FACULTY OF ENGINEERING
UNIVERSITY OF PORTO
Department of Informatics Engineering

Roberta Barbosa Oliveira

Pattern Recognition in Dermoscopy Images using
Ensemble Methods

Thesis proposal submitted to Department of Informatics Engineering of Faculty of Engineering, University of Porto, in fulfilment of the doctoral program requirements.

Supervisor: Prof. João Manuel R. S. Tavares (FEUP)
Co-supervisor: Prof. Aledir Silveira Pereira (UNESP)

Porto, Portugal
2015
ABSTRACT

Skin cancer is considered one of the most common types of cancer in several countries and its incidence rate has increased in recent years. Melanoma cases have presented an increasing number of deaths worldwide, since this type of skin cancer is the most aggressive compared to other types. For dealing with this issue, computer-aided diagnosis (CAD) systems based on pattern recognition have been developed to assist dermatologists in early diagnosis of skin cancer.

Pattern recognition of dermoscopy images is a challenging task to assist in skin lesion diagnosis. Another problematic issue related to this task is in identifying the presence of global patterns, since such patterns have not been well studied in research for automatic diagnosis. Currently, classifiers have been proposed for pattern recognition of dermoscopy images. The search for better performing classifiers is a relevant issue of this approach. One challenge that affects performance of classifiers includes defining which features are meaningful to describe the patterns of interest. The feature selection and ensemble of classifiers may be combined to achieve superior performance for the skin lesion pattern recognition.

We propose developing a homogeneous ensemble method by using a feature selection for input feature manipulation to overcome the results of skin lesion pattern recognition from dermoscopy images. To achieve this goal, we will first do the image pre-processing and identify the region of interest (ROI) using a segmentation algorithm. In addition, features based on colour and texture properties will be extracted from the ROI. In order to manipulate the input features to generate the ensemble model we intend to use different feature vectors obtained by using a feature selection scheme.

With the proposed approach in this thesis we aim to contribute to the dermoscopy CAD area with a method that provides information concerning skin lesion patterns to assist dermatologists in diagnosing lesions. In addition, our proposed approach provides an effective and robust classification for pattern recognition of pigmented skin lesions, and may even be used in real environment applications.

Keywords: pattern recognition, classification, pattern analysis, segmentation, dermoscopy, pigmented skin lesions.
LIST OF FIGURES

Figure 1.1: Four examples of dermoscopy images of the current database .................. 10
Figure 2.1: Examples of dermoscopy (a and c) and macroscopic (b and d) images .... 14
Figure 2.2: Application of smoothing filters ............................................................ 16
Figure 2.3: Segmentation results after applying Canny’s edge detector to a dermoscopy image (a and c), and to a macroscopic image (b and d) ................................... 19
Figure 2.4: Segmentation results after applying Otsu’s method (a) to the dermoscopy image shown in Figure 2.3a, and (b) to the macroscopic image shown in Figure 2.3b . 20
Figure 2.5: Segmentation results after applying the Mumford-Shah method .......... 21
Figure 2.6: Segmentation results obtained after applying the fuzzy c-means method... 23
Figure 2.7: Segmentation results obtained after applying the Chan-Vese model ...... 25
Figure 2.8: Distribution of the reviewed methods for the segmentation of skin lesions, according to the applied principle ................................................................. 26
Figure 2.9: Examples of global patterns in dermoscopy images .............................. 37
Figure 2.10: Examples of local patterns in a dermoscopy image ............................. 38
Figure 3.1: Pipeline of the proposed approach for skin lesion pattern recognition..... 59
Figure 3.2: Pattern analysis process ................................................................. 62
Figure 3.3: Pattern recognition by using ensemble models................................. 65
Figure 4.1: Segmentation results ................................................................. 69
Figure 4.2: Examples of feature descriptors for skin lesions ............................. 70
## LIST OF TABLES

Table 2.1: Research that has been performed on the image segmentation of skin lesions................................................................. 17

Table 2.2: Descriptors that have been used in the feature extraction of skin lesions in images based on the ABCD(E) rule and texture analysis............................................ 28

Table 2.3: Descriptors that have been used in the feature extraction of skin lesions according to the pattern analysis, seven-point checklist and Menzies’ method.............. 29

Table 2.4: Criteria of the ABCD(E) rule for the diagnosis of pigmented skin lesions from clinical and dermoscopy analysis ........................................................................ 30

Table 2.5: The pattern analysis in dermoscopy images according to the literature........ 36

Table 2.6: Diagnostic criteria included the seven-point checklist and Menzies’ methods........................................................................ 39

Table 2.7: Summary of advantages and disadvantages of feature selection models..... 43

Table 2.8: Classification methods applied to discriminate skin lesions from images and their advantages and disadvantages................................................................. 46

Table 2.9: Results of recent studies focused on the global pattern classification ....... 50

Table 2.10: Results of recent studies focused on the skin lesion classification .......... 50

Table 3.1: Number of new cases of skin cancer, according to gender, in the USA, UK and Brazil ........................................................................................................... 56

Table 3.2: A brief description of the current image database........................................ 60

Table 4.1: Classification results of malignant and benign lesions according to feature selection filters............................................................... 71

Table 4.2: Task timeline according to the proposed approach steps......................... 72
LIST OF ACRONYMS

ABCD(E): asymmetry, border, colour, diameter or differential structures, evolution or elevation
AI: artificial intelligence
AMSFCM: anisotropic mean shift algorithm
ANNs: artificial neural networks
ASM: angular second moment
AUC: area under an ROC curve
BCM: box-counting method
C&R: classification and regression
CART: classification and regression trees
CFS: correlation-based feature selection
CIE: international commission on illumination
CMY: cyan, magenta, yellow
CSLM: confocal scanning laser microscopy
DOG: derivatives of Gaussian
ELM: epiluminescence microscopy
FCM: fuzzy c-means
FFT: fast fourier transform
FN: false negative
FP: false positive
GA: genetic algorithm
GLCM: grey-level co-occurrence matrix
GP: genetic programming
GRFS: gain ratio feature selection
GSFS: generalized sequential feature selection
GVF: gradient vector flow
HCA: hill-climbing algorithm
HNB: hidden Naïve Bayes
HSV: hue, saturation, value
HVC: hue, value, chroma
IFFT: inverse fast Fourier trans
KNN: k-nearest neighbour
L*C*H: lightness, chroma, hue
LMT: logistic model tree
MFHs: multidimensional receptive field histograms
MIFS: mutual information-based feature selection
MLP: multilayer perceptron
MRF: Markov random fields
MRI: magnetic resonance imaging
NBL: Naïve Bayes multinomial
NBTree: naïve bayes/decision tree
OCT: optical coherence tomography
OPF: optimum-path forest
PCA: principal component analysis
RAC: region-based active contour
RBF: radial basis function
RF: random forest
RGB: red, green, blue
RMS: root mean squared
ROC: receiver operating characteristics
ROI: region of interest
SMOTE: synthetic minority oversampling technique
SRM: statistical region merging
SVMs: support vector machines
TDS: total dematoscopy score
TN: true negative
TP: true positive
UK: United Kingdom
USA: United States of America
XVAL: cross-validation
CONTENTS

1 INTRODUCTION .................................................................................................................. 9

2 STATE-OF-THE-ART REVIEW .......................................................................................... 13
  2.1 Imaging techniques ........................................................................................................... 13
  2.2 Image pre-processing ...................................................................................................... 15
  2.3 Image segmentation ........................................................................................................ 16
    2.3.1 Edge-based segmentation ....................................................................................... 18
    2.3.2 Thresholding-based segmentation .......................................................................... 18
    2.3.3 Region-based segmentation ................................................................................... 20
    2.3.4 Segmentation based on artificial intelligence ....................................................... 21
    2.3.5 Segmentation based on active contours .................................................................. 23
    2.3.6 Discussion ............................................................................................................... 25
  2.4 Feature extraction .......................................................................................................... 27
    2.4.1 ABCD(E) rule .......................................................................................................... 30
    2.4.2 Texture analysis ...................................................................................................... 34
    2.4.3 Pattern analysis ...................................................................................................... 35
    2.4.4 Seven-point checklist ............................................................................................. 38
    2.4.5 Menzies’ method .................................................................................................... 40
    2.4.6 Discussion ............................................................................................................... 40
  2.5 Skin lesion classification ................................................................................................. 40
    2.5.1 Feature selection ..................................................................................................... 41
    2.5.2 Methods for classification ....................................................................................... 45
    2.5.3 Evaluating the performance of classification ......................................................... 47
    2.5.4 Skin lesion classification and its features ............................................................... 49
    2.5.5 Discussion ............................................................................................................... 51
  2.6 Summary ........................................................................................................................ 53
3 THESIS STATEMENT ................................................................................. 56
  3.1. Motivation and problems ..................................................................... 56
  3.2. Objectives and expected contributions .............................................. 58
  3.3. Proposed Approach ........................................................................... 58
    3.3.1. Image database ........................................................................... 59
    3.3.2. Image enhancement and noise removal ......................................... 60
    3.3.3. ROI extraction ........................................................................... 60
    3.3.4. Pattern analysis ........................................................................... 61
    3.3.5. Pattern recognition ..................................................................... 64
    3.3.6. Validation process ........................................................................ 67
  3.4. Summary ............................................................................................ 68

4 THESIS PLANNING ................................................................................. 69
  4.1. Previous work .................................................................................... 69
  4.2. Task planning ..................................................................................... 72

5 IMPLICATIONS OF RESEARCH .............................................................. 74

REFERENCES .......................................................................................... 75
1 INTRODUCTION

Computational systems for skin cancer diagnosis have been proposed in order to aid dermatologists in early assessment of such diseases, or even to monitor pigmented skin lesions [40, 85, 143, 145, 153]. Such lesions, which may be classified as benign or malignant, are mainly caused by an abnormal production of melanocyte cells, caused by factors such as excessive sun exposure. Melanocyte cells are responsible for creating the substance melanin, one of the functions of which is to provide pigmentation in the skin.

Benign lesions display a more organized behaviour than malignant lesions, since the former are not able to proliferate into other tissues. Seborrheic keratosis (Figure 1.1b) and nevus (Figure 1.1a), such as melanocytic, blue, halo, spitz, dysplastic, Clark, recurrent, reed, congenital, compound, intradermal, junctional, dermal, are examples of benign lesions. In the case of malignant lesions, i.e., skin cancer, the cells split quickly, and may invade other parts of the body. Indeed, these cells do not die, as generally occurs with normal cells. Skin cancer may be divided into two categories: melanoma (Figure 1.1c) and non-melanoma (Figure 1.1d). Melanoma is the most aggressive form of skin cancer, and also the one with the highest mortality rate, due to its high levels of metastasis [57]. Basal cell carcinoma and squamous cell carcinoma are two examples of non-melanoma skin cancer (NMSC), the most common of all skin cancers. Moreover, these types of cancer have a higher chance of cure than melanoma, since they have a reduced capacity to spread (metastasis) to other parts of the body. The number of skin cancer cases has increased in recent years, and consequently an increasing number of deaths caused by this disease has been reported, particularly by melanoma [82, 163, 172]. Thus, pigmented skin lesions have been a cause of global concern, since some types of benign lesions may become skin cancer, such as dysplastic nevus.
Figure 1.1: Four examples of dermoscopy images of the current database: (a) junctional nevus, (b) seborrheic keratosis, (c) squamous cell carcinoma, and (d) melanoma.

Dermoscopy images [22, 40, 69, 86, 142, 159], as illustrated in Figure 1.1 are widely applied for image analysis of pigmented skin lesions. Such images may be acquired from dermatoscope devices or specific cameras to provide a better visualization of the pigmentation pattern on the skin surface. The identification of regions of the lesion in such images may be performed in order to assist in the process of analysis. Segmentation is an important step that allows the extraction of regions of interest (ROI) within a dermoscopy image [37]. Previous studies [110, 144, 159, 179, 184, 188] have shown that computational methods for image segmentation may provide suitable results for the identification of skin lesions in images. Frequently, the images under analysis are pre-processed for image enhancement and artefact removal, so that more robust segmentations may be achieved [6, 24].

Image computational analysis of pigmented skin lesions is usually based on the intensities of the pixels within the segmented ROIs. Furthermore, the extraction of representative features of the ROIs under analysis is an important step for the efficient classification of the segmented lesions. In this step, a few common difficulties are: 1) identification of the features to be used; 2) to confirm that the number of selected features is sufficient to describe the classification problem; 3) the number of selected features is too large, which requires high computational resources; and 4) there are unnecessary and/or irrelevant features that should be removed from the features set. Techniques to reduce the dimensionality of the data may be used to solve these problems according to the following reduction strategies: feature transformation (well-known as feature
extraction in literature concerning pattern recognition [66, 176]), and feature selection [73].

The feature extraction strategy allows the modification of all of the data of the image, in order to emphasize the most effective features, ensuring the correct separation of the classification classes [66]. Such strategy is based on the generation of a new feature space, which may expand or reduce, according to the adopted strategy. The new features may be extracted by means of discovery of missing information from relationships among the features, or even by means of searching for a new feature space with smaller dimensions, through functional mapping. In contrast, new features are not created in the feature selection strategy, meaning that a subset from the original features is selected by using this approach. Both approaches may also be combined in order to achieve a better representation of the features. For example, in cases in which the feature extraction step increases the number of features, feature selection techniques could provide an automatic reduction of such excessive features. Furthermore, a larger feature space may include redundant or irrelevant data [103, 118]. For example, the principal component analysis (PCA) [76], which is a technique for space dimensionality reduction based on basic linear transforms of the input features, has been applied to both feature extraction [99, 101] or feature selection [85, 86, 111, 142, 156].

Several clinical methods are used by dermatologists in order to diagnose pigmented skin lesions based on their own features [9, 27, 87]. The ABCD(E) rule is a commonly used method to classify such lesions based on Asymmetry, Border, Colour, Diameter (or Differential Structures in the case of dermoscopy images), and Evolution (or Elevation) criteria. Additionally, texture analysis may be performed in order assess the superficial roughness of the lesions. Clinical methods based on pattern analysis are also used to classify features in pigmented skin lesions specifically visible in dermoscopy images. The Seven-point checklist and Menzies’ methods were proposed to simplify the clinical approaches based on pattern analysis methods [87]. Several computational solutions [46, 84-86, 111, 116, 168] have been proposed for feature extraction and selection of pigmented skin lesions, in order to represent them according to a certain criteria. Such features may be used for the classification process, in order to provide dermatologists with a computer-aided diagnosis of the pigmented skin lesions [50, 69, 145, 183].

Classifiers are important techniques to assist in the computational diagnosis of skin lesions from dermoscopy images. The skin lesion classification systems should have a high performance, considering that they will be used to assist in dermatological diagnosis. The evaluation and improvement of the performance of the classifier is an essential characteristic of pattern recognition [176]. A relevant problem that affects the performance of classifiers is the definition of which features are meaningful, in order to adequately represent the classes. Therefore, the application of several descriptions may be necessary, considering the large quantity of features present in images. A solution for this problem is the application of techniques of feature selection to define the more appropriate features from images [40, 69, 86]. These techniques permit the removal of redundant and/or irrelevant features. Moreover, these techniques reduce the
dimensionality of data. As a result, they reduce the time of feature extraction, training and testing, decrease the complexity of classification, and improve the classification accuracy rate. Another solution to improve the performance of the classifier is with ensemble models [12, 22, 142].

Ensemble models may be constructed of either only one induction algorithm, classified as homogeneous [139] or several induction algorithms, classified as heterogeneous [60, 185]. Several methods for constructing homogeneous ensembles have been developed through data manipulation, such as manipulating the training samples and the input features [58]. Algorithms for manipulating the training samples allow the generation of multiple hypotheses, in which the learning algorithm is applied to different subsets of the training samples. The traditional algorithms applied to manipulate the training samples are Bagging and Boosting [58]. Random Forest and AdaBoost are also popular ensemble methods. Random Forest [30] is a variation of the Bagging algorithm that is used to create individual decision trees, whereas AdaBoost [191] is a popular boosting algorithm that maintains a set of weighting systems over the training samples. Algorithms for manipulating the input features allow the generation of ensembles based on different features available to the learning algorithm [58]. This process may be, for example, the splitting of a set of features into subsets [97] or by using feature selection for ensembles [170, 185, 191].

The objective of this thesis is the development of a new approach based on ensemble model and feature selection for skin lesion pattern recognition in dermoscopy images, in order to provide information that may assist the dermatologist, as well as to contribute to the CAD area with an increased performance of the automatic diagnosis. The thesis is organized as follows: In Chapter 2, an overview of the main computational vision methods that have been applied in image-based analysis and classification of skin cancers and other pigmented skin lesions, is addressed. In addition, a current discussion about the application of such methods, as well as future trends concerning computational methods for skin lesion image analysis, is also provided. In Chapter 3, an approach for skin lesion pattern recognition in dermoscopy images is proposed. The goals and objective of the achievement of this project, as well as the proposed approach steps are described. In Chapter 4, a brief description of the previous work is presented, along with the outline of task planning for the following years of this thesis. Finally, the main implications and expected contributions of this thesis are presented.
2 STATE-OF-THE-ART REVIEW

This section reviews computational vision methods proposed for analysing pigmented skin lesions in images, particularly in clinical and dermoscopy images. This review aims to present the current methods, and outline a comparative analysis with regards to several of the fundamental steps of image analysis, such as image acquisition, pre-processing, segmentation, feature extraction and classification. The techniques employed in segmentation step are explained, and their strengths and weaknesses are identified. Moreover, several of the reviewed image processing techniques are applied to macroscopic and dermoscopy images, in order to exemplify their results. We introduce and discuss descriptors applied to extract the features of such lesions according to several clinical methods, including the ABCD rule, texture analysis, seven-point checklist, pattern analysis and Menzies’ method. In addition, methods suggested for classifying lesions from their features are identified, including the feature selection step, classifiers used, evaluation procedures adopted and their advantages and disadvantages. Finally, the properties and performance of some of the reviewed methods are summarized. Furthermore, the discussion of the each step and future trends are outlined.

2.1. Imaging techniques

Different non-invasive imaging techniques have been employed to assist dermatologists in the diagnosis of skin lesions. Dermoscopy, confocal scanning laser microscopy (CSLM), optical coherence tomography (OCT), ultrasound, magnetic resonance imaging (MRI), and spectroscopic imaging are examples of these techniques [146, 162, 175]. Macroscopic images, commonly known as clinical images [12, 35, 179], and images acquired by epiluminescence microscopy (ELM), also called dermoscopy or dermatoscopy images [42, 69, 123, 159, 184, 188, 190], are normally used in the computational analysis of skin lesions. Figure 2.1 presents examples of dermoscopy and macroscopic images.
Clinical images are usually obtained using common digital video or image cameras. However, the imaging conditions are frequently inconsistent; for example, images are acquired from variable distances or/and under different illumination conditions. Furthermore, the images may have poor resolution, which may constitute an obstacle when the size of the lesion is small. An additional problem with clinical images is related to the presence of artefacts, such as hair, reflections, shadows and skin lines, which may hinder the adequate analysis of the imaged skin lesions.

Essentially, ELM is a non-invasive technique for image acquisition, where the lesion may be immersed in oil, and subsequently a dermatoscope device (which includes a specific camera) acquires the images. This technique allows a better visualization of the pigmentation pattern on the skin surface. Besides the non-polarised imaging modality due to the oil immersion, there are two other modalities of ELM that may be used: cross-polarization and transillumination, also called side or epi-transillumination. In these modalities, the images are acquired via a nevoscope device, which allows the acquisition of images with a variable amount of transillumination or cross-polarized surface light. Both modalities highlight the surface pigmentation, but the transillumination modality has the advantage of highlighting the subsurface vasculature and blood flow. However, hairs and air bubbles must be subsequently removed from the images, to allow for a better recognition of the skin lesions.
2.2. Image pre-processing

The image pre-processing step is an important aspect for the effective identification and analysis of pigmented skin lesions in images. Effective approaches based on colour space transformation [3, 7], contrast enhancement [3, 7, 152] and artefact removal [3, 8] as a pre-processing step have been proposed in order to improve the segmentation accuracy. As mentioned earlier, the images under analysis may contain several artefacts, such as hairs, reflections, shadows, skin lines and air bubbles, which may affect the accuracy of the image segmentation step. Nevertheless, these artefacts may be reduced by the application of image filters, such as the median, average and anisotropic diffusion filters, thus improving the segmentation accuracy, without losing relevant information about the lesions.

The median filter [71], which is a non-linear image filtering method, is commonly applied on noisy images, and has shown successful results. Unlike linear filters, such as the average filter [71], this type of filter allows the smoothing of the original image without blurring edges and thin details. The median filter has often been applied to smooth images of skin lesions, as well as to remove artefacts, maintaining the edges of the lesion, which is imperative for an adequate segmentation [39, 40, 70, 102, 159]. To establish the best median filtering mask for the smoothing of skin lesion images, Celebi et al. [39] established a theory, which considers that, for an effective smoothing, the size of the filtering mask should be proportional to the size of the input image.

Anisotropic diffusion [137] has also been used for smoothing skin lesion images [24, 138]. This filter is applied iteratively, such that the number of iterations is determined according to the amount of noise present in the input image. However, relevant edges may be removed when the number of iterations is too large. Improvements have been proposed, in order to enhance the results of the anisotropic diffusion filter. For example, Barcelos et al. [23] proposed an enhancement to the anisotropic diffusion algorithm originally suggested by Perona and Malik [137]. The improved algorithm not only aims at smoothing very noisy images without removing relevant edges, but also considers the improvements proposed by Alvarez et al. [13] and Nordström [122] to enhance the edges.

Results of the application of the median [71], average [71] and anisotropic diffusion [137] filters are shown in Figure 2.2. A matrix, using a 9-by-9 neighbourhood, was applied, both to perform the median and average filtering, since different masks did not provide a smoother image nor lower the image noise level. The resultant smoothed image obtained with the application of the median filter has reduced the presence of hairs. In addition, unlike the smoothed image obtained using the average filter, the lesion edges were enhanced. The application of the anisotropic diffusion filter led to good results with respect to noise removal. The smoothing process stopped when the maximum number of iterations was reached (150 iterations). Nevertheless, this filter did not enhance the lesion edges like the median filter.
Morphological filtering [164], which is based on set theory, may also be used to enhance skin lesions in images [26]. For example, one may refer to the work of Beuren et al. [26]; they used colour morphological filtering to enhance the regions of the lesions. Moreover, morphological filtering has been applied in order to include in the lesion region areas of low contrast borders [123, 124], and to remove image noise [124, 159].

Unlike most methods proposed in the literature for reducing the influence of hairs on images of skin lesions, Abbas et al. [6] suggested an effective pre-processing method for the reduction of different artefacts and consequently, a better detection of lesion borders. Essentially, this method consists of three steps: 1) specular reflection reduction by applying homomorphic filtering [10], fast Fourier transform (FFT) and high pass filtering, in order to modify the illumination and reflectance, thus obtaining high contrasted skin lesions, 2) the reduction of dermoscopy-gel or air bubbles, based on an adaptive and recursive weighted median filter, and 3) hair, blood vessel and skin line detection and reduction, using a line detection procedure based on the two-dimensional (2D) derivatives of Gaussian (DOG) [141], and the exemplar-based inpainting technique [53].

2.3. Image segmentation

Segmentation allows the extraction of the ROI of an image. Bearing in mind that the skin lesion is the ROI in the image under analysis, the segmentation process should not cease until the lesion is fully detached from the image background, or until some other outcome is
reached. Some artefacts, such as hairs, reflections, shadows, skin lines and bubbles, may influence the result of the segmentation process, making it a complex computational task. Nonetheless, as mentioned previously, pre-processing techniques may be applied to the original images, with the purpose of facilitating the segmentation process, and improving the resultant accuracy.

In general, the segmentation process is based on the discontinuity and similarity of some properties of the ROIs to be segmented [71]. The segmentation methods may be edge-based, i.e., the approaches are based on information about the image edges, more specifically, they search for abrupt changes, i.e., discontinuities, in the intensity of the image pixels relative to their neighbours. Edge detectors are the most common examples of such methods. In addition, the segmentation process may depend on similarity criteria, such as similar grey-levels, colours or textures. Thresholding- and region-based segmentation are some examples of methods that use similarity criteria to identify skin lesions in images.

In the following sections, we will discuss the applicability of some methods commonly used in the literature for the segmentation of pigmented skin lesions in images, such as the edge-, thresholding- and region-based methods, and methods based on artificial intelligence (AI) and active contours. The reviewed research is summarized in Table 2.1. Research that combines different methods [34, 110, 142, 184] and that compare segmentation methods [159] is also included in Table 2.1.

Table 2.1: Research that has been performed on the image segmentation of skin lesions.

<table>
<thead>
<tr>
<th>Segmentation method</th>
<th>Technique</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edge-based</td>
<td>Edge detectors</td>
<td>[24, 68, 138]</td>
</tr>
<tr>
<td></td>
<td>Otsu’s thresholding</td>
<td>[7, 40, 70, 101, 123, 124]</td>
</tr>
<tr>
<td></td>
<td>Fuzzy logic</td>
<td>[184]</td>
</tr>
<tr>
<td></td>
<td>Renyi’s entropy</td>
<td>[26]</td>
</tr>
<tr>
<td></td>
<td>Adaptive thresholding</td>
<td>[22, 159]</td>
</tr>
<tr>
<td></td>
<td>Iterative thresholding</td>
<td>[142, 152]</td>
</tr>
<tr>
<td></td>
<td>Ensemble</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
<td>[182]</td>
</tr>
<tr>
<td>Thresholding-based</td>
<td>Region growing</td>
<td>[40, 85, 102, 110, 159]</td>
</tr>
<tr>
<td></td>
<td>Statistical region merging</td>
<td>[39, 68]</td>
</tr>
<tr>
<td></td>
<td>Iterative stochastic region merging</td>
<td>[179]</td>
</tr>
<tr>
<td>Region-based</td>
<td>Neural networks</td>
<td>[34, 152]</td>
</tr>
<tr>
<td></td>
<td>Evolutionary computation</td>
<td>[144]</td>
</tr>
<tr>
<td></td>
<td>Fuzzy logic</td>
<td>[34, 110, 142, 159, 184, 190]</td>
</tr>
<tr>
<td>AI-based</td>
<td>Adaptive snake</td>
<td>[159]</td>
</tr>
<tr>
<td></td>
<td>Gradient vector flow</td>
<td>[159, 187-189]</td>
</tr>
<tr>
<td></td>
<td>Level set</td>
<td>[106]</td>
</tr>
<tr>
<td></td>
<td>Region-based active contour algorithm</td>
<td>[3, 6]</td>
</tr>
<tr>
<td></td>
<td>Active contour model without edges</td>
<td>[159]</td>
</tr>
<tr>
<td></td>
<td>Expectation-maximization level set</td>
<td>[159]</td>
</tr>
</tbody>
</table>
2.3.1. Edge-based segmentation

In an image to be segmented, changes in the pixel intensity may be determined based on the magnitude of the gradient used to detect the edges of the ROI [71]. The Prewitt, Sobel, Roberts, Laplacian [71] and Canny [32] operators are common examples of edge detectors that lead to image segmentation based on edges. According to Sonka et al. [164], edge detectors may only achieve partial image segmentation. Therefore, the application of another segmentation method is needed to improve the final segmentation result. In particular, edge detectors present the following problems [164]: 1) the detection of an edge where no real border exists, 2) the non-detection of an edge where a real border exists, 3) the possibility generating of double edges, and 4) the large sensitivity to image noise.

Generally, the edge detector developed by Canny [32] is applied on skin lesion images [24, 138], due to its advantages compared to other edge detectors: 1) it provides a good edge detection with a low error probability, 2) it allows a good location of the edge pixels, and 3) it avoids the detection of double edges. Barcelos and Pires [24] and Pires and Barcelos [138] employed Canny’s edge detector after the application of an anisotropic diffusion smoothing filter [23], and the results showed that the unwanted edges were removed. However, some regions of the skin lesions were not included in the detected edge map and the edges were not completely closed.

Figure 2.3 presents the segmentation results of applying Canny’s edge detector to two original skin lesion images [32]. Customarily, a median filter [71] is applied before the edge detector, in order to smooth the original image. The result of this filter is the enhancement of the lesions, and the reduction of the image noise level. However, the edges generated by Canny’s edge detector are not satisfactory, both for dermoscopy and macroscopic images. Although the lesions were identified by the detector, the generated edges are discontinuous; thus, the boundaries of the lesions were not fully detected. In addition, there was a large sensitivity to the noise, which generated boundaries that are not part of the lesions.

2.3.2. Thresholding-based segmentation

The thresholding technique has been commonly used in several skin lesion segmentation methods proposed in literature [22, 26, 40, 42, 70, 101, 123, 124, 142, 152, 159, 182, 184]. This technique is based on the histogram of the input image, which represents the distribution of the input image pixels in terms of each possible intensity level. Thus, the thresholding technique entails the selection of one or multiple threshold values to separate the ROIs in the input images.
Among the various techniques proposed in literature to define the threshold value(s), we may cite Otsu's method [134], which has many applications in image segmentation of skin lesions [7, 40, 70, 101, 123, 124]. This method is based on a normalized histogram built in order to set the optimal threshold value, which separates the pixels of the input image into two homogeneous classes with minimal variance between them: one class for the ROI and the other for the image background. However, Otsu's method has revealed some problems, such as: (1) the segmented lesions tend to be smaller than their real size; and (2) it may lead to very irregular lesion edges. Yuksel and Borlu [184] proposed a method using the type-2 fuzzy logic technique [114] to solve such problems, which automatically determines the threshold value to segment dermoscopy images. This technique exhibits good performance in dealing with fuzzy values, by determining whether a specific image intensity level belongs to lesion regions or belongs to the background skin.

Figure 2.4 presents the segmentation results after the application of Otsu's method [134] to a dermoscopy and a macroscopic image. A median filter [71] was employed before the segmentation step to reduce the noise in the original images. Although several lesion boundaries are correctly detected, several other regions, such as edges with low contrast, were not identified as part of the lesions. Furthermore, this technique is very sensitive to artefacts and, therefore, because of reflections, some interior regions of the lesions were wrongly identified as belonging to the lesions.
A thresholding method based on Renyi’s entropy [151] has also been applied to define the threshold value, leading to segmentations that preserve the lesions’ geometry and shape [26]. Based on Renyi’s entropy, two probability distributions are defined using the pixels’ intensity: one for the object, i.e., the skin lesion, and another for the image background, i.e., the skin background.

2.3.3. Region-based segmentation

The region growing algorithm [71], splitting and merging operations [71] and Mumford-Shah method [119] are examples of region-based techniques that have been used to segment skin lesion images. The region growing algorithm consists in grouping similar neighbouring pixels, or grouping sub-regions, into larger homogeneous regions according to a growing criterion. For example, in a given region of an image, pixels with similar properties, such as grey-level, colour or texture, are grouped [36, 40, 85]. The splitting and merging operations are region-based techniques applied to group similar regions [110, 159]. Thus, the same intensity is attributed to all input pixels that have similar intensity, in agreement with the grouping criterion. On the other hand, the Mumford-Shah method divides the original image into several regions, merging the close regions by analysing their pixel intensities. The active contour without edges model [45] is based on this method and has been used for the image segmentation of skin lesions [6, 159].

Examples of the results obtained by the Mumford-Shah method applied to skin lesion images are presented in Figure 2.5. The method was employed on two images that were previously smoothed using the median filter [71]. Observation of the resultant images, shows that the lesions were completed identified, including the lesion regions with considerable colour variation. However, some regions were erroneously identified as belonging to the lesions due to artefacts. In addition, the method may not adequately identify lesion regions that show low contrast relative to the skin background.
Figure 2.5: Segmentation results after applying the Mumford-Shah method: (a) to the dermoscopy image shown in Figure 2.3a, and (b) to the macroscopic image shown in Figure 2.3b.

Celebi et al. [39] and Ganzeli et al. [68] employed the statistical region merging (SRM) algorithm [121] to detect the edges in images of skin lesions. This algorithm is a technique developed to segment colour images based on region growing and merging. Simplicity, computational efficiency and excellent performance are the main advantages reported for the SRM algorithm. Image quantization and colour space transformation steps, that are commonly applied to the original images before their segmentation, are unnecessary when this algorithm is used to segment skin lesion images.

A method to segment skin lesion images through iterative stochastic region merging has been proposed by Wong et al. [179] based on the SRM algorithm [121]: each image pixel is assigned to a single region, which is subsequently merged with other regions in a stochastic way, based on a probability function of regions fusion. This process is characterized by a multi-path refining of the results, in order to achieve the best final segmentation. This method has been shown to be robust to image artefacts, and to perform successfully in cases where several skin lesions, structural lesion variations, varying illuminations and colour variations are present in the input images. In addition, it generates successful segmentation in cases where there is a low contrast between the lesion and the skin background near the lesion boundaries.

2.3.4. Segmentation based on artificial intelligence

Techniques based on AI have also been proposed for the image segmentation of skin lesions, in which pixels are classified as belonging to the ROIs or to the background of the images. Neural networks, evolutionary computation and fuzzy logic are some examples of these techniques, which aim at performing similar tasks to humans, based on learning, natural evolution and human reasoning. These techniques may be combined among themselves, or with other traditional image processing techniques, to improve segmentation performance.

Artificial neural networks (ANNs) [80], which are parallel distributed systems composed of simple processing units, have been applied to segment images with skin lesions [152], with the purpose of obtaining similar results to the human brain. The segmentation performance of ANNs may be improved through the application of Genetic Algorithms (GAs) [79], which are computational techniques for searching and optimization. GAs are based on
natural evolution and biological genetics, which aim at finding the best solution for a given problem, e.g. GAs may be employed to optimize the ANN parameters.

Roberts and Claridge [144] presented a method to segment skin lesion images through genetic programming (GP) [95], which is a technique based on natural evolution to solve problems following the concepts of genetic algorithms. The proposed method consists in creating a random population of programs from the function and terminal sets. The function set is built from the image processing operations, such as image thresholding, morphological operations, edge detection and merging. The terminal set is built from information in the input image data, such as the intensity and coordinate values of the pixels. This method showed good generalization with a very small set of training samples. Furthermore, the system learns by example, thus increasing the amount of problems in which it is applicable. However, this method has some disadvantages regarding the complexity of its implementation and the presence of unnecessary steps, requiring high computational effort.

Fuzzy logic deals with uncertain and imprecise values. Many algorithms based on fuzzy logic have been proposed to segment skin lesions in images [34, 110, 142, 159, 184, 190]. This method allows the representation of intermediate values within an interval; in other words, the input data is qualitatively analysed (linguistic values). Frequently, the fuzzy method is applied together with other segmentation techniques. In Maeda et al. [110] and Silveira et al. [159] the fuzzy method, combined with both splitting and merging techniques, was used to segment dermoscopy images. This combination, originally proposed Maeda et al. [108] and Maeda et al. [109], generates an algorithm for unsupervised perceptual segmentation of natural colour images using a fuzzy-based homogeneity measure, which performs the fusion of colour and texture features. The algorithm includes four steps: simple splitting, local merging, global merging and boundary refinement.

The fuzzy method was also used to define a threshold value from fuzzy intensity, by applying the type-2 fuzzy logic technique [114]; the idea was to determine whether a specific intensity belongs to a ROI or to the image background [184]. Another method, named neuro-fuzzy approach [34], combines fuzzy logic with neural networks to segment dermatological images. In addition, fuzzy logic, combined with clustering techniques, has been employed in the image segmentation of skin lesions, e.g. the fuzzy c-means (FCM) algorithm [142, 190]. The basic idea behind the FCM algorithm is to find the centre of each cluster, similarly to the traditional c-means algorithm. Nevertheless, this process is more flexible, since partial membership may be introduced in the clusters.

Figure 2.6 presents the segmentation results obtained by applying the fuzzy c-means method to two images, which have been previously smoothed by applying the median filter [71]. The resultant images indicate that the lesions were successfully segmented. However, some lesion pixels with low contrast were not clustered into the lesion groups.
Figure 2.6: Segmentation results obtained after applying the fuzzy c-means method: (a) to the dermoscopy image shown in Figure 2.3a, and (b) to the macroscopic image shown in Figure 2.3b.

Zhou et al. [190] proposed a new mean shift approach, based on the FCM algorithm, called the anisotropic mean shift algorithm (AMSFCM), to segment dermoscopy images. The AMSFCM algorithm [174] is more effective than the FCM algorithm, and requires less computational time than the traditional mean shift technique. Furthermore, it provides superior segmentation results. Mean shift-based techniques [51] allow the estimation of local density gradients of similar pixels, by using radially symmetric kernels. However, these kernels may not adequately deal with the presence of irregular structures and noise in the input image. On the other hand, the AMSFCM algorithm provides improved performance in these cases, since it uses an anisotropic kernel.

2.3.5. Segmentation based on active contours

Algorithms based on active contours have been used for segmenting skin lesion images [3, 6, 159, 188]. In these algorithms, the initial curves move toward the boundaries of the objects of interest through appropriate deformation. A deformable model may be classified as parametric [33, 88, 180, 181] or geometric [43-45, 133, 173, 186], according to the technique used to track the curve movement.

Parametric models include the traditional active contour models, namely, snake models [88]. Typically, in these models, the curve deformation is guided by energy forces, in which an internal energy determines the edges’ smoothness by the definition of the curve elasticity and rigidity; in other words, it controls the degree of shrinkage or expansion of the curve, in order to avoid over-deformations. An external energy is also included in the models, which has the function of driving the curve to the desired boundary. This energy may be defined by the user or through an automatic process. An image-based energy may also be defined, which drives the curve toward interesting image features, such as those based on image intensity, and gradient, or line segments and corners. However, this model has some limitations [45, 180, 181]: 1) the curve initialization must be near the object’s boundary, 2) it has difficulty in dealing with boundaries with large curvatures, 3) the stop criterion of the curve deformation depends on the image gradient, which may cause bad edge localization when the gradient value is not
high enough, and 4) it has difficulty in dealing with topological changes during the curve evolution.

The gradient vector flow (GVF) [181] is another parametric model used in the segmentation of skin lesions [159, 188]. Xu and Prince [181] proposed a new external energy for the active contour models, which is computed by a linear partial differential equation, and extends the gradient vectors at the image edges to the whole image. The goal of the new model was to overcome some of the main problems of the traditional snake model, in particular, the curve initialization and the convergence onto boundary regions with large curvature. On the other hand, Zhou et al. [188], Zhou et al. [189] and Zhou et al. [187] proposed a new type of dynamic energy for the segmentation of skin lesions, that combines the classical GVF model [181] and the mean shift algorithm [49]. This algorithm was designed to find the most similar edges to the true boundaries, by calculating the distance between the centroid of the curve and the true boundary of the object of interest. Thus, the curve evolution towards the ROI is generated by the gradient vector flow, as well as by the mean shift of the pixels contained within the curve. This combination makes the model versatile and flexible, because the successful calculation of the image-based energies is guaranteed, even in very noisy images.

Geometric models are characterized by the topological changes that the curve may experience during the segmentation process, and are less dependent on the initial curve conditions. Level set [133] and active contour model without edges, known as Chan-Vese model [45], are examples of geometric models. The level set method was originally proposed by Osher and Sethian [133] to handle topological changes during the curve evolution, which is one of the limitations of the traditional parametric models. The curve evolution is implicitly tracked by a level set function, which allows the easy identification of a pixel: whether an image pixel is located inside, outside or on the curve. The geometric properties of the curve may be easily computed by the level set function.

The active contour model without edges proposed by Chan and Vese [43, 45] is based on the average of the intensities of the image pixels, and not on the image gradient. Therefore, the model uses the concepts of the Mumford-Shah [119] and level set [133] segmentation techniques. Essentially, the Chan-Vese model considers a "fitting" term for the energy minimization, which is calculated by means of energy functionals, to identify whether the object of interest is inside or outside the curve. The energy minimization allows the deformation of the curve toward the boundary of the object, where the inside and outside intensities are constant and similar. This model has been used in the segmentation of skin lesions [6, 159], due to its advantages when compared with other segmentation techniques based on the active contour model [45], such as: 1) the initial curve may be defined more freely in the image, 2) the inner contours are automatically detected without the need to introduce a second curve in the image, 3) the object detection is carried out even in the presence of varying intensities, very smooth boundaries and where the boundaries may not be successfully defined by the gradient, a situation which is not effectively handled by the traditional active contour model, and 4) it provides effective detection of object boundaries even on noisy images, without the necessity to previously smooth the original images.
Abbas et al. [3] presented an improved perceptually-oriented region-based active contour (RAC) scheme [98], where the segmentation concept is based on the Chan-Vese model [45] to determine the edges of lesions. The authors suggested this model due to its ability to simultaneously define multiple regions, to separate heterogeneous objects, to deal successfully with image noise, and because of the automatic convergence of the modelled curve.

Figure 2.7 presents the segmentation results obtained by applying the traditional Chan-Vese model [45] to two images, which have been previously smoothed using a median filter [71]. The segmentation process stopped when the edges were on the lesion boundaries, or when the maximum number of iterations was reached. From the resultant images, one may confirm that this model provided good segmentation results, having identified low contrast boundaries and overcome the image noise.

![Figure 2.7: Segmentation results obtained after applying the Chan-Vese model: (a) to the dermoscopy image shown in Figure 2.3a, and (b) to the macroscopic image shown in Figure 2.3b.](image)

2.3.6. Discussion

In general, the segmentation results are post-processed, in order to improve the accuracy of the obtained lesion edges. In many cases, morphological filters are used to smooth the edges, to remove the isolated regions and/or even to fill the interior of the segmented lesion regions [39, 123, 124, 142, 159, 190]. The final contours obtained for the lesions may be compared with ground truths defined by one or more specialists. Additionally, the accuracy of the edge detection results may be measured using statistical metrics in order to estimate the associated precision and recall, sensitivity and specificity, error probability and exclusive disjunction (XOR) [7, 41, 70]. The accuracy of the segmentation depends on the model and techniques used to solve the problem. Figure 2.8 illustrates the distribution of the methods reviewed in this article, according to the applied principle, which have been developed to segment pigmented skin lesions in images.
Threshold-based techniques have been widely used, mainly because of their simplicity, computational efficiency and good performance. Algorithms based on the active contour model have been frequently proposed for the segmentation of skin lesions. Nevertheless, parametric models have difficulty in dealing with topological changes and large curvatures. On the other hand, geometric models do not present such problems, but their computational complexity may be prohibitive. The wide use of techniques based on AI is justified by the advantages it offers, such as the possibility of learning from sample cases provided by the ANNs, the search and optimization for the best segmentation results provided by algorithms based on GAs, and the capability to deal with imprecise values that are provided by fuzzy logic. Region-based methods have also been used, since such methods have shown successful performance with several issues, such as illumination and colour variation. Usually, edge-based segmentation techniques are not applied independently, since these techniques may not completely identify the edges of the lesions, which is imperative in the analysis of skin lesions in images.

The hill-climbing algorithm (HCA) is a technique based on the clustering of points on an image, which is also applied to detect the ROI of skin lesion images [7]. This algorithm takes an image and the number of histogram bins in each dimension as input parameters, and returns a labelled image, whereas in the traditional k-means algorithm, the numbers of clusters (k) are specified manually by the users. Image segmentation based on such a technique consists of a simple, fast and non-parametric algorithm. In Abbas et al. [4] and Abbas et al. [1], a new segmentation method based on dynamic programming was proposed in order to overcome the limitation of thresholding, region-growing and clustering, as well as level set-based segmentation methods. This method is a general optimization solution, and due to its edge-based segmentation capabilities, to solve the local minima or overlapping problems, computational efficiency, and excellent performance to border detection in dermoscopy images.

In order to obtain enhanced segmentation results, both from dermoscopy and macroscopic images, pre-processing methods, such as the median filter [71], are usually applied to the original images to reduce the noise. In addition to the noise reduction, such methods may even enhance the lesions. For example, the anisotropic diffusion filter proposed by Barcelos et al. [23] has been used in order to reduce the noise and enhance the lesions in clinical images.
With regard to the segmentation step, Canny’s edge detector is not an indicated technique to segment skin lesions, since it may produce segmentations with edges not completely closed. On the other hand, Otsu, Mumford-Shah and FCM segmentation techniques may completely identify the lesions in the images. However, lesion boundaries with low contrast are generally not detected by such techniques. Moreover, these techniques are susceptible to image artefacts. Other techniques based on entropy or fuzzy logic, to define the threshold value [26, 184] may sometimes achieve superior segmentation results. The region-based approach proposed by Wong et al. [179] has a better segmentation performance, even in the presence of boundaries with low contrast. In addition, such a method allows in dealing with structural variations, varying illuminations and colour variations. Other techniques were also suggested to convert the FCM segmentation method into a more effective method for segmenting skin lesions in images [190]. Using these methods, better segmentation results may be achieved, even in the presence of irregular lesions and image noise. The Chan-Vese model [45] is a good option for the segmentation of skin lesions, since it adequately deals with varying intensities, low contrast boundaries and noisy images. Nevertheless, the Chan-Vese model also has disadvantages; for example, the segmentation result depends on the suitability of the initial curve.

2.4. Feature extraction

The feature extraction step in the computational analysis of pigmented skin lesions in images is usually based on methods commonly used by dermatologists in their clinical diagnosis. Several descriptors used in such a clinical diagnosis have been proposed for feature extraction in order to characterize skin lesions by computational methods. In general, these descriptors are categorized in shape features, border irregularity, colour variation and pattern analysis [40, 94, 150].

The ABCD (or ABCDE) rule, texture analysis, seven-point checklist, pattern analysis and Menzies’ method, are examples of approaches employed in the literature for the computational analysis of skin lesions in images [15, 29, 87]. The first two methods may be employed to extract features from both clinical and dermoscopy images, whereas the other methods are usually applied to dermoscopy images in order to identify more detailed pattern features on the surfaces of the lesions. Some suggested descriptors for the feature extraction of skin lesions based on the ABCD(E) rule and texture analysis are summarized in Table 2.2. On the other hand, the descriptors reported in the literature based on pattern analysis, the seven-point checklist and Menzies’ method, are indicated in Table 2.3.
Table 2.2: Descriptors that have been used in the feature extraction of skin lesions in images based on the ABCD(E) rule and texture analysis.

<table>
<thead>
<tr>
<th>Clinical method</th>
<th>Feature</th>
<th>Descriptor</th>
<th>References</th>
</tr>
</thead>
</table>
| Asymmetry (A)   |         | Asymmetry index descriptors based on axis of symmetry; | [46, 156] 1  
|                 |         | Size functions; | [40, 69] 2  
|                 |         | Statistical descriptors based on centre of gravity and distribution; | [54] 2  
|                 |         | Colour asymmetry descriptor; | [40] 2  
|                 |         | Geometrical descriptors. | [35] 1  
|                 |         |                          | [12, 40, 69, 84-86] 2  
| Border (B)      |         | Geometrical descriptors based on best-fit ellipse’s axes; | [46, 156] 1  
|                 |         | Statistical descriptors based on border gradient and periphery regions; | [35] 1  
|                 |         | Harmonic wavelet transform descriptor; | [50] 2  
|                 |         | Boundary-series analysis in histogram and the three-level wavelet transform. | [69] 2  
| ABCD(E) rule    |         | Statistical descriptors based on colour models; | [35, 46, 156] 1  
|                 |         | Amount of colour pixels; | [12, 40, 46, 84-86, 111, 142, 156] 2  
| Colour (C)      |         | Relative colour descriptors; | [46] 1  
|                 |         | Colour asymmetry; | [40] 2  
|                 |         | Centroid and LUV histogram distances (from the L*u*v* colour model); | [40] 2  
|                 |         | Border’s colour gradient; | [84-86] 2  
|                 |         | Quantized colour space. | [84-86] 2  
| Diameter (D)    |         | Semi-major axis of the best-fit ellipse. | [156] 1  
| Differential structures (D) | | Multidimensional receptive fields histograms. | [169] 2  
| Elevation (E)   |         | Thickness descriptors based on 3D surface image. | [27, 62, 77]  
| Evolution (E)   |         | Registration approach based on bipartite graph matching; | [81] 1  
|                 |         | Registration procedure based on stochastic gradient descent. | [161] 2  
| Texture analysis |         | Grey-level co-occurrence matrix; | [12, 40, 59, 84-86, 111, 142]  
|                 |         | Box-counting method; | [59, 69, 167]  
|                 |         | Wavelet-based descriptors; | [69, 167]  
|                 |         | Histogram-based descriptor; | [22]  
|                 |         | Statistical descriptor based on intensities of the pixels inside the lesion regions; | [35]  
|                 |         | Gabor filter banks; | [183]  
|                 |         | Auto-regression. | [183]  

1 According to the ABCD(E) rule for clinical images.
2 According to the ABCD(E) rule of dermoscopy.
Table 2.3: Descriptors that have been used in the feature extraction of skin lesions according to the pattern analysis, seven-point checklist and Menzies’ method.

<table>
<thead>
<tr>
<th>Clinical method</th>
<th>Feature</th>
<th>Descriptor</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pattern analysis</strong></td>
<td>Shape</td>
<td>Measures based on connected components;</td>
<td>[116, 168]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Edge detection;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mathematical morphology;</td>
<td>[83]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statistical descriptors;</td>
<td>[55, 86]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geometrical descriptors;</td>
<td>[55, 86]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Single instance learning method.</td>
<td>[160]</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Statistical descriptors based on colour models;</td>
<td>[86]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amount of colour pixels;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Border’s colour gradient;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantized colour space;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Texture</td>
<td>Multiscale steerable pyramids transform decomposition;</td>
<td>[5]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grey-level co-occurrence matrix;</td>
<td>[86, 168]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intensity histogram;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Differential statistical;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fourier power spectrum;</td>
<td>[168]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Run-length matrix.</td>
<td></td>
</tr>
<tr>
<td><strong>7-point checklist</strong></td>
<td>Shape</td>
<td>Irregularity ratio descriptor based on threshold.</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
<td>Principal component analysis;</td>
<td>[99, 101]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2D histogram construction;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peaks picking method;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statistical descriptors based on histogram;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brown pigmentation descriptor based on threshold;</td>
<td>[25]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absolute and relative colour descriptors.</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>Texture</td>
<td>Local discontinuity descriptors based on median filter and close-opening operation;</td>
<td>[99, 101]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High-pass filtering;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suitable thresholding;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fourier descriptors;</td>
<td>[25, 99, 101]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grey-level co-occurrence matrix.</td>
<td>[38]</td>
</tr>
<tr>
<td><strong>Menzies’ method</strong></td>
<td>Colour</td>
<td>Jeffries-Matusita and transformed divergence separability measures.</td>
<td>[158]</td>
</tr>
</tbody>
</table>
2.4.1. ABCD(E) rule

The ABCD(E) rule provides the following five criteria for diagnosis of pigmented skin lesions from clinical or dermoscopy images: asymmetry (A); border (B); colour (C); diameter or differential structures (D); and evolution or elevation (E) [9, 27, 166], Table 2.4. Such a rule has been widely used in the image-based analysis of skin lesions [12, 46, 156] (Table 2.2). The clinical diagnosis usually considers all the aforementioned criteria to diagnose the malignancy of the lesions, whereas in the computational diagnosis, the criteria are used individually [158] or combined among them [156], according to the goals to be achieved.

Table 2.4: Criteria of the ABCD(E) rule for the diagnosis of pigmented skin lesions from clinical and dermoscopy analysis.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Clinical</th>
<th></th>
<th>Dermoscopy 1</th>
<th></th>
<th>Weight factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign lesion</td>
<td>Malign lesion</td>
<td>Definition</td>
<td>Score</td>
<td></td>
</tr>
<tr>
<td>Asymmetry (A)</td>
<td>Shape is symmetric</td>
<td>Shape is asymmetric</td>
<td>Border, colours or structures are asymmetric in 0, 1, or 2 perpendicular axes</td>
<td>0-2</td>
<td>1.3</td>
</tr>
<tr>
<td>Border (B)</td>
<td>Border is regular or well-defined</td>
<td>Border is irregular or ill-defined</td>
<td>Abrupt cut-off of network at the border in 0-8 segments</td>
<td>0-8</td>
<td>0.1</td>
</tr>
<tr>
<td>Colour (C)</td>
<td>Colours are uniform</td>
<td>Colours are non-uniform</td>
<td>Presence of six possible basic colours</td>
<td>1-6</td>
<td>0.5</td>
</tr>
<tr>
<td>Diameter (D)</td>
<td>Size &lt; 6 mm</td>
<td>Size ≥ 6 mm</td>
<td>Presence of five differential structural components</td>
<td>1-5</td>
<td>0.5</td>
</tr>
<tr>
<td>Evolution (E)</td>
<td>Non-change</td>
<td>Changes in size, shape or shades of colour features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevation (E)</td>
<td>Smooth surface</td>
<td>High surface</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total dermoscopy score (TDS) = (A score x 1.3) + (B score x 0.1) + (C score x 0.5) + (D score x 0.5). Diagnosis: TDS<4.75, benign melanocytic lesion; TDS of 4.75-5.45, suspicious lesion; TDS>5.45, lesion highly suspicious for melanoma.

A. Asymmetry

The asymmetry criterion may be examined according to dividing the region of the lesion under analysis into two sub-regions by an axis of symmetry, in order to analyse the similarity of the area by overlapping the two sub-regions of the lesion along the axis. The lesion may be considered symmetric when the two sub-regions are highly similar, which is prevalent in benign lesions. Otherwise, the lesion is considered asymmetric, which is associated with malignant lesions. From such an axis, the asymmetry index may be calculated by the difference between the two sub-regions of the lesion; for example, by applying the XOR operation between them [46]. In some studies, the axis of symmetry is defined based on the principal axis of inertia [46], major axis of the best-fit ellipse [156], major and minor axis orientation [40, 69] and longest or shortest diameter [120].
In order to obtain other features related to the asymmetry criterion, geometrical measures are calculated from the segmented lesion area. These measures allow the assessment of asymmetry of the shape, border and differential structure on skin lesions, specifically in dermoscopy images. Such measures include the area of the lesion (computed as the number of pixels inside the lesion region [12, 69] or by applying the bit quads method [140]); aspect ratio [12, 40]; compactness [40]; perimeter [12, 69]; greatest diameter [40, 69]; shortest diameter [69]; equivalent [12, 40]; convex hull [12]; eccentricity [12, 40]; solidity [12, 40]; rectangularity [12, 40]; entropy measures [12]; circularity index (namely thinness ratio) [12, 69, 84-86]; and irregularity index [69]. The asymmetry of differential structures inside the lesion in dermoscopy images may also be considered [27], such as solid pigments of the lesions computed according to Chang et al. [46]. In addition, colour asymmetry may also be measured to identify malignancy from dermoscopy images, as proposed by Celebi et al. [40].

In other studies [84-86] asymmetry is considered, according to pre-defined regions inside the lesion under analysis, as: the size ratio between the pre-defined region and original lesion, area circularity, differences of centre of gravity between the region and the original lesion, and standard deviation and skewness of the distribution of mass of the region. In addition, size functions [65] were applied by d'Amico et al. [54] for representing the features in a qualitative manner, in order to evaluate the asymmetry between the boundary shape, colour and distribution of mass, as well as to compare the two halves of a lesion by computing the distance between the size functions.

B. Border

A lesion’s border is represented by pixels comprising the lesion's boundary, obtained as a result of the lesion segmentation process [110, 144, 159, 179, 184, 188]. The border criterion is basically measured by the irregularity of its shape, which may be regular (usually related to in malignant lesions) or irregular (usually related to benign lesions) in clinical images, or even an abrupt cut-off of network at the border of the lesion in dermoscopy images. Several methods have been proposed for assessing the border’s irregularity; for example, based on geometrical measures [46, 156] such as lesion’s perimeter and area, or even on the average and variance of the gradient magnitude of the associated edges [12].

In order to identify the sharp transition between inside and outside regions of a lesion concerning its border, Iyatomi et al. [85] and Iyatomi et al. [86] divided the lesion region into eight equiangular regions. For each region, the ratio of the colour intensity inside and outside the lesion and the gradient of the colour intensity were computed in particular colour channels, according to a pre-defined window centred at the border of the lesion. Whereas, Celebi et al. [40] computed the differences and ratios of two statistics (mean and standard deviation) over a particular colour channel, considering the following regions: lesion and inner and outer peripheral regions relative to the border of the lesion.

The harmonic wavelet transform descriptor is another method that may be applied to extract border related features [50]. Such a descriptor provides information across multiple scales of roughness, by calculating harmonic wavelet coefficients of the border, which define the distribution of energy for several analysis levels. In another study [69], the border features
are derived from constructing a boundary-series model of the lesion border by computing the distance between each border’s pixel and the centroid of the lesion. Consequently, the obtained boundary-series is analysed by using statistical measures in both the spatial and frequency domains. The former is computed by histogram of the boundary-series, and the latter is computed by applying a three-level wavelet transform.

C. Colour

The colour criterion consists of analysing the change of tonality of pigmented skin lesions in order to identify malignant lesions, since the colours present in such lesions are usually non-uniform [9]. Malignant lesions, specifically in dermoscopy images, may also be diagnosed according to the number of the possible basic colours, such as white, red, light-brown, dark-brown, blue-grey and black, present in the skin lesions. For example, Alcón et al. [12] computed the number of pixels within the segmented area for each one of the six possible basic colours of the lesion, based on the nearest colours to each basic colour by using the Euclidean distance.

The RGB (red, green, blue) colour model is commonly employed to represent the colours of skin lesions [12, 40, 46, 84-86, 156]. Other colour models have also been applied in order to obtain more specific information about a lesion’s colours, such as:

- rgb (normalized RGB) [40];
- HSV (hue, saturation, value) [22, 40, 84, 86];
- HVC (hue, value, chroma) [142];
- CMY (cyan, magenta, yellow) [111];
- YUV (Y-luminance, UV-chrominance channels) [111];
- I1/2/3 (Ohta space) [40];
- Opp (Opponent colour space) [22];
- L*C*H (lightness, chroma, hue) [46]; and
- CIE (international commission on illumination) colour space, for instance $L^*a^*b^*$ ($L^*$-lightness, $a^*,b^*$-chrominance channels) [22, 111] and $L^*u^*v^*$ ($L^*$-lightness, $u^*,v^*$-chrominance channels) [22, 40].

Statistical measures are widely applied to the feature extraction from skin lesion images [12, 40, 46, 84-86, 111, 142, 156]. The minimum, maximum, average, standard deviation, skewness and variance are examples of such measures, which may be computed for each colour channel of the lesion region by using one or several colour models. Furthermore, these measures may also be applied to other regions associated with the lesion’s border, in order to identify a sharp transition between them, which indicates malignancy. The background skin (normal skin), and surrounding skin (inner or outer peripheral regions) are examples of such regions, which may be considered from the lesion. Peripheral regions may be defined by a recursive erosion process [84-86], a fast Euclidean distance transform algorithm [40], or a circular region with centre point upon the lesion centroid [46]. In addition, such regions may
reduce the effects of peripheral inflammation and errors caused by automatic border detection, as proposed by Celebi et al. [40].

Skin lesion features based on relative colours have been proposed [38, 40, 46, 84-86], in order to assess colour features from the different regions associated with the lesion. The relative colour consists of comparing each pixel value of the lesion to the average colour value of the surrounding skin. Furthermore, this feature may present advantages such as compensating the variation of colour of the image caused by illumination, and equalizing variations in skin colour among individuals [38]. In addition, other colour features may be extracted, such as:

- the percentage of melanoma and non-melanoma colour pixels within the ROI [46];
- the clustering ratio of melanoma colour pixels within the ROI [46];
- the variance, relative chromaticity and ratio for each RGB channel [46];
- the difference in colour and \( L^*C^*H \) between the average of the lesion values and the surrounding skin values [46];
- the number of colours in the lesion area and inner peripheral region in the RGB and HSV colour models from the quantized image [84-86];
- the average colour of normal skin [84-86];
- the average colour differences between the inner peripheral region and the lesion area [84-86]; and
- colour asymmetry, centroid distances, and LUV histogram distances (from the \( L^*u^*v^* \) colour model) [40].

D. Diameter or differential structures

The diameter criterion is examined according to the size of the lesion, which is defined by the greatest distance between any two points of a lesion's edge. As such, a diameter equal to, or greater than, six millimetres may indicate malignancy [9]. This criterion is not commonly applied in the computational analysis of skin lesions due to its great dependence on the image resolution [40], since the image size affects the number of pixels for each segmented lesion’s region. An application of this criterion is presented in She et al. [156], in which the diameter of a lesion is calculated considering the semi-major axis of the best-fit ellipse. Thus, the scale is converted from pixels to millimetres based on the knowledge of parameters of the image’s pixels and spatial relation at a particular magnification.

The differential structures of skin lesions may be also examined, more specifically in the dermoscopy images, i.e., pigment network, black dots, globules, structureless areas, homogeneous areas and branched streaks. For example, in Torre et al. [169] multidimensional receptive field histograms (MFHs) were obtained by means of Gaussian derivatives and a Laplacian Gaussian operator, in order to reproduce features of the local differential structures of skin lesions.
E. Elevation or evolution (E)

The "E" criterion may be examined in order to measure the lesion’s elevation [166], since the malignant lesions usually present a higher elevation than benign lesions. The elevation criterion is a morphological feature that may be measured considering its surface. Furthermore, such a criterion may also represent the historical evolution of the lesion in order to diagnose it, including changes in its shape, size, shades of colour, or surface features. Hence, changes in lesion features in the last 3 months or more is a strong index of malignant lesions [27]. To the best of our knowledge, few previous image analysis systems of skin lesions surveyed in the literature have used such criteria [62, 77, 81]. One of the reasons may be related to the complexity of feature extraction from the elevation criterion, or even the unavailability of a database with at least two images of the same lesion that must be taken over time to assess its evolution.

Three-dimensional digital imaging may be designed to extract information about the elevation feature of skin lesions. For example, Hani et al. [77] and Fadzil et al. [62] proposed a method to measure the thickness of some skin lesion types from the 3D surface image. Lesion’s thickness is the elevation present between the base and the surface of the lesion. In addition, registration techniques may be applied to skin lesion images to detect changes in their structure over time, as the algorithm introduced by Huang and Bergstresser [81]. The authors proposed a new technique for the melanoma registration, based on bipartite graph matching, in order to find sufficiently good correspondences between successive images of multiple skin lesions. The authors used the Voronoi cells and distances between points to transfer the point registration problem in images to a bipartite graph-matching problem.

2.4.2. Texture analysis

Texture analysis is frequently considered for image analysis of skin lesions, since it assists in discriminating between benign and malignant lesions by measuring the roughness of their structure. Some studies have also considered using the texture to replace the differential structure (D) criterion regarding the ABCD rule of dermoscopy [12, 35]. Among the various texture descriptors applied in the literature, as detailed in Table 2.2, the grey-level co-occurrence matrix (GLCM) proposed by Haralick et al. [78] has been one of the most commonly employed [12, 40, 59, 84-86, 111, 142, 168]. This method is a statistical measure that computes the joint probability of occurrence of grey-levels considering two pixels spatially separated by a fixed vector. Several measures may be computed based on the grey-level co-occurrence matrix, such as variance, entropy, dissimilarity, correlation, contrast, energy, maximum probability, inverse difference, inverse difference moment, angular second moment (ASM), mean, standard deviation and homogeneity.

In order to discriminate the shape texture between the benign and malignant lesions, Dobrescu et al. [59] combined statistical measures derived from the GLCM and fractal dimension measures derived from several manners by using the box-counting method (BCM). Image-based fractal dimension [11] is a procedure for splitting the image in several quadrants to quantify the irregularity level or self-similarity of the image's fractals. Several methods have
been proposed to measure the fractal dimension of skin lesion images [59, 69, 167], with the BCM being one of the most commonly used methods, since it is simple and effective [11]. This method defines a grid over the image, i.e., it divides the image into several squares of the same size. The process is iterative, in which the image is divided into quadrants and the size of each square decreases at each iteration. Therefore, the amount of squares that covered some part of the image (fractal) is counted in each iteration. Basically, the final fractal dimension is equal to the sum of all the fractal dimensions obtained in each iteration divided by the total amount of fractal dimensions.

Wavelet-based descriptors [69, 167] are also proposed for texture feature extraction of skin lesion images. Such descriptors allow the application of several measures in order to extract information from wavelet coefficients, which may include energy, average, minimum, maximum, median, standard deviation, variance, skewness, kurtosis, norm, entropy and average-energy. Computational methods based on histograms are also used by some researchers to represent texture features. For example, Tanaka et al. [168] computed some aforementioned statistical measures based on the intensity histogram, whereas Barata et al. [22] applied gradient histograms, such as the gradient amplitude and gradient orientation, to represent the texture feature. In order to compute the image gradient, the authors applied a Gaussian filter to the grey-level image for further computation of the gradient vector at each pixel using the well-known Sobel filter.

2.4.3. Pattern analysis

The pattern analysis assists in pigmented skin lesion diagnosis by determining the presence of specific patterns visible in dermoscopy images, which may be divided into global and local patterns [17], as detailed in Table 2.5. Global patterns are represented by textured structures present in most of the lesions, such as reticular, globular, cobblestone, homogeneous, starburst, parallel and multi-component, or even may be non-specific, when aforementioned patterns are absent. Some examples of such patterns are illustrated in Figure 2.9. Local patterns are dermoscopic structures, such as a pigmented network, dots/globules, streaks, blue-whitish veil, pigmentation, hypo-pigmentation, regression structures, and vascular structures. Such patterns may be present or absent, as well as presenting irregular/regular or atypical/typical structures, as shown in Table 2.5, which may define the type of lesion or whether it is benign or malignant. Examples of such patterns are illustrated in Figure 2.10.

The pattern analysis consists of examining the size, uniformity and distribution of the aforementioned patterns. The benign lesion structures are usually uniform; in other words, the lesions do not present several patterns in their structure. Therefore, the presence of at least three (multicomponent), parallel or nonspecific global patterns indicates a higher probability of being a melanoma (malignant lesion). Furthermore, the presence of local patterns, such as blue-whitish veil and regression structures, or even some patterns considered atypical, irregular or asymmetric may identify a melanoma [17]. Shape, colour and texture information may be extracted to identify such patterns [5, 83, 86, 116, 147, 155, 160, 168], as summarized in Table 2.3. For example, Situ et al. [160] extracted several features concerning pattern analysis criteria
in order to detect melanoma. The multicomponent global pattern features were considered by the authors, since such a feature may be an indication of melanoma. Local binary patterns as low level features were used to represent this feature, as well as the standard deviation and entropy of the histograms defined from wavelet and colour moments, respectively. The lesion is divided into five segments to analyse the presence of local pattern features that include irregular dot/globular, atypical pigmented network, and blue-whitish veil, which are obtained by using instance prototypes from the single instance learning method.

Table 2.5: The pattern analysis in dermoscopy images according to the literature.

<table>
<thead>
<tr>
<th>Global feature</th>
<th>References</th>
<th>Local feature</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reticular</td>
<td>[2, 5, 83, 116, 147, 155, 168]</td>
<td>Pigmented network (present or absent/ typical or atypical)</td>
<td>[14, 21, 25, 64, 100, 101, 147, 149, 157, 160, 161, 177]</td>
</tr>
<tr>
<td>Globular</td>
<td>[2, 5, 83, 116, 147, 155, 168]</td>
<td>Dots/globules (present or absent/ regular or irregular)</td>
<td>[64, 101, 160, 161]</td>
</tr>
<tr>
<td>Cobblestone</td>
<td>[2, 5, 116, 147, 155]</td>
<td>Steaks (present or absent/ regular or irregular)</td>
<td>[25, 61, 101, 117, 148]</td>
</tr>
<tr>
<td>Homogeneous</td>
<td>[2, 5, 116, 147, 155, 168]</td>
<td>Blue-whitish veil (present or absent)</td>
<td>[19, 38, 61, 83, 101, 160]</td>
</tr>
<tr>
<td>Starburst</td>
<td>[2, 5]</td>
<td>Blotches or pigmentation (present or absent/ regular or irregular)</td>
<td>[89, 101, 107, 165]</td>
</tr>
<tr>
<td>Parallel</td>
<td>[2, 5, 86, 116, 147, 155]</td>
<td>Hypopigmentation (present or absent)</td>
<td>[55]</td>
</tr>
<tr>
<td>Multicomponent (combination of three or more global patterns)</td>
<td>[5, 160]</td>
<td>Regression structures (present or absent)</td>
<td>[61, 101]</td>
</tr>
<tr>
<td>Non-specific (absent patterns)</td>
<td>–</td>
<td>Vascular structures (present or absent)</td>
<td>–</td>
</tr>
</tbody>
</table>

1 The references of research about local features also include the papers focused on the seven-point checklist method.

In order to identify global patterns, such as reticular, globular, cobblestone, homogeneous and parallel in dermoscopy images, Mendoza et al. [116] defined a set of measures for the connected component representation of every pattern sample, which include the area, perimeter, area-perimeter quotient, and roundness quotient. In addition, the median, average, and variance were computed for the connected components, from the aforementioned measures. Some of these measures are scale invariant in order to analyse how such measures in the connected component pattern sample are best suited for pattern discrimination.

Model-based methods have also been proposed in order to analyse the global patterns such as reticular, globular, cobblestone, homogeneous and parallel of dermoscopy images [147, 155]. For example, Sadeghi et al. [147] proposed a novel approach to identify the aforesaid global patterns based on texture models. Such models consist of representing the texture for each pattern by mean frequency histograms, which are modelled by the joint probability distribution of filter responses, according to a set of different filter banks. This distribution is represented by texton (cluster centre) frequencies, and thus, textons and texture models are learnt from training images. Therefore, texton-based classification in the $L^*a^*b^*$ colour space and the grey-level image applying several filter banks is performed, in which a new image is mapped to a texton distribution, in order to compare its distribution to the learnt models.
Figure 2.9: Examples of global patterns in dermoscopy images: (a) reticular, (b) globular, (c) cobblestone, (d) homogeneous, (e) parallel and (f) starburst [16].
Figure 2.10: Examples of local patterns in a dermoscopy image: (a) atypical pigmented network, (b) irregular dots/globules, (c) blue-whitish veil, (d) irregular pigmentation and (e) irregular streaks (adapted from Celebi et al. [38]).

Due to the low number of criteria to be extracted, the seven-point checklist and Menzies’ methods were introduced for skin lesion diagnosis from dermoscopy images in order to simplify the common pattern analysis [87]. The criteria of such methods are detailed in Table 2.6.

### 2.4.4. Seven-point checklist

The seven-point checklist has been applied in the literature to achieve better accuracy for the computational diagnosis of dermoscopy images [25, 99, 101]. This method consists basically of seven criteria based on local patterns that may be applied to diagnose the malignancy in pigmented skin lesions, particularly melanomas, which are divided into major and minor criteria [17]. The presence of the atypical pigmented network, blue-whitish veil and atypical vascular pattern represents the major criteria, whereas the presence of the irregular streaks, irregular pigmentation, irregular dots/globules, and regression structures represents the minor criteria. A total score of three or more points is more likely to be melanoma, for which the presence of each major criterion receive two points and each minor criterion receives one point [87].
Table 2.6: Diagnostic criteria included the seven-point checklist and Menzies’ methods.

<table>
<thead>
<tr>
<th>Major criteria</th>
<th>Minor criteria</th>
<th>Colour of lesion</th>
<th>Symmetry of pattern</th>
<th>Positive feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical pigmented</td>
<td>Irregular streaks</td>
<td>One colour</td>
<td>Symmetrical pattern</td>
<td>Blue-whitish veil</td>
</tr>
<tr>
<td>network</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue-whitish veil</td>
<td>Irregular pigmentation</td>
<td>More than one colour</td>
<td>Asymmetrical pattern</td>
<td>Multiple brown dots</td>
</tr>
<tr>
<td>Atypical vascular pattern</td>
<td>Irregular dots/globules</td>
<td></td>
<td></td>
<td>Pseudopods</td>
</tr>
<tr>
<td></td>
<td>Regression structures</td>
<td></td>
<td></td>
<td>Radial streaming</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Scarlike depigmentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Peripheral black dots/globules</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Multiple colours (5 or 6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Multiple blue/grey dots</td>
</tr>
<tr>
<td>Broad pigment network</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Seven-point total score < 3 = non-melanoma or ≥ 3 = melanoma.
2 Major criteria receive 2 points.
3 Minor criteria receive 1 point.
4 Diagnosis for benign lesions (symmetrical pattern and one colour) and malignant lesions (asymmetrical pattern, more than one colour and at least one positive feature).

Previous studies [25, 38, 101] focused on the detection of one or several criterions of the seven-point checklist method. For example, Celebi et al. [38] extracted colour and texture features in order to define the blue-white veil area in the lesions. The colour features were extracted based on the concept of the absolute and relative colour, which were computed according to the chromaticity coordinates, relative RGB ratio, normalized relative RGB ratio, relative RGB difference, and a normalized relative RGB difference. In addition, the texture features were measured based on the GLCM.

Algorithms based on estimative chromatic and shape parameters, and texture have been proposed by Betta et al. [25] to detect the occurrence of two criteria of the seven-point checklist. The irregular streaks are identified when an irregular contour and brown pigmentation are found in the same sub-image. For this purpose, the original image is divided into 16 parts to evaluate the irregularity ratio of the border, as well as the presence of the brown colour for each one these parts, which are compared to a pre-defined threshold. The atypical pigment network is identified according to the analysis of the spectral content obtained by applying a fast Fourier transform (FFT).

Likewise, features concerning colour and texture of skin lesions are extracted by Leo et al. [101] in order to classify five criteria of the seven-point checklist method, since all criteria are related to chromatic or morphological features. The authors segmented the lesion colour by using the PCA, 2D histogram construction, peak-picking algorithm, and histogram and lesion
partitioning, in order to detect a blue-whitish veil, irregular pigmentation, and regression structures. In addition, the authors combined structural and spectral techniques [71] to extract texture features, such as medial filter, close-opening operation, FFT, high-pass filtering, inverse fast Fourier transform (IFFT) and suitable thresholding, in order to detect the atypical pigment network, and irregular streaks.

### 2.4.5. Menzies’ method

Menzies’ method allows for identifying colour patterns within the lesion and the asymmetry along any axis drawn through the centre of the lesion, as well as the amount of positive features, such as blue-white veil, multiple brown dots, pseudopods, radial streaming, scar-like depigmentation, peripheral black dots/globules, multiple colours (five or six), multiple blue/grey dots, and broad pigment network [87]. In malignant lesions, particularly melanomas, an asymmetric pattern, more than one colour, and at least one positive feature are usually presented, whereas the benign lesions present a symmetric pattern and only one colour [87]. Silva and Marcal [158] proposed an alternative approach based on Menzies’ method in order to analyse the presence of six colour classes (white, red, light-brown, dark-brown, blue grey and black) for dermoscopy images. The Jeffries-Matusita and transformed divergence separability measures were applied to evaluate the separability degree between the colour classes.

### 2.4.6. Discussion

Dermoscopy images have been widely used for diagnosis of pigmented skin lesions [22, 40, 69, 86, 142, 159], since they allow suitable visualization with more details of pigmentation patterns on the surface of the lesion. Furthermore, previous clinical studies have addressed an increase of sensitivity of the melanoma diagnosis by dermoscopy compared to diagnosis by clinical image [113]. Among the several skin lesion diagnostic methods using dermoscopy images [87], the ABCD rule has commonly been applied to extract features for computational analysis [12, 40, 84-86, 111, 142]. This rule allows for easy understanding and provides simplicity of application while showing reliable results for the melanoma diagnosis. In contrast, previous clinical studies [15] reported that methods based on pattern analysis performed better than the ABCD rule for the diagnosis of melanocytic skin lesions. In recent years, descriptors mainly based on colour and texture have been proposed to identify and classify patterns in skin lesion images, as well as to discriminate benign and malignant lesions. Pattern analysis of pigmented skin lesions has shown promising results and may continue to be the focus of intense research in the coming years [5, 38, 101, 158, 160].

### 2.5. Skin lesion classification

The classification step consists of recognizing and interpreting the information about the pigmented skin lesions based on features extracted from images. The classification process generally occurs by randomly dividing the available image samples in training and test sets. The training step consists of developing a classification model to be used by one or more
classifiers based on the samples of the training set. Each sample is composed of features extracted from a given image and its corresponding class value, which are applied as input data to the classifier for the learning process. The testing step consists of measuring the accuracy of the model learned by the training step over the test set. In addition, such a process may present several problems concerning the dataset, such as features containing different ranges, unbalanced dataset regarding the number of samples, and/or a large number of features. Therefore, this process may require pre-processing of data, in which several techniques may be applied to overcome these problems.

Feature normalization is a pre-processing step, in which techniques may be applied in order to solve the problem of different ranges. Such techniques consist of adjusting the value range of each feature so that the values remain in short intervals, e.g. between 0 and 1, or -1 and 1. Therefore, this procedure prevents the feature with range of values greater than other features from influencing the results, since several classifiers may not deal properly with different ranges. The z-score transformation is a common method employed for data normalization, which allows transforming all numeric features in values within the same range, as discussed by Celebi et al. [40].

Unbalanced datasets concerning the number of samples in each class is a classification problem that may decrease the accuracy of the evaluation result, since the classifiers tend to be based on classes with the highest occurrence. Techniques such as sampling methods [47, 178] may be applied to solve such a problem. Celebi et al. [40] combined sampling methods to balance a dataset containing 88 samples of melanomas and 476 samples of benign lesions using the random under-sampling and synthetic minority oversampling technique (SMOTE) [48]. Random under-sampling method removes samples from the majority class in an arbitrary way, whereas the SMOTE adds synthetic samples to the minority class based on its k-nearest neighbours. Another method to solve the unbalanced dataset problem was used by Barata et al. [22], in which the dataset is composed of 25 samples of melanoma and 151 samples of nevi. The authors repeated the melanoma features belonging to each training set until the same number of samples for both classes was obtained. Furthermore, they added Gaussian noise to each repeated feature set in order to prevent equal samples in the training set.

Feature selection [72] is a pre-processing step in machine learning that may be addressed to deal with datasets contain a large number of features for skin lesion classification. Such a step is commonly addressed in the literature [46, 84-86, 111, 116, 168], since the set of features used may influence the results of the classification. The feature selection process is presented with details in the following section. The most popular methods for skin lesion classification, as well as its evaluation procedures, are discussed. Furthermore, some results of recent studies for classification of skin lesion and its features are also provided.

### 2.5.1. Feature selection

A feature selection step may be employed for the skin lesion classification in order to select the most relevant features and reduce the dimensionality of the feature space so that
irrelevant and/or redundant features are removed. Moreover, such features may influence the performance of the classification process, i.e., render it a slower process [75]. Several benefits are associated with the application of feature selection schemes, such as [40]:

- to reduce the feature extraction time;
- to decrease the classification complexity;
- to improve the classification accuracy rate;
- to decrease training and testing time; and
- to simplify the data understanding and visualizing.

Essentially, the feature selection process has the following steps: A) feature subset selection, B) feature subset evaluation, C) stopping criterion, and D) validation procedure [56]. Search strategies may be applied to define candidate subsets from extracted features of skin lesions, which are evaluated and compared to the previous best subset until a given stopping criterion is reached. This process is iterative, and it only finishes when it reaches the established stop criterion. Thus, the selected best subset should be verified for the specific problem, i.e., the skin lesion classification.

A. Feature subset selection

Feature subset selection consists of finding features through a given process of heuristic searches in order to identify a candidate feature subset for evaluation. Several search algorithms, such as best-first [111], ranker [69, 111], incremental stepwise [46, 84-86, 116, 168] and random [40, 46], have been employed for the feature subset selection process. Exhaustive and genetic searches are other examples of such algorithms that may be applied [178]. These algorithms influence the search direction and execution time of the selection process depending on the adopted search strategy, which may be complete, sequential, and random [56, 104]. For example, Mendoza et al. [116] adopted a sequential search strategy (incremental stepwise) that combines the forward and backward selection, since both techniques may achieve good results in several situations. The forward selection process starts with an empty set, and the best features are gradually added to the set according to the performance obtained from the previous steps, whereas the backward selection process starts with a full set, i.e., with all features and the worst attributes are removed at each iteration; in other words, the features considered irrelevant or redundant are removed. Therefore, such a combination represents a bidirectional selection, in which the features may be simultaneously added and removed [72]. Another model to establish a feature subset is applying embedded methods such as decision-tree algorithms, which incorporate the feature selection in its training process [75].

B. Evaluation function

In this step, the selected feature subset is then evaluated according to the type of search algorithm applied before, and the evaluation process may be based on models such as filter [104], wrapper [92], hybrid [104] and embedded [75], for which the advantages and disadvantages of each feature selection model are summarized in Table 2.7.
Table 2.7: Summary of advantages and disadvantages of feature selection models.

<table>
<thead>
<tr>
<th>Feature selection model</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Filter</strong></td>
<td>The model is simpler and faster compared to wrapper and hybrid; Independent evaluation criteria; It may overcome over-fitting; Computationally efficient.</td>
<td>The selected features may not be the most relevant for the application.</td>
</tr>
<tr>
<td><strong>Wrapper</strong></td>
<td>It searches for features the most relevant to improve mining performance.</td>
<td>It is more computationally expensive; It does inherit any bias of a mining algorithm; It may occur over-fitting.</td>
</tr>
<tr>
<td><strong>Hybrid</strong></td>
<td>It inherits the advantages of filter and wrapper models; It deals with large data sets; It avoid the pre-specification of a stopping criterion.</td>
<td>It inherits the disadvantages of wrapper models; Model is complex.</td>
</tr>
<tr>
<td><strong>Embedded</strong></td>
<td>The model is simple and incorporates the feature selection as part of the training process; It makes better use of the available data; It achieves a faster solution.</td>
<td>The selection depends solely on applied classification method.</td>
</tr>
</tbody>
</table>

The filter model allows for evaluating the goodness of selected features without using any mining algorithms. Each candidate subset is evaluated by means of applying an independent criterion, which may be based on distance\(^1\), information\(^2\), dependency\(^3\), or consistency\(^4\) measures, in order to compare it with the best current subset previously established. If the evaluated subset is considered the best, it becomes the best current subset. This model has several advantages, such as: the filter model is simpler and faster compared to wrapper and hybrid models; independent evaluation criteria without involving any mining algorithm; it may be used to reduce space dimensionality and overcome over-fitting; and it is also computationally efficient [72, 104]. Examples of filter methods applied in the literature based on the aforementioned independent measures are: gain ratio feature selection (GRFS) [69], information gain measure [69], chi-squared [69], correlation-based feature selection (CFS) [12, 40, 46, 69, 111], ReliefF [40, 69], mutual information-based feature selection (MIFS) [40] and generalized sequential feature selection (GSFS).

The evaluation of feature subsets based on the wrapper model is similar to the filter model. The main difference between these two models is the use of mining algorithms to

---

1 These measures try to find the feature that may separate the classes as far as possible by greater distance between them.
2 These measures establish the information gain from a feature.
3 These measures are also known as correlation measures applied to evaluate the ability to predict the value of one feature from the value of another.
4 These measures consist in finding a minimum number of features that may separate classes as consistently as the full set of features may.
evaluate each candidate subset in order to determine the most relevant subset, for which the mining algorithm tends to perform better when searching for such a subset [104]. The k-nearest neighbours (KNNs) [71], Bayesian learning [52], support vector machines (SVMs) [31] and artificial neural networks (ANNs) [80] are examples of such algorithms that may be applied to evaluate the candidate subsets. Several wrapper algorithms may be generated by varying the application of search strategies and mining algorithms in order to identify the algorithm that better represents the features for a given problem. Nevertheless, this model is only viable when the algorithm shows a better performance of the classification by using the selected feature subset instead of using the original feature set. Furthermore, this model is not widely applied to select features, since it is shown to have high computational time [40].

The hybrid model combines properties of filter and wrapper models to evaluate feature subsets in order to consider the advantages of both models, as well as to deal with large data sets. Generally, the process of such a model begins the search from an empty set by using the forward selection algorithm to establish feature subsets at each increasing cardinality, and then an independent measure and a mining algorithm are applied to evaluate the feature subsets. The independent measure is employed for defining the best subsets at each increasing cardinality, considering all possible subsets, and then, the mining algorithm is employed to select the final best subset among the best subsets across different cardinalities.

The embedded model has a built-in mechanism to perform the feature selection; it incorporates the feature selection as part of the training process. The decision tree induction methods, such as classification and regression trees (CART), are examples of such a model. Consequently, when such methods are applied to the feature selection, a tree is constructed from given data and thus, the features not appearing in the tree during the processing are considered irrelevant, whereas the other form the selected feature subset. The methods based on this model may render the process more efficient in several respects: they make better use of the available data, since splitting the training data into a training and validation set is not necessary, and they achieve a faster solution by avoiding beginning the classification process from scratch for every analysed feature subset [72].

C. Stopping criterion

The stopping criterion determines the situation in which the feature selection process must stop. Some examples of such a criterion occur when: 1) the search is complete, 2) the established minimum number of features is achieved, 3) the setting iteration maximum number of the process is achieved, and 4) addition or removal of any feature occurs that worsens the outcome of the best found subset until that moment [104].

D. Validation procedure

The result validation step consists of verifying the best feature subset established by the previous steps. Hence, the validation process may be performed upon applying classifiers from a new set of features in order to measure the classification performance or error rate of the selected feature subset. For example, Maglogiannis and Doukas [111] applied the Bayes Net, SVMs and CART classification methods to evaluate the obtained subsets by using several feature selection algorithms such as the CFS, PCA and GSFS. Furthermore, the achieved results
are compared to the results obtained from all features without applying any feature selection algorithm. The authors concluded that the application of feature selection algorithms may reduce the complexity of the classification. On the other hand, the performance is not always good, and is highly dependent upon the classifier. Therefore, they opted to use all features for the skin lesion classification.

2.5.2. Methods for classification

Classification methods, such as nearest neighbours [22, 142], decision trees [38, 46, 50, 69, 99, 101, 111], Bayesian learning [69, 111], ANNs [50, 85, 111, 158, 167] and SVMs [22, 40, 69, 111, 142, 167, 183], have been commonly applied to discriminate skin lesions in images. A brief description of the main advantages and disadvantages of such classifiers, as well as their techniques applied to the learning objective, are presented in Table 2.8.

Classifiers based on nearest neighbours, known as lazy algorithms [178], are commonly applied due to their simplicity of implementation and their facility to deal with the existence of correlated features. However, they are sensitive to the existence of irrelevant features, and they require a great deal of time for classifying large datasets. The KNN method classifies the data based on the distance of the k-nearest neighbours, in which k is usually an odd number and represents the amount of neighbours. For example, Barata et al. [22] compared several techniques, such as Euclidean, Kolmogorov, and Kullback–Leibler in order to measure the distance of nearest neighbours from different k values. The authors concluded that it is not clear which of these three employed distances is the best for such a problem, since all were considered to be the best for certain test situations. On the other hand, Rahman et al. [142] used the Bhattacharyya distance measure, since such a measure is based on the correlation between the colours and may perform better than the traditional Euclidean distance.

A decision tree [75] has a structure similar to a flowchart, in which each internal node (non-leaf) represents a test of a feature, each branch represents a result of the test, and each external node (leaf) indicates a prediction of the class. Understanding such a structure, as well as ease of rule generation, is quite straightforward. However, the excess of adjustments (overfitting) and the difficulties in dealing with correlated features are the major drawbacks of decision trees. Several algorithms based on decision trees have been applied to classify skin lesions. Examples of such algorithms include Naïve Bayes/decision trees (NBTree) [111], classification and regression trees (CART) [111], C5.0 [50], decision tree learner (J48) [12], classification and regression (C&R) [50], logistic model tree (LMT), [12, 69, 99, 101], decision stump [12], C4.5 [38, 46] and random forest (RF) [69].

Bayesian methods [52] compute the probability of a given set of features to belong to each class, assuming that the features are independent. The Bayes networks [12, 111], Naïve Bayes multinomial (NBL) [111] and Hidden Naïve Bayes (HNB) [69] are examples of Bayesian learning-based methods applied to classify skin lesions. Although Bayesian methods provide fast training and no sensitivity to irrelevant features, they assume that the features need to be independent.
Table 2.8: Classification methods applied to discriminate skin lesions from images and their advantages and disadvantages.

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Technique</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest neighbour</td>
<td>Simplicity of implementation; Dealing well with the existence of correlated features.</td>
<td>Sensitive to the existence of irrelevant features; Long classification time from large datasets.</td>
<td>KNN</td>
<td>[22, 142]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KStar,</td>
<td>[111]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LWL</td>
<td>[111]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NBTree</td>
<td>[111]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CART</td>
<td>[111]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C5.0</td>
<td>[50]</td>
</tr>
<tr>
<td>Decision tree</td>
<td>Simplicity of the structure understanding and visualization; Facilitating rule generation.</td>
<td>Possibility to occur excess of adjustment; It do not deal well with correlated features.</td>
<td>C&amp;R</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LTM</td>
<td>[12, 69, 99, 101]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>J48</td>
<td>[12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decision Stump</td>
<td>[12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C4.5</td>
<td>[38, 46]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RF</td>
<td>[69]</td>
</tr>
<tr>
<td>Bayesian learning</td>
<td>Training is fast; It is not sensitive to irrelevant features.</td>
<td>Assuming that the features are independents.</td>
<td>Bayes Networks</td>
<td>[12, 111]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NBL</td>
<td>[111]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HNB</td>
<td>[69]</td>
</tr>
<tr>
<td>ANN</td>
<td>Capability and flexibility to solve several non-separable problems.</td>
<td>Long training time.</td>
<td>MLP architecture</td>
<td>[50, 85, 158, 167]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RBF network</td>
<td>[111]</td>
</tr>
<tr>
<td>SVM</td>
<td>Good generalization properties; Kernel functions simplify the process of separation of non-linear data.</td>
<td>It is sensitive to the noise; The classification process is binary.</td>
<td>RBF</td>
<td>[22, 40, 69, 111, 142, 167]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Polynomial</td>
<td>[183]</td>
</tr>
</tbody>
</table>

ANN: artificial neural network; SVM: support vector machine; KNN: k-nearest neighbour; LWL: locally weighted learning; NBTree: naïve Bayes/decision tree; CART: classification and regression trees; C&R: classification and regression; LTM: logistic model tree; J48: decision tree learner; RF: random forest; NBL: naïve Bayes multinomial; HNB: hidden naïve Bayes; MLP: multilayer perceptron; RBF: radial basis function.

ANNs [80] are parallel distributed systems composed of layers of input and output elements linked by weighted connections. During the learning phase, the weights are adjusted to predict the correct class based on the input samples. The multilayer perceptron (MLP) is one of the most applied architectures of ANNs [50, 158, 167], since such architecture presents good capability and flexibility to solve several non-separable problems. This architecture may include one or more layers of processing, also called hidden layers, placed between the input and output layers. The back-propagation is a supervised learning algorithm widely used in the MLP architecture [85], which consists of forward and backward processes applied to adjust the weight values of the connections. The forward process allows the distribution of input data through the network, layer-by-layer, and their weights remain fixed. Afterwards, the backward
process propagates the obtained output, layer-by-layer toward the input layer, considering the computed error signal based on the input values and corresponding output in order to adjust the weights. Although ANNs have been proposed to solve many pattern recognition problems, these classifiers may have long training time depending on the size of the training set.

SVMs [31] involve a method based on statistical learning applied to building a hyper-plane to separate the data according to the defined classes. This classifier has been commonly applied to classify skin lesions due to its good generalization properties. Furthermore, kernel functions simplify the process of separating the non-linear data by using a simple hyper-plane in a high dimension feature space. The radial basis function (RBF) [22, 40, 69, 111, 142, 167] and polynomial [183] kernels have been commonly used in several studies. Some kernels are based on properties of other learning methods, such as the RBF kernel that is based on functions of ANNs. However, this classifier is sensitive to noise and the classification process is based on a binary class. Celebi et al. [40] adopted the RBF kernel due to the several advantages compared to other kernels, such as: greater stability compared to the polynomial kernel and reduced number of hyper-parameters that need to be established compared to the polynomial and sigmoid kernels.

Other methods employed for the skin lesion classification include: Gaussian maximum likelihood (Gaussian ML) [142], linear models [84, 86, 156], classification via regression [111], or even ensemble learning [12, 22, 69, 142]. The ensemble learning [178], which have recently been adopted, consists of combining the results of several classification models in order to develop a more robust system that provides more accurate results than by using a single classifier. Average, weighted average, sum, product, maximum, minimum and median are some examples of integration strategies based on the outputs of classifiers. Voting methods from the candidates of a rank may also be employed for this same purpose. Rahman et al. [142] combined the output of SVM, Gaussian ML and KNN classifiers by using product, sum, maximum and average rules. The strategy that obtained the best combination of classifiers was the sum rule. Alcón et al. [12] and Barata et al. [22] compared the meta-classifier AdaBoost to other classifiers to find the best classification result. The AdaBoost is a popular boosting algorithm, which uses the weighted majority voting in order to rank higher the classifiers that perform best during the training step. Barata et al. [22] obtained better classification results with the AdaBoost compared to the KNN and SVM, whereas Alcón et al. [12] obtained the best results by using an individual LMT classifier.

2.5.3. Evaluating the classification performance

The main objective of the classification process of skin lesions is to achieve good results for distinguishing between different classes. In order to fulfil this purpose, several classification models based on different feature subsets, samples and classifiers are evaluated by using test sets. Therefore, new samples are classified and the predicted class is compared to the known class to evaluate the classification performance.

Among several evaluation procedures, the cross-validation (XVAL) procedure [18, 178] is the most commonly used in the literature to evaluate the results of skin lesion
classification. The k-fold cross-validation [38, 40, 50, 69, 86, 111, 116, 142, 167] are examples of cross-validation methods proposed for classifying skin lesions in images. The k-fold cross-validation consists of splitting the training set in k subsets of equal size, with the method being repeated k times. Each subset is employed once as a test set, and the others are employed as the training set. The best model is chosen according to its performance, which is measured by averaging the accuracy obtained from each trial. The leave-one-out cross validation considers all samples for the training step, except one. The number of times the trainings will be done is the same as the number of samples, since each sample will be out of training at least once. The final performance of such an evaluation procedure is obtained by averaging the individual precision of each training.

The half-and-half test is another evaluation procedure, which was applied by Iyatomi et al. [86]. This procedure divides all samples into training and testing in halves, and the performance is achieved by averaging both combinations. The authors evaluated the performance of classifiers using 10-fold cross-validation, leave-one-out cross-validation and half-and-half tests, and they concluded that the results are almost equivalent and may be considered reasonable.

Performance metrics [63] are computed to compare the performance of one or several classification models according to the outcomes of classifiers. Some possible outcomes of classifiers based on the predicted class and known class are: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). These outcomes represent the amount of correct (true) and incorrect (false) classification for each class (positive and negative). For example, in a classification process between two classes, one class may be considered positive and another negative. Usually, the positive samples represent the most important class to classify (e.g., skin cancer), and benign lesion stands for the negative samples. Therefore, the TP rate is the amount of correctly classified positive samples, the TN rate is the amount of correctly classified negative samples, the FP rate is the amount of incorrectly classified negative samples, and the FN rate is the amount of incorrectly classified positive samples. These rates may be represented by a confusion matrix, which is the basis for several metrics used by researchers to measure the performance of the classification, such as: 1) the precision that is the percentage of correctly classified samples for each given class with respect to its true and false predictions; 2) the recall that is the percentage of correctly classified samples for each given class, but only with respect to its true prediction; and 3) the accuracy that is the percentage of correctly classified samples based on the precision of all classes [12, 40, 50, 69, 111, 142, 158, 168]. Another performance metric applied is the root mean squared (RMS) that is the difference between predicted value and known value [111].

The sensitivity (based on TP rate), specificity (based on FP rate) and area under the ROC curve (AUC) are additional terms associated with the receiver operating characteristics (ROC) graph [63], which are also employed to compare the performance of the classification [12, 22, 38, 40, 50, 69, 84, 85, 111, 158, 167]. These two rates are applied to generate the ROC curve, which is a two-dimensional graphic representation of the performance of the classification, in order to show the trade-off between benefits (TP rate plotted on the Y axis) and costs (FP rate plotted on the X axis), without considering the class distribution. From ROC
curve, it is possible to measure the AUC, which represents the expected predictive performance as a single scalar value between 0 and 1.0. Therefore, the higher the AUC value, the better will be the performance of the classifier.

### 2.5.4. Classification performance of recent studies

For the classification process, one or several methods have been evaluated to achieve the best results. The performance of such a process depends on several issues, such as the segmented image, and extracted or selected features, as well as the classification method used. The output of the skin lesion classification may be binary or ternary, and includes different classes according to the classification goal, e.g., malignancy of the lesions (benign versus malignant) [50, 69, 183], and distinct types of skin lesions (melanoma versus nevus [85, 111], melanocytic versus non-melanocytic [84], and dysplastic versus non-dysplastic versus melanotic [111]). Furthermore, skin lesion features, such as border features (regular versus irregular [50]), presence of main colours existing in malignant lesions [158], presence of features of the seven-point checklist [25, 39, 99, 101], and presence of global patterns [2, 5, 83, 116, 147, 155, 168] and local patterns [25, 101, 160, 161] are also classified.

The results of recent studies regarding global pattern classification based on the average performance calculated from the accuracy, sensitivity or specificity metrics for each pattern are summarized in Table 2.9. To the best of our knowledge, few methods dealing with the classification of all global patterns of the skin lesions have been done.

Table 2.10 summarizes the best results of recent studies concerning skin lesion classification, which includes performance metrics such as accuracy, sensitivity, specificity and AUC. For example, for the discrimination between benign lesions and melanomas, Celebi et al. [40] proposed the SVM, whereas Iyatomi et al. [85] proposed the ANN classifier. The performance of several classifiers has also been compared [167]. In the study of Surówka [167], the SVM (RBF kernel) performed better than ANN. Furthermore, ANNs are also compared to other classifiers in Clawson et al. [50], such as C5.0 and C&R, in which a C5.0-classifier performed best. Maglogiannis and Doukas [111] compared the performance in terms of accuracy, RMS, TP and FP rates, and AUC of the following classifiers: Bayesian networks, SVM, ANN with multilayer perceptron, NBTree, CART, RBF networks, NBL, MLR, KStar, LWL and classification via regression. The authors classified different skin lesion classes by performing three experiments: 1) the first experiment was composed of the classification between melanoma and nevus, in which MLR, SVM, LWL and CART were considered the best classification methods; 2) the second experiment was determined by the classification between dysplastic and non-dysplastic lesions. In this case, the multilayer perceptron, SVMs, and Bayes networks performed better than other classifiers, regarding the corresponding ROC curves; and 3) finally, another experiment was based on the classification between melanotic, dysplastic and non-dysplastic, in which the SVM was the most accurate classifier.
### Table 2.9: Results of recent studies focused on the global pattern classification.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>Reti. (%)</th>
<th>Glob. (%)</th>
<th>Cob. (%)</th>
<th>Homo. (%)</th>
<th>Starb. (%)</th>
<th>Paral. (%)</th>
<th>Mult. (%)</th>
<th>Average (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5]</td>
<td>2013</td>
<td>87.11 (SE) 97.96 (SP)</td>
<td>86.25 (SE) 97.21 (SP)</td>
<td>87.76 (SE) 93.23 (SP)</td>
<td>90.47 (SE) 95.10 (SP)</td>
<td>89.62 (SE) 90.14 (SP)</td>
<td>85.25 (SE) 89.50 (SP)</td>
<td>98.50 (SE) 93.11 (SP)</td>
<td>89.28 (SE) 93.75 (SP)</td>
</tr>
<tr>
<td>[2]</td>
<td>2012</td>
<td>90 (SE) 93 (SP)</td>
<td>90.50 (SE) 94.99 (SP)</td>
<td>91.70 (SE) 93.60 (SP)</td>
<td>89.45 (SE) 92.30 (SP)</td>
<td>91.71 (SE) 95.10 (SP)</td>
<td>93.43 (SE) 98.26 (SP)</td>
<td>-</td>
<td>93.08 (SE) 91.45 (SP)</td>
</tr>
<tr>
<td>[147]</td>
<td>2012</td>
<td>Non-reported</td>
<td>Non-reported</td>
<td>Non-reported</td>
<td>Non-reported</td>
<td>-</td>
<td>Non-reported</td>
<td>-</td>
<td>86.8 (A)</td>
</tr>
<tr>
<td>[83]</td>
<td>2011</td>
<td>95 (A) 89 (A)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[160]</td>
<td>2011</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Non-reported</td>
<td>-</td>
</tr>
<tr>
<td>[116]</td>
<td>2009</td>
<td>90 (A) 100 (A)</td>
<td>85 (A)</td>
<td>100 (A)</td>
<td>-</td>
<td>95 (A)</td>
<td>-</td>
<td>94 (A)</td>
<td></td>
</tr>
<tr>
<td>[155]</td>
<td>2009</td>
<td>90 (A) 80 (A)</td>
<td>80 (A)</td>
<td>90 (A)</td>
<td>-</td>
<td>90 (A)</td>
<td>-</td>
<td>86 (A)</td>
<td></td>
</tr>
<tr>
<td>[86]</td>
<td>2008</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>93.1 (SE) 97.7 (SP)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[168]</td>
<td>2008</td>
<td>92.7 (A) 93.4 (A)</td>
<td>-</td>
<td>95.5 (A)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94 (A)</td>
<td></td>
</tr>
</tbody>
</table>

Ref.: reference; Reti.: reticular; Glob.: globular, Cob.: cobblestone; Homo.: homogeneous; Starb.: starburst; Paral.: parallel; Mult.: multicomponent; A: accuracy; SE: sensitivity; SP: specificity.

Ensemble methods were also employed and they performed better than individual classifiers in the studies proposed by Rahman et al. [142] and Barata et al. [22]. However, Alcón et al. [12] obtained the best results in both the individual LMT classifier and AdaBoost ensemble method. Meanwhile, the authors considered the LMT classifier more useful due to the complexity computation of the ensemble model. Consequently, there is no ideal method to solve all problems in skin lesion classification, as may be observed in findings in the literature. The performance of the classification relies on several conditions, mainly on good extracted features, as previously discussed.
Table 2.10: Results of recent studies focused on the skin lesion classification.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Sample</th>
<th>Feature selection</th>
<th>Classifier (extracted features/selected features)</th>
<th>Classification</th>
<th>A (%)</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>[22]</td>
<td>2013</td>
<td>176 dermoscopy images</td>
<td>Selection for each type of feature</td>
<td>AdaBoost Global method (8/1)</td>
<td>Melanoma/nevus</td>
<td>-</td>
<td>96</td>
<td>80</td>
<td>-</td>
</tr>
<tr>
<td>[22]</td>
<td>2013</td>
<td>176 dermoscopy images</td>
<td>Selection for each type of feature</td>
<td>Local method (8/1)</td>
<td>Melanoma/nevus</td>
<td>-</td>
<td>100</td>
<td>75</td>
<td>-</td>
</tr>
<tr>
<td>[69]</td>
<td>2012</td>
<td>289 dermoscopy images</td>
<td>GRFS</td>
<td>RF (35 455/23)</td>
<td>Malignant/benign</td>
<td>91.26</td>
<td>-</td>
<td>-</td>
<td>0.937</td>
</tr>
<tr>
<td>[84]</td>
<td>2010</td>
<td>655 dermoscopy images</td>
<td>Incremental Stepwise method</td>
<td>Linear model (42N/2)</td>
<td>Melanocytic/non-melanocytic</td>
<td>-</td>
<td>97.99</td>
<td>86.64</td>
<td>-</td>
</tr>
<tr>
<td>[12]</td>
<td>2009</td>
<td>152 clinical images</td>
<td>CFS</td>
<td>LMT (45/5)</td>
<td>Melanoma/nevus</td>
<td>86</td>
<td>94</td>
<td>68</td>
<td>0.890</td>
</tr>
<tr>
<td>[50]</td>
<td>2009</td>
<td>30 dermoscopy images</td>
<td>-</td>
<td>C5.0 (non-reported/-)</td>
<td>Malignant/benign</td>
<td>93.30</td>
<td>80</td>
<td>96</td>
<td>-</td>
</tr>
<tr>
<td>[111]</td>
<td>2009</td>
<td>3639 dermoscopy images</td>
<td>-</td>
<td>MLR/ SVM/ LWL/ CART (31/-)</td>
<td>Melanoma/nevus</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td>[111]</td>
<td>2009</td>
<td>3639 dermoscopy images</td>
<td>-</td>
<td>Multilayer perceptron/ SVM/ Bayes networks (31/-)</td>
<td>Dysplastic/non-dysplastic</td>
<td>73.20/ 76.08/ 68.94</td>
<td>-</td>
<td>-</td>
<td>0.688/ 0.607/ 0.663</td>
</tr>
<tr>
<td>[111]</td>
<td>2009</td>
<td>3639 dermoscopy images</td>
<td>-</td>
<td>SVM (31/-)</td>
<td>Melanotic/dysplastic/non-dysplastic</td>
<td>77.06</td>
<td>-</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td>[85]</td>
<td>2008</td>
<td>1258 dermoscopy images</td>
<td>Incremental Stepwise method</td>
<td>Back-propagation ANN (428/72)</td>
<td>Melanoma/nevus</td>
<td>94.10</td>
<td>85.90</td>
<td>86.0</td>
<td>0.928</td>
</tr>
<tr>
<td>[86]</td>
<td>2008</td>
<td>199 dermoscopy images</td>
<td>Incremental Stepwise method</td>
<td>Linear model (482/10)</td>
<td>Melanoma/nevus</td>
<td>-</td>
<td>100</td>
<td>95.9</td>
<td>0.993</td>
</tr>
<tr>
<td>[142]</td>
<td>2008</td>
<td>558 dermoscopy images</td>
<td>PCA</td>
<td>Ensemble learning (126/10)</td>
<td>Malignant/benign/dysplastic</td>
<td>75.69</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[167]</td>
<td>2008</td>
<td>39 dermoscopy images</td>
<td>-</td>
<td>SVM (231/-)</td>
<td>Malignant/benign</td>
<td>94.70</td>
<td>95</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[40]</td>
<td>2007</td>
<td>564 dermoscopy images</td>
<td>CFS</td>
<td>SVM (473/18)</td>
<td>Melanoma/benign</td>
<td>-</td>
<td>92.34</td>
<td>93.33</td>
<td>0.966</td>
</tr>
<tr>
<td>[156]</td>
<td>2007</td>
<td>36 dermoscopy images</td>
<td>PCA</td>
<td>Linear model (8/2)</td>
<td>Melanoma/nevus</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.940</td>
</tr>
</tbody>
</table>

A: accuracy; SE: sensitivity; SP: specificity; AUC: area under the ROC curve; GRFS: gain ratio feature selection; CFS: correlation-based feature selection; PCA: principal component analysis; BoF: bag of features; RF: random forest; LTM: logistic model tree; MLR: multinomial logistic regression; SVM: support vector machine; LWL: locally weighted learning; CART: classification and regression trees; ANN: artificial neural network.

### 2.5.5. Discussion

The classification process of skin lesion images must perform very well, since they are often used to assist dermatologists in the diagnosis by means of CAD systems. The evaluation and improvement of the performance of classifiers are essential for the pattern recognition research field [176]. A relevant problem that affects the performance of classifiers is the definition of the meaningful features for representing the classes. Consequently, the feature
representation step is very important to achieve better performance for the computational diagnosis of skin lesion images. The application of several descriptions may be required considering the large amount of features extracted from images. For dealing with this issue, feature selection methods have been applied to establish the most relevant features \[40, 69, 86\], since these methods allow removing the redundant and/or irrelevant features. As a consequence, the feature extraction time, the training and testing computational load and the classification complexity are all reduced, while the classification performance may be improved. The result of the feature selection process depends on the search strategy and evaluation model applied as well as their established parameters.

Feature selection methods based on a filter model \[104\] are more often preferred to other models due to the following advantages: computationally efficient, simpler and faster methods, independent evaluation criteria, and ability to overcome over-fitting \[12, 40, 69, 85, 116, 142, 168\]. Nevertheless, the features selected by using a filter model may not be the most relevant for the application, whereas the wrapper model \[92\] may be applied to search for the most relevant features based on mining algorithms to improve the performance of the feature selection. The wrapper model is not commonly applied due to the high computational time, as demonstrated by Celebi et al. \[40\]. However, efficient search strategies may be proposed for this model to avoid the time-consuming task of classifying skin lesions. Although the hybrid model inherits the advantages of both filter and wrapper models, this model may be complex and also inherits the disadvantages of wrapper model. Methods based on an embedded model provide simplicity and a faster solution for the feature selection step compared to methods based on the filter model. However, the methods based on the filter have greater advantages that justify their use \[72\]. Another means of determining the most discerning features based on colour and texture features was addressed by Barata et al. \[22\], who compared the features performed by using each individual feature, all the colour features, both texture and colour features, and the best texture and colour features. The authors concluded that the colour features provide better results than the use of texture features when used individually.

In regard to the classification process, the performance depends on several factors, such as the extracted and selected features, established parameters and chosen classification method. The classification methods should be chosen based on the classification problem and available data regarding advantages and disadvantages corresponding to each method. A brief description regarding advantages and disadvantages of the main classification methods applied to the pattern recognition problems are presented in Table 2.8. Hence, when the objective is to have a fast and simple system, the nearest neighbours \[178\], Bayesian leaning \[52\] and decision tree \[75\] methods may be appropriate options. However, the use of the nearest neighbour method is not recommended, when the dataset is large or has irrelevant features, since it results in a lengthier classification process time.

As observed from the state-of-the-art in this Chapter, techniques based on a decision tree were used by many authors for the skin lesion classification \[12, 38, 46, 50, 69, 99, 101, 111\]. The simplicity of the structure in terms of ease of understanding and visualization, as well as the easy rule generation, is one of the important advantages of this method. However, the difficulty of dealing with correlated features and the possibility for an excess of adjustment to
occur are the major drawbacks of such a method. Despite the long training time, ANNs have been proposed in various studies [50, 85, 158] to deal with complex pattern recognition problems. The SVM classifier [31] has also been applied to discriminate skin lesions, due to its good generalization and simplification of the non-linear data separation by means of kernel functions [22, 40, 69, 111, 142, 167, 183]. Several classification methods have also been compared to achieve the best results to identify skin lesions in images or to identify their features. The SVM performed better than other classifiers in several studies [111, 167]. Ensemble models [178], which aim to combine the strengths of different classifiers, have also been proposed to improve the performance of the classification of skin lesions. Ensemble models have performed better than an individual classifier [22, 142].

In addition to methods discussed above, other classifiers may be applied for recognizing the lesions or patterns. For example, the application of the classifier based on an optimum-path forest (OPF) proposed by Papa et al. [136] shows great potential for achieving superior classification results. The OPF is a supervised classifier applied to solving pattern recognition problems as a graph based on prototypes to represent each class by one or more optimum-path trees, considering some key samples. The training samples are nodes of a complete graph, whose arcs are the link of all pairs of nodes. The arcs are weighted by the distances between the feature vectors of their corresponding nodes. The classification of a new sample is defined according to the strong connectivity of the path between the sample and the prototype. Therefore, the path with minimum-cost among all paths is considered the optimum one.

The OPF was compared to SVM, ANN-MLP, and k-NN using several datasets and descriptors based on shape, colour, and texture [136]. The accuracy of the classification based on OPF was usually better than ANN-MLP and k-NN and similar to SVM. Nevertheless, the OPF classifier showed excellent computational time, which is important for large datasets, and some interesting properties, such as quickness, simplicity, ability to deal with multiclass classification and overlapping between classes, parameter independence and assumption not based on the shape of the classes.

2.6. Summary

Image computational analysis of pigmented skin lesions is an area of great research interest due to its importance in skin cancer prevention, as well as in the early diagnosis. This chapter provided an overview of current developments of computational vision methods for skin lesion image analysis. The main steps of image computational analysis used in the literature, such as image acquisition, pre-processing, segmentation, feature extraction and classification, as well as the techniques used in each one were considered. Techniques used to acquire and pre-process images, with a focus on their subsequent segmentation were introduced. The reviewed segmentation techniques were classified into five categories: edge-, thresholding-, region-, AI- and active contour-based. We have presented and discussed results obtained with some of these techniques applied to dermoscopy and macroscopic images of skin lesions. The Chan-Vese model, which is based on active contour model without edges, provides good results.
on images with colour variation and low contrast of the lesion boundaries. Therefore, such a model is a good option for the segmentation of skin lesions. However, other methods, with improvements, or in combination with other techniques, may also provide good lesion detection. Studies specifically addressing techniques applied to the feature extraction step based on several clinical methods were also presented in this review. The skin lesion classification step was addressed by including feature selection methods and classifiers, as well as their evaluation procedures and some performance results.

From this review, one may conclude that several methodologies focused on skin lesion image analysis have been proposed for use in CAD systems. Such systems aim at an effective computational diagnosis of pigmented skin lesions to assist the dermatologists in their diagnosis. Dermoscopy images should be more commonly used in the computational diagnosis of skin lesions, since these images present less artefacts and more detailed features, which may lead to a more adequate lesion segmentation and analysis. Nevertheless, techniques to remove or reduce the artefacts are usually necessary to obtain robust segmentation results. Most studies involve extraction of several features from dermoscopy images and comparison of two or more classification methods to identify benign and malignant lesions. However, some studies used feature selection methods to achieve a better classification performance. Detection and classification of skin lesion features have also been the goal in several studies. Recently, global and local pattern recognition has been of great interest to researchers.

In conclusion, future trends regarding image computational analysis of pigmented skin lesions involve searching for new methods aiming to develop more efficient and effective systems for the computational diagnosis based on dermoscopy images. Hence, several issues may be addressed to achieve this goal, in particular:

- Image segmentation: the development of new segmentation methods to search for superior accuracy in terms of the detection of the lesion edges, as well as to take into account other issues in the development of computational solutions, such as computational performance, automaticity level, image noise smoothing and removal, and image enhancement;

- The ABCDE rule: the evolution criterion (E) may be better explored in order to develop methods to analyse changes in size, shape, shades of colour and surface features on skin lesions. Extracted features based on evolution criterion along with the other criteria may complement the diagnosis;

- Global pattern analysis: the development and evaluation of new computational methods to identify the presence of global patterns, such as reticular, globular, cobblestone and parallel, as well as starburst and multicomponent patterns, since few studies have explored such patterns. Moreover, these methods could identify the absence of such patterns. To the best of our knowledge, no previous study has addressed this issue;

- Local pattern analysis: the lack of computational methods for local pattern analysis needs to be overcome to identify the presence of important patterns, such as pigmented network, dots/globules, streaks, blue-whitish veil, blotches,
hypopigmentation, regression structures and vascular pattern. Proposed algorithms to detect the irregularity of some skin lesion patterns also are important to assist in diagnosis. In addition, such algorithms may be combined in order to classify skin lesions based on a seven-point checklist;

- Menzies’ method: to the best of our knowledge, an approach based on the method was proposed in only one study. Therefore, the development of new approaches for colour and asymmetry patterns, and positive feature analysis is important for future applications of this method for computational diagnosis of skin lesions;

- Feature selection: in order to find more relevant features for the given problem, different feature selection models may be compared. In addition, efficient search strategies may be proposed for a wrapper model, as well as new classifier to evaluate the selected subsets, in order to solve some of its limitations;

- Classification performance: the evaluation of new classifiers, ensemble models and parameter optimisation need to be addressed in order to classify skin lesions and to improve on the current results.

Computational methods based on these issues may perform better and more effectively in diagnosing skin lesions. In addition, such methods may cover several problems regarding skin lesion analysis, which convert CAD systems into more complete systems for diagnosing such lesions based on dermoscopy images.

The next chapter will discuss a proposed approach for this research, aiming to deal with some of the issues noted in the literature. Further aspects of this proposal, such as motivation to the development of this work, encountered problems, research hypothesis, goals and expected results will also be addressed on the following sections.
3 THESIS STATEMENT

In this chapter, a new approach is proposed, based on ensemble methods and feature selection for pattern recognition in dermoscopy images. The motivation and problems in regards to a given area, according to the analysis from state-of-the-art practices (see Chapter 2), are presented in the following sections. Furthermore, the objectives and expected contributions and achievements of this project are described. Afterwards, the proposed approach steps are outlined.

3.1. Motivation and problems

Skin cancer is one of the most common cancers worldwide [172]. Table 3.1 presents recent data regarding skin cancer in the United States of America (USA), the United Kingdom (UK) and Brazil, according to gender.

Table 3.1: Number of new cases of skin cancer, according to gender, in the USA, UK and Brazil.

<table>
<thead>
<tr>
<th>Country</th>
<th>Type of skin cancer</th>
<th>Year</th>
<th>Number of new cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Male</td>
</tr>
<tr>
<td>USA¹</td>
<td>Melanoma</td>
<td>2014</td>
<td>43,890</td>
</tr>
<tr>
<td></td>
<td>Non-melanoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK²</td>
<td>Melanoma</td>
<td>2010</td>
<td>6,201</td>
</tr>
<tr>
<td></td>
<td>Non-melanoma</td>
<td></td>
<td>55,747</td>
</tr>
<tr>
<td>Brazil³</td>
<td>Melanoma</td>
<td>2014</td>
<td>2,960</td>
</tr>
<tr>
<td></td>
<td>Non-melanoma</td>
<td></td>
<td>98,420</td>
</tr>
</tbody>
</table>

¹ Estimated number, based on 1995-2010 incidence rates.
² Confirmed cases in 2010.
³ Estimated number in 2014, and valid also for 2015.

In the USA, 76,100 new cases of melanoma were estimated to be diagnosed in 2014. This estimate does not include NMSC, since it is not required that this form of skin
cancer be reported to cancer registries. For the same year, 9,710 deaths from melanoma were estimated. Another interesting point concerns melanoma incidence rates, which have increased during the last 30 years; for example, the incidence rates from 2006 to 2010 have increased by 2.7% per year [163]. In the UK, melanoma was the 15th most common cancer in 2010, with approximately 12,800 new cases of this disease. Consequently, melanoma was the 18th most common cause of death from cancer. In 2011, there were 2,209 deaths from melanoma, and 590 deaths from NMSC. Of these deaths from melanoma, 59% were in men, and 41% were in women [172]. In Brazil, NMSC will be the most common form of cancer, since approximately 182,000 new cases are estimated for 2014 and 2015. Although NMSC has a lower mortality rate, it has a higher incidence than melanoma [82]. Melanoma was the 19th most common cancer worldwide in 2008, with an approximate estimation of 200,000 new cases. The incidence rate of melanoma is the highest in Australia/New Zealand, Northern America and Northern Europe, and the lowest in South-Central Asia [172].

As previously reported, the alarming increase in the number of skin cancer cases worldwide has raised interest concerning the computational systems for automatic diagnosis to assist in preventing the development of skin cancer. Several methods for image analysis focused on dermoscopy imagery have been proposed for CAD systems [40, 85, 153]. Such systems aim at an effective automated diagnosis to assist dermatologists in their diagnosis. Particularly, CAD systems may be used to monitor benign skin lesions to prevent the development of malignancy. Moreover, malignant lesions may be diagnosed at an early stage, during which the patient has a higher probability of cure, and more favourable conditions to be properly treated.

Although this research topic has been addressed in several studies, resulting in successful applications, new methods may be proposed fill gaps that still have not been fully addressed, as well as to improve the performance of existing methods. Most studies involve only the extraction of several features from dermoscopy images and the comparison between two or more classification methods in order to discern benign and malignant lesions [50, 69, 183]. Performance assessment of classifiers is an important issue for the pattern recognition process [176]. Therefore, some studies proposed feature selection methods [69, 116, 168] or ensemble models [22, 142] to achieve better classification performance. Detection and recognition of skin lesion features have also been a target of search in several studies [50, 158]. Recently, global and local pattern recognition has been of great interest to researchers [5, 25, 39, 99, 101, 116, 147, 155].

Pattern recognition of dermoscopy images is a challenging task to assist in discriminating between benign and malignant skin lesions. One concern in this task is in identifying the presence of global patterns, since few studies have been done on such patterns in automatic diagnosis of skin lesions [5, 83, 86, 116, 147, 155, 160, 168]. To the best of our knowledge, only one study dealing with the classification of all global patterns of skin lesions has been proposed [5], and no previous study has addressed the issue to identify the absence of such patterns. Models based on textons [147], Markov random fields (MRF) [155], AdaBoost.MC [5], discriminant analysis [168], SVM [2], feature-space nearest neighbour [116] and linear classifier [86] are some methods that have been
proposed for global pattern recognition. Another difficulty involves defining what features are meaningful to describe the patterns of interest.

We hypothesized that the performance of skin-lesion pattern-recognition may be improved by using ensemble models based on feature selection for manipulating their input features.

3.2. Objectives and expected contributions

The main objective of this PhD project is to develop a new method for pattern recognition in dermoscopy images based on ensemble learning models. In this approach, the feature selection and pattern classification is one of the factors most important for identifying the malignancy in pigmented skin lesions, since the combination of several learning models and appropriate features may improve the performance of classification. Some important issues concerning the main objective to be achieved are:

- using dermoscopy images: the method should be able to segment dermoscopy images and extract the features from identified ROI based on pattern analysis;
- ensemble learning: the method should generate a homogeneous ensemble of base classifiers through changing the set of features in order to combine the most relevant features of several base classifier models for pattern recognition;
- pattern recognition: the method should be able to identify global patterns in dermoscopy images, and even to classify the malignancy of skin lesions;
- real environment application: the method should be effective and robust enough for automatic diagnosis of pigmented skin lesions in real environment scenarios.

The proposed approach for this PhD project aims to contribute to overcoming these issues and to provide information that may assist dermatologists in their diagnosis. We hope to contribute to the dermoscopy CAD area, in order to improve automatic diagnosis by means of pattern recognition of pigmented skin lesions. In addition, we intend to construct a dataset with several extracted features from dermoscopy images, and make it available to researchers, which might allow for its application in future studies concerning machine learning and pattern recognition.

3.3. Proposed Approach

In this section, an approach to solve the previously discussed problems concerning pattern recognition of dermoscopy images is proposed. Figure 3.1 illustrates the pipeline concerning the proposed approach, which is composed of the following steps: 1) original image enhancement and noise removal, 2) ROI extraction of the actual image,
3) pattern analysis of the extracted ROI, and 4) pattern recognition of the lesion in question. The first step is mainly applied to deal with noisy images to be analysed. The second step consists of lesion identification of the image being studied. The third step is applied to the feature extraction of an identified lesion based on pattern analysis. Finally, the last step is applied to pattern and lesion classification according to extracted features.

Figure 3.1: Pipeline of the proposed approach for skin lesion pattern recognition.

We believe that the results of this work will contribute to the pattern recognition of dermoscopy images, for which global patterns may be identified and a better result may be achieved in the pattern and lesion classification. In the next sections, the image database and each step of the proposed approach, as well as the validation process are outlined.

3.3.1. Image databases

The databases that will be used to develop this project is composed of dermoscopy images, since such images allow suitable visualization of pigmented patterns on the surface of the lesion, whose identification is the focus of this work. In addition, previous clinical studies observed an increase of sensitivity to melanoma with diagnosis by dermoscopy relative to the diagnosis by clinical image [113]. Examples of dermoscopy images are shown in Figure 1.1.

A brief description of the current databases is shown in Table 3.2. The current databases have a total of 1,536 dermoscopy images, obtained from the EDRA Atlas [16], PH² database [115], area researchers [42] and Dermoscopy Atlas [28]. Of these, 1,011 images are benign lesions (nevus, seborrheic keratosis, dermatofibroma, melanosis, miscellaneous and vascular lesion) and 525 images are malignant lesions (melanoma and non-melanoma). Examples of benign and malignant images are shown in Figure 1.1.
Table 3.2: A brief description of the current image databases.

<table>
<thead>
<tr>
<th>Database</th>
<th>Number of images</th>
<th>Manual border</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>EDRA atlas</td>
<td>1040</td>
<td>No</td>
<td>728</td>
</tr>
<tr>
<td>PH2 database</td>
<td>200</td>
<td>Yes</td>
<td>160</td>
</tr>
<tr>
<td>Area's researchers</td>
<td>90</td>
<td>Yes</td>
<td>67</td>
</tr>
<tr>
<td>Dermoscopy atlas</td>
<td>206</td>
<td>No</td>
<td>56</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1536</strong></td>
<td><strong>290</strong></td>
<td><strong>1011</strong></td>
</tr>
</tbody>
</table>

1 344 reticular, 82 globular, 30 cobblestone, 78 homogeneous, 46 starburst, 47 parallel, 332 multicomponent, 64 non-specific and 17 other pattern images.

The benign lesions included in the nevus category are melanocytic, blue, Clark, combined, compound, congenital, dysplastic (atypical), halo, intradermal, junctional, dermal, recurrent, reed and Spitz. The malignant lesions included in the melanoma category are superficial spreading, nodular, malign lentigo, acral lentiginous, amelanotic, recurrent, invasive and in situ, whereas the lesions in the non-melanoma category are basal cell carcinoma and squamous cell carcinoma. Moreover, a great deal of information concerning diagnosis of lesions by expert dermatologists is available in this database, including among them, the results of manual borders and diagnostics concerning the lesions and global patterns, such as reticular, globular, cobblestone, homogeneous, parallel, starburst, multicomponent and non-specific patterns, which are requirements for the development and evaluation of this work. Examples of global pattern images are shown in Figure 2.9.

### 3.3.2. Image enhancement and noise removal

Dermoscopy images may present some noise, such as hair, veins and bubbles that may influence the outcome of the segmentation process. Therefore, a technique for image pre-processing will be employed in this project to remove or reduce noise, as well as to enhance the lesion’s boundaries. We intend to use the median filter for this purpose, since such a filter is commonly applied to noisy images and it has shown good results when applied on skin lesions [39, 40, 70, 159]. From the experimental results, it was observed that by using a median filter to obtain a smoothed image might reduce the hairs, as well as to enhance the lesion, which is an important requirement for skin lesion segmentation. The result of a smoothed image after applying median filter is presented in Figure 2.2(b).

### 3.3.3. ROI extraction

The ROI may be extracted by a segmentation process using a pre-processed image. The results of such a process should be effective, so information of the lesion may be extracted with more precision. In addition, the accuracy of such a process directly
influences the feature extraction step, which is required to represent the lesion with more precision to be employed in the classification process. Therefore, appropriate segmentation techniques are important to obtain good results for the problem in question. From the critical analysis of methods presented in the state-of-the-art, we do not intend to employ methods based on edges, since such methods showed several disadvantages [24, 138], i.e., they may not identify the lesion's edges completely in some cases. Moreover, the edges are an important feature for the skin lesion analysis process. In the work of Silveira et al. [159], the authors compared several techniques for melanoma diagnosis from dermoscopy images, considering several methods, such as the thresholding-, active contour-, region- and fuzzy-based methods. The best results were obtained by means methods based on active contour. Nevertheless, these methods were semi-supervised, since it was necessary to insert a curve.

We have already studied skin lesion segmentation using the active contour model without edges (Chan-Vese model) [125]. However, we only applied it to the clinical images. We also compared the segmentation results between the Otsu method [134] and Chan-Vese model [45] (presented in section 4.1). The Chan-Vese model provided good segmentation results according to evaluation by expert dermatologists. This model allowed for identifying lesions with low contrast boundaries and some artefacts. On the other hand, the Otsu method for segmentation based on thresholding has shown several problems in the literature, such as irregular regions and lesion edges may not be identified [184].

The database of this work has several types of lesions and some may have the boundary with little contrast (as presented in Figure 1.1(a)) or other artefacts. Therefore, we intend to use an active contour algorithm based on a geometric model, since such algorithms presented good results for the skin lesion segmentation [159], and even may render the process fully automatic for curve definition. Although Chan-Vese model has not been used for skin lesion segmentation, we believe that such a model may improve the results of the dermoscopy image segmentation, which will be evaluated in the validation process (presented in section 3.3.6). The results of this step will provide the identification of the points corresponding to the region of the lesion in the image. We will use such points to determine some information about the lesion, in order to extract its features. The result of a segmented dermoscopy image after applying Chan-Vese model is presented in Figure 2.7(a).

3.3.4. Pattern analysis

After the ROI extraction, the next step is to extract lesion features based on pattern analysis in order to numerically describe its properties. The pattern analysis consists of examining the presence of specific patterns visible in dermoscopy images, which the uniformity and distribution of such patterns assist in the skin lesion diagnosis (presented in section 2.4.3). We intend to use the pattern analysis method, since such a method may present higher clinical diagnosis accuracy compared with diagnosis by using the ABCD rule [15, 113]. In addition, it is our objective to develop an effective
classification model for global patterns, since the current state of this approach still lacks of developments. For example, the multicomponent pattern is underexploited and the non-specific pattern has not been studied, and the presence of such patterns indicate a higher probability of being a melanoma. Therefore, we believe that global pattern analysis should be most explored.

The features extracted from skin lesion images should be effective for class separability. Several methods to extract colour [5, 83, 86, 160, 165] and texture [5, 86, 168] features for the pattern analysis are proposed in the literature, since such features allow a good representation of skin lesion patterns. Nevertheless, few of them compare or combine different techniques to extract features. For the pattern analysis process, we intend to extract features based on colour and texture properties by using different techniques in order to dispose of several features that may be relevant for the pattern discrimination. The pattern analysis process will be performed by two methods: 1) colour feature extraction and 2) texture feature extraction. This process is illustrated in Figure 3.2.

![Pattern analysis process](image)

Figure 3.2: Pattern analysis process.

Colour and texture features will be extracted from the image according to the extracted ROI. From these features, a dataset is constructed with a set of \( n \) samples \( (x_i) \), according to number of images of the database. Each sample \( (x_i) \) is composed of \( m \) features \( (x_{im}) \) and the class to which it belongs \( (y_i) \). Such samples vectors are used for the pattern recognition process.

**Colour feature extraction**

For representing the colour features, the original images will be transformed for different colour spaces, such as \( HSV \) [22, 40, 84, 86], \( CIE L^*a^*b^* \) [22, 111] and \( CIE L^*u^*v^* \) [22, 40], since such models are widely used for the colour feature extraction in a skin lesion image. Furthermore, such a model allows the representation of colours based
on human perception. These CIE models are perceived as uniform colour space; thus, a unified colour space may simplify the identification of colour properties, since it is easier to preserve colour-difference ratios [102]. Afterwards, we intend to compute statistical measures using the aforementioned colour models and lesion peripheral regions, and even, to identify the presence and amount of basic colours of the skin lesions, in particular:

- Statistical measures such as minimum, maximum, average, standard deviation, skewness, and variance [12, 40, 46, 84-86, 111, 142, 156];

- Relative colours from the surrounding region (inner/outer peripheral regions). These features may present advantages such as compensating colour variations caused by illumination used, and equalizing skin colour variations among individuals [38]. In addition, such regions may reduce peripheral inflammation effects and errors caused by automatic border detection algorithms [40]; and

- Basic colours such as white, red, light-brown, dark-brown, blue-grey and black. The identification of such colours is an important requirement in diagnosing melanoma [87].

**Texture feature extraction**

Texture analysis methods are usually categorized as structural, statistical, model-based and transform [112]. Although the structural approach provides a good symbolic description, some extracted feature may be more useful for synthesis than an analysis task. Among the various statistical methods, the co-occurrence matrix has presented a potential for effective texture discrimination of dermoscopy images [12, 38, 40, 59, 84-86, 111, 142, 168]. Fractal dimension is a model-based method, which is also potentially useful for the texture analysis [59, 69, 167]. Fourier [25, 99, 101, 168], Gabor [183] and wavelet [69, 160, 167] transforms are also applied to extract skin lesion features. Methods based on Fourier transform may present poor performance, due to its lack of spatial localization, whereas a Gabor filter allows a superior spatial localization. Therefore, the wavelet transform presents several advantages compared with the Gabor transform, i.e., varying the spatial resolution allows it to represent textures at the most suitable scale, and there are even several choices for the wavelet function, in which it is possible to choose wavelets best suited for a given application [112]. For representing the texture features, we intend to employ different texture analysis methods, in order to find the best features to represent the skin lesion patterns. First, we will convert the original RGB image into a grey-level image. Afterwards, texture features will be extracted based on the aforementioned methods, namely:

- Statistical measures based on a grey-level co-occurrence matrix such as variance, entropy, dissimilarity, correlation, contrast, energy, maximum probability, inverse difference, inverse difference moment, angular second moment, average, standard deviation and homogeneity [78];

- Fractal dimension by using a box-counting method [11]; and
• Measures such as energy and entropy based on a wavelet transform [154].

We believe that by using descriptors discussed above, the extracted features will be able to better represent the lesion’s patterns based on colour and texture properties. In addition, this step will achieve another objective: the construction of a more complete dataset with pigmented skin lesion features obtained from a dermoscopy image to be used in pattern recognition problems in future studies.

3.3.5. Pattern recognition

After constructing the set of features, the next step is the pattern recognition of skin lesion images based on extracted features. The result of the pattern recognition process of skin lesion images should have a high performance, since they are often used to assist dermatologists in their diagnosis. The evaluation and improvement of the performance of classifiers are essential for the pattern recognition research field [176]. An ensemble of methods has commonly been applied in several studies to overcome the results of classification for several problems, i.e., multi-class problem [67, 132, 185, 191]. The ensemble learning process consists of integrating several classification models for a given problem. The classification with a single model chooses the finest performance between different datasets, i.e. using cross-validation, whereas the ensemble of classifiers may use them all and combine the results [178]. For construction of ensemble models, it is necessary to determine which sets of learning models should be generated and how the results of such sets should be integrated [170].

The classification performance depends on several factors, for example the dataset, sample size and dimensionality of data [176]. A relevant problem that affects the performance of classifiers is to choose meaningful features to represent the classes. Manipulating the input features by using different feature subsets, or a so-called ensemble feature selection process, is one of the methods applied to generate ensembles and solve such a problem [58]. This process aims to find a set between the feature subsets that entails a disagreement between the classifiers of each model, whereas the traditional algorithm aims to find the best feature subset. Likewise the traditional model, the ensemble feature selection is typically composed of a search strategy and fitness function [170].

We intend to develop a new homogeneous ensemble method using a feature selection for the input feature manipulation. Although ensemble models have often shown superior performance in classification than the individual classifiers, some models may show a greater computational cost. Therefore, the proposed ensemble method aims to obtain a good performance, as well as to have a lower computational cost compared to the popular ensemble methods. This method involves the following steps: 1) construing ensemble models by manipulating the input features for the base classifier, 2) ensemble model selection by using a measure of diversity, and 3) combining the ensemble model
by using an integration technique. The proposed ensemble learning process is shown in Figure 3.3.

Figure 3.3: Pattern recognition by using ensemble models.

Here, the pattern recognition is divided into two steps: pattern classification and skin lesion classification. The pattern classification consists of identifying the presence of global patterns on skin lesions, such as reticular, globular, cobblestone, homogeneous, parallel, starburst, multicomponent and non-specific patterns. The skin lesion classification consists of identifying whether the lesions are malignant or benign. Hence, there is a multi-class and single-class classification problem. We intend to use the same extracted features (presented in section 3.3.4) and ensemble models (presented in Figure 3.3) for both pattern and skin lesion classification, since the features are important for both classification problems.

The problem of unbalanced datasets (presented in Table 3.2) will also be considered as a data pre-processing step. Sampling methods, such as over and under-sampling have been used to solve this problem [47]. Nevertheless, random under-sampling may remove important samples, and random over-sampling may lead to over-fitting. Hence, we intended to apply SMOTE [48], which is an over-sampling technique for overcoming the over-fitting and expand the decision region of minority class samples, which generate synthetic samples. The k-fold cross-validation evaluation procedure [178] may be applied to avoid over-fitting while testing the capacity of the classifier to generalize, since such a procedure is widely employed in the literature [38, 40, 50, 69, 86, 111, 116, 142, 167] and has shown good results compared with other procedures [86, 91].
Ensemble models

This process aims to generate an ensemble model with the best feature subsets. For generating ensemble models, we propose applying the OPF classifier based on input feature manipulation for each set of training data established by the cross-validation procedure [178]. The OPF classifier [136] has been proposed for solving pattern recognition problems as a graph based on prototypes that represent each class by one or more optimum-path trees. This classifier has been applied in several studies for pattern recognition and medical applications [135]. The OPF classifier has performed well with large datasets and has shown to be able to deal with multi-class classification, natively [136]. In addition, such a classifier is quick, simple, and a non-parameter. Ensembles of OPF classifiers for reducing the size of the training set by using under-sampling have been proposed in Ponti Jr and Rossi [139].

In order to manipulate the input features to generate ensembles, we intend to apply sets of parameter variations of the genetic algorithm (GA) as a search strategy to generate different feature vectors. The GA is very useful for large-scale problems, since it presents a low computation time, and high possibility of finding better solutions than other selection algorithms, such as forward and backward sequential selection, while such algorithms are faster, but are poorer in performance [97]. Furthermore, GAs have shown better results in feature selection for ensembles [170].

Ensemble model selection

Several studies have used measures of diversity for ensemble [20, 96, 97, 170, 171]. Two classifiers are considered diverse if they differ in some input [58]. Therefore, the combination of the output of several classifiers is interesting if there is disagreement between base classifiers of an ensemble. We consider measures of diversity in order to select the best ensemble model generated by each iteration of the training process, thus the model with the best feature subsets. We intend to apply several diversity measures in order to determine the best measure to represent the diversity of the ensemble models.

There are two types of measures that have been applied to analyse the diversity of ensembles: pairwise and non-pairwise [97]. Pairwise measures (i.e. the Q statistic, correlation, disagreement and double-fault) allow comparing differences in predictions of two base classifiers. The final diversity by using pairwise measures usually is the averaged over all possible pairs in a base classifier set. Non-pairwise diversity measures (i.e. the entropy of the votes, Kohavi-Wolpert variance, inter-rater agreement, difficulty index, generalized diversity and coincident failure diversity) are employed to measure differences in predictions of sets of more than two base classifiers. Non-pairwise measures are typically more complex than the pairwise measures. Nonetheless, the interpretation of pairwise measures may become less clear when a measure is established for the entire ensemble. Several measures have been compared in the literature, and different results were obtained in the studies [97, 105, 170].
Model integration

Applying a good integration method based on diversity of the base models is also important for the performance of the ensemble model [170]. The challenging problem of integration is to determine which one(s) of the classifiers to choose or how to combine the results produced by the base classifiers. Voting methods have commonly been suggested to combine the classification results in order to generate an ensemble model [170]. Such methods analyse which class received the majority vote based on the results produced by all base classifiers. For the integration of ensemble models, we intend to compare several methods in order to use the best method for a given problem.

We believe that the proposed model by using an ensemble of classifiers by feature selection to manipulate the input features will be able to deliver a better result for the recognition of global patterns and skin lesions, which will be evaluated in the validation process (presented in Section 3.3.6).

3.3.6. Validation process

The evaluation of segmentation and pattern recognition steps will be considered in this process by using quantitative validation. For evaluating the proposed segmentation method we will use the statistical metrics most commonly used [41, 70], such as precision, recall, sensitivity, specificity, error probability and XOR measures, in order to compare our results with ground-truth models obtained by specialists. In addition, we intend to compare the result of our segmentation against other results obtained by current algorithm in the literature, such as the methods presented in Section 2.3, by using our dataset. The segmentation evaluation needs to have good results, since the classification step depends on the segmentation result.

For evaluating the proposed pattern recognition model we will use performance measures such as sensitivity, specificity and AUC based on the ROC curve, since the ROC curve is a very useful tool for visualizing and evaluating classifiers. Currently, such measures are commonly used and are able to provide a richer measure of classification performance than other evaluation measures [63]. Such measures will be employed to evaluate new data sets for both global pattern classification and skin lesion classification obtained by the classification model in the training process. Furthermore, we intend to compare the final results of our classification method against the ones obtained using classifiers commonly used and ensemble methods, such as bagging [58], boosting [58], random forest [30] and AdaBoost [191], or even the classification by using an individual OPF classifier [136].

The proposed method will also be evaluated by using qualitative validation. For qualitative validation, we intend to apply our system in a real environment in order to be evaluated by specialists.
3.4. Summary

In this chapter, an approach was presented for skin lesion pattern recognition based on colour and texture properties. The motivation and problems for the development of a given approach, as well as the objectives and expected contributions were outlined. The proposed approach is composed of the following steps: image enhancement and noise removal, ROI extraction, pattern analysis, and pattern recognition. Each step of the proposed approach, as well as the database and validation process was described. For preprocessing the image, we proposed the median filter to enhance the image and remove the noise. For the ROI extraction, we proposed an active contour model without edges in order to segment the image. For the pattern analysis, we proposed the colour and texture feature extraction to represent lesion features. Finally, for the pattern recognition, we proposed a new ensemble model of OPF classifiers based on feature selection for the input feature manipulation.
4 THESIS PLANNING

In this chapter, a brief description of the previous work is presented. Thereafter, the task planning for the following years of this thesis is outlined.

4.1. Previous work

It is important to mention that some work concerning pigmented skin lesion segmentation and classification has already been performed [125-131]. However, such work was based on the ABCD rule, and texture analysis using clinical images. In this context, various algorithms that may be used for this thesis were implemented. For example, a computational method for extracting contours of skin lesions from clinical images was proposed [125], which employed anisotropic diffusion filter [23] to smooth the original images, and active contour model without edges [45] to segment the smoothed images. Experimental tests were performed to compare the results obtained with the proposed method, and definition of the threshold, by using Otsu method. The lesion final contour obtained by both segmentation methods were evaluated by a specialist. The result of the application of such techniques in a given image is presented in Figure 4.1. The evaluation results revealed that the proposed method is effective in detecting skin lesions and extracting their contours from clinical images, since in several cases, the contour defined the given lesion in a more effective way. The proposed method adequately deals with low contrast boundaries, as well as noisy images, such as hairs, shadows and reflections.

![Figure 4.1: Segmentation results](image)

(a) (b) (c)

Figure 4.1: Segmentation results: (a) original image, (b) result after applying Otsu’s method and (c) result after applying the Chan-Vese model.
Another experimental test performed consisted of analysing whether the feature extraction based on the ABCD rule provided a good result for the distinction between benign and malignant lesions [129]. The ratios between the distances of the first half-line (P1) and second half-line (P2), which represent the perpendicular line for each point of longest diagonal (D), as well as the standard deviation from distances, were extracted to represent the asymmetry criterion (Figure 4.2a). The amounts of peaks, valleys and straights were extracted by vector product and by inflexion points by means of one-dimensional edges (Figure 4.2b) in order to represent the border criterion. The mean, variance and standard deviation values for each RGB component were extracted to represent the colour criterion. From the classification results, it was concluded that the extracted features are promising for skin lesion differentiation.

![Figure 4.2: Examples of feature descriptors for skin lesions: (a) two half-lines (P1 and P2) that represent a perpendicular of longest diagonal (D) and (b) one-dimensional edge of a lesion.](image)

A skin lesion characterization method from the ABCD rule, and texture analysis, was also analysed [127]. Besides the asymmetry, border and colour features extracted by using the above-mentioned descriptors, the texture features are also established by fractal dimension, using the box-counting method [11], which is applied as input data to the SVM classifier [31]. The study aims to classify the asymmetry of lesions into symmetric or asymmetric [128] and the texture of lesions classified into smooth or rough [126]. The goal is to provide dermatologists with information concerning extracted features in order to assist them in their diagnosis.

Research regarding feature selection and classification of pigmented skin lesions was performed, in order to analyse whether an automatic feature selection scheme will improve the performance of classifiers, in regard to the utilization of all features of a set. For example, the combining of some search and learning algorithms based on the wrapper model [72] for feature selection was studied [130]. The model utilized allowed the selection of relevant features from the training, set and the discarding of irrelevant features, for the problem of classification. Thus, an improvement in skin lesion classification was obtained. The search algorithms applied to generate the subset were: sequential forward search [104], sequential backward search [104] and genetic algorithms [79]. Afterward, each defined feature subset, by using such algorithms, was evaluated by learning algorithms to define the most relevant subset. The learning algorithms employed were: KNNs [71], naïve Bayes [52] and ANNs [80]. Genetic algorithms and naïve Bayes
classifier were the models that allowed the selection of the features with more precision and recall for the skin lesion classification, considering three classes: melanoma, nevus and seborrheic keratosis.

The filter model [72] for the feature selection was also evaluated [131], since it has been widely studied for skin lesion feature selection, and has shown suitable results. The feature selection filters considered were: CFS [74], ReliefF [93] and GRFS [75]. The Naive Bayes [52], KNN [178], ANNs [80], SVMs (polynomial and RBF kernels) [31], J48 decision tree method [178] and RF [30] were the classifiers employed by means of several parameters. The best results for each model are shown in Table 4.1. The feature normalization and unbalanced dataset problems were also considered. The z-score transformation [90] was applied to normalize the feature values. The combination of stratified remove folds [178] and SMOTE method [48], was applied to unbalanced dataset classes. The 10-fold cross-validation method [91] was applied to evaluate the skin lesion classification. The AUC value [63] was computed, to measure the performance of the lesion classification. The feature selection filters were evaluated according to performance of classifiers, in regard to classifications obtained without the application of the feature selection schema. The higher the AUC value, better the performance of the classifier. The combination of the ReliefF method [93] and multilayer perceptron classifier [80] provided the best performance for the classification (AUC=0.878) and thus the best feature subset (red value in Table 4.1). In general, the feature selection scheme obtained better classification results than when all features were employed for the classification (bold value in Table 4.1).

Table 4.1: Classification results of malignant and benign lesions according to feature selection filters.

<table>
<thead>
<tr>
<th>Classifier \ Feature selection</th>
<th>CFS (GA)</th>
<th>ReliefF (Ranker)</th>
<th>GRFS (Ranker)</th>
<th>No features selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>0.771</td>
<td><strong>0.858</strong></td>
<td>0.752</td>
<td>0.842</td>
</tr>
<tr>
<td>KNN</td>
<td><strong>0.840</strong></td>
<td>0.804</td>
<td>0.794</td>
<td>0.712</td>
</tr>
<tr>
<td>ANN (multilayer perceptron)</td>
<td>0.848</td>
<td><strong>0.878</strong></td>
<td>0.846</td>
<td>0.846</td>
</tr>
<tr>
<td>SVM (polynomial kernel)</td>
<td>0.842</td>
<td>0.802</td>
<td>0.802</td>
<td><strong>0.869</strong></td>
</tr>
<tr>
<td>SVM (RBF kernel)</td>
<td><strong>0.851</strong></td>
<td>0.640</td>
<td>0.621</td>
<td>0.783</td>
</tr>
<tr>
<td>J48</td>
<td>0.723</td>
<td>0.739</td>
<td><strong>0.794</strong></td>
<td>0725</td>
</tr>
<tr>
<td>RF</td>
<td>0.818</td>
<td>0.761</td>
<td><strong>0.832</strong></td>
<td>0765</td>
</tr>
</tbody>
</table>

KNN: k-nearest neighbour; ANN: artificial neural network; SVM: support vector machine; J48: decision tree learner; RBF: radial basis function; RF: random forest; CFS: correlation-based feature selection; GRFS: gain ratio feature selection; GA: genetic algorithm; AUC: area under the ROC curve.

In recent years, a review of the literature concerning image analysis of pigmented skin lesions has been done. The survey includes mainly the image segmentation (presented in Section 2.3), feature extraction (presented in Section 2.4) and the classification process (presented in Section 2.5). However, such a review is constantly being updated. In addition, algorithms for the colour and texture analysis, as well as pattern recognition, were also reviewed in order to find new techniques to solve the
problems mentioned previously. In parallel with the literature review, the authors of this study have been experimenting with segmentation techniques, in which the resulting images will be used for the pattern recognition process.

4.2. Task planning

From the proposed approach steps (presented in Section 3.3), the tasks that will be developed in this work are: 1) definition of the image database to be employed, 2) image enhancement and noise removal, 3) ROI extraction from images, 4) pattern analysis, 5) pattern recognition, 6) result validation, and 7) writing of the thesis. The proposed approach timeline for the aforementioned tasks is presented in Table 4.2 and the each task to be performed is summarized as follows.

Table 4.2: Task timeline according to the proposed approach steps.

<table>
<thead>
<tr>
<th>Task \ Trimester</th>
<th>Year 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Definition of the image database</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Image enhancement and noise removal</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 ROI extraction</td>
<td>✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Pattern analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Pattern recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Result validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Writing of the thesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Definition of the image database: this task aims to search and organize the dermoscopy images available by researchers to be employed for the pattern recognition of pigmented skin lesions. These images are used in the next task. This task has already been performed (✓);

2. Image enhancement and noise removal: this task will allow the application of proposed smoothing technique to reduce noise and enhance the regions of the lesions from original images in order to achieve a better segmentation. The smoothed images are used in the next task. This task has already been performed (✓);

3. ROI extraction: this task consists of the application of the proposed segmentation method by using pre-processed images in order to define the ROI of each image. This information will be necessary in the next step. This task has already been performed (✓). Evaluation of segmentation results by using statistical analysis and comparison with other works is also included here.
4. Pattern analysis: this task consists of implementing and applying several algorithms in order to extract features to represent the lesion patterns. The result of this step is the generation of a dataset with feature vectors that will be used in the next task;

5. Pattern recognition: this task aims to implement the algorithms according to proposed ensemble model and to apply them to the pattern recognition in dermoscopy images. In addition, the implemented algorithms will be evaluated based on statistical analysis and compared with current studies;

6. Result validation: this task aims to validate the results by the performance analysis, and by the application of the system in a real environment;

7. Writing of the thesis: this task consists of writing the thesis report of the proposed approach concerning all the experimental tests, evaluation and validation of the developed and implemented algorithms.

During the development of this Doctoral project, results are expected to be divulged by means of papers. The papers will be submitted to both national and international conferences, as well as to relevant journals, which are relevant to the topic of this project, such as machine learning, artificial intelligence, pattern recognition and medical image analysis.
5. IMPLICATIONS OF RESEARCH

There are several approaches in the literature for skin lesion pattern recognition from dermoscopy images. Nevertheless, most of studies involve only extraction of several features and comparing two or more classifiers in order to discriminate benign and malignant lesions. Recently, the global pattern recognition has been of interest to researchers. However, the recognition of such patterns is still little explored in research on the automatic diagnosis and most studies do not deal with the classification of all global patterns.

We intend to develop a new approach for the automatic diagnosis of dermoscopy images based on pattern analysis techniques and ensemble learning models to improve skin lesion classification results. In addition, the approach to be developed will allow for identifying lesion patterns based on possible global patterns in order to obtain more detail concerning the lesion, since such a pattern is important for the diagnosis. The proposed approach, if sufficiently accurate, will provide dermatologists better information concerning lesions in order to assist them in the early diagnosis of skin cancer.

The proposed ensemble method will combine classification models based on more meaningful features to describe the patterns of interest. Therefore, the method may obtain a superior performance, present a lower computational cost, and may also be applied to other pattern recognition problems. We expect to make an important scientific contribution to the area of automated dermoscopy image diagnosis. In addition, we hope that the developed computational system contributes to the CAD area in order to enhance performance automatic classification as a supporting tool for diagnosis.

Finally, we hope to contribute with an available dataset composed of several features extracted from dermoscopy images, which might allow its application in future studies concerning machine learning and pattern recognition. Furthermore, by the end of this thesis, we hope we have inspired our search group to explore other dermoscopy image analysis problems, in order to analyse more information about pigmented-skin lesions that are important for the diagnosis.
REFERENCES


59. Dobrescu, R.; Dobrescu, M.; Mocanu, S.; Popescu, D. Medical images classification for skin cancer diagnosis based on combined texture and fractal


100. Leo, G. D.; Liguori, C.; Paolillo, A.; Sommella, P. An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin
lesions, in IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems. 2008: IEEE.


