structures present in the lesion. The table below describes the criteria used in the ABCD rule.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
<th>Score</th>
<th>Weight Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymmetry</td>
<td>In 0, 1, or 2 axes; assess not only contour, but also colors and structures</td>
<td>0-2</td>
<td>1.3</td>
</tr>
<tr>
<td>Border</td>
<td>Abrupt ending of pigment pattern at the periphery in 0-8 segments</td>
<td>0-8</td>
<td>0.1</td>
</tr>
<tr>
<td>Color</td>
<td>Presence of up to six colors 1-6 (white, red, light-brown, dark-brown, blue-gray, black)</td>
<td>1-6</td>
<td>0.5</td>
</tr>
<tr>
<td>Differential Structures</td>
<td>Presence of network, structureless or homogeneous areas, streaks, dots, and globules</td>
<td>1-5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The value of TDS which is calculated based on these criteria will allow the correct diagnosis of melanocytic lesions. It should be noted that TSD values lower than 4.75 refers to a benign melanocytic lesion while values higher than 5.45 indicates that there is a high probability of being a melanoma. For values between the two thresholds, it is considered a suspicious lesion.

### 2.2.2 7-point checklist

The 7-point checklist [1] is another method widely used among the medical community to diagnose skin melanocytic lesions. As a reliable technique, it has been used even by the less experienced clinicians who have been able to detect melanomas with a high success rate [1]. It is based on specific morphological features which are common among melanomas. These standard criteria can be divided into two groups: the major and minor criteria. Those features which belong to the first class are very important in the diagnosis of melanoma, contributing to the final score with 2 points. On the
other hand, the minor features contribute with 1 point if they are present or 0 in case of absence. A summary of the major and minor criteria is represented in the following tables.

Table 2.3: Standard Major Criteria in 7-point checklist (Adapted from [1] [10]).

<table>
<thead>
<tr>
<th>Major Criteria</th>
<th>Definition</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical pigment network</td>
<td>Black, brown, or gray network with irregular meshes and thick lines</td>
<td>2</td>
</tr>
<tr>
<td>Blue-whitish veil</td>
<td>Confluent, gray-blue to whitish-blue diffuse pigmentation associated with pigment network alterations, dots/globules and/or streaks</td>
<td>2</td>
</tr>
<tr>
<td>Atypical vascular pattern</td>
<td>Linear, dotted or globular vessels not clearly combined with regression structures and associated with pigment network alterations, dots/globules and/or streaks</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.4: Standard Minor Criteria in 7-point checklist (Adapted from [1] [10]).

<table>
<thead>
<tr>
<th>Minor Criteria</th>
<th>Definition</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irregular streaks</td>
<td>Irregular, more or less confluent, linear structures not clearly combined with pigment network lines</td>
<td>1</td>
</tr>
<tr>
<td>Irregular pigmentation</td>
<td>Black, brown, and/or gray pigmented areas with irregular shape and/or distribution</td>
<td>1</td>
</tr>
<tr>
<td>Irregular dots/globules</td>
<td>Black, brown, and/or gray round to oval structures irregularly distributed within the lesion</td>
<td>1</td>
</tr>
<tr>
<td>Regression structures</td>
<td>White areas (white scarlike areas) and blue areas (gray-blue areas, peppering, multiple blue-gray dots) may be associated</td>
<td>1</td>
</tr>
</tbody>
</table>

The color is the main feature analysed in the major criterion: blue-whitish veil and in the minor criteria: irregular pigmentation and regression structures, reinforcing its importance in the classification of skin lesions [10].

To compute the final score, it is necessary to sum all the scores assigned to each criterion and analyse the result. If the value obtained is higher than or equal to 3, the lesion is classified as a melanoma, otherwise it is considered a nevu. Therefore, the lesion must contain at least one of the following combinations of features to be a melanoma: one major feature and one minor feature, two major features or three minor features. An example of a skin lesion containing specific morphological features which led to its classification as being a melanoma is present in the figure 2.1.
2.3 Automatic Systems for Diagnosis

The development of CAD systems emerged in order to overcome the problems of subjectivity, reliability and reproducibility inherent to the human clinical inspection. These systems provide not only the possibility of non-experienced dermatologists to use them but also constitute a second diagnostic tool to assist the experts in the classification of dermoscopic lesions. Since 1987, when Cascinelli et al. carried out a study related to the automated skin lesion classification, this area has been widely investigated in order to improve its accuracy [15].

An automatic system for diagnosis usually follow a standard approach based on three steps to classify dermoscopic lesions. Firstly, an automatic segmentation of the lesion is performed. From this segmentation results a binary image which identifies the pixels belonging to the lesion. The following step includes both the feature extraction and selection. In order to reduce data dimensionality, the feature selection choose a subset of the most relevant features which constitute the input of the classifier. The classification step includes two phases. Initially, the classifier is trained based on a set of images previously diagnosed by an experienced dermatologist which constitute the training set. Then, the classifier is used in order to label each image from the test set as a melanoma (1) or a nevu (0). Figure 2.2 summarizes the general steps included in automatic systems.

![Figure 2.2: Scheme illustrating an automatic system for diagnosis.](image)

2.3.1 Segmentation

In figure 2.2, the segmentation of the lesion is the first step. The concept of segmentation is related to the partition of the image into different regions. The pixels belonging to each group share specific properties.
characteristics as color, texture, intensity, among others. Thus, this step is very important since errors that may occur at this level can affect the posterior results leading to a misclassification of lesions.

K. A. Norton et al. [25] investigated a method to segment melanocytic and non-melanocytic lesions based on the Otsu thresholding applied to the green channel of RGB images. It resulted in a binary image since each pixel was labeled with the value 0 or 1 when compared to a threshold. Two other steps were followed in order to improve the result: correction of non-uniform lightning and noise reduction. The proposed algorithm was tested in 107 non-melanocytic and 319 melanocytic lesions and achieved sensitivity and specificity values of 84.5% and 88.5% when applied to non-melanocytic lesions and 93.9% and 93.8% for melanocytic lesions which proved to be an accurate segmentation technique.

Other methods of segmentation were developed in order to become the classification more efficient. A comparative study conducted by S. S. Varshney et al. [26] revealed that K-means Clustering is a more efficient segmentation algorithm when compared to Edge-detection, Multiple Thresholding and Region-based methods. K-means algorithm iteratively analyses each pixel and assigned it to one of the K clusters that minimizes the variance between its centroide and the pixel. On the other hand, the Edge-detection method recognizes pixels which represent discontinuities or describe abrupt changes in image brightness. The Region-based method considers an initial pixel which belongs to the lesion. A final region is created by grouping new pixels to the initial one while the homogeneity is verified.

The study performed by A. Parolin et al. [16] used snake segmentation technique whose flexibility leads to an accurate contour while S. O. Skrøvseth et al. [11] developed three segmentation algorithms from which stands out the histogram analysis. To conclude, an article described by M. E. Celebi et al. [15] revealed that JSEG algorithm introduced by Y. Deng et al. [27] is an efficient method to segment lesions since in only 32 images, out of a dataset with 596, it produced a less accurate contour.

2.3.2 Feature Extraction

The segmentation process is followed by the feature extraction. Some interesting studies use approaches related to medical procedures in order to classify the lesions. J. F Alcón et al. [21] followed the ABCD rule and describe an extraction of asymmetry, border, color and texture features. G. D. Leo et al. [10] focused on specific features related to the texture, color and structure of the lesion to create a CAD sytem based on the 7-point checklist method. In [18], Q. Abbas et al. based on color symmetry and multiscale texture analysis in order to measure the color, architectural order, symmetry of pattern and homogeneity of lesions and follow another medical approach called C.A.S.H.

The literature also describes an extraction of features related to the color, shape, structure and texture of lesions which do not intend to follow the medical approaches [13] [15] [17]. Although H. Iyatomi et al. [13] and M. E. Celebi et al. [15] extracted the same types of features, their purposes are different. In [13] the goal is to differentiate between melanocytic and non-melanocytic lesions while in [15] the intend is to detect melanomas which belongs to melanocytic class. Relatively to texture,
both studies focused on co-occurrence matrices which express the relation between pixels and their neighbors. On the other hand, while in [15], the authors used four different features to characterize the shape of lesions: area, aspect ratio, asymmetry and compactness, in [13] only the shape asymmetry is used to classify its database. K. Ramlakhan et al. [17] described in their study an extraction of other features related to the shape of the lesion: significant defects, a circularity index and a hull/contour ratio which measure the border irregularity. The color features extracted in the previous studies will be discussed in detail in the following section.

2.3.3 Extraction of color features

This thesis focuses on a classification of melanocytic lesions based on color features. Thus, a review of the state of the art must be performed in order to understand which features of color were studied in previous works. Despite the importance of color in classification of dermoscopic images, it was studied by many authors along with other common features.

The information provided by different color spaces can be crucial in a successful classification process. Thus, M. E. Celebi et al. [15] studied six different color spaces: RGB, rgb (RGB normalized), HSV, CIELuv, I1/2/3 (Ohta space) and I1/2/3 which try to complement each other in three specific properties: invariance to illumination intensity, decoupling of chrominance and luminance and perceptual uniformity. Considering these six color spaces, four color features were extracted. The mean and standard deviation and centroidal distances were extracted from all color spaces. The color asymmetry feature was only extracted from RGB color space and CIELuv histogram distances computed from histograms coarsely quantized were performed in CIELuv color space. Later on, D. Nie et al. [20] proposed a system based on RGB histogram features in order to describe color. Those features were extracted from each relative color image which result from a subtraction between the average RGB values of its skin-only pixels and its tumor-only pixels.

On the other hand, J. F. Alcón et al. [21] quantified the color information of their CAD system based on the ABCD rule by computing the Euclidean distance between each pixel and each of the six colors considered by this rule. The color space analysed was RGB and each of those colors presents specific values of R, G and B channels. To enhance the system, the intensity values of each pixel in the three channels were stored in a vector and specific measures were computed: mean, variance, maximum and minimum values. Also, A. Parolin et al. [16] created an automatic system based on the ABCDE rule whose color was analysed in RGB and CIELab color spaces. Three color features were extracted: RGB color variance, CIELab mean values and relative chromaticity which was computed for each RGB color channel. The relative chromaticity measures the color of the lesion relative to the surrounding skin.

F. Ercal et al. [29] described the color features extracted from three different color spaces. Regarding the RGB color space, the variance and relative chromaticity were computed for each channel. Based on the equations to transform RGB into spherical coordinates, the brightness and two angles were computed for each pixel in the lesion and an average was taken. Also, an average of the lightness, hue and chroma of the tumor area was considered.
2.3.4 Feature Selection

A selection step usually precedes the classification process in order to remove redundant and irrelevant features as well as reduce the dimensionality of data. Different criteria for the selection of the most important features have been used. One of the most common methods is the filter which rely on general properties of data, not involving any learning algorithm. On the other hand, wrapper method uses the prediction performance of a learning algorithm in order to evaluate the relevance of features. In [15], three algorithms were chosen to perform a filter method of feature selection: ReliefF, Mutual information based Feature Selection and Correlation-based Feature Selection (CFS). This last selection technique is one of the most promising, simple and fast, so its use is too common. J. F. Alcón et al. [21] proved that the use of CFS in conjunction with different classifiers leads to promising results. For this reason, it is the method for feature selection used in this thesis and it will be described in detail in section 3.5. In another study, H. Iyatomi et al. [13] used a statistical F-test to perform the selection of features while K. Tabatabaie et al. [19] explore the T-test method. Both F-test and T-test are included in filter methods.

2.3.5 Classification

The final step of an automatic system for diagnosis is the classification. The decision process is based on machine learning techniques in which a set that include the labels of previously classified images is trained. This decision can be based on a binary system [10] [13] [15] [18] [17] [21] [28] [29] or a multi-label algorithm [18]. However, most of the studies are dedicated to a two-class label classification, namely the distinction between melanomas and benign lesions. There are a variety of methods used to classify lesions and all of them have shown robustness and successful performances: k-Nearest Neighbors (kNN) [17], Support Vector Machine (SVM) [15] [28], Artificial Neural Network (ANN) [29], Decision Trees (Decision Stump, Logistic Model Trees (LMT) and J48) [10] [21], Bayesian Network [16] [21] and AdaBoost [18] [21].

Over the recent years, promising studies were performed in order to improve CAD systems. M. E. Celebi et al. [15] achieved a specificity of 92.34% and a sensitivity of 93.33% by testing a database with 564 images, containing 88 melanomas, on a SVM classifier. On the other hand, J. F. Alcón et al. [21] tested different classifiers and concluded that a feature selection method (CFS) along with a LMT classifier presents an efficient performance. A sensitivity of 94% and a specificity of 68% was achieved when applied to 152 images (45 clark nevi (benign) and 107 melanomas).

G. D. Leo et al. [10] obtained sensitivity and specificity results of 83% and 76% through a LMT classifier applied on a dataset with 287 images, 175 of which were melanomas, following a 7-point checklist approach. A system based on ABCDE rule was developed by A. Parolin et al. [18] in order to distinguish between melanomas and benign lesions. Considering a database with 290 images (151 melanomas and 139 benign lesions), a total accuracy of 82.55% was achieved by a Bayesian classifier.

K. Ramhlakhan et al. [17] decided to use a more simple classifier (kNN) in order to label 83
images (46 malignant lesions), achieving a sensitivity of 60.7% and a specificity of 80.5%. Another successful approach developed by Q. Abbas et al. [18] proved that an adaptative boosting multi-class label algorithm (AdaBoost) is efficient in classification as it obtains a sensitivity of 89.28% and a specificity of 93.75% when applied to a dataset with 350 images. This dataset was divided into 7 groups of classes containing 50 images with specific features associated: reticular, globular, cobblestone, homogeneous, parallel ridge, starburst and multicomponent patterns.

A summary of the studies mentioned above and others performed in dermoscopy area is presented in table 2.5 which is organized in different sections. It refers the authors as well as the extracted features, classifier, database and performance measures. All of these researches show great performances and promising results. Nonetheless, it should be noted that the datasets used to compute the classifier performance measures vary in each study. Thus, the final results may differ depending on the database used in the different approaches.

The most accurate systems use larger databases with a variety of images provided by different sources. Nevertheless, authors try to select the more appropriate images so as not to impair the results achieved. The contrast between the lesion and the skin, the presence of too much artifacts (hairs) and the lesion not entirely within the dermoscopic image [15] are the mainly requirements defined by them.

**Table 2.5:** Summary of important studies based on Automatic Diagnostic Systems (Dataset Section refers to the number of images analysed and the number of melanomas considered in parenthesis).

<table>
<thead>
<tr>
<th>Authors</th>
<th>Features</th>
<th>Classifier</th>
<th>Database</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. E. Celebi et al. [15]</td>
<td>Color; Texture; Shape</td>
<td>SVM</td>
<td>564 (88)</td>
<td>SP=92.34%; SE=93.33%</td>
</tr>
<tr>
<td>J. F. Alcón et al. [21]</td>
<td>ABCD Rule: Asymmetry; Border; Color; Differential Structures</td>
<td>LMT</td>
<td>152 (107)</td>
<td>SP=68%; SE=94%</td>
</tr>
<tr>
<td>G. D. Leo et al. [10]</td>
<td>7-point checklist: pigment network; Blue-whitish veil; Atypical vascular pattern; Irregular streaks; Regression structures; Irregular pigmentation; Irregular dots/globules</td>
<td>LMT</td>
<td>287 (175)</td>
<td>SP=76%; SE=83%</td>
</tr>
<tr>
<td>A. Parolin et al. [16]</td>
<td>ABCDE Rule: Asymmetry; Border; Color; Diameter; Evolution</td>
<td>Bayesian</td>
<td>290 (151)</td>
<td>ACC=82.55%</td>
</tr>
<tr>
<td>K. Ramlakhan et al. [17]</td>
<td>Color; Shape</td>
<td>kNN</td>
<td>83 (46)</td>
<td>SP=80.5%; SE=60.7%</td>
</tr>
<tr>
<td>Q. Abbas et al. [18]</td>
<td>Color; Texture</td>
<td>AdaBoost</td>
<td>350</td>
<td>SP=93.75%; SE=89.28%</td>
</tr>
<tr>
<td>F. Ercal et al. [29]</td>
<td>Color; Shape</td>
<td>ANN</td>
<td>240 (160)</td>
<td>ACC=86%</td>
</tr>
<tr>
<td>H. Iyatomi et al. [13]</td>
<td>Color; Texture; Shape</td>
<td>Linear Classifier</td>
<td>548 (107)</td>
<td>SP=86.6%; SE=98%</td>
</tr>
<tr>
<td>N. Karami et al. [28]</td>
<td>Color; Structure</td>
<td>SVM</td>
<td>160 (80)</td>
<td>ACC=94%</td>
</tr>
</tbody>
</table>
Feature Extraction and Selection

Contents

3.1 Preprocessing ................................................................. 18
3.2 Color Spaces ............................................................... 21
3.3 Color Density Estimation using Histograms ......................... 26
3.4 Color Density Estimation using Gaussian Mixture Models ........ 27
3.5 Feature Selection .......................................................... 31
3.1 Preprocessing

In this work, features describing the color distribution in an image were extracted in three different ways: considering only the lesion region, considering both the lesion and the surrounding but separately and finally, considering the entire image except for the dark corners since they do not give any relevant information to the study. The following figure illustrates the three regions of feature extraction considered.

![Figure 3.1: Figure illustrating (a) the original image and the three specific regions of feature extraction: (b) lesion, (c) skin and (d) the entire image except for the dark corners.](image)

The segmentation of the lesion was manually performed by an expert dermatologist which allowed to detach the lesion used in the extraction. On the other hand, the extraction performed in the skin surrounding the lesion and in the entire image required the removal of the corners through a threshold based on Otsu’s method [30].

3.1.1 Otsu’s Method

Considering a grayscale image, this method automatically selects an optimal threshold that minimizes the intraclass variance and maximizes the separability of the two resultant classes [30].

For the purpose, an histogram containing $L$ gray levels is regarded as a probability function, $p_i$:

$$\sum_{i=1}^{L} p_i = 1 \ , \ p_i \geq 0$$

(3.1)

This normalization results in a global mean level of the image:

$$\mu_T = \mu(L) = \sum_{i=1}^{L} i p_i$$

(3.2)

Dividing the pixels into two classes $C_0$ and $C_1$ with a threshold at level $k$, $C_0$ denotes the pixels with levels $[1, \ldots, k]$ and $C_1$ denotes the pixels with levels $[k+1, \ldots, L]$. The probabilities of class occurrence, $\omega_0$ and $\omega_1$, are given by:

$$\omega_0 = Pr(C_0) = \sum_{i=1}^{k} p_i = \omega(k)$$

(3.3)

$$\omega_1 = Pr(C_1) = \sum_{i=k+1}^{L} p_i = 1 - \omega(k)$$

(3.4)

where
\[ \omega_0 + \omega_1 = 1 \]  

The class mean levels are given by equations 3.6 and 3.7 which use the notation \( Pr(i|C_0) \) for the probability of the gray level \( i \) given that the pixel is classified as \( C_0 \):

\[
\mu_0 = \sum_{i=1}^{k} i \cdot Pr(i|C_0) = \sum_{i=1}^{k} \frac{p_i}{\omega_0} = \frac{\mu(k)}{\omega(k)} \tag{3.6}
\]

\[
\mu_1 = \sum_{i=k+1}^{L} i \cdot Pr(i|C_1) = \sum_{i=k+1}^{L} \frac{p_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \tag{3.7}
\]

where

\[
\omega(k) = \sum_{i=1}^{k} p_i \quad \text{and} \quad \mu(k) = \sum_{i=1}^{k} i \cdot p_i \tag{3.8}
\]

The following relation is verified for any value of \( k \) chosen:

\[ \omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T \tag{3.9} \]

The class variances are defined by:

\[
\sigma_0^2 = \sum_{i=1}^{k} (i - \mu_0)^2 \cdot Pr(i|C_0) = \sum_{i=1}^{k} (i - \mu_0)^2 \cdot \frac{p_i}{\omega_0} \tag{3.10}
\]

\[
\sigma_1^2 = \sum_{i=k+1}^{L} (i - \mu_1)^2 \cdot Pr(i|C_1) = \sum_{i=k+1}^{L} (i - \mu_1)^2 \cdot \frac{p_i}{\omega_1} \tag{3.11}
\]

Based on the previous equations, Otsu defined three parameters in order to evaluate the threshold: the within-class variance, \( \sigma_W^2 \), the between-class variance, \( \sigma_B^2 \) and the total variance of levels, \( \sigma_T^2 \).

\[
\sigma_W^2(k) = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2 \tag{3.12}
\]

\[
\sigma_B^2(k) = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \tag{3.13}
\]

\[
\sigma_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2 \cdot p_i \tag{3.14}
\]

where

\[
\sigma_W^2(k) + \sigma_B^2(k) = \sigma_T^2 \tag{3.15}
\]

Considering these variances, three measures of class separability were introduced:

\[
\gamma = \frac{\sigma_B^2(k)}{\sigma_W^2(k)} , \quad \kappa = \frac{\sigma_T^2}{\sigma_W^2(k)} , \quad \eta = \frac{\sigma_B^2(k)}{\sigma_T^2} \tag{3.16}
\]
Since those measures are related by equation 3.15, only one needs to be maximized. Considering the $\eta$ or equivalently $\sigma_B^2$, the maximization leads to:

$$
\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k) [1 - \omega(k)]}
$$

(3.17)

Finally, the optimal threshold $k^*$ is given by:

$$
\sigma_B^2(k^*) = \max_{1 \leq k \leq L} \sigma_B^2(k)
$$

(3.18)

The removal of the corners was performed based on the method described. Regarding the red channel of each RGB image, a threshold was computed in order to separate the dark corners from the remaining image. Figure 3.2 illustrates a grayscale image histogram marked at the selected threshold.

![Grayscale image histogram marked at the selected threshold computed based on the Otsu’s method.](image)

**Figure 3.2:** Grayscale image histogram marked at the selected threshold computed based on the Otsu’s method.

Based on this threshold, it is possible to convert the image into a binary one. All the intensity values of pixels which exceed the threshold are replaced by 1 and the other are replaced by 0.

The following step was the elimination of the four binary components which are connected to the corners. A flood-fill operation was performed starting in four selected pixels belonging to each of the corners. This will change those connected pixels defined as 0 by Otsu’s methods to 1 until boundaries are reached. Then, a logical conjunction followed by a logical complement were performed in order to detach the dark corners. The figure below illustrates a comparison between the original image with corners and the result of applying the preprocessing method for their removal.
Due to the fact that not all of the images had corners, it was necessary to use an effectiveness metric defined by Otsu in order to evaluate the separability of the classes.

$$\eta(k^*) = \frac{\sigma_B^2(k^*)}{\sigma_T^2}$$  \hspace{1cm} (3.19)

It represents the effectiveness thresholding of the input image and varies in the range \([0,1]\). It is an important measure due to its invariance with respect to affine transformations of gray level scale \([30]\). This presents lower values for pictures which have constant gray levels while the upper bound is achieved only by a two-valued image. Therefore, it was considered that in the case of this parameter achieves a lower value than 0.7, the elimination of the corners would not be performed since more constant gray levels are presented and therefore, the image does not have corners.

### 3.2 Color Spaces

The analysis of lesion colors is crucial in the classification of dermoscopic images. One of the main reasons is that color, contrary to other image properties, is not dependent on the size, direction and angle of the image, increasing its robustness \([31]\).

The color study accomplished in this thesis focuses on four different color spaces which were used to extract features. Although RGB color space based on primary colors red, green and blue is the most widely used, there are other well-known color spaces such as HSV, CIELab and Opponent which could give useful and important information about images. Therefore, color intensity of images from these color spaces were quantified by histograms and Gaussian Mixture Models. The features extracted were subsequently used for color density estimation, as explained in sections 3.3 and 3.4.

#### 3.2.1 RGB color space

The original set of images has its color information stored in the traditional RGB color space. The importance of this color space falls on the fact that all color spaces result from its linear or non-linear transformation. Nevertheless, human perception of color is not equal to changes that can occur in RGB space which means that small modifications in color points can result in great changes in color \([32]\).

This color space is based on three independent chromaticities red, green and blue and with these primary colors any other color can be obtained. This is illustrated in the following equation where any
color $F$ can be obtained by the addition of different combinations of these three colors. Thus, RGB is an additive color model [33] [32].

$$F = Re_1 + Ge_2 + Be_3 = R \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + G \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + B \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$ (3.20)

In this equation, $R$, $G$ and $B$ correspond to the red, green and blue channels, respectively. The combination of these colors at maximum values results in white color, otherwise it results in black color. It is important to refer another way of representing RGB color space which is a cube, as illustrated in figure 3.4. This geometry associated with the cartesian coordinate system has the x-axis, y-axis and z-axis as referring to the red, green and blue channels, respectively. All coordinates in the cube lie in the range [0,255].

$$RGB_{coordinates} = (R_i, G_i, B_i), \quad 0 \leq R_i, G_i, B_i \leq 255$$ (3.21)

![RGB color cube](image)

**Figure 3.4:** Representation of RGB color cube (Adapted from [3]).

RGB color space allows secondary colors to be produced by the primary colors [34]. As it is represented in the figure, the combination of two primary colors at its full intensities and the other at its minimum value results in a secondary color, cyan, magenta or yellow. Each color $F$ is a point in the cube and any change in one of its three coordinates results in a change of the $F$ value.

### 3.2.2 HSV color space

One important cylindrical-coordinate color model is HSV. It has an advantage over RGB color space as it considers the perception of human vision [4] [35]. The cubic geometry of RGB color space was restructured in order to become more intuitive and relevant in terms of perception as it is shown in figure 3.5. HSV model can be represented not only by a cone but also by a cylinder or an hexagonal cone which are too common.
HSV color space decomposes the color according to physiological criteria as hue (H), saturation (S) and value (V), revealing different information about color variations when compared to RGB. These three coordinates characterize each color which can be individually analyzed. Hue refers to the pure spectrum colors and corresponds to the dominant color as perceived by a human which allows the distinction between different colors. Saturation relates to the higher or lesser degree of color intensity and value refers to the brightness of a specific color. In figure 3.5, it can be observed HSV geometry which presents an angle around a central vertical axis which corresponds to the hue. The distance from the axis represents saturation and the distance along the axis is the value. The black color is in the geometry corner and the white color lies on the base as well as primary and secondary colors.

HSV model results from a non-linear transformation of RGB color space, as it is represented:

\[
\begin{align*}
    h &= \begin{cases} 
        0, & \text{if } \max = \min \\
        \left(60^\circ \frac{g - b}{\max - \min} + 360^\circ \right) \mod 360^\circ, & \text{if } \max = r \\
        \left(60^\circ \frac{b - r}{\max - \min} + 120^\circ\right), & \text{if } \max = g \\
        \left(60^\circ \frac{r - g}{\max - \min} + 240^\circ\right), & \text{if } \max = b
    \end{cases} \\
    s &= \begin{cases} 
        0, & \text{if } \max = 0 \\
        \frac{\max - \min}{\max}, & \text{if } \max = r
    \end{cases} \\
    v &= \max
\end{align*}
\]

These equations show red channel as r, green channel as g and blue channel as b normalized in the range [0,1]. These definitions amount to the base’s representation of HSV color space which has each color (both primary and secondary) separated by a 60° arc.

The hue component of HSV color space has values in a range between 0 and 360°, where 0 corresponds to red color, 120° to green color and 240° to blue color [32]. The loop ends up in red color at 360°. About saturation and value, they fit into the range [0,100]. A completely saturated color is called pure and has its saturation at value 100 whereas an image with saturation 0 will be on gray level. Also, a color presenting the highest level of brightness has its value at 100.
3.2.3 CIELab color space

CIELab is described as one of the most complete color spaces as it considers all the colors visible to the human eye. As the other color spaces discussed, CIELab is also a three-dimensional space with components $L^*$, $a^*$ and $b^*$. Its geometry is represented in figure 3.6.

![Figure 3.6: Representation of CIELAB color space geometry](image)

This color space has three axis perpendicular to each other. The component L varies from 0 to 100 [31]. The value 0 corresponds to the black color and the value 100, at the top, refers to the white color. On the other hand, the a-axis and b-axis are the opponent-color dimensions expressing color variations. A-axis represents the color position between red (positive direction) and green (negative direction) and b-axis reveals color variations between yellow (positive direction) and blue (negative direction). It is important to refer that this geometry takes into account that red color opposes to green color and yellow color is contrary to blue one, so one color can not have two colors that are opposed to each other.

Considering a specific value of luminance $L^*$, an "a-b" plane is obtained [5]. Thus, each color is represented by two coordinates in this plane, one representing a and other representing b. In figure 3.7, it is illustrated the uniformity and graduality of color changes. Each pair of coordinates represents a pure color, with a luminance that characterizes its lightness or darkness [5].

![Figure 3.7: An "a-b" plane defined by a specific value of luminance L](image)
The database used in this work is stored in RGB format, so it is necessary to convert it to CIELab color space. This conversion makes use of some equations which are explained below. The first step is to convert RGB image to XYZ, because the CIELab model is defined in relation to a white point of XYZ color space. This conversion required a transformation matrix which considers an CIE illuminant to define the white point as it depends on the illumination. A CIE illuminant is a theoretical source of visible light which is defined by its power spectrum. Thus, it was used the standard D65 illuminant which corresponds to R=G=B=100 in equation 3.25. It is one of the most common CIE illuminants of D series and tries to describe a natural daylight environment [36] [37].

\[
\begin{bmatrix}
X \\ Y \\ Z
\end{bmatrix} =
\begin{bmatrix}
0.412453 & 0.357580 & 0.180423 \\
0.212671 & 0.715160 & 0.072169 \\
0.019334 & 0.119193 & 0.950227
\end{bmatrix}
\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\] (3.25)

After the conversion from RGB to XYZ, the following equations can be used to convert between XYZ and CIELab color spaces:

\[
X_1 = X/X_n, \quad Y_1 = Y/Y_n, \quad Z_1 = Z/Z_n
\] (3.26)

\[
X_1 = \begin{cases} 
X_1^{1/3}, & \text{if } X_1 > 0.008856 \\
7.787 X_1 + 16/116, & \text{if otherwise}
\end{cases}
\] (3.27)

\[
Y_1 = \begin{cases} 
Y_1^{1/3}, & \text{if } Y_1 > 0.008856 \\
7.787 Y_1 + 16/116, & \text{if otherwise}
\end{cases}
\] (3.28)

\[
Z_1 = \begin{cases} 
Z_1^{1/3}, & \text{if } Z_1 > 0.008856 \\
7.787 Z_1 + 16/116, & \text{if otherwise}
\end{cases}
\] (3.29)

Finally, \(X_1, Y_1\) and \(Z_1\) are used to compute \(L^*, a^*\) and \(b^*\) and a converted image is obtained.

\[
L^* = 116 Y_1, \quad a^* = 500 (X_1 - Y_1), \quad b^* = 200 (Y_1 - Z_1)
\] (3.30)

### 3.2.4 Opponent color space

The Opponent color space also studied in this work is based on a linear transformation of RGB color space [38]. It is also a three-dimensional color space with one achromatic component called luminance and two chromatics. The first component is obtained by the addiction of the three RGB channels, red, green and blue expressing intensity differences.

\[
O_1 = R + G + B
\] (3.31)

The other two components, \(O_2\) and \(O_3\), describe green-red and yellow-blue variations providing the color information, as it is represented below.

\[
O_2 = R - G, \quad O_3 = 2B - R - G
\] (3.32)
3.3 Color Density Estimation using Histograms

The first set of features studied was the color density estimated with histograms which describe the frequency of color that occur in an image. This is a global feature descriptor since no information about the spatial arrangement of the pixels is encoded in an histogram. Thus, it only provides statistical information.

A non-parametric approach is used as the number of pixels of each distinct color is calculated. Each extraction region is analysed and the bin corresponding to each pixel color is incremented by one. In the end, the histogram is normalized so that its values represent the frequency of each color.

Figure 3.8 exemplifies an image stored in RGB color space. Regarding this image, it is possible to illustrate the color distribution of each color channel along its lesion region as in figure 3.9.

![Figure 3.8: Image stored in RGB color space.](image)

![Figure 3.9: Example of the color distribution of each color channel along the lesion region of an RGB color space.](image)

Nevertheless, the features extracted consider the information provided by the three color channels of each image in each color space simultaneously which was used to generate 3D histograms, providing relevant information about the distribution of the color data. The histograms created for each color space have different dimensions:

- **RGB histograms**: Considering RGB color space, each color component is represented by a byte so there are 256 different values for each component which means that exist $256^3$ different colors, resulting in an histogram with dimension $256 \times 256 \times 256$. 
- **HSV histograms**: Relatively to HSV color space, it was created histograms with dimension 361x101x101 as the range of hue's values is [0,360] and saturation and value varies between 0 and 100.

- **CIELAB histograms**: The first component of CIELAB color space, L*, varies in a range between 0 and 100 and a* and b* can have both negative and positive values resulting in histograms with dimension 101x261x261.

- **Opponent histograms**: Finally, Opponent color space led to the largest histograms with dimensions 766x200x470.

All of the histograms obtained for each image in different color spaces were discretized using bins with different sizes so the influence of the number of bins in histograms could be analyzed. The discretization of information will result in a reduction of intervals number and can reveal different data information. The step chosen for histogram discretization was 4, 8, 16 and 32 leading to histograms with different number of bins in each color space as represented in the following table.

Table 3.1: Dimension of histograms used in this work considering the color space and the number of bins.

<table>
<thead>
<tr>
<th>Step of Discretization</th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>64x64x64</td>
<td>91x26x26</td>
<td>26x66x66</td>
<td>192x50x118</td>
</tr>
<tr>
<td>8</td>
<td>32x32x32</td>
<td>46x13x13</td>
<td>13x33x33</td>
<td>96x25x59</td>
</tr>
<tr>
<td>16</td>
<td>16x16x16</td>
<td>23x7x7</td>
<td>7x17x17</td>
<td>48x13x30</td>
</tr>
<tr>
<td>32</td>
<td>8x8x8</td>
<td>12x4x4</td>
<td>4x9x9</td>
<td>24x7x16</td>
</tr>
</tbody>
</table>

A large number of bins can result in a sharp histogram leading to an histogram that does not fit data in a correct way. On the other hand, a small number of bins results in an excessive smoothing of histograms. So, an intermediate number of bins will be appropriate [7].

### 3.4 Color Density Estimation using Gaussian Mixture Models

Another approach to analyse color intensities for different color spaces is the parametric method called Gaussian Mixture Model (GMM). Contrary to histograms, it is a model governed by specific parameters which were generated for each image.

Regarding the histograms illustrated in the previous section, it is possible to conclude that they do not resemble a single Gaussian distribution which motivates the choice of a more complex model, such as the Gaussian Mixture Model. This is defined as a linear superposition of Gaussians. Considering an initial dataset that is thought to have been generated by a mixture of $K$ components, the aim is to estimate the mean and the covariance matrix associated with each Gaussian as well as the weight that specifies how likely each Gaussian is to be chosen [7] [39]. A Gaussian Mixture takes the form:

$$ p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x | \mu_k, \Sigma_k) $$

(3.33)
It is important to refer that \( P(x) \) corresponds to the Gaussian mixture as well as each Gaussian component is represented by \( \mathcal{N}(x|\mu_k, \Sigma_k) \), being characterized by its own mean, \( \mu_k \), and covariance matrix, \( \Sigma_k \). The other components of the mixture, \( \pi_k \), are called mixing coefficients and are related to the weights previously mentioned. These parameters have to follow specific conditions. Considering that both \( P \) and Gaussian components are normalized, varying between 0 and 1, it is verified that:

\[
\sum_{k=1}^{K} \pi_k = 1
\]  

(3.34)

In addition, due to the fact that \( P \geq 0 \) and \( \mathcal{N}(x|\mu_k, \Sigma_k) \geq 0 \), it results that:

\[
\forall k: \pi_k \geq 0
\]  

(3.35)

To accomplish the main purpose which is to estimate the parameters that describe the Gaussian Mixture, \( P(x) \), \( \mu_k \) and \( \Sigma_k \) of the defined Gaussians, Maximum Likelihood can be applied to the GMM represented in equation 3.33. Considering a dataset \( X \) with \( N \) observations, the logarithm of the likelihood function is represented below.

\[
\ln p(X|\pi, \mu, \Sigma) = \sum_{n=1}^{N} \ln \left\{ \sum_{k=1}^{K} \pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k) \right\}
\]  

(3.36)

One complex problem in solving this equation arises due to the presence of singularities. Considering that one component of Gaussian Mixture has its mean value equal to one data point, it will result in a contribution term that in a limit to zero will go to infinite. As a consequence, the logarithm of the likelihood function will also tend to infinite.

Another problem appears when maximizing the same equation. As the sum from \( k = 1 \) to \( K \) appears inside the logarithm, it does not interact with the Gaussian itself. So, when the derivative of the equation are set to zero, the solution will not achieve a closed form.

In order to determine the parameters of the Gaussian Mixture, an effective iterative algorithm called Expectation-Maximization (EM) was used as an optimization method.

### 3.4.1 Expectation Maximization Algorithm

As mentioned above, Expectation-Maximization (EM) is an iterative algorithm which aims to calculate the Maximum Likelihood estimator of a specific parameter, in this case \( \mu \) and \( \Sigma \), given a probability model [7] [39]. Based on equation 3.36, the derivative of the likelihood was computed with respect to the mean Gaussian component, \( \mu_k \). Then, the expression was set to 0, resulting in the following equation:

\[
0 = -\sum_{n=1}^{N} \sum_{j} \pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k) \sum_{j} \pi_j \mathcal{N}(x_n|\mu_j, \Sigma_j) \frac{x_n - \mu_k}{\sum_{j} \pi_j \mathcal{N}(x_n|\mu_j, \Sigma_j)}
\]  

(3.37)

This mathematical operation leads to a term, \( \gamma(z_k) \), which represents the responsibility that component \( k \) takes for explaining the observation \( x \):
\( \gamma(z_k) = \frac{\pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k)}{\sum_j \pi_j \mathcal{N}(x_n|\mu_j, \Sigma_j)} \)  

(3.38)

Rearranging equation 3.37 and multiplying by the inverse of \( \Sigma_k \), it is possible to obtain:

\[
\frac{\delta \ln p(X|\pi, \mu, \Sigma)}{\delta \mu_k} = 0 \Rightarrow \mu_k = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) x_n, \text{ where } N_k = \sum_{n=1}^{N} \gamma(z_{nk})
\]

(3.39)

This equation expresses \( N_k \) as being the effective number of points belonging to the Gaussian \( k \).

The same reasoning can be applied to find the derivative of equation 3.36 with respect to \( \Sigma_k \). By equating it to zero, it results in:

\[
\frac{\delta \ln p(X|\pi, \mu, \Sigma)}{\delta \Sigma_k} = 0 \Rightarrow \Sigma_k = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) (x_n - \mu_k)(x_n - \mu_k)^T
\]

(3.40)

The last derivative with respect to the mixing coefficients has to be computed based on the Lagrange Multiplier formulation followed by the maximization of the expression:

\[
0 = \sum_{n=1}^{N} \ln \left( \sum_{k=1}^{K} \pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k) \right) + \lambda \left( \sum_{k=1}^{K} \pi_k - 1 \right)
\]

(3.41)

The value of \( \lambda \) can now be estimated, \( \lambda = -N \), and using it, mixing coefficients can be calculated by:

\[
\pi_k = \frac{N_k}{N}, \text{ where } N \text{ represents data points}
\]

(3.42)

In conclusion, the iterative method of EM require the equations described above. It has an initialization step, followed by two steps, Expectation (E step) and Maximization (M step), that update the variables iteratively, evaluating the logarithm of Maximum Likelihood in order to converge the process.

- **Initialization**: To initiate the method, the initial values of means, \( \mu_k \), covariance matrices, \( \Sigma_k \), and mixing coefficients, \( \pi_k \), have to be defined and the logarithm of Maximum Likelihood computed.

- **E Step**: The parameters estimated are used to evaluate the responsability terms, \( \gamma(z_{nk}) \). Thus, equation 3.38 will be used.

- **M Step**: Based on the responsibilities evaluated in the previous step, parameters will be re-estimated. Equations 3.39, 3.40 and 3.42 will allow the computation of \( \mu_k, \Sigma_k \) and \( \pi_k \), respectively.

- **Evaluation**: At the end, the logarithm of Maximum Likelihood will be evaluated in order to verify if the convergence was achieved. If the change of the evaluation function is lesser than a threshold, the process finishes once the convergence occured. Otherwise, the process goes back to E and M steps.
EM algorithm always ensures convergence but may be slow and require too many iterations. Furthermore, the convergence could end in a local maximum which is not the highest of local maxima. It is also important to refer that EM depends on its initialization and different initialization methods could lead to different results.

This thesis studies two other features:

- **Covariance Matrices**: The EM algorithm was applied to the database of dermoscopic images using a mixture of two Gaussians. One parameter that describes a Gaussian is the covariance matrix. As each image has three channels, it is characterized by two covariance matrices with dimension 3x3, $\Sigma_k$, one for each Gaussian $k = 1, 2$. A covariance matrix is a symmetrical statistical measure which contains a diagonal that describes the variance of data and nondiagonal entries that express the correlation among it \[40\]. Depending on the relation between data points, it can have negative, positive or zero values \[41\]. It also has a lower dimension when compared with other features, such as the intensity histograms. Therefore, its use allows the reduction of feature dimensionality, decreasing computational time and correlating all image color information. Another advantage is its independence regarding the image size.

- **Histograms**: The parameters, $\mu_k$ and $\Sigma_k$ for each Gaussian $k = 1, 2$, generated by the EM algorithm were used to create a probability density model, $p(x)$ of the three-dimensional distribution which is the mixture of Gaussians:

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$ \hspace{1cm} (3.43)

The third feature considered in this study is $p(x)$ which is normalized in a range between 0 and 1. To allow the comparison with the results acquired previously, these normalized histograms were rearranged in order to obtain the same dimensions as the histograms described in section 3.3.

Figure 3.10 compares the color density estimated with histograms in the lesion region with the probability density model estimated with the EM algorithm for the three channels of the same RGB image illustrated in figure 3.8. The histogram is shown in blue and the probability density model is shown in red. The two curves are similar although the GMM shows a higher smoothness than the histogram.
3.5 Feature Selection

The feature selection is an important step in this thesis due mainly to the high dimensionality of the 3D histograms used in the classification process which increases the computational time and complexity of the problem. Furthermore, images do not present all color intensities in each channel so that many entries of the histograms are null and irrelevant for the study.

Therefore, it is necessary to use an efficient method for feature selection in order to choose from the histograms, a subset of color intensities that can predict well the class to which each image belongs, decreasing the dimension of features space.

In machine learning, features are usually represented by a vector. Each image has associated a discrete vector, $F = \{f_1, \ldots, f_n\}$, where $n$ represents the total number of features, which contains all image features that were extracted \[43\]. This vector constitutes an input of the classifier.

The task of feature selection, that will look for the subset of $F$ which is more relevant, is widely used and many approaches can be followed. There are two main methods to select features: wrapper and filter methods. In this thesis, a filter method based on correlation coefficients is used for selection.

3.5.1 Correlation-Based Method

The Correlation-based Feature Selection (CFS) \[44\] is a filter method used to extract the more relevant features from an initial set. Considering a matrix $M$ with a number of rows equal to the number of images, each column of $M$ represents a feature. Thereby, for each feature the Pearson correlation is calculated between the vector of values it takes for each image of the training set, $X$, and the vector of labels, $Y$. The Pearson Correlation describes a linear relationship between two variables, in this case between features and the labels assigned. It fits in a range between -1 and +1. A correlation of +1 shows a perfect linear relationship while a value of -1 also reveals a perfect relationship but with a negative slope. When it equals 0, it refers to a total data points dispersion. The evaluation function is represented in the following equation, where $E$ refers to the expectation value and $\sigma$ to the standard deviation.
\[ \rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y} = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sigma_x \sigma_y} \quad (3.44) \]

The coefficients computed allow the association of each feature with the corresponding correlation. The following step is the descending arrangement of absolute correlation values in order to understand which features are highly correlated with labels and more relevant to the classification process. It is necessary to set a threshold which defined the minimum value chosen for correlation to get a good predictive classification. In this work, it was defined a variable \( \text{max} \) that expresses the maximum value achieved by the CFS. All the features that had a correlation value that fell in the range between 0.60 \( \text{max} \) and \( \text{max} \) were considered highly correlated and used to generate the model for classification.
4 Classification

Contents

4.1 Classifiers ........................................ 34

33
In the previous sections, methods of feature extraction and selection were described. In order to classify each image based on those extracted features, a classification method is required to distinguish between malignant lesions, melanomas, and benign lesions, nevi. Three different classifiers were considered: k-Nearest Neighbors (kNN), Support Vector Machine (SVM) and Gaussian Naive Bayes (GNB).

The classification of the unknown samples is based on a training set previously classified in a process called supervisioned learning. A set in the form \((x_i, t_i), i = 1, ..., N\) is required to train the classifier: \(x_i\) contains the features extracted from image \(i\), \(t_i \in \{0, 1\}\) refers to the label of the image \(i\) assigned previously by an expert dermatologist in which 1 refers to melanomas and 0 to nevi and \(N\) is the number of images present in the training set. Based on this set, each image from the test set will be classified. The comparison between the output of the classifier and the label of each image defined by the dermatologist allows the evaluation of the classifier performance.

4.1 Classifiers

4.1.1 k-Nearest Neighbors (kNN)

The k-Nearest Neighbors (kNN) is one of the most simple classifiers used to evaluate data. It considers a distance function which is computed between the features belonging to the image in the test set and all the images in the training set.

When the simplest form of this method is applied which means that only one neighbor is considered (\(k=1\)) for the classification, the image to be tested is classified with the same label as the image in the training set which has the lower value of distance \([37]\). Thus, this calculus is expressed by equation 4.1:

\[
y(x) = t_q, \text{ where } q = \arg\min_i \|x - x_i\|
\]  

(4.1)

This equation refers to \(x\) as the set of features of the image to be tested and \(x_i\) as the set of features of the training image \(i\).

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

There are different distances that can be applied to compare features. For the comparison of the histograms, the Chi-square distance measure was used which is sensitive to quantization effects \([45]\). This distance has been proving to be effective for histogram comparison in different areas which include dermoscopy \([46]\). The Euclidean and Bhattacharyya are also two distances commonly used in order to classify dermoscopic images using kNN classifier.
The Chi-squared distance is defined by equation 4.2. This considers two histogram features \( x_1 = (x_{11}, x_{12}, ..., x_{1F}) \) and \( x_2 = (x_{21}, x_{22}, ..., x_{2F}) \) which correspond to two different images. Each image contains \( F \) features.

\[
\chi^2(x_1, x_2) = \frac{1}{2} \sum_{j=1}^{F} \frac{(x_{1j} - x_{2j})^2}{(x_{1j} + x_{2j})}
\]  

(4.2)

As a metric, the Chi-square distance obey to specific axioms defined by:

1. \( \chi^2(x_1, x_2) \geq 0 \) and \( \chi^2(x_1, x_2) = 0 \) only if \( x_1 = x_2 \)

2. \( \chi^2(x_1, x_2) = \chi^2(x_1, x_3) = \chi^2(x_2, x_3) \)

3. \( \chi^2(x_1, x_3) \leq \chi^2(x_1, x_2) + \chi^2(x_2, x_3) \)

When comparing covariance matrices generated by the GMM, it was used a metric proposed by [47] which measure the dissimilarity between two matrices as follows:

\[
\rho(x_1, x_2) = \sqrt{\sum_{p=1}^{T} \ln^2 \lambda_p(x_1, x_2)}
\]  

(4.3)

In this equation, the generalized eigenvalues of \( x_1 \) and \( x_2 \) are calculated by \( \lambda_p(x_1, x_2) \), where \( p = 1, ..., T \) and \( T \) represents the total number of eigenvalues obtained by the equation:

\[
\lambda_p x_1 z_p - x_2 z_p = 0, \quad p = 1, ..., T
\]  

(4.4)

The generalized eigenvectors are represented by \( z_p \neq 0 \). It is also important to refer that this distance also satisfies the same axioms defined for chi-square distance.

The following general example illustrates a data point classified by the three points of the database that are nearer which means that it is applied a 3-Nearest Neighbor classifier.

Figure 4.1: Representation of data points being classified by k-Nearest Neighbors classifier. The parameter k equals a value of 3 and the point takes the label of the class that is in majority which is the red one [7].

In this figure, the point to be evaluated is closer to two red points and only one blue point. Thus, it will be classified as belonging to the red class which is the majority class.

The main problem when using this classifier is to find the optimal k value which can only be determined experimentally. Thus, it is necessary to find a range of odd k values. Then, some error rates are estimated in order to find which k optimizes the process by minimizing the error. These types of errors will be explained later.
4.1.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a much more complex classifier when compared to k-Nearest Neighbors. Its popularity has been growing lately due to its use in some classification problems in different areas which include dermoscopy where some lesions had been classified using it [15] [48].

The SVM is a decision machine that does not take into account any posterior probabilities [7]. Considering a process of data classification into two classes, a linear model of SVM takes the form:

\[ y(x) = w^T \phi(x) + b \]  

(4.5)

In this equation, \( \phi(x) \) represents a general feature-space transformation, \( w \) is a weight vector and \( b \) is a bias parameter.

For a better understanding of how this method works, instead of assuming class 0 to nevu and class 1 to melanoma, the two classes will be represented by -1 and +1, respectively. The new data \( x \) will be classified according to the sign of \( y(x) \).

The advantages of SVM ranging from a good generalization ability to a capability of finding an unique solution. In case of multiple possible solutions, SVM can solve the problem by the concept of margin, which is defined as the perpendicular distance between the closest sample and a decision boundary or hyperplane. On the other hand, the decision boundary or hyperplane is defined as the one that maximizes the margin. Both the margin and the decision boundary are represented in the figure 4.2.

Figure 4.2: Definition of margin and decision boundary represented by the red line in Support Vector Machine [7].

Considering that the classes are linearly separable, a set of parameters \( w \) and \( b \) will satisfy the conditions:

\[ y(x_i) > 0 , \text{for data with label } t_i = +1 \]  
(4.6)

\[ y(x_i) < 0 , \text{for data with label } t_i = -1 \]  
(4.7)

These equations imply that \( t_i y(x_i) > 0 \) for all data points which means that all data is properly classified. The parameter \( w \) is perpendicular to the hyperplane and the displacement of the decision boundary from the origin is given by \( -b/\|w\| \). Moreover, the perpendicular distance between a data point \( x \) and the hyperplane is given by \( |y(x)|/\|w\| \). It is illustrated in figure 4.3.

The next step is to parameterize both the margin and the hyperplane. As all data points have to satisfy the statement \( t_i y(x_i) > 0 \), the margin can be defined as:
Figure 4.3: Figure representing the decision boundary and its relation with the SVM parameters and a data point \( x \). [7]

As the hyperplane results from the maximization of the margin, the goal is to find the parameters \( w \) and \( b \) that satisfies:

\[
\arg\max_{w,b} = \max_{\|w\|} \min_i [t_i(w^T \phi(x_i) + b)]
\]  

(4.9)

This equation is too complex to achieve a direct solution, so some mathematical simplifications were applied. At the end, the result is a canonical representation of the hyperplane that define the constraints to all data:

\[
t_i(w^T \phi(x_i) + b) \geq 0
\]  

(4.10)

However, the common cases of overlapping classes imply the need of using a new variable which quantifies the error and is expressed by \( \xi_i \geq 0 \), \( i = 1, \ldots, N \). All \( N \) images from the training set will have a \( \xi \) associated. If data points are on or beyond the correct side of margin boundary, \( \xi_i \) is equal to 0, otherwise a value of \( \xi_i = |t_i - y(x_i)| \) will be associated to the data point. Based on this equation, it is easy to predict that only the support vectors which lie closer to the hyperplane, inside the margins, are used to estimate parameters \( w \) and \( b \) used to parameterize the margin and consequently to evaluate the sign of \( y(x_i) \).

The constraints defined above need to be modified in order to allow data points belonging to one class to be in the the wrong side of the decision boundary. So, equation 4.10 is replaced by:

\[
t_i y(x_i) \geq 1 - \xi_i
\]  

(4.11)

Thus, SVM classifier aims to obtain the hyperplane with the maximum margin in order to separate the classes taking into account the penalizing values. This is achieved by minimizing the quantity:
This equation uses a parameter \( C > 0 \) which corresponds to the weight associated to the sum of all errors, \( \xi_i \). This parameter is controlled by the user. To solve equation 4.12, the Lagrangian formulation is used in order to simplify the calculations and take into account the constraints defined above. Two Lagrange multipliers, \( \{a_i \geq 0\} \) and \( \{\mu_i \geq 0\} \), appear in the following formulation:

\[
L(w, b, a) = \frac{1}{2} ||w||^2 + C \sum_{i=1}^{N} \xi_i - \sum_{i=1}^{N} a_i \{t_i y(x_i) - 1 + \xi_i\} - \sum_{i=1}^{N} \mu_i \xi_i
\] (4.13)

Based on equations 4.5 and 4.13, the derivatives with respect to \( w, b \) and \( \xi_i \) are defined and equaled to 0 as follows:

\[
\frac{\delta L}{\delta w} = 0 \Rightarrow w = \sum_{i=1}^{N} a_i t_i \phi(x_i)
\] (4.14)

\[
\frac{\delta L}{\delta b} = 0 \Rightarrow b = \sum_{i=1}^{N} a_i t_i
\] (4.15)

\[
\frac{\delta L}{\delta \xi_i} = 0 \Rightarrow a_i = C - \mu_i
\] (4.16)

These conditions were replaced in equation 4.12 and a dual representation of Lagrangian was obtained:

\[
\tilde{L}(a) = \sum_{i=1}^{N} a_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{M} a_i a_j t_i t_j k(x_i, x_j)
\] (4.17)

The maximization of \( \tilde{L}(a) \) guided by the conditions \( 0 \leq a_i \leq C \) and \( \sum_{i=1}^{N} a_i t_i = 0 \) is required to solve the problem. The Lagrangian formulation is based on a kernel function \( k(x_i, x_j) \) which is defined as an inner product in feature space that evaluates data to allow both a linear or a non-linear classification of classes. However, in this thesis, the simplest linear kernel was used which is defined by:

\[
k(x_i, x_j) = x_i^T x_j
\] (4.18)
4.1.3 Gaussian Naive Bayes (GNB)

The Naive Bayes classifier is another supervised learning method of classifying data based on a training set whose choice is explained by the great success it has been achieved in recent years when applied to dermoscopy [16].

It assumes that features extracted from images are conditionally independent and makes use of the Bayes theorem which explains how to compute the conditional probability, \( p(t|x) \), taking into account the probability of \( x \) given \( t \), \( p(x|t) \), and the probability of \( x \) and \( t \), \( p(x) \) and \( p(t) \) [8].

\[
p(t|x) = \frac{p(x|t) p(t)}{p(x)} \quad (4.19)
\]

A general explanation of this classifier is illustrated in the figure below.

![Figure 4.5: Scheme of Naive Bayes classifier, assuming the conditional independence of features \((x_1, x_2, ..., x_F)\) given a class \(t\) [Adapted from [8]].](image)

The two possible classes in which data can be classified are 1 and 0 when it is considered a melanoma or a nevu, respectively. Thus, it is possible to define the conditional probability of each class, \( t = \{0, 1\} \), given a set of features \( x \) of a particular image belonging to the test set to conclude which is the most probable class. \( x \) is composed of the values assumed by each feature of the image being analysed, as \( x = (x_1, x_2, ..., x_F) \), for \( F \) features.

\[
p(t = 1|x) = \frac{p(x|t = 1) p(t = 1)}{p(x)} \quad (4.20)
\]

\[
p(t = 0|x) = \frac{p(x|t = 0) p(t = 0)}{p(x)} \quad (4.21)
\]

The next step is to compute \( p(x|t = 1) \) and \( p(x|t = 0) \) in order to solve equations 4.20 and 4.21. As each feature value is conditionally independent and each \( x \) is defined as a group of the assigned values \((x_1, x_2, ..., x_F)\), it follows that:

\[
p(x|t = 1) = \prod_j p(x_j|t = 1) \quad (4.22)
\]

\[
p(x|t = 0) = \prod_j p(x_j|t = 0) \quad (4.23)
\]

Actually, the denominator of equations 4.20 and 4.21 is not calculated because the main goal is to compare \( p(t = 1|x) \) and \( p(t = 0|x) \) to conclude if \( p(t = 1|x) > p(t = 0|x) \) is verified or the opposite. If this statement is confirmed, the test image will be labeled with the class 1, otherwise it will be labeled with the class 0. Moreover, it is just a normalizing factor to make sure that the probability ranges from 0 to 1.
In this thesis, it was assumed that the individual features are normally distributed within each class. Thus, for each feature in the training set, it is possible to find a mean, $\mu$, and a standard deviation, $\sigma$, for the feature set of the images belonging to class 1 and 0, separately. Thus, a model with a probability density function defined for a Gaussian distribution can be used to classify data. The classifier becomes Gaussian Naive Bayes (GNB).

Considering the features of an image from the test set, $x$, it is possible to compute the probability of this image to belong to each class. Based on the parameters estimated, it results:

$$
p(x|t=0) = \prod_{j} N(x_{j}|\mu_{0j}, \sigma_{0j}^{2})
$$  \hspace{1cm} (4.24)

$$
p(x|t=1) = \prod_{j} N(x_{j}|\mu_{1j}, \sigma_{1j}^{2})
$$  \hspace{1cm} (4.25)

For each feature $j$:

$$
p(x_{j}|t) = N(x_{j}; \mu, \sigma), \text{ where } N(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^{2}}{2\sigma^{2}}}
$$  \hspace{1cm} (4.26)

Equation 4.24 refers to $\mu_{0j}$ as the mean of feature $j$ in nevi class and $\sigma_{0j}$ as the standard deviation of feature $j$ in the same class. The same reasoning is applied to melanomas class ($t=1$). These probability values will be compared and the higher value defines the label of the image.
5

Results and Discussion

Contents

5.1 Database ................................................................. 42
5.2 Performance Assessment............................................. 43
5.3 Classification Results................................................. 45
The following section aims to present the results achieved by the proposed system considering a specific database of dermoscopic images. An analysis of the parameters chosen for each classifier and a discussion about the relevance of the features extracted as well as the extraction region and color spaces used will be performed in order to understand which are the most relevant and conclude about the combination that performs best. The performance measures of k-Nearest Neighbors (kNN), Support Vector Machine (SVM) and Gaussian Naive Bayes (GNB) classifiers will be estimated to allow an efficient assessment of the results.

5.1 Database

This work was performed using a database of 221 dermoscopic images provided by Hospital Pedro Hispano in Matosinhos. Each image was classified previously as nevu or melanoma by an expert who analysed each one individually. Figure 5.1 illustrates examples of lesions classified as melanomas and nevi.

![Figure 5.1: Illustration of four dermoscopic images: the lesions depicted on the right side are melanomas while the other belong to nevi class.](image)

Although the images have been analysed using different color spaces, the original set was stored in RGB color space. However, not all images meet the requirements for a good evaluation of the final results and some could influence them negatively. Thus, a previous review of the database was performed in order to exclude the images that did not belong to melanocytic class and those which did not contain the entire lesion and had too many artefacts, such as brightness caused by the use of dermoscopic gel and hair. This analysis led to a new database containing 148 images as described by the following table.

<table>
<thead>
<tr>
<th>Melanocytic Lesions</th>
<th>Melanomas</th>
<th>Nevi</th>
</tr>
</thead>
<tbody>
<tr>
<td>148</td>
<td>14</td>
<td>134</td>
</tr>
</tbody>
</table>

The lack of images belonging to the positive class required the creation of more melanomas so that their number could equal nevi number, improving and becoming more consistent the results obtained
in classification. Therefore, melanomas were submitted to random rotational transformations in a range between $1^\circ$ and $359^\circ$ in order to generate new melanomas which were added to the initial set of images. This new set with 268 images was used for different studies in this work instead of the initial one. It should be noted that the rotated versions of images belonging to the training set have never been included in the test set. These transformations are illustrated in figure 5.2.

![Figure 5.2: Original image of a melanoma and the new one obtained by random rotational transformation.](image)

The number of images in the database for each class is presented in table 5.2.

<table>
<thead>
<tr>
<th>Melanocytic Lesions</th>
<th>Melanomas</th>
<th>Nevus</th>
</tr>
</thead>
<tbody>
<tr>
<td>268</td>
<td>134</td>
<td>134</td>
</tr>
</tbody>
</table>

### 5.2 Performance Assessment

The algorithms used to classify dermoscopic images generate a prediction label of what is the class to which each image from the test set belongs. It is important to refer that the samples used in the test phase were not involved in the train of the classifier. Also, this classification process is binary which means that the result for each analysed image will be 1 or 0 depending on the class to each it belongs, melanoma or nevus, respectively.

The performance assessment of the classifiers used in this study was based on a comparison between the label assigned by the expert and the result achieved. Thus, the measures applied include a count of True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN).

- **#TP**: Number of images correctly classified as melanomas.
- **#FP**: Number of images misclassified as melanomas.
- **#TN**: Number of images correctly classified as nevi.
- **#FN**: Number of images misclassified as nevi.

The explanation of these measures can be simplified by the concept of confusion matrix which can illustrate the number of correct predictions in relation to the expected label. While each column represents the label in a predicted class, each row consists in the actual class diagnosed by the expert. The following table presents a confusion matrix.
Table 5.3: General Confusion Matrix.

<table>
<thead>
<tr>
<th>Actual Class</th>
<th>Predicted Class</th>
<th>Melanoma (1)</th>
<th>Nevu (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melanoma (1)</td>
<td>TP</td>
<td></td>
<td>FP</td>
</tr>
<tr>
<td>Nevu (0)</td>
<td>FN</td>
<td></td>
<td>TN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensibility</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
</table>

Throughout the study, each set of features has been used by the three classifiers to label each image from the test set. Then, an evaluation based on these measures was performed by computing the Sensitivity (SE), Specificity (SP) and Accuracy (ACC). While SE is related to the ability of the classifier to identify positive results, SP refers to the negative results. The ACC measures the proportion of both true positives and negatives in the total dataset, which means that an accuracy of 100% refers to a system that classifies all images according to the prediction performed by the expert.

Although the detection of melanomas is tough and challenging due to melanoma’s ability of mimicking benign lesions and the existence of only 14 melanomas in the original database, the accuracy was computed with a prevalence of 0.5 for SE and SP. This means that TP and TN are both important and the cost of misclassification of melanomas is not emphasized.

\[
SE = \frac{TP}{TP + FN} \quad (5.1)
\]

\[
SP = \frac{TN}{TN + FP} \quad (5.2)
\]

\[
ACC = \frac{TP + TN}{TP + FN + TN + FP} \quad (5.3)
\]

The analysis of some performance measures required the creation of Receiver Operating Characteristic (ROC) curves which allow the evaluation of a binary classifier by varying a discrimination threshold. It plots the SE (True Positive Rate (TPR)) in relation to 1 minus SP (False Positive Rate (FPR)). Therefore, a perfect prediction method will result in a SE and SP of 100%. It constitutes an efficient performance measure as it is independent of the class distribution, error costs and decision threshold [15]. Usually, a diagonal divides the ROC space into two parts. Points above the diagonal represents a better classification when compared to random.

A common way of analysing the curve is estimating the Area Under Curve (AUC) based on the trapezoidal rule. This method computes the definite integral to find the area between the curve and the x axis by approximating the region under the graph to a trapezoid. The higher the value associated with the area, better the prediction.

As the database of dermoscopic images is relatively small and since each image can not be included in both training and test set at the same time, a procedure called leave-one-out cross-validation was applied. In this procedure, each image is used once as the test set using all other images as the training set. So, all images will constitute training and test sets but in different experiments. This type of cross-validation constitutes a special case of k-fold cross validation. Instead of creating k
subsets in which one of this subset will constitute the test set and the other k-1 will be the training set, leave-one-out cross-validation creates a number of folds equal to the number of observations.

The result of the classification is the label of each image from the test set which is compared to the classification performed by an expert and different statistical measurements are computed.

5.3 Classification Results

The following section aims to present the results achieved by the classification of pigmented skin lesions using the proposed system and the three classifiers described in chapter 4:

- k-Nearest Neighbors (kNN);
- Support Vector Machine (SVM);
- Gaussian Naive Bayes (GNB).

For the purpose, three types of color features were considered:

- Color densities estimated with histograms;
- Color densities estimated using GMM;
- Covariance matrices obtained by GMM.

The feature extraction methods were applied not only in the region of interest (lesion) but also in the skin surrounding the lesion and the entire image except for the dark corners. Thus, three sets of features were used by each classifier:

- Lesion features;
- All image features (except corners);
- Lesion and Skin features extracted separately.

The histogram features were generated considering a different number of bins both in parametric and non-parametric methods as illustrated in table 5.4. As expected, those with a step of discretization of 16 revealed better results since the intermediate number of bins studied does not result in a peaked histogram or in an excessive smoothness leading to a more efficient classification.

**Table 5.4:** Dimension of histograms used in this work considering the color space and the number of bins.

<table>
<thead>
<tr>
<th>Step of Discretization</th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>64x64x64</td>
<td>91x26x26</td>
<td>26x66x66</td>
<td>192x50x118</td>
</tr>
<tr>
<td>8</td>
<td>32x32x32</td>
<td>46x13x13</td>
<td>13x33x33</td>
<td>96x25x59</td>
</tr>
<tr>
<td>16</td>
<td>16x16x16</td>
<td>23x7x7</td>
<td>7x17x17</td>
<td>48x13x30</td>
</tr>
<tr>
<td>32</td>
<td>8x8x8</td>
<td>12x4x4</td>
<td>4x9x9</td>
<td>24x7x16</td>
</tr>
</tbody>
</table>

The color density was estimated using histograms and a Gaussian Mixture Model generated by 2 Gaussians. A different number of Gaussians, namely five, seven and nine Gaussians were used
to generate other Gaussian mixture models. However, the results obtained with the different models hardly changed which explains the choice of 2 Gaussians used in this work. The following figure illustrates how the probability density model created by the mixture of two Gaussians suits the color histogram of the red channel of an RGB image for different quantizations and decreases the associated noise. Once again, the histograms with a step of discretization of 16 revealed to be the most appropriate.

Figure 5.3: Probability density function estimated using Gaussian Mixture Model generated by 2 Gaussians compared to color histogram for the red channel of an RGB image considering different quantizations.

The covariance matrices generated by the GMM method were also used as features.

The first step in this study was trying to find evident differences between melanomas and nevi in each color space analysed (RGB, HSV, CIELab, Opponent). The illustration of the color densities estimated using 3D histograms is too complicated to interpret so three separate graphs representing the color distribution of the database in each channel are shown in the next figures in order to understand the main color differences between the two classes: red color refers to melanomas and blue color to nevi.
Figure 5.4: Color distribution of dermoscopic images stored in RGB color space, considering the three channels separately. The red color refers to melanomas and blue to nevi.

Figure 5.5: Color distribution of dermoscopic images stored in HSV color space, considering the three channels separately. The red color refers to melanomas and blue to nevi.

Figure 5.6: Color distribution of dermoscopic images stored in CIELab color space, considering the three channels separately. The red color refers to melanomas and blue to nevi.
Analysing each graph, it is clear that a perfect separation of classes is not achieved. However, the color distribution of melanomas tends to be concentrated in a particular region of each graph. Thus, although some nevi have a color distribution overlapped to some melanomas, it is possible to separate both classes extracting color features. Opponent and RGB appear to be the color spaces that performs a more efficient differentiation of the two classes.

A discussion of the results obtained by each classifier with the leave-one-out procedure is performed in the following sections which are presented in terms of SE, SP, ACC and AUC.

5.3.1 Classification results of k-Nearest Neighbors (kNN)

The k-Nearest Neighbors (kNN) is the most simple classifier used throughout this thesis. It depends on the number of neighbors $k$ used in the classification process which was considered the discriminative parameter. Due to the fact that the optimal $k$ has to be estimated experimentally, different odd values were tested $k \in \{1, 3, 5, \ldots, 55\}$ in order to conclude which $k$ achieved higher performance measures in terms of SE, SP and ACC in the classification based on two classes.

For the purpose, each set of features (color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM) extracted from different color spaces (RGB, HSV, CIELab and Opponent) were evaluated considering the different extraction regions described in section 3.1. It should be noted that kNN classifier uses different metrics in order to compare features. While the histograms are compared using the Chi-square distance, the classification based on the covariance matrices obtained by GMM use a different metric to measure the dissimilarity between them.

Color densities estimated with histograms obtained by the non-parametric method were the first feature to be studied. Tables 5.5, 5.6 and 5.7 describe the best results achieved by the features extracted from the region of interest (lesion), the entire image except for the dark corners removed by the preprocessing method and those extracted from the lesion and skin extracted separately.

Based on the $k$ values tested, it was verified that, in most of the cases, when $k$ increases, the percentage of correctly classified lesions increases for the three sets of features as the classifier becomes less specific in pattern recognition. The results also show that higher values of accuracy
are obtained when sensitivity and specificity are both increased.

**Table 5.5:** Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated with histograms in the region of interest (Lesion).

<table>
<thead>
<tr>
<th>Color densities estimated with histograms</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>5</td>
<td>78.36</td>
<td>96.27</td>
<td>87.31</td>
</tr>
<tr>
<td>HSV</td>
<td>3</td>
<td>78.36</td>
<td>95.52</td>
<td>86.94</td>
</tr>
<tr>
<td>CIELab</td>
<td>27</td>
<td><strong>85.82</strong></td>
<td><strong>89.55</strong></td>
<td><strong>87.69</strong></td>
</tr>
<tr>
<td>Opponent</td>
<td>5</td>
<td>78.36</td>
<td>95.52</td>
<td>86.94</td>
</tr>
</tbody>
</table>

**Table 5.6:** Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated with histograms in the entire image.

<table>
<thead>
<tr>
<th>Color densities estimated with histograms</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>5</td>
<td>85.82</td>
<td>88.06</td>
<td>86.94</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>75.37</td>
<td>95.52</td>
<td>85.45</td>
</tr>
<tr>
<td>CIELab</td>
<td>39</td>
<td><strong>91.79</strong></td>
<td><strong>91.79</strong></td>
<td><strong>91.79</strong></td>
</tr>
<tr>
<td>Opponent</td>
<td>7</td>
<td>92.54</td>
<td>87.31</td>
<td>89.93</td>
</tr>
</tbody>
</table>

**Table 5.7:** Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated with histograms in the lesion and in the skin extracted separately.

<table>
<thead>
<tr>
<th>Color densities estimated with histograms</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>79.10</td>
<td>98.51</td>
<td>88.81</td>
</tr>
<tr>
<td>HSV</td>
<td>3</td>
<td>79.10</td>
<td>91.79</td>
<td>85.45</td>
</tr>
<tr>
<td>CIELab</td>
<td>39</td>
<td><strong>92.54</strong></td>
<td><strong>90.30</strong></td>
<td><strong>91.42</strong></td>
</tr>
<tr>
<td>Opponent</td>
<td>7</td>
<td>92.54</td>
<td>88.06</td>
<td>90.30</td>
</tr>
</tbody>
</table>

Based on the results, it is clear that for all color spaces analysed the performance measures of sensitivity, specificity and accuracy are fairly good, achieving values always above than 70%.

Although the original database format is RGB color space, this was not the one which showed better results. CIELab achieved the highest values of each assessed measure in the three different extraction regions considering an increased number of neighbors.

An increase in the performance of the classifier considering the information extracted from the skin is also verified. This leads to one important conclusion that not only the lesion is a region of interest in the process of classification but also the surrounding regions become useful to differentiate these two classes.

Considering the results obtained for the lesion, HSV and Opponent color space exhibit the worst results and once more, CIELab was the most successful. It should be noted that Opponent color space improves its rates of classification by adding skin color features which does not occur when considering HSV.

The experimental value of $k$ parameter that optimizes the classification process also differs depending on the color space studied. The best results were obtained for CIELab color space considering 39 neighbors ($k = 39$) which achieved a sensitivity of 91.79%, a specificity of 91.79% and an accuracy of 91.79%. It is important to refer that this performance considered an extraction of features from the entire image. Nevertheless, an extraction of features from lesion and skin separately also
led to a high sensitivity of 92.54% along with a specificity of 90.30% and an accuracy of 91.42% for the same number of neighbors, which means that the information provided by the skin contributes to an improvement of the classification results.

The color densities estimated using GMM were also used in the classification process leading to the following results.

Table 5.8: Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated using GMM in the region of interest (Lesion).

<table>
<thead>
<tr>
<th>Color densities estimated using GMM</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>100.00</td>
<td>91.79</td>
<td>95.90</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>100</td>
<td>91.04</td>
<td>95.52</td>
</tr>
<tr>
<td>CIELab</td>
<td>5</td>
<td>76.12</td>
<td>80.60</td>
<td>78.36</td>
</tr>
<tr>
<td>Opponent</td>
<td>1</td>
<td><strong>98.51</strong></td>
<td><strong>94.03</strong></td>
<td><strong>96.27</strong></td>
</tr>
</tbody>
</table>

Table 5.9: Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated using GMM in the entire image.

<table>
<thead>
<tr>
<th>Color densities estimated using GMM</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>100</td>
<td>97.76</td>
<td>98.88</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>100</td>
<td>88.81</td>
<td>94.40</td>
</tr>
<tr>
<td>CIELab</td>
<td>3</td>
<td>100</td>
<td>85.82</td>
<td>92.91</td>
</tr>
<tr>
<td>Opponent</td>
<td>1</td>
<td><strong>99.25</strong></td>
<td><strong>91.04</strong></td>
<td><strong>95.15</strong></td>
</tr>
</tbody>
</table>

Table 5.10: Results of SE, SP and ACC achieved by the kNN classifier considering the color densities estimated using GMM in the lesion and in the skin extracted separately.

<table>
<thead>
<tr>
<th>Color densities estimated using GMM</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>100</td>
<td>97.01</td>
<td>98.51</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>100</td>
<td>92.54</td>
<td>96.27</td>
</tr>
<tr>
<td>CIELab</td>
<td>7</td>
<td>78.36</td>
<td>70.90</td>
<td>74.63</td>
</tr>
<tr>
<td>Opponent</td>
<td>1</td>
<td><strong>99.25</strong></td>
<td><strong>95.52</strong></td>
<td><strong>97.39</strong></td>
</tr>
</tbody>
</table>

These features reveal to be the most efficient features used to separate the two classes, melanomas and nevi, achieving values of sensitivity and specificity too high.

However, CIELab which appeared to be the most relevant color space in the classification process achieved the lowest performance measures in comparison with the other color spaces studied. On the other hand, RGB color space presents the more successful rates with a sensitivity of 100%, a specificity of 97.76% and an accuracy of 98.88% for an optimal value of $k = 1$.

Once more, the information provided by skin features becomes the classification more accurate. Nevertheless, the set of features extracted from the lesion also led to good results with Opponent as being the most successful color space with a sensitivity of 98.51%, a specificity of 94.03% and an accuracy of 96.27%.

Similarly, the results of the classification using the covariance matrices generated by GMM as features were obtained considering the three extraction regions. The best results for each color space is represented in the following tables.
Table 5.11: Results of SE, SP and ACC achieved by the kNN classifier considering the covariance matrices in the region of interest (Lesion).

<table>
<thead>
<tr>
<th>Covariance matrices obtained by GMM</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>15</td>
<td>88.06</td>
<td>76.12</td>
<td>82.09</td>
</tr>
<tr>
<td>HSV</td>
<td>7</td>
<td><strong>97.76</strong></td>
<td><strong>80.60</strong></td>
<td><strong>89.18</strong></td>
</tr>
<tr>
<td>CIELab</td>
<td>15</td>
<td>98.51</td>
<td>76.87</td>
<td>87.69</td>
</tr>
<tr>
<td>Opponent</td>
<td>47</td>
<td>97.76</td>
<td>64.18</td>
<td>80.97</td>
</tr>
</tbody>
</table>

Table 5.12: Results of SE, SP and ACC achieved by the kNN classifier considering the covariance matrices in the entire image.

<table>
<thead>
<tr>
<th>Covariance matrices obtained by GMM</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>11</td>
<td>88.81</td>
<td>82.09</td>
<td>85.45</td>
</tr>
<tr>
<td>HSV</td>
<td>5</td>
<td>91.79</td>
<td>79.10</td>
<td>85.45</td>
</tr>
<tr>
<td>CIELab</td>
<td>9</td>
<td><strong>96.27</strong></td>
<td><strong>76.87</strong></td>
<td><strong>86.57</strong></td>
</tr>
<tr>
<td>Opponent</td>
<td>15</td>
<td>92.54</td>
<td>76.12</td>
<td>84.33</td>
</tr>
</tbody>
</table>

Table 5.13: Results of SE, SP and ACC achieved by the kNN classifier considering the covariance matrices in the lesion and in the skin extracted separately.

<table>
<thead>
<tr>
<th>Covariance matrices obtained by GMM</th>
<th>Optimal k</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>11</td>
<td>97.76</td>
<td>72.39</td>
<td>85.07</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>74.63</td>
<td>83.58</td>
<td>79.10</td>
</tr>
<tr>
<td>CIELab</td>
<td>3</td>
<td><strong>97.76</strong></td>
<td><strong>83.58</strong></td>
<td><strong>90.67</strong></td>
</tr>
<tr>
<td>Opponent</td>
<td>5</td>
<td>91.79</td>
<td>77.61</td>
<td>84.70</td>
</tr>
</tbody>
</table>

The results show that covariance matrices allow an increase of sensitivity rate close to 100% in detriment of a significant decrease in specificity values when compared to the other features analysed. The Opponent color space achieved the lowest performance measures in the three extraction regions while CIELab stands out for the good results, achieving the highest performance when the set of features extracted from the lesion and skin separately are used. Thus, a sensitivity of 97.76%, a specificity of 83.58% and an accuracy of 90.67% is obtained when the classification is performed based on 3 neighbors ($k = 3$).

Considering the results achieved previously, it appears that information concerning the skin does not become so crucial since the results seem to be quite similar. HSV revealed higher assessed measures when considering the lesion region with an accuracy of 89.18% along with a sensitivity of 97.76% and a specificity of 88.60%, decreasing its performance by adding skin color information. Although the other color spaces improve their performances using skin features, this increase is not as significant as in the previous features.

The optimal $k$ for color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM are also different. The higher number of neighbors is explained by a poorly differentiation between the two classes and the lack of examples belonging to melanomas class. On the other hand, color densities estimated using GMM characterize dermoscopic images in a much efficient way as the classifier’s performance is maximized by a few number of neighbors.
A summary of the best results achieved by the kNN classifier considering the optimal number of $k$, the color spaces, the extraction region and the performance measures can be observed in the following table.

**Table 5.14**: Best results obtained by the kNN classifier using color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM as color features.

<table>
<thead>
<tr>
<th>Features</th>
<th>Optimal $k$</th>
<th>Color Space</th>
<th>Extraction Local</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color densities estimated with histograms</td>
<td>39</td>
<td>CIELab</td>
<td>All Image Features</td>
<td>91.79</td>
<td>91.79</td>
<td>91.79</td>
</tr>
<tr>
<td>Color densities estimated using GMM</td>
<td>1</td>
<td>RGB</td>
<td>All Image Features</td>
<td>100</td>
<td>97.76</td>
<td>98.88</td>
</tr>
<tr>
<td>Covariance Matrices obtained by GMM</td>
<td>3</td>
<td>CIELab</td>
<td>Lesion and Skin Features</td>
<td>97.76</td>
<td>83.58</td>
<td>90.67</td>
</tr>
</tbody>
</table>

According to the table, it appears that all features achieve high performance measures contributing to a good separation of classes and a successful detection of melanomas. Color densities estimated using GMM are the most relevant features for the classification of pigmented skin lesions using kNN classifier due to the highest accuracy rate achieved. Moreover, the most important color space in this context is RGB as it leads to successful performance measures associated with $k$ lower values using the three sets of features extracted from the three different extraction regions. The extraction performed in the entire image seems to be the most important for the differentiation between the two classes as the features extracted from the skin generally improve the results obtained for the lesion region.

The best result achieved by the kNN was used in order to illustrate some examples of cases which result from the classification. This was accomplished by the color densities estimated using GMM extracted from the entire image when considering the RGB color space. The sensitivity value of 100% means that all melanomas were accurately detected. Therefore, as no FN cases exist, only TN, TP and FP cases are represented.

![Figure 5.8: Cases of nevi correctly classified (TN) by kNN classifier.](image)
5.3.2 Classification results of Support Vector Machine (SVM)

The three color features extracted (color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM) were also used as inputs of the Support Vector Machine (SVM) classifier with a linear kernel in order to classify the dermoscopic images. For the purpose, the parameter $C$ was defined as discriminative due to the fact that it controls the cost associated to the misclassification by varying the soft margins defined by the classifier. Different values of this parameter were tested in a range of variation $C \in \{1, 2, 3, 4, 7, 9, 10, 20, 30, 40, 50\}$. The aim was to find the parameter value that optimizes the process by maximizing the accuracy measure. The higher values of accuracy are related to an increased sensitivity and specificity.

Contrary to kNN classifier, SVM shows a more constant values of performance rates with respect to the discriminative parameter which reflects the good linear separation of the two classes originally performed by the classifier.

Once more, color densities estimated with histograms were the first feature to be studied. The best results of sensitivity, specificity and accuracy achieved by the variation of parameter $C$ are presented in the following tables for each color space and extraction region.
Table 5.15: Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated with histograms in the region of interest (Lesion).

<table>
<thead>
<tr>
<th>Color densities estimated with histograms</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>92.54</td>
<td>97.76</td>
<td>95.15</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>91.04</td>
<td>96.27</td>
<td>93.68</td>
</tr>
<tr>
<td>CIELab</td>
<td>3</td>
<td>91.79</td>
<td>88.81</td>
<td>90.30</td>
</tr>
<tr>
<td>Opponent</td>
<td>1</td>
<td>91.79</td>
<td>100</td>
<td>95.90</td>
</tr>
</tbody>
</table>

Table 5.16: Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated with histograms in the entire image.

<table>
<thead>
<tr>
<th>Color densities estimated with histograms</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>94.03</td>
<td>99.25</td>
<td>96.64</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>91.04</td>
<td>99.25</td>
<td>95.15</td>
</tr>
<tr>
<td>CIELab</td>
<td>2</td>
<td>84.33</td>
<td>97.01</td>
<td>90.67</td>
</tr>
<tr>
<td>Opponent</td>
<td>2</td>
<td>92.54</td>
<td>99.25</td>
<td>95.90</td>
</tr>
</tbody>
</table>

Table 5.17: Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated with histograms in the lesion and in the skin extracted separately.

<table>
<thead>
<tr>
<th>Color densities estimated with histograms</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>1</td>
<td>92.54</td>
<td>100</td>
<td>96.27</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>93.28</td>
<td>98.51</td>
<td>95.90</td>
</tr>
<tr>
<td>CIELab</td>
<td>1</td>
<td>83.58</td>
<td>96.27</td>
<td>89.93</td>
</tr>
<tr>
<td>Opponent</td>
<td>1</td>
<td>93.28</td>
<td>99.25</td>
<td>96.27</td>
</tr>
</tbody>
</table>

By inspecting each table, it is possible to conclude that the skin has also importance in the classification process as the best result is achieved for the features extracted from the entire image. The most relevant color spaces are RGB and Opponent with the higher performances achieved in the three extraction regions. However, RGB color space reveals to be the most successful with a sensitivity of 94.03%, a specificity of 99.25% and an accuracy of 96.64%, unchanged with parameter C. The worst results are obtained when the three types of features are extracted from the CIELab color space in the three different extraction regions.

Regarding the features extracted from the lesion region, the Opponent color space stands out for the good performance achieving a sensitivity of 91.79% along with a specificity of 100% and an accuracy of 95.90%.

The color densities estimated using GMM were also input of SVM classifier leading to the results of SE, SP and ACC represented in tables 5.18, 5.19 and 5.20.

Table 5.18: Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated using GMM in the region of interest (Lesion).

<table>
<thead>
<tr>
<th>Color densities estimated using GMM</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>2</td>
<td>88.06</td>
<td>99.25</td>
<td>93.66</td>
</tr>
<tr>
<td>HSV</td>
<td>2</td>
<td>85.82</td>
<td>96.27</td>
<td>91.04</td>
</tr>
<tr>
<td>CIELab</td>
<td>50</td>
<td>67.91</td>
<td>97.01</td>
<td>82.46</td>
</tr>
<tr>
<td>Opponent</td>
<td>9</td>
<td>92.54</td>
<td>88.81</td>
<td>90.67</td>
</tr>
</tbody>
</table>
Table 5.19: Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated using GMM in the entire image.

<table>
<thead>
<tr>
<th>Color densities estimated using GMM</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>2</td>
<td>90.30</td>
<td>100</td>
<td>95.15</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>88.06</td>
<td>97.01</td>
<td>92.54</td>
</tr>
<tr>
<td>CIELab</td>
<td>50</td>
<td>85.07</td>
<td>96.27</td>
<td>90.67</td>
</tr>
<tr>
<td>Opponent</td>
<td>30</td>
<td>92.54</td>
<td>93.28</td>
<td>92.91</td>
</tr>
</tbody>
</table>

Table 5.20: Results of SE, SP and ACC achieved by the SVM classifier considering the color densities estimated using GMM in the lesion and in the skin extracted separately.

<table>
<thead>
<tr>
<th>Color densities estimated using GMM</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>7</td>
<td>92.54</td>
<td>98.51</td>
<td>95.52</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>88.06</td>
<td>97.01</td>
<td>92.54</td>
</tr>
<tr>
<td>CIELab</td>
<td>1</td>
<td>79.10</td>
<td>96.27</td>
<td>87.69</td>
</tr>
<tr>
<td>Opponent</td>
<td>9</td>
<td>90.30</td>
<td>95.52</td>
<td>92.91</td>
</tr>
</tbody>
</table>

Despite the high performance level, the color densities estimated using GMM do not perform best as occurred with kNN classifier. A sensitivity of 92.54%, a specificity of 98.51% and an accuracy of 95.52% were the best measures obtained with the features extracted from RGB color space and a discriminative parameter $C = 7$. In comparison with the optimal values achieved in the previous results, a more rigid margin is created and the classifier does not allow as many training errors as with a lower parameter $C$.

CIELab color space shows the less relevant results in the three extraction regions with an accuracy never above 91%. Once more, those results are obtained with the information provided by the skin surrounding the lesion and in the first two cases, with a cost of misclassification too high.

Considering only the features extracted from the lesion, RGB continues to stand out with a sensitivity of 88.06%, a specificity of 99.25% and an accuracy of 93.66%.

The covariance matrices obtained by GMM were also analysed, leading to the results expressed in the following tables:

Table 5.21: Results of SE, SP and ACC achieved by the SVM classifier considering the covariance matrices obtained by GMM in the region of interest (Lesion).

<table>
<thead>
<tr>
<th>Covariance matrices obtained by GMM</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>2</td>
<td>74.63</td>
<td>84.33</td>
<td>79.48</td>
</tr>
<tr>
<td>HSV</td>
<td>2</td>
<td>89.55</td>
<td>82.84</td>
<td>86.19</td>
</tr>
<tr>
<td>CIELab</td>
<td>1</td>
<td>85.07</td>
<td>94.78</td>
<td>89.93</td>
</tr>
<tr>
<td>Opponent</td>
<td>50</td>
<td>74.63</td>
<td>85.07</td>
<td>79.85</td>
</tr>
</tbody>
</table>

Table 5.22: Results of SE, SP and ACC achieved by the SVM classifier considering the covariance matrices obtained by GMM in the entire image.

<table>
<thead>
<tr>
<th>Covariance matrices obtained by GMM</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>30</td>
<td>78.36</td>
<td>95.52</td>
<td>86.94</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>77.61</td>
<td>84.33</td>
<td>80.97</td>
</tr>
<tr>
<td>CIELab</td>
<td>3</td>
<td>70.90</td>
<td>91.04</td>
<td>80.97</td>
</tr>
<tr>
<td>Opponent</td>
<td>1</td>
<td>74.63</td>
<td>97.76</td>
<td>86.19</td>
</tr>
</tbody>
</table>
Table 5.23: Results of SE, SP and ACC achieved by the SVM classifier considering the covariance matrices obtained by GMM in the lesion and in the skin extracted separately.

<table>
<thead>
<tr>
<th>Covariance matrices obtained by GMM obtained by GMM</th>
<th>Optimal C</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB</td>
<td>50</td>
<td>85.07</td>
<td>97.76</td>
<td>91.42</td>
</tr>
<tr>
<td>HSV</td>
<td>1</td>
<td>84.33</td>
<td>94.03</td>
<td>89.18</td>
</tr>
<tr>
<td>CIELab</td>
<td>1</td>
<td><strong>88.81</strong></td>
<td><strong>96.27</strong></td>
<td><strong>92.54</strong></td>
</tr>
<tr>
<td>Opponent</td>
<td>20</td>
<td>86.57</td>
<td>94.03</td>
<td>90.30</td>
</tr>
</tbody>
</table>

By analysing each table, it appears that these features show an increase in the accuracy rate when considering CIELab. The best result is achieved by the same color space considering the set of features extracted from the lesion and skin separately with an accuracy of 92.54% for \( C = 1 \). Values of sensitivity and specificity of 88.81% and 96.27% are obtained. CIELab color space also shows the highest accuracy rate regarding the features from the lesion region, achieving a sensitivity of 85.07%, a specificity of 94.78% and an accuracy of 89.93%. It should be noted that covariance matrices used as features, in some cases, required a higher cost associated to the parameter \( C \) to become the classifier more accurate.

RGB and Opponent color spaces improve significantly their results when considering the information provided by the skin features extracted, reinforcing the importance of the skin in the classification process.

A summary of the best results achieved by SVM classifier is described in the following table.

Table 5.24: Best results obtained by the SVM classifier using color densities estimated with histograms, color densities estimated using GMM and covariance matrices obtained by GMM as color features.

<table>
<thead>
<tr>
<th>Features</th>
<th>Optimal C</th>
<th>Color Space</th>
<th>Extraction Local</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>ACC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color densities estimated with histograms</td>
<td>1</td>
<td>RGB</td>
<td>All image Features</td>
<td>94.03</td>
<td>99.25</td>
<td>96.64</td>
</tr>
<tr>
<td>Color densities estimated using GMM</td>
<td>7</td>
<td>RGB</td>
<td>Lesion and Skin Features</td>
<td>92.54</td>
<td>98.51</td>
<td>95.52</td>
</tr>
<tr>
<td>Covariance Matrices obtained by GMM</td>
<td>1</td>
<td>CIELab</td>
<td>Lesion and Skin Features</td>
<td><strong>88.81</strong></td>
<td>96.27</td>
<td><strong>92.54</strong></td>
</tr>
</tbody>
</table>

Three important conclusions can be drawn from the table: the most relevant color space, the more useful type of features and the set of features that best differentiates the two classes. Color densities estimated with histograms generated from the non-parametric method achieved the highest accuracy rate, 96.64% along with a sensitivity of 94.03% and a specificity of 99.25%. These results are achieved with a minimal cost of misclassification expressed by \( C = 1 \).

Furthermore, the features extracted from the lesion do not ensure an optimal classification which means that the information provided by the skin should be considered in order to improve the results as mentioned before.

Regarding the color densities estimated with histograms extracted from the RGB color space considering the entire image which is the set that performs best, there are cases of TN, FN, TP and FP that should be illustrated.
5.3.3 Classification results of Gaussian Naive Bayes (GNB)

The classification based on Gaussian Naive Bayes (GNB) used only two types of features extracted: color densities estimated with histograms and color densities estimated using GMM. The covariance matrices obtained by GMM were not used by this classifier. This model is not appropriate as covariance matrices do not have a Gaussian distribution.

Following the same sequence of the previous results, the first feature used to classify the database into one of the two classes were color densities estimated with histograms. The results of SE, SP and ACC obtained for the three regions of extraction are presented in tables 5.25 and 5.26.
Table 5.25: Results of SE and SP achieved by GNB classifier for color densities estimated with histograms.

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SE (%)</td>
<td>SP (%)</td>
<td>SE (%)</td>
<td>SP (%)</td>
</tr>
<tr>
<td>Lesion</td>
<td>92.54</td>
<td>92.54</td>
<td>64.93</td>
<td>99.25</td>
</tr>
<tr>
<td>Entire image</td>
<td>91.79</td>
<td>92.54</td>
<td>92.54</td>
<td>92.54</td>
</tr>
<tr>
<td>Lesion-Skin</td>
<td>94.78</td>
<td>92.54</td>
<td>85.82</td>
<td>92.54</td>
</tr>
</tbody>
</table>

Table 5.26: Results of ACC achieved by GNB classifier for color densities estimated with histograms.

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACC (%)</td>
<td>ACC (%)</td>
<td>ACC (%)</td>
<td>ACC (%)</td>
</tr>
<tr>
<td>Lesion</td>
<td>92.54</td>
<td>82.09</td>
<td>88.06</td>
<td>92.16</td>
</tr>
<tr>
<td>Entire image</td>
<td>92.16</td>
<td>92.54</td>
<td>91.04</td>
<td>89.18</td>
</tr>
<tr>
<td>Lesion-Skin</td>
<td>93.66</td>
<td>89.18</td>
<td>88.43</td>
<td>93.66</td>
</tr>
</tbody>
</table>

The results were analysed by creating ROC curves which are illustrated in figure 5.15 considering the three different extraction regions: only the lesion, the entire image except for the dark corners and the lesion and the skin but separately. Based on this threshold, it is possible to conclude about the highest sensitivity and specificity that can be achieved as described in table 5.27.

Figure 5.15: ROC curves obtained by GNB classifier for color densities estimated with histograms in the three extraction regions. The symbol (o) refers to the output result of the classifier for each color space expressed in tables 5.25 and 5.26. Color blue refers to RGB; color green to HSV; color red to CIELab and color black to Opponent.

Table 5.27: Maximum values of SE and SP achieved for color densities estimated with histograms.

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SE (%)</td>
<td>SP (%)</td>
<td>SE (%)</td>
<td>SP (%)</td>
</tr>
<tr>
<td>Lesion</td>
<td>92.54</td>
<td>92.54</td>
<td>94.78</td>
<td>85.82</td>
</tr>
<tr>
<td>entire image</td>
<td>92.54</td>
<td>92.54</td>
<td>97.01</td>
<td>92.54</td>
</tr>
<tr>
<td>Lesion-Skin</td>
<td>91.04</td>
<td>99.25</td>
<td>96.27</td>
<td>92.54</td>
</tr>
</tbody>
</table>

The ROC curves also allow the estimation of the Area Under Curve (AUC) which results are described in table 5.28.
Table 5.28: Results of AUC achieved by GNB classifier for color densities estimated with histograms.

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC (%)</td>
<td>AUC (%)</td>
<td>AUC (%)</td>
<td>AUC (%)</td>
</tr>
<tr>
<td>Lesion</td>
<td>94.12</td>
<td>95.81</td>
<td>93.91</td>
<td>94.52</td>
</tr>
<tr>
<td>entire image</td>
<td>94.04</td>
<td>93.75</td>
<td>95.56</td>
<td>94.72</td>
</tr>
<tr>
<td>Lesion-Skin</td>
<td>98.05</td>
<td>97.32</td>
<td>91.07</td>
<td>96.86</td>
</tr>
</tbody>
</table>

Analysing the successful results achieved by the color densities estimated with histograms, it is possible to conclude that the previous assumption of features independence can be used in order to classify images based on GNB.

Regarding the information provided by tables 5.27 and 5.28, it is clear that the highest performance is achieved by RGB color space considering those features extracted from the lesion and skin separately. This can be confirmed by the ROC curve associated. A specific decision threshold allow a sensitivity of 91.04% along with a specificity of 99.25%. The measures achieved by the Opponent color space reveals that it is also a relevant color space in this context.

On the other hand, HSV is the most useful color space achieving the highest value of AUC when considering only the features extracted from the lesion. An AUC of 95.81% along with a sensitivity of 94.78% and a specificity of 85.82% were obtained.

The color densities estimated using GMM were the second set of features studied. Tables 5.29 and 5.30 refer to the outputs generated by GNB with respect to the SE, SP and AUC achieved. Figure 5.16 illustrates the ROC curves considering the three extraction regions. The optimized results are represented in the table 5.31 as well as the AUC is estimated in the table 5.32.

Table 5.29: Results of SE and SP achieved by GNB classifier for color densities estimated using GMM.

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SE (%)</td>
<td>SP (%)</td>
<td>SE (%)</td>
<td>SP (%)</td>
</tr>
<tr>
<td>Lesion</td>
<td>85.82</td>
<td>98.51</td>
<td>91.79</td>
<td>85.82</td>
</tr>
<tr>
<td>entire image</td>
<td>89.55</td>
<td>91.79</td>
<td>92.54</td>
<td>87.31</td>
</tr>
<tr>
<td>Lesion-Skin</td>
<td>88.06</td>
<td>98.51</td>
<td>95.52</td>
<td>57.46</td>
</tr>
</tbody>
</table>

Table 5.30: Results of ACC achieved by GNB classifier for color densities estimated using GMM.

<table>
<thead>
<tr>
<th></th>
<th>RGB</th>
<th>HSV</th>
<th>CIELab</th>
<th>Opponent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACC (%)</td>
<td>ACC (%)</td>
<td>ACC (%)</td>
<td>ACC (%)</td>
</tr>
<tr>
<td>Lesion</td>
<td>92.16</td>
<td>88.81</td>
<td>76.12</td>
<td>75.37</td>
</tr>
<tr>
<td>entire image</td>
<td>90.67</td>
<td>89.93</td>
<td>83.58</td>
<td>66.04</td>
</tr>
<tr>
<td>Lesion-Skin</td>
<td>93.28</td>
<td>76.49</td>
<td>73.13</td>
<td>74.25</td>
</tr>
</tbody>
</table>
By inspecting the previous tables, it is possible to conclude that the color densities estimated using GMM also lead to successful results.

Once more, RGB stands out for the good performance achieved. An AUC of 96.39% along with a maximum sensitivity of 88.06% and specificity of 98.51% is obtained through a decision threshold used to create the ROC curve. The information provided by the skin continues to be crucial to ensure an efficient detection of melanomas. On the other hand, Opponent color space reveals a decrease in its performance while CIELab improve its assessed measures when compared to the other studies performed. Those results confirm the conclusion drawn by the color densities estimated with histograms that an independency of features can be considered for the differentiation of the two classes.

The best results of GNB are summarized in the following table concerning the maximum sensitivity and specificity achieved as well as the AUC.
Table 5.33: Best results obtained by the GNB classifier using color densities estimated with histograms and color
densities estimated using GMM.

<table>
<thead>
<tr>
<th>Features</th>
<th>Color Space</th>
<th>Extraction Local</th>
<th>SE(%)</th>
<th>SP(%)</th>
<th>AUC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color densities estimated with histograms</td>
<td>RGB</td>
<td>Lesion and Skin Features</td>
<td>91.04</td>
<td>99.25</td>
<td>98.05</td>
</tr>
<tr>
<td>Color densities estimated using GMM</td>
<td>RGB</td>
<td>Lesion and Skin Features</td>
<td>88.06</td>
<td>98.51</td>
<td>96.39</td>
</tr>
</tbody>
</table>

Analysing the table 5.33, GNB appears to be a successful approach in the classification of images based on color features by considering their independence. From all color spaces, RGB achieves the best results when considering color densities estimated with histograms with an AUC of 98.05%. The extraction region seems to play an important role in the learning procedure ensuring the improvement of the final results similarly to what occurred with the other classifiers. The following figures illustrate cases of TN, FN, TP and FP considering the most successful result achieved by the intensity color histograms.

Figure 5.17: Cases of nevi correctly classified (TN) by GNB classifier.

Figure 5.18: Melanoma misclassified as a nevu (FN) by GNB classifier.

Figure 5.19: Cases of melanomas correctly classified (TP) by GNB classifier.
5.3.4 Discussion

In order to perform a final assessment of the best results achieved by the three different classifiers, a statistical study is presented. This includes the most significant color space as well as the most effective set of features and extraction region. Figure 5.21 illustrates a bar graph which plots the optimal performance measures achieved by each set of features used by kNN, SVM and GNB represented in table 5.34. Those results were obtained considering the information provided by the skin which proved to be crucial to achieve high rates of accuracy and efficiently detect melanomas.

Table 5.34: Optimal performance measures achieved by each set of features used by kNN, SVM and GNB.

<table>
<thead>
<tr>
<th>Feature</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>ACC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color densities estimated with histograms kNN</td>
<td>91.79</td>
<td>91.79</td>
<td>91.79</td>
</tr>
<tr>
<td>Color densities estimated using GMM kNN</td>
<td>100.00</td>
<td>97.76</td>
<td>98.88</td>
</tr>
<tr>
<td>Covariance matrices obtained by GMM kNN</td>
<td>97.76</td>
<td>83.58</td>
<td>90.67</td>
</tr>
<tr>
<td>Color densities estimated with histograms SVM</td>
<td>94.03</td>
<td>99.25</td>
<td>96.64</td>
</tr>
<tr>
<td>Color densities estimated using GMM SVM</td>
<td>92.54</td>
<td>98.51</td>
<td>95.52</td>
</tr>
<tr>
<td>Covariance matrices obtained by GMM SVM</td>
<td>88.81</td>
<td>96.27</td>
<td>92.54</td>
</tr>
<tr>
<td>Color densities estimated with histograms GNB</td>
<td>91.04</td>
<td>99.25</td>
<td>95.15</td>
</tr>
<tr>
<td>Color densities estimated using GMM GNB</td>
<td>88.06</td>
<td>98.51</td>
<td>93.29</td>
</tr>
</tbody>
</table>
The most common studies performed the extraction of features only in the region of interest considering that only the lesion contains the information required for a successful differentiation of melanocytic lesions. To perform a comparison with the previous figure, a bar graph describes only the optimal measures achieved in this region regarding each set of features used by each classifier.

**Table 5.35:** Optimal performance measures achieved by each set of features used by kNN, SVM and GNB considering only the lesion region.

<table>
<thead>
<tr>
<th>Feature</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>ACC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color densities estimated with histograms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kNN</td>
<td>85.82</td>
<td>89.55</td>
<td>87.69</td>
</tr>
<tr>
<td>GMM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color densities estimated using GMM kNN</td>
<td>98.51</td>
<td>94.03</td>
<td>96.27</td>
</tr>
<tr>
<td>Covariance matrices obtained by GMM kNN</td>
<td>97.76</td>
<td>80.60</td>
<td>89.18</td>
</tr>
<tr>
<td>Color densities estimated with histograms SVM</td>
<td>91.79</td>
<td>100.00</td>
<td>95.90</td>
</tr>
<tr>
<td>Color densities estimated using GMM SVM</td>
<td>88.06</td>
<td>99.25</td>
<td>93.66</td>
</tr>
<tr>
<td>Covariance matrices obtained by GMM SVM</td>
<td>85.07</td>
<td>94.78</td>
<td>89.93</td>
</tr>
<tr>
<td>Color densities estimated with histograms GNB</td>
<td>94.78</td>
<td>85.82</td>
<td>90.30</td>
</tr>
<tr>
<td>Color densities estimated using GMM GNB</td>
<td>86.57</td>
<td>98.51</td>
<td>92.54</td>
</tr>
</tbody>
</table>

**Figure 5.22:** Bar graph of the performance measures achieved by each classifier used to study the color features considering only the region of interest (Lesion).

Based on the two graphs, it is possible to confirm that when the skin features are considered, the classifier's accuracy and the number of melanomas correctly classified increases. The color densities estimated using GMM show the best results in both approaches which proves that these color features are the most significant regardless of the extraction region. Another interesting point is to conclude about the importance of the color space. Regarding the lesion region, the best results were achieved when the histogram features from the parametric method were extracted from the Opponent color space. On the other hand, the best results achieved by the global system consider the RGB color space, which means that both color spaces are the most relevant depending on the extraction site.

In the previous sections, examples of dermoscopic lesions correctly classified and misclassified
were presented for each classifier. An interesting point is to understand if there are lesions diagnosed as melanomas that all classifiers label as nevi (FN) or lesions diagnosed as nevi and classified as melanomas (FP).

Regarding the kNN classifier, the best result achieved efficiently detect all melanomas. On the other hand, SVM and GNB misclassified distinct lesions as being nevi. However, there is a specific lesion which is an example of FP for all classifiers which is illustrated in figure 5.23.

Figure 5.23: Example of a nevus misclassified as a melanoma by all classifiers.
6

Conclusions and Future Work
The aim of this thesis was the development of an automatic system for the classification of pigmented skin lesions based on color features. This automatic system includes different steps: a feature extraction, a feature selection and a classification based on three different classifiers, kNN, SVM and GNB. The extraction of features based on both parametric and non-parametric methods was performed regarding the region of the lesion and skin.

The non-parametric method includes the color densities estimated with histograms while the parametric approach considers the color densities estimated using Gaussian Mixture Models (GMM) and the covariance matrices generated by the same method. The irrelevance and redundancy inherent to some features required an algorithm based on the correlation method in order to select the most important.

This thesis addresses three problems:

- What is the most significant color space?
- Among the three types of features studied which are the most important?
- The extraction of features considering only the segmented lesion ensures an optimal result? What is the role of the skin surrounding the lesion?

The results obtained were quite promising and successful, proving that color features play a decisive role in the classification of skin lesions. The color densities estimated using GMM extracted from the RGB color space reveal to be the most important features when considering the information provided by the lesion and skin. The performance measures achieved were SE=100%, SP=97.76% and an ACC=98.88% by the kNN classifier which means that all the melanomas were correctly detected. The non-parametric approach also leads to good results. The best performances achieves a SE=94.03%, a SP=99.25% and an ACC=96.64% by the SVM considering the same color space and an extraction from the entire image.

Those results prove that this system is an useful tool for the classification of pigmented skin lesions.

However, the classification of skin lesions based on CAD systems is still an area in developing. Thus, a discussion of some suggestions for future work will be performed:

- The use of new color features to analyse pigmented skin lesions or the combination of other types of features in this system in order to improve the results.
- The original database consider only a few number of images belonging to melanomas class. Therefore, a database with more examples of melanomas will contribute to a better learning of the classifier and, consequently, to an improvement of the final results.
- As different classifiers lead to different classification rules, the ensemble of them might improve the performance measures.
Bibliography


