Pattern Recognition in Pigmented Skin Lesion Images using Ensemble Methods

Roberta Barbosa Oliveira

Porto, Portugal
March 2017

Roberta Barbosa Oliveira

Pattern Recognition in Pigmented Skin Lesion Images using Ensemble Methods

Dissertation submitted to the Faculdade de Engenharia da Universidade do Porto in fulfilment of the requirements for the degree of Doctor in Informatics Engineering

Supervisor

Professor Doctor João Manuel R. S. Tavares
Faculdade de Engenharia da Universidade do Porto

Co-supervisor

Professor Doctor Aledir Silveira Pereira
Universidade Estadual Paulista

Faculdade de Engenharia da Universidade do Porto

Porto, Portugal
March 2017

ACKNOWLEDGEMENT

This Thesis represents the achievement of goals through a lot of effort and dedication. I would like to take this opportunity to thank all of those who have been with me at this important moment of my life, who have contributed directly or indirectly to this work. This journey would not have been possible without the support of many people that I will always be very grateful for.

Firstly, I thank God for always being present in my life, helping me to overcome difficult moments and giving me the strength to overcome obstacles. Also, for giving me so much happiness and so many opportunities, but above all, for giving me health and a wonderful family, and friends in my life.

To my parents, Selma and João Batista, for the affection, trust, understanding, support, structure and education that they gave me, and for teaching me to be a better person every day. My sisters, Renata Daniele and Maria Júlia, who encouraged me in the development of this work. And also to all my relatives for the support and understanding in my departure for the development of this Thesis.

I would like to thank my supervisor, Dr. João Manuel R. S. Tavares, for receiving me, believing in my work, giving me the opportunity to develop this research, and for always being willing to pass me knowledge. I will be eternally grateful for your attention, dedication, supervision and encouragement for the development of this Thesis.

To my co-supervisor, Dr. Aledir Silveira Pereira, who has become an example and a father figure for me. I would like to thank him for the concern, understanding, advice and encouragement to continue to pursue my goals, for having believed in my potential and encouraged me to take the doctorate degree.

I am grateful to the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for the financial support, granted through the PhD scholarship, and also the *Programa Ciência sem Fronteiras* (CsF) for this great opportunity.

To the *Faculdade de Engenharia da Universidade do Porto* (FEUP) and *Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial* (INEGI) for the structure and space offered for the development of this research.

To the professors of FEUP, especially Dr. Eugénio da Costa Oliveira, Dr. António Augusto de Sousa, Dr. Gabriel de Sousa Torcato David, Dr. Carlos Manuel Milheiro de Oliveira Pinto Soares, Dr. João Pedro Carvalho Leal Mendes Moreira and Dr. Rui Carlos Camacho de Sousa Ferreira da Silva, who received me with affection, and who influenced my personal and professional progress.

To the professor of FEUP, Dr. Jaime Santos Cardoso, and the professor of the *Universidade do Minho*, Dr. Luis Paulo Reis, for the suggestions and collaborations in the presentation of the thesis plan, who contributed to improve the work, as well as to the many anonymous reviewers of my papers, for their constructive comments.

To the employees of FEUP, especially Sandra Reis, for the attention and affection with which they received me and for the support offered over this period.

To the professors, Dr. Norian Marranghello, Dr. João Paulo Papa and Dr. Rodrigo Capobianco Guido of the *Universidade Estadual Paulista* (UNESP), as well as Dr. Edgard de Macedo Silva of the *Instituto Federal da Paraíba* and Dr. Pedro Pedrosa Rebouças Filho of the *Instituto Federal de Educação, Ciência e Tecnologia do Ceará*, for the attention, availability, and contributions throughout the study.

To my co-workers and companions of this journey, Tiago Lourenço, Yunior Luis, Alex Araújo, Daniel Moura, Jessica Codesso, Carlos Gulo, André Pilastre, Renata Oliveira, Benedito Júnior, Andreia Zanella, Carla Martins, Maria Vasconcelos, Nuno Vasconcelos, Erico Leão, Fabrício Sperandio, Hélio Cavalcanti, Zhen Ma, Rui Carreira, Carlos Penichet, André Santos, Pedro Cunha, Rozilene Aroeira, Ana Ferreira, Sepideh Haeri, Matteo Spezialetti and Gabriela Yanagihara. Friendships that were built throughout this period and that also contributed to the development of this work. I will remember them with affection in my heart, not only for the contributions in the development of my thesis, but mainly for the companionship, understanding, advice, moments of relaxation and fun inside and outside the work environment.

Special thanks to some friends, Catarina Ferreira, Paula Voabil, Kristina Lutsik, Filipa Rocha, Susana Silva, Marília Carmo, Cláudia Faceira, Beatriz Carvalhinho, Lúcia Bandeirinha, Carol Miranda, Nathasa Araújo and Fábio Pinto, who were very important in my life during this period and who have always been with me in good and bad times, who supported me, understood my absence and made my days happier.

“Destiny is not a matter of chance; it is a matter of choice. It is not a thing to be waited for; it is a thing to be achieved.”

William Jennings Bryan

ABSTRACT

Pigmented skin lesions have been a cause for global concern in efforts to prevent the development of skin cancer. In addition, skin cancer can be diagnosed at an early stage, during which the patient has a higher probability of cure, and more favourable conditions for being properly treated. To deal with this problem, computer-aided diagnosis (CAD) systems have been developed to assist dermatologists in early diagnosis of skin cancer. Pattern recognition in macroscopic and dermoscopic images is a challenging task in skin lesion diagnosis. Thus, in an initial step of this project, the current computational methods suggested for skin lesion diagnosis in such images were reviewed according to the fundamental steps of image computational analysis. In addition, the strengths and weaknesses of the reviewed methods were discussed. The search for better performing classification has been a relevant issue for pattern recognition in images. One challenge that affects performance of classification includes defining which features are meaningful to describe skin lesion patterns. Hence, this project was particularly focused on feature extraction and skin lesion classification, especially in macroscopic and dermoscopic images.

For the pattern recognition in macroscopic images, a computational approach was developed to detect skin lesion features according to the asymmetry, border, colour and texture properties, as well as to diagnose types of skin lesions, i.e., nevus, seborrheic keratosis and melanoma. In this approach, an anisotropic diffusion filter was applied to enhance the input image and an active contour model without edges was used in the segmentation of the enhanced image. Finally, a support vector machine (SVM) classifier was adopted to classify each feature property according to their clinical principles, and also for the classification between different types of skin lesions. Experiments were performed regarding the segmentation and classification of pigmented skin lesions in macroscopic images, with the results obtained being very promising.

Given that the evaluation and improvement of the classification performance is an essential requirement of pattern recognition, the ensemble methods are seen as promising methods to be used for skin lesion classification due their ability to integrate several classification models to provide a more robust system. In addition, appropriate features and ensemble methods can be combined to achieve superior performance for skin lesion

classification. For the pattern recognition in dermoscopic images, firstly, a combination of features was analysed based on shape properties, colour variation, and texture analysis by using different feature extraction methods. Furthermore, several colour spaces were used for the feature extraction related to both colour and texture features. For evaluating the proposed feature extraction, different categories of classifiers were adopted, and different feature selection algorithms were compared. The combination of features presented is effective for skin lesion classification.

For the skin lesion computational diagnosis in dermoscopic images, ensemble classification models based on input feature manipulation were proposed. The extracted features were used to ensure the diversity of the ensemble classification models. The input feature manipulation process was based on subset selection models: 1) a feature subset selection model based on specific feature groups, 2) a correlation-based subset selection model, and 3) a subset selection model based on feature selection algorithms. Each ensemble classification model was generated by using an optimum-path forest (OPF) classifier and integrated with a majority voting strategy. Promising results were achieved with the proposed ensemble classification models. The performed experiments allowed to analyse the effectiveness of the developed approaches, as well as to verify the aspects that should be improved.

***Keywords*:** Macroscopic and dermoscopic images; Image processing and analysis; Image segmentation; Image classification; Feature extraction and selection; Classifiers; Ensemble classification models.

RESUMO

As lesões de pele pigmentadas têm sido motivo de preocupação global no que diz respeito aos esforços para prevenir o desenvolvimento de câncer de pele. Além disso, o câncer da pele pode ser diagnosticado em uma fase precoce, durante a qual o paciente tem uma maior probabilidade de cura, e condições mais favoráveis para ser devidamente tratada. Para lidar com esse problema, sistemas de diagnóstico auxiliado por computador (CAD) têm sido desenvolvidos para auxiliar os dermatologistas no diagnóstico precoce de câncer de pele. O reconhecimento de padrões em imagens macroscópicas e dermatoscópicas é uma tarefa desafiadora no diagnóstico de lesões de pele. Assim, em uma etapa inicial deste projeto, métodos computacionais que têm sido recomendados para diagnóstico de lesões de pele nessas imagens foram identificados e revistos de acordo com as etapas fundamentais de análise computacional de imagens. Além disso, os pontos fortes e fracos dos métodos foram identificados e discutidos. A busca por melhor desempenho da classificação tem sido uma questão relevante para o reconhecimento de padrões em imagens. Um desafio que afeta o desempenho da classificação inclui definir quais as características que são significativas para descrever padrões das lesões de pele. Deste modo, este projeto focou-se particularmente na extração de características e na classificação das lesões de pele, especialmente em imagens macroscópicas e dermatoscópicas.

Para o reconhecimento de padrões em imagens macroscópicas, foi desenvolvida uma abordagem computacional para detetar características de lesão de pele de acordo com as propriedades de assimetria, borda, cor e textura, bem como para diagnosticar tipos de lesões de pele, isto é, nevo, queratose seborréica e melanoma. Nesta abordagem, um filtro de difusão anisotrópico foi aplicado para melhorar a imagem de entrada e um modelo de contorno ativo sem bordas foi usado na segmentação da imagem melhorada. Finalmente, foi adotado um classificador de máquina de vetores de suporte (SVM) para classificar cada propriedade de característica de acordo com seus princípios clínicos, e também para a classificação entre diferentes tipos de lesões de pele. Foram realizados experimentos relacionados a segmentação e classificação de lesões de pele pigmentadas em imagens macroscópicas, sendo os resultados obtidos muito promissores.

Dado que a avaliação e melhoria do desempenho da classificação é um requisito essencial para o reconhecimento de padrões, os métodos de *ensembles* são vistos como

métodos promissores para a classificação de lesões de pele devido à sua capacidade de integrar vários modelos de classificação para fornecer um sistema mais robusto. Além disso, características adequadas e métodos de *ensembles* podem ser combinados para alcançar um desempenho superior para a classificação de lesões de pele. Para o reconhecimento de padrões em imagens dermatoscópicas, primeiramente, uma combinação de características foi analisada neste projeto com base em propriedades de forma, variação de cor e análise de textura usando diferentes métodos de extração de características. Além disso, vários espaços de cores foram usados para a extração de características relacionadas com cor e textura. Para avaliar a extração de características propostas, foram adotadas diferentes categorias de classificadores e comparados diferentes algoritmos de seleção de características. A combinação de características apresentou ser eficaz na classificação das lesões de pele.

Para o diagnóstico computacional de lesões de pele em imagens dermatoscópicas, foram propostos modelos de classificação de *ensembles* baseados na manipulação das características de entrada. As características extraídas foram utilizadas para garantir a diversidade dos modelos de classificação de *ensembles*. O processo de manipulação das características de entrada foi baseado em modelos de seleção de subconjuntos: 1) modelo de seleção de subconjuntos de características a partir de grupos de características específicas, 2) modelo de seleção de subconjuntos usando correlação e 3) modelo de seleção de subconjuntos baseado em algoritmos de seleção de características. Cada modelo de classificação de *ensemble* foi gerado usando um classificador de floresta de caminho ótimo (OPF) e integrado com uma estratégia de voto majoritário. Resultados promissores foram obtidos com os modelos de classificação de *ensembles* propostos. Os experimentos realizados permitiram analisar a efetividade das abordagens desenvolvidas, bem como verificar os aspetos a serem melhorados.

Palavras-chave: Imagens macroscópicas e dermatoscópicas; Processamento e análise de imagens; Segmentação de imagens; Classificação de imagens; Extração e seleção de características; Classificadores; Modelos de classificação de *ensembles*.

CONTENTS

PART A

THESIS REPORT

1. Introduction

Skin cancer is one of the most commonly diagnosed cancers worldwide and melanoma is the most aggressive form of skin cancer, as well as the one with the highest mortality rate. For example, in the United States, 87,110 new cases of melanoma were estimated to be diagnosed in 2017. Furthermore, 9,730 deaths from melanoma were estimated for the same year [1]. Computational systems have been proposed in order to assist dermatologists in skin cancer diagnosis, or even to monitor skin lesions [2]. Image acquisition, pre-processing, segmentation, feature extraction and selection, and classification are fundamental steps commonly found in computational systems for diagnosing skin lesions. Macroscopic and dermoscopic images are examples of images acquired from non-invasive imaging techniques [3], which have been widely used in such systems.

This PhD project is focused on feature extraction and skin lesion classification in both macroscopic and dermoscopic images. Macroscopic images are usually obtained by using common digital cameras, while dermoscopic images are acquired by a dermatoscope device that allows a more detailed visualization of the lesion patterns on the skin surface. Figures 1 and 2 show some examples of such images considered in this project. Macroscopic images in Figure 1 (a-c) show three examples of pigmented skin lesions, i.e., nevus, seborrheic keratosis, and melanoma, respectively. Nevus and seborrheic keratosis are benign lesions, and melanoma is a malignant lesion. The difficulty in distinguishing such types of skin lesions has become a challenging research area. Hence, the differentiation between these types of skin lesions and their features were explored in this project. Dermoscopic images are shown in Figure 2 (a-b), which presents benign and malignant lesions, respectively. The search for better classification performance in dermoscopic images is a challenging task for skin lesion computational diagnosis. Thus, using skin lesion computational analysis in dermoscopic images to improve the diagnosis for benign and malignant lesions was the main goal of this research project.

The feature extraction process is usually based on the intensities of the pixels within the regions of interest (ROI). Segmentation is an important step that allows the extraction of the ROI within an image. Previous studies have shown that computational methods for image segmentation can provide suitable results for the identification of skin lesions in images [4]. Customarily, the images under analysis can be pre-processed for image

enhancement and artefact removal, so that more robust segmentations can be achieved [5]. The extraction of representative features of the ROI under analysis is essential for efficient classification of the skin lesions. The extracted features are usually based on clinical approaches used by dermatologists in diagnosing skin lesions. The ABCD rule is a commonly used method to classify such lesions in macroscopic and dermoscopic images according to asymmetry, border, colour and diameter (or differential structures in the case of dermoscopic images) criteria [6,7]. Additionally, texture analysis can be performed to assess the surface roughness of the lesions to assist in discriminating between benign and malignant lesions. Several computational solutions have been proposed for feature extraction of pigmented skin lesions, in order to represent them according to such clinical approaches [8-10].

Figure 1: Examples of pigmented skin lesions in macroscopic images: (a) nevus, (b) seborrheic keratosis and (c) melanoma [11].

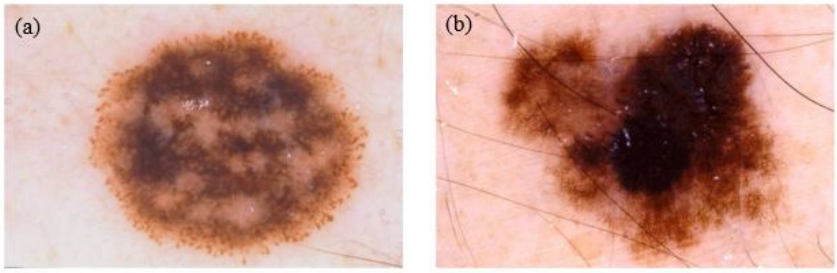

Figure 2: Examples of pigmented skin lesions in dermoscopic images: (a) benign lesion and (b) malignant lesion [12].

Skin lesion classification systems should demonstrate high performance and efficacy, considering that they will be used to assist in dermatological diagnosis. The evaluation and improvement of the classification performance are essential requirements of pattern recognition [13]. Some relevant challenges that should be considered are: 1) defining what features are meaningful and essential to represent the skin lesion patterns, 2) analysing whether there are redundant and irrelevant features that should be removed from the feature set, and 3) confirming whether the number of selected features is sufficient to describe the classification problem. The application of several descriptions

may be needed, considering the enormous quantity of information present in images. Nevertheless, a larger feature space can include redundant and irrelevant data. One solution to this problem is the application of feature selection algorithms to define the most appropriate features from images, since such algorithms permit the removal of redundant and irrelevant features [8,14]. Moreover, these algorithms reduce the dimensionality of data. As a result, they decrease the time of feature extraction, training, testing, and the complexity of classification, and they still can improve the classification accuracy rate.

Another solution to improve classification performance is ensemble methods [15]. Several studies have recently proposed ensemble classification models to achieve better performance of skin lesion classification from dermoscopic images [16,17]. Ensemble methods consist of integrating several classification models in order to develop a more robust system that provides more accurate results than by using a single classifier. Such models can be composed of either only one learning algorithm, classified as homogeneous or several learning algorithms, classified as heterogeneous. Several algorithms for constructing homogeneous ensembles have been developed through data manipulation, such as manipulating the training samples or the input features [18,19]. 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. Algorithms for manipulating the input features allow the generation of ensembles based on different features available to the learning algorithm. This process can involve, for example, the splitting of a set of features into subsets. Therefore, the use of ensemble methods to classify skin lesion images has presented promising perspectives, and this PhD project was particularly dedicated to exploring such perspectives.

1.1. Main goals of the project

The main objective of this PhD project was to develop algorithms for pattern recognition in skin lesion images in order to assist dermatologists in diagnosis of skin cancer. Hence, some important goals to be achieved in this project were:

- To provide a review of the most relevant computational methods that have been developed to assist pigmented skin lesion diagnosis from macroscopic and dermoscopic images. Essentially, this review is highly valuable for the design and

implementation of competent expert systems for the skin lesion pattern recognition in images;

- To segment skin lesions and extract their features, based on clinical approaches commonly used by dermatologists in pigmented skin lesion diagnosis, for the classification process;
- To analyse a combination of features by using different techniques for feature extraction; then, compare several feature selection algorithms and classifiers to evaluate the classification performance from the extracted features, in order to identify several features that may be relevant for skin lesion computational diagnosis;
- To develop new classification models based on ensemble methods and input feature manipulation for pattern recognition in skin lesion images. For these models, the feature subsets and ensemble methods are one of the most important factors for identifying malignancy in pigmented skin lesions, since the combination of several classification models and appropriate features can improve the performance of classification. This last goal was selected as the main objective of this project.

By fulfilling the aforementioned goals, this project aimed to contribute substantially to the Biomedical and Computational area, as well as to provide information that can assist dermatologists in obtaining more efficient diagnoses with the developed computational solutions.

1.2. Main contributions achieved

Five papers related to our research have been submitted to international journals during the development of this project. Of these, three papers have already been published [20-22], and two papers are currently under review. In addition, one full paper [23] and two abstracts [24,25] related to the developed work have been included in conference proceedings. One full paper [26] and one abstract [27] has also been accepted in conferences.

The main achievements and contributions obtained from this research can be summarized as follows:

- A comprehensive and current review concerning the most relevant computational methods to assist skin lesion diagnosis in both macroscopic and dermoscopic images was presented. This review includes methods that have been suggested for the steps of image pre-processing and segmentation, feature extraction and selection, and image classification, in which their advantages and disadvantages are identified. In addition, several of the reviewed segmentation methods are applied to macroscopic and dermoscopic images, in order to exemplify and discuss their applications;
- An effective computational approach was developed for the segmentation and classification of pigmented skin lesions in macroscopic images. The developed approach is based on asymmetry, border, colour and texture analysis for extracting skin lesion features and it allowed for distinguishing between some types of skin lesions;
- An effective approach for the combination of skin lesion features by using different feature extraction algorithms was proposed. The combination of features presented is relevant for skin lesion computational diagnosis in dermoscopic images. The main contribution of this approach was the texture analysis based on several colour channels, as well as the application of different texture-based feature extraction algorithms. In addition, experiments to analyse these features were performed for the image classification in benign or malignant lesions by using different classifiers and feature selection algorithms that ensure the effectiveness of this approach;
- Novel effective classification models based on ensemble methods and input feature manipulation to improve the skin lesion computational diagnosis from dermoscopic images were developed. The main contribution of these models is the feature subset selection based on specific feature groups and feature selection algorithms for the input feature manipulation, which allowed the finest generation of diversity for the ensemble models and better classification accuracy.

1.3. Organization of the thesis

This thesis is divided into two parts; namely, Part A and Part B. A summary description of the research developed, the conclusions drawn for the PhD project, and possible future studies about skin lesion classification are provided in Part A. In Part B, the research developed is detailed in five articles. In these articles, the computational methods for skin

lesion pattern recognition in both macroscopic and dermoscopic images that were developed to fulfil the presented objectives are described. The articles included in this part are:

- **Article 1:**

  Title: Computational methods for the image segmentation of pigmented skin lesions: A Review

  Authors: Roberta B. Oliveira, Mercedes E. Filho, Zhen Ma, João P. Papa, Aledir S. Pereira and João Manuel R. S. Tavares

  Published in Journal: Computer Methods and Programs in Biomedicine, 131:127-141, 2016

- **Article 2:**

  Title: Computational methods for pigmented skin lesion classification in images: Review and future trends

  Authors: Roberta B. Oliveira, João P. Papa, Aledir S. Pereira and João Manuel R. S. Tavares

  Published in Journal: Neural Computing and Applications, 27:1-24, 2016

- **Article 3:**

  Title: A computational approach for detecting pigmented skin lesions in macroscopic images

  Authors: Roberta B. Oliveira, Norian Marranghello, Aledir S. Pereira and João Manuel R. S. Tavares

  Published in Journal: Expert Systems with Applications, 61:53-63, 2016

- **Article 4:**

  Title: Computational diagnosis of skin lesions from dermoscopic images using a combination of features

  Authors: Roberta B. Oliveira, Aledir S. Pereira and João Manuel R. S. Tavares

  Submitted to an international journal (under review), 2017

- **Article 5:**

  Title: Skin lesion computational diagnosis of dermoscopic images: Ensemble models based on input feature manipulation

  Authors: Roberta B. Oliveira, Aledir S. Pereira and João Manuel R. S. Tavares

  Submitted to an international journal (under review), 2017

2. Brief description of the developed work

Skin lesion pattern recognition in macroscopic and dermoscopic images has become a challenging research area due to the difficulty in discerning some patterns of skin lesions. In order to fulfil the goals of this project, the following steps were defined: 1) reviews of computational method for pigmented skin lesion analysis, 2) pattern recognition in macroscopic images, and 3) pattern recognition in dermoscopic images. The first step includes the revision of the current state of the art concerning the segmentation and the classification of pigmented skin lesions in both macroscopic and dermoscopic images, which is provided in Articles 1 and 2 included in Part B. The second step reviews both the segmentation and the classification problems in macroscopic images, which is described in Article 3 included in Part B. Finally, the third step mainly outlines the feature extraction and classification problems in dermoscopic images, which is addressed in Articles 4 and 5 included in Part B. The overview of the developed work for the skin lesion pattern recognition is presented in Figure 3. In this section, a brief description of the developed work is presented.

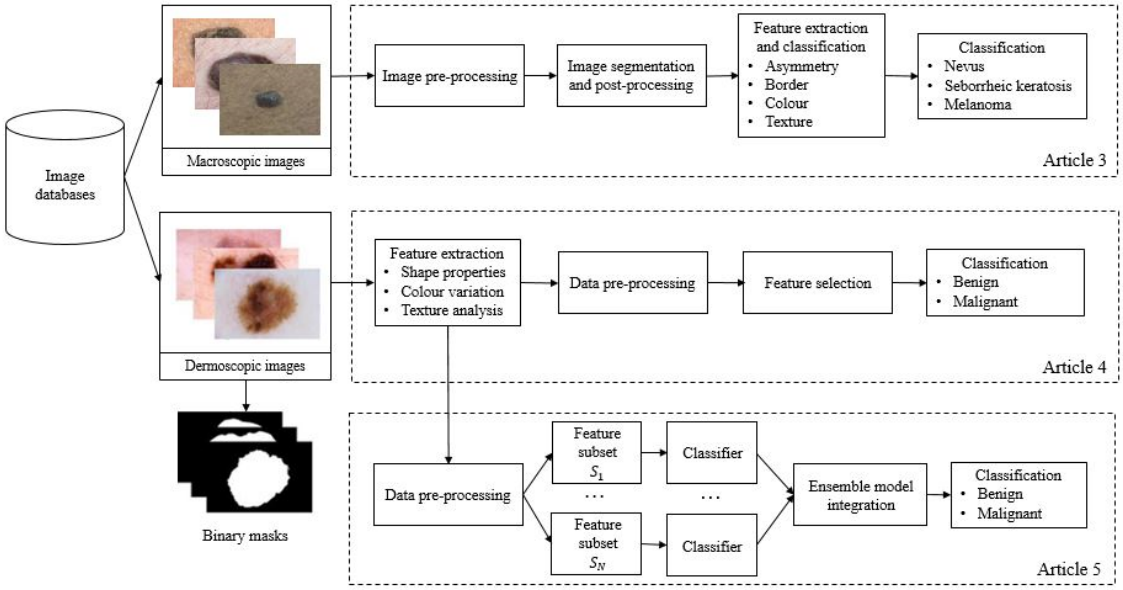

Figure 3: Overview of our research on skin lesion pattern recognition. (In addition to the article indicated in this scheme, two articles were provided with the revision about the segmentation and the classification of pigmented skin lesions in both macroscopic and dermoscopic images.)

2.1. Reviews of computational methods for pigmented skin lesion analysis

In order to analyse the current methods that have been proposed for skin lesion computational diagnosis in both macroscopic and dermoscopic images, two detailed reviews with regards to several of the fundamental steps of image computational analysis were carried out. One review focused on the image segmentation process, and other one on the classification process. These reviews allowed for defining the main guidelines for the development of this work concerning the pattern recognition of skin lesion images.

For the revision of image segmentation computational methods, three fundamental steps of image processing were taken; namely, image acquisition, pre-processing and segmentation. For each step, the techniques commonly used were explained, and their strengths and weaknesses were identified. The image segmentation techniques were classified into five categories according to their segmentation principle. In addition, several of the reviewed techniques are applied to both macroscopic and dermoscopic images in order to exemplify their results. This review is presented in Article 1 included in Part B.

For the revision of computational methods that have been developed for skin lesion classification, the steps of feature extraction and selection, and image classification were addressed. Feature extraction methods were introduced based on clinical approaches commonly used by dermatologists in skin lesion diagnosis. The feature selection process was described according to the following steps: 1) feature subset selection, 2) subset evaluation, 3) stopping criterion, and 4) validation procedure. The image classification step was addressed by including classifiers and evaluation procedures, as well as some performance results for pattern and disease classification. In addition, a discussion about the advantages and disadvantages of the reviewed methods, as well as future trends are also provided. This review is presented in Article 2 included in Part B.

2.2. Pattern recognition in macroscopic images

For the identification and classification of pigmented skin lesions in macroscopic images a new computational approach was developed based on the asymmetry, border, colour and texture analysis. The developed approach involves briefly the following steps: 1) image pre-processing, 2) image segmentation and post-processing, 3) feature extraction and classification, and 4) classification of type of skin lesions. The first step intends to

enhance input images corrupted by noise and is based on an anisotropic diffusion filter [28]. The second step identifies the lesion presented in the enhanced image by using an active contour model without edges (Chan-Vese model [29]), and for post-processing the segmented region is identified based on morphological filtering to improve the quality of the segmentation result. In the third step, features are extracted from the post-processed region, which include the asymmetry, border, colour, and texture properties. In addition, a support vector machine (SVM) classifier is applied to classify each feature property into two categories, according to their clinical principles. Finally, the last step concerns the binary classification between different types of skin lesions, i.e., nevus, seborrheic keratosis, and melanoma, based on the SVM classifier [30] and by using a histogram intersection kernel [31].

For assessing the asymmetry properties, features are obtained from the ratios between each pair of the semi-lines that represents the perpendicular lines by overlapping the two sub-regions of the lesion along an axis. For assessing the border properties, a number of peaks, valleys and straight lines of the border are computed by using the vector product and inflexion point descriptors from a one-dimensional border. For extracting the colour properties, statistical measures, i.e., average, variance and standard deviation, are computed for each colour channel of the *RGB* colour space. For the texture analysis, a fractal dimension method [32] is adopted based on grey-level images; by using the methods mentioned earlier, 44 features were extracted to represent the lesion. These feature extraction methods were also important for pattern recognition in dermoscopic images considered in the next sub-section.

A subjective evaluation was performed to analyse the obtained segmentation results. Hence, the visual assessment by a specialist of the segmented regions classified whether the lesions presented in 408 images were correctly segmented or not. The segmentation results obtained by the proposed approach were compared against the threshold-based segmentation results achieved by using a well-known method proposed by Otsu [33]. Afterwards, 385 correctly segmented images were used for the classification process, in which the results obtained by the proposed approach were compared with the results achieved by the SVM classifier by using a kernel commonly used, i.e., the radial basis function (RBF) kernel [30]. The proposed approach is detailed in Article 3 included in Part B.

2.3. Pattern recognition in dermoscopic images

Performance improvement of classifiers has become a challenging research topic for skin lesion pattern recognition in dermoscopic images. The defining of what features are meaningful to describe the skin lesion patterns is an important issue for the classification process. Thus, this project analysed a combination of features based on shape properties, colour variation and texture analysis by using different feature extraction methods. This approach involves the following steps: 1) feature extraction, 2) data pre-processing, 3) feature selection, and 4) image classification.

For the feature extraction step, 510 features related to shape, colour and texture were extracted from the skin lesion images. The shape properties include the lesion area, border perimeter, equivalent diameter, compactness, circularity, solidity, rectangularity, aspect ratio, eccentricity, lesion asymmetry, and border irregularity. The *RGB*, *HSV*, *CIE Lab* and *CIE Luv* colour spaces were used for extracting both colour and texture properties. The average, variance, standard deviation, minimum and maximum colours, and colour skewness were computed for each colour channel of the colour spaces. The texture features were also extracted for each colour channel of the colour spaces by using the fractal dimension analysis [32], discrete wavelet transform [34] and co-occurrence matrix [35] methods. For the data pre-processing step, a normalization method and a resampling procedure are applied to the data [36] in order to scale all numeric values to within the same interval and distribute the samples evenly for each class.

The feature selection and image classification steps consist of evaluating the effectiveness of the proposed combination of features in the benign or malignant lesion classification. Different feature selection algorithms were adopted, which include the Relief-F [37], information gain-based feature selection [38], gain ratio-based feature selection (GRFS) [36], Pearson's correlation coefficient-based feature selection [38], correlation-based feature selection (CFS) [39] and principal-component analysis (PCA) [40]. In addition, different categories of classifiers were compared; namely, the k-nearest neighbours (KNN) [41], Bayes networks [42], C4.5 decision tree [43], multilayer perceptron (MLP) [44], SVM [30] and optimum-path forest (OPF) [45]. The developed approach was applied to a set of 1104 dermoscopic images by using a cross-validation procedure [36]. This proposed approach is presented in Article 4 included in Part B.

From the analysed set of features, ensemble classification models based on input feature manipulation were proposed in order to improve the skin lesion computational diagnosis in dermoscopic images. Feature subsets from the shape properties, colour variation and texture analysis were selected to generate diversity for the ensemble models. The feature subsets are based on specific feature groups, subset correlation (CFS), and different feature selection algorithms, such as the algorithms mentioned earlier. Each ensemble classification model is constructed by using an OPF classifier [45] and integrated with a majority voting strategy [15]. For the OPF classifier, the Euclidean, Chebyshev and Manhattan distance functions [36] were used to measure the distances between the feature subsets.

The effectiveness of the feature groups used and feature selection algorithms was individually evaluated to define the best feature subsets for the benign or malignant lesion classification process. For the performance evaluation of the ensemble classification models, the data was pre-processed as discussed previously and the proposed models were applied to a set of 1104 dermoscopic images by using a cross-validation procedure [36]. The classification results achieved by the proposed ensemble classification models were compared against the ones obtained using ensemble algorithms that are commonly used in the literature; namely, bagging [15], AdaBoost [46] and random forest [47]. Furthermore, the proposed ensemble models were also compared to the individual OPF classifier [45] to analyse the efficacy of the ensemble algorithms. These ensemble classification models are considered in Article 5 included in Part B.

3. Conclusion and future works

Skin lesion pattern recognition is an area of great research interest due to its importance in skin cancer prevention, as well as in early diagnosis. This PhD project approached both macroscopic and dermoscopic image pattern recognition, and promising results were obtained for skin lesion diagnosis. Pattern recognition in macroscopic images are still little explored in research on pattern recognition field, and most studies do not deal with the classification of all features considered in present project. The approach using segmentation and classification of pigmented skin lesions in macroscopic images presented in this project, as described in Article 3 included in Part B, allowed identifying lesion features and the distinguishing between some types of skin lesions. Although the proposed approach achieved good segmentation results, mainly with noisy images, it

cannot perform well on images with too low contrast boundaries, shadows and reflections. Both feature and disease classifications presented significant results. However, some classification results were not expressive, e.g., the colour and texture feature classifications. Unbalanced databases regarding the number of samples for each class may have decreased the accuracy of the classification results, since the classifier tends to be based on classes with the highest occurrence.

Dermoscopic images have been widely applied for pattern recognition of pigmented skin lesions, since such images allow suitable visualization with more details of pigmentation patterns on the surface of the lesion. Pattern recognition in dermoscopic images has involved searching for new methods aiming to develop more effective systems for better skin lesion computational diagnosis. An approach for the combination of features, based on shape properties, colour variation and texture analysis by using different feature extraction methods, was presented in Article 4 included in Part B. The colour and texture analysis, based on several colour spaces combined with shape properties and by using different feature extraction methods, provided very promising results for skin lesion pattern recognition. In addition, the effectiveness of the combination of features was very important for the constructing ensemble classification models to improve the skin lesion computational diagnosis from dermoscopic images.

The ensemble classification models described in Article 5 in Part B were based on input feature manipulation from the shape properties, colour variation and texture analysis. The extracted features obtained successful results for the proposed ensemble models. Specific feature groups and feature selection algorithms were defined to find the best features for the classification process, as well as to generate diversity for the ensemble classification models. The best classification results were obtained by the feature subset selection model based on specific feature groups. The feature manipulation process based on such feature subsets allowed the best generation of diversity for the ensemble classification model. Although the ensemble models developed have achieved good results, there are some issues that can be enhanced, such as the ensemble integration method to combine the classification results produced by the base classifiers.

In conclusion, future studies, regarding the methods developed in this Thesis by using both macroscopic and dermoscopic images, should include searching for new methods aiming to improve the results for better skin lesion pattern recognition. Therefore, recommendations for future research can be summarized as follows:

- To improve the segmentation process, presented in Article 3 in Part B, the pre-processing process should be better addressed, since it significantly affects the skin lesion segmentation results. Therefore, the development of algorithms for dealing with reflections and shadows, as well as hair removal can be considered to solve the previously discussed problems concerning the image segmentation step. In addition, the definition of an initial automatic curve based on approximate lesion localization for the segmentation by using active contours should be addressed in order to improve the segmentation accuracy. Another relevant issue to be approached is the application of the segmentation algorithm in colour images. These improvements can make the segmentation method more accurate and thus, it still can be combined with the classification method described in Article 5 in Part B, since the lack of a segmentation process is a limitation of such a study.
- To enhance the ensemble classification models, described in Article 5 in Part B, the challenging problem of integration is an objective to be addressed, which can improve the classification results achieved with such ensemble models. In order to approach other problems concerning the dermoscopic image diagnosis, the proposed ensemble models should be taken into account in future studies; for example, to identify the presence of global and local patterns. Pattern analysis [48] is a challenging task in discriminating between benign and malignant skin lesions.
- Deep learning architectures [49] should be taken into account in future studies related to the skin lesion classification in both macroscopic and dermoscopic images. The classification results can be improved by using such architectures, since these architectures have revealed one can stress the capacity of learning from enormous amounts of data in an unsupervised way. In addition, these architectures have been recently used for skin lesion computational diagnosis with promising results [50-56].

References

[1] American Cancer Society (2017). *Cancer Facts & Figures 2017*. American Cancer Society, Atlanta.

[2] Scharcanski, J.; Celebi, M. E. (2013). *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, Berlin, Heidelberg.

[3] Smith, L.; MacNeil, S. (2011). State of the art in non-invasive imaging of cutaneous melanoma. *Skin Research and Technology*, 17 (3):257-269.

[4] Silveira, M.; Nascimento, J. C.; Marques, J. S.; Marcal, A. R. S.; Mendonca, T.; Yamauchi, S.; Maeda, J.; Rozeira, J. (2009). Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):35-45.

[5] Abbas, Q.; Fondón, I.; Rashid, M. (2011). Unsupervised skin lesions border detection via two-dimensional image analysis. *Computer Methods and Programs in Biomedicine*, 104 (3):e1-e15.

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

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

[8] Garnavi, R.; Aldeen, M.; Bailey, J. (2012). Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 16 (6):1239-1252.

[9] Cavalcanti, P. G.; Scharcanski, J. (2011). Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics*, 35 (6):481-491.

[10] Kasmi, R.; Mokrani, K. (2016). Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. *IET Image Processing*, 10 (6):448-455.

[11] Diepgen, T. L.; Yihune, G. (2012). *Dermatology Online Atlas*. Dermatology Information System - DermIS. http://www.dermis.net/dermisroot/en/home/index.htm. Accessed August 2012.

[12] Gutman, D.; Codella, N. C. F.; Celebi, E.; Helba, B.; Marchetti, M.; Mishra, N.; Halpern, A. C. (2016). *Skin Lesion Analysis toward Melanoma Detection: A Challenge* at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC), arXiv preprint arXiv:1605.01397.

[13] Webb, A. R. (2003). *Statistical pattern recognition*. 2 edn. John Wiley & Sons, England.

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

[15] Dietterich, T. G. (2000). *Ensemble methods in machine learning*. In: Multiple Classifier Systems, vol 1857. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 1-15.

[16] Grzesiak-Kopeć, K.; Ogorzałek, M.; Nowak, L. (2016). *Computational Classification of Melanocytic Skin Lesions*. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L. A., Zurada, J. M. (eds) 15th International Conference on

Artificial Intelligence and Soft Computing, Zakopane, June 12-16 2016. Springer, pp 169-178.

[17] Abbas, Q.; Sadaf, M.; Akram, A. (2016). Prediction of Dermoscopy Patterns for Recognition of both Melanocytic and Non-Melanocytic Skin Lesions. *Computers*, 5 (3):13.

[18] Schaefer, G.; Krawczyk, B.; Celebi, M. E.; Iyatomi, H. (2014). An ensemble classification approach for melanoma diagnosis. *Memetic Computing*, 6 (4):233-240.

[19] Barata, C.; Emre Celebi, M.; Marques, J. S. (2015). *Melanoma detection algorithm based on feature fusion*. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society Milan, August 25-29 2015. IEEE, pp 2653-2656.

[20] Oliveira, R. B.; Papa, J. P.; Pereira, A. S.; Tavares, J. M. R. S. (2016). Computational Methods for Pigmented Skin Lesion Classification in Images: Review and Future Trends. *Neural Computing and Applications*, 27:1-24.

[21] Oliveira, R. B.; Marranghello, N.; Pereira, A. S.; Tavares, J. M. R. S. (2016). A computational approach for detecting pigmented skin lesions in macroscopic images. *Expert Systems with Applications*, 61:53-63.

[22] Oliveira, R. B.; Filho, M. E.; Ma, Z.; Papa, J. P.; Pereira, A. S.; Tavares, J. M. R. S. (2016). Computational methods for the image segmentation of pigmented skin lesions: a review. *Computer Methods and Programs in Biomedicine*, 131:127-141.

[23] Oliveira, R. B.; Tavares, J. M. R. S.; Marranghello, N.; Pereira, A. S. (2013). *An Approach to Edge Detection in Images of Skin Lesions by Chan-Vese Model*. In: 8th Doctoral Symposium in Informatics Engineering - DSIE'13, Porto, Portugal, January 24-25 2013.

[24] Oliveira, R. B.; Tavares, J. M. R.; Pereira, A. S. (2014). *Features selection for the classification of skin lesion from images*. In: 11th World Congress on Computational Mechanics (WCCM XI), 5th European Conference on Computational Mechanics (ECCM V) and 6th European Conference on Computational Fluid Dynamics (ECFD VI), Barcelona, July 20-25 2014.

[25] Oliveira, R. B.; Tavares, J. M. R.; Pereira, A. S. (2013). *Selection of Image Features based on a Wrapper Model for the Classification of Skin Lesions*. In: 12th US National Congress of Computational Mechanics, Raleigh, North Carolina, July 22-25 2013.

[26] Oliveira, R. B.; Pereira, A. S.; Tavares, J. M. R. S. (2017). *Pattern Recognition in Macroscopic and Dermoscopic Images for Skin Lesion Diagnosis*. In: VipIMAGE 2017 - VI ECCOMAS Thematic Conference on Computational Vision and Medical Image processing, Porto, Portugal, October 18-20 2017.

[27] Oliveira, R. B.; Marranghello, N.; Pereira, A. S.; Tavares, J. M. R. S. (2017). *Segmentation of Skin Lesion Images based on an Active Contour Model*. In: Congress on Numerical Methods in Engineering, Valencia, Spain, July 3-5 2017.

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

[29] Chan, T. F.; Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10 (2):266-277.

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

[31] Barla, A.; Odone, F.; Verri, A. (2003). *Histogram intersection kernel for image classification*. In: International Conference on Image Processing, Italy, September 14-17 2003. IEEE, pp 513-516.

[32] Al-Akaidi, M. (2004). *Fractal speech processing*. Cambridge university press, New York.

[33] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9 (1):62-66.

[34] Scheunders, P.; Livens, S.; Van de Wouwer, G.; Vautrot, P.; Van Dyck, D. (1998). Wavelet-based texture analysis. *International Journal on Computer Science and Information Management*, 1 (2):22-34.

[35] Haralick, R. M.; Shanmugam, K.; Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3 (6):610-621.

[36] Witten, I. H.; Frank, E.; Hall, M. A. (2011). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco.

[37] Kononenko, I. (1994). *Estimating attributes: Analysis and extensions of RELIEF*. In: Bergadano, F., De Raedt, L. (eds) Machine Learning: ECML-94, vol 784. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 171-182.

[38] Guyon, I.; Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157-1182.

[39] Hall, M. A. (2000). *Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning*. In: Proceedings of the 17th International Conference on Machine Learning, San Francisco, USA, June 29 - July 02 2000. Morgan Kaufmann Publishers Inc., pp 359-366.

[40] Hand, D.; Mannila, H.; Smyth, P. (2001). *Principles of Data Mining*. The MIT Press, London.

[41] Cover, T.; Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13 (1):21-27.

[42] Congdon, P. (2007). *Bayesian statistical modelling*, vol 704. 2 edn. John Wiley & Sons, Chichester.

[43] Quinlan, J. R. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., USA.

[44] Haykin, S. S. (1999). *Neural networks: a comprehensive foundation*. Prentice Hall, Englewood Cliffs, USA.

[45] Papa, J. P.; Falcao, A. X.; Suzuki, C. T. (2009). Supervised pattern classification based on optimum-path forest. *International Journal of Imaging Systems and Technology*, 19 (2):120-131.

[46] Freund, Y.; Schapire, R. E. (1997). A desicion-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55 (1):119-139.

[47] Breiman, L. (2001). Random forests. *Machine learning*, 45 (1):5-32.

[48] Argenziano, G.; Soyer, H. P.; Chimenti, S.; Talamini, R.; Corona, R.; Sera, F.; Binder, M.; Cerroni, L.; De Rosa, G.; Ferrara, G.; Hofmann-Wellenhof, R.; Landthaler, M.; Menzies, S. W.; Pehamberger, H.; Piccolo, D.; Rabinovitz, H. S.; Schiffner, R.; Staibano, S.; Stolz, W.; Bartenjev, I.; Blum, A.; Braun, R.; Cabo, H.; Carli, P.; De Giorgi, V.; Fleming, M. G.; Grichnik, J. M.; Grin, C. M.; Halpern, A. C.; Johr, R.; Katz, B.; Kenet, R. O.; Kittler, H.; Kreusch, J.; Malvehy, J.; Mazzocchetti, G.; Oliviero, M.; Özdemir, F.; Peris, K.; Perotti, R.; Perusquia, A.; Pizzichetta, M. A.; Puig, S.; Rao, B.; Rubegni, P.; Saida, T.; Scalvenzi, M.; Seidenari, S.; Stanganelli, I.; Tanaka, M.; Westerhoff, K.; Wolf, I. H.; Braun-Falco, O.; Kerl, H.; Nishikawa, T.; Wolff, K.; Kopf, A. W. (2003). Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet. *Journal of the American Academy of Dermatology*, 48 (5):679-693.

[49] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2 (1):1-127.

[50] Lequan, Y.; Chen, H.; Dou, Q.; Qin, J.; Heng, P. A. (2016). Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. *IEEE Transactions on Medical Imaging*:1-11.

[51] Codella, N.; Nguyen, Q.-B.; Pankanti, S.; Gutman, D.; Helba, B.; Halpern, A.; Smith, J. R. (2017). Deep Learning Ensembles for Melanoma Recognition in Dermoscopy Images. *IBM Journal of Research and Development*, 61 (4/5):1-28, arXiv preprint arXiv:1610.04662v04662.

[52] Esteva, A.; Kuprel, B.; Novoa, R. A.; Ko, J.; Swetter, S. M.; Blau, H. M.; Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542:115-118.

[53] Demyanov, S.; Chakravorty, R.; Abedini, M.; Halpern, A.; Garnavi, R. (2016). *Classification of dermoscopy patterns using deep convolutional neural networks*. In: 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, April 13-16 2016. IEEE, pp 364-368.

[54] Sabbaghi, S.; Aldeen, M.; Garnavi, R. (2016). *A deep bag-of-features model for the classification of melanomas in dermoscopy images*. In: 38th Annual International

Conference of the Engineering in Medicine and Biology Society (EMBC), Orlando, Florida, August 16-20 2016. IEEE, pp 1369-1372.

[55] Kawahara, J.; BenTaieb, A.; Hamarneh, G. (2016). *Deep features to classify skin lesions*. In: 13th International Symposium on Biomedical Imaging (ISBI) Prague, Czech Republic, April 13-16 2016. IEEE, pp 1397-1400.

[56] Sun, X.; Yang, J.; Sun, M.; Wang, K. (2016). *A Benchmark for Automatic Visual Classification of Clinical Skin Disease Images*. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) 14th European Conference on Computer Vision, Amsterdam, October 11-14 2016. Springer, pp 206-222.

PART B - ARTICLE 1

COMPUTATIONAL METHODS FOR THE IMAGE SEGMENTATION OF PIGMENTED SKIN LESIONS: A REVIEW

Roberta B. Oliveira, Mercedes E. Filho, Zhen Ma, João P. Papa, Aledir S. Pereira and João Manuel R. S. Tavares

Published in: Computer Methods and Programs in Biomedicine, 131:127-141, 2016

Abstract

Background and Objectives: Because skin cancer affects millions of people worldwide, computational methods for the segmentation of pigmented skin lesions in images have been developed in order to assist dermatologists in their diagnosis. This paper aims to present a review of the current methods, and outline a comparative analysis with regards to several of the fundamental steps of image processing, such as image acquisition, pre-processing and segmentation. *Methods:* Techniques that have been proposed to achieve these tasks were identified and reviewed. As to the image segmentation task, the techniques were classified according to their principle. *Results:* The techniques employed in each step are explained, and their strengths and weaknesses are identified. In addition, several of the reviewed techniques are applied to macroscopic and dermoscopy images in order to exemplify their results. *Conclusions:* The image segmentation of skin lesions has been addressed successfully in many studies; however, there is a demand for new methodologies in order to improve the efficiency.

Keywords: Image acquisition; Image pre-processing; Image segmentation; Pigmented skin lesion images.

1. Introduction

Pigmented skin lesions, which may be classified as benign or malignant, are mainly caused by an abnormal production of a group of cells in some specific regions. Benign lesions have a more organized behaviour than malignant lesions, since the former do not proliferate into other tissues. Nevus, such as melanocytic, blue, halo, sptiz and dysplastic (Figure 1a), and seborrheic keratosis (Figure 1b), 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 1c) and non-melanoma (Figure 1d). Basal cell carcinoma and squamous cell carcinoma are two examples of non-melanoma skin cancer (NMSC) and are 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. Melanoma is the most aggressive form of skin cancer, and the one with the highest mortality rate, due to its high levels of metastasis [1].

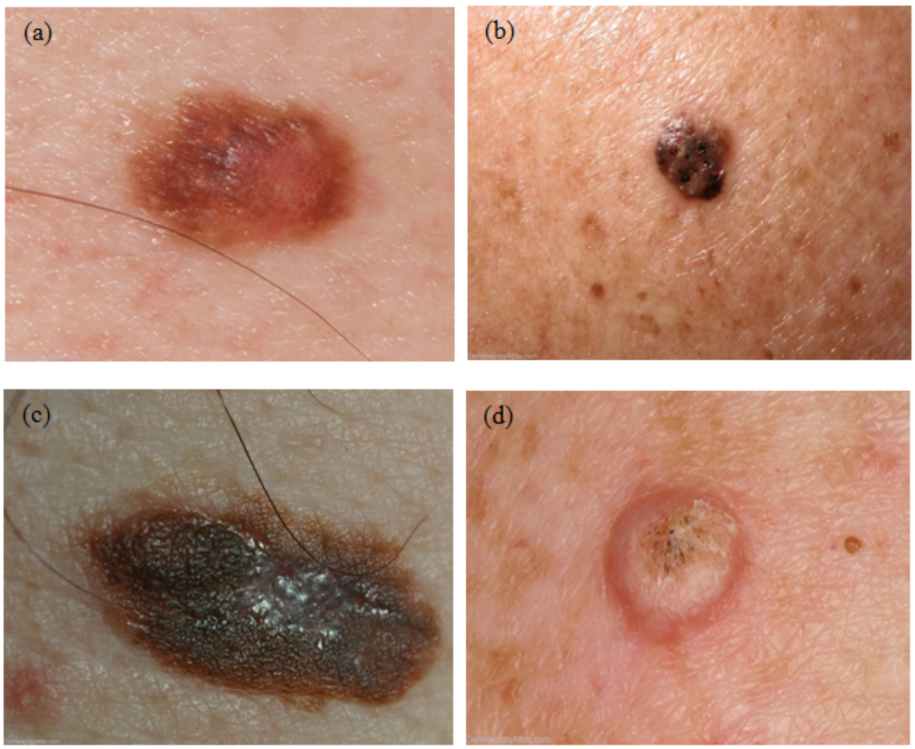

Figure 1: Four examples of skin lesions: (a) dysplastic nevus, (b) seborrheic keratosis, (c) melanoma, and (d) squamous cell carcinoma (images publicly available in [2]).

Melanoma was the 19th most common cancer worldwide in 2008, with an approximate estimation of 200,000 new cases, and with the highest incidence rate in Australia/New Zealand, Northern America and Northern Europe, and the lowest in South-Central Asia [3]. Table 1 presents recent data regarding skin cancer in the United States of America (USA), the United Kingdom (UK) and Brazil, according to gender. In the USA, 76,100 new cases of melanoma were estimated to be diagnosed in 2014 [4]. This estimate does not include NMSC, since this form of skin cancer is not required to 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. In the UK, melanoma was the 15th most common cancer in 2010, with approximately 12,800 new cases of this disease [3]. As a result, melanoma was the 18th most common cause of death from cancer in the UK. In 2011, there were 2,209 deaths from melanoma, and 590 deaths from NMSC in the UK. Of these deaths from melanoma, 59% of the deaths were male patients, and 41% of the deaths were female patients. In Brazil, NMSC will be the most common form of cancer, since approximately 182,000 new cases are estimated in 2014 and 2015 [5]. Although NMSC has a lower mortality rate, it has a higher incidence than melanoma.

Table 1: Number of new cases of skin cancer, according to gender, in the USA, UK and Brazil.

Country	Type of skin cancer	Year	Number of new cases	
			Male	Female
USA[a]	Melanoma	2014	43,890	32,210
UK[b]	Melanoma	2010	6,201	6,617
	Non-melanoma		55,747	43,802
Brazil[c]	Melanoma	2014	2,960	2,930
	Non-melanoma		98,420	83,710

[a] Estimated number, based on 1995-2010 incidence rates.
[b] Confirmed cases in 2010.
[c] Estimated number in 2014, and valid also for 2015.

Recently, there has been a great interest in the development of computer-aided diagnosis (CAD) systems for the detection and analysis of pigmented skin lesions from images [6-9], which can assist the dermatologist in preventing the development of malignant lesions. Particularly, CAD systems may be used to monitor benign skin lesions, in order 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 for being properly treated.

On the other hand, there is also a great interest concerning the image segmentation step of the CAD systems. This step allows for a better representation of the lesion under study, and extraction of its features. Image segmentation has, therefore, a critical role in the effectiveness of the CAD systems. Previous studies [10-15] 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 [16,17]. An overview of lesion border detection methods, which addresses the pre-processing, segmentation and post-processing steps, is presented in Celebi et al. [18]. In addition, the authors also discuss performance evaluation issues, and propose guidelines for future studies. However, they primarily focus on dermoscopy images of pigmented skin lesions, and the segmentation methods were classified according to the images to be segmented. In this review, we introduce some of the most relevant solutions that have been developed to assist the diagnosis of skin lesions from images, including those concerning the steps of image acquisition, pre-processing and segmentation. In particular, we comprehensively review the computational techniques that have been suggested for the image segmentation of pigmented skin lesions. In the following sections, these techniques are classified into five classes according to their

segmentation principle, specifically, based on edges, thresholding, regions, artificial intelligence techniques, and the ones based on active contours. In addition, several of the reviewed techniques are applied to macroscopic and dermoscopy images, in order to exemplify and discuss their applications.

The paper is organized as follows: in Section 2, a review of the current state-of-the-art concerning the image segmentation of pigmented skin lesions is provided. In addition, smoothing and segmentation results by using several methods are presented. In Section 3, the properties of some of the reviewed computational methods are discussed, and their advantages/disadvantages are identified. Finally, in Section 4, the conclusions of the review and future trends are outlined.

2. Image segmentation of pigmented skin lesions

2.1. Imaging techniques

Different non-invasive imaging techniques have been employed to assist dermatologists in the diagnosis of skin lesions. Dermoscopy, photography, confocal scanning laser microscopy (CSLM), optical coherence tomography (OCT), ultrasound, magnetic resonance imaging (MRI), and spectroscopic imaging are examples of these techniques [19-21]. Macroscopic images, commonly known as clinical images [13,22,23], and images acquired by epiluminescence microscopy (ELM), also called dermoscopy or dermatoscopy images [12,14,15,24-27], are normally used in the computational analysis of skin lesions. Figure 2 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 cause complications 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.

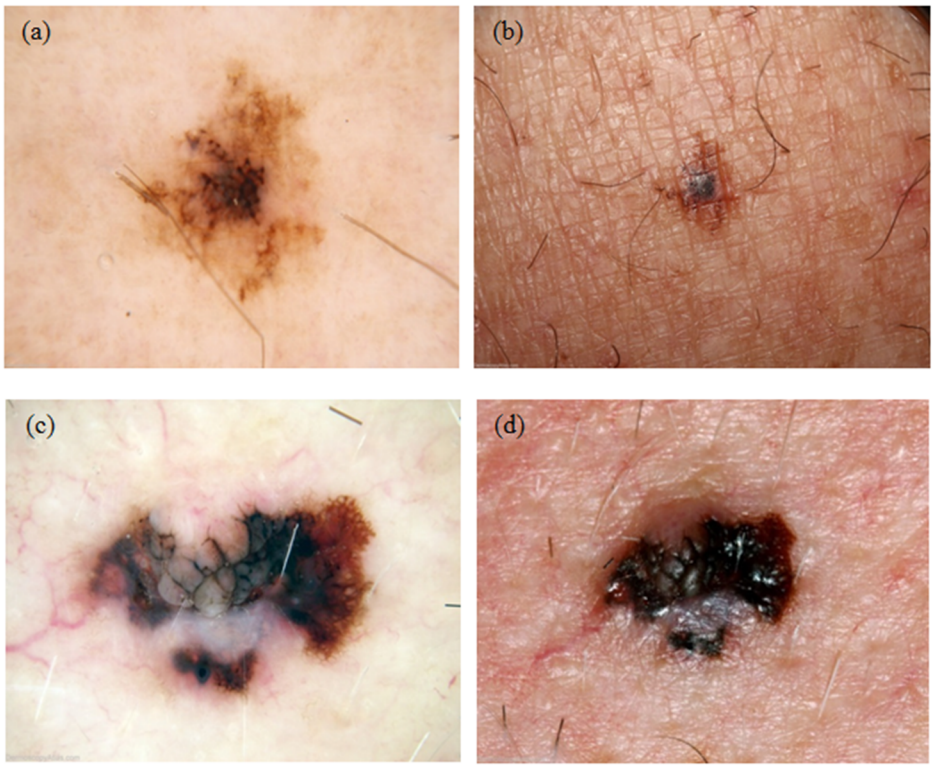

Figure 2: Examples of dermoscopy (a and c) and macroscopic (b and d) images: (a) and (b) are images of melanoma in situ, and (c) and (d) are of invasive melanoma (these images are publicly available in [2]).

Essentially, ELM is a non-invasive technique for image acquisition, where the lesion is 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. 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. Effective methods based on colour space transformation [28-30], illumination correction [31,32], contrast enhancement [28,29,33,34] and artefact removal [28,35], as a pre-processing step have been proposed in order to improve the segmentation accuracy.

In order to pre-process both macroscopy and dermoscopy images, the original *RGB* (red, green, blue) colour image may be used. The application may adopt scalar (single channel) or vector (multichannel) processing. In scalar processing, the colour image is converted into a scalar image such as, for example, a grey-level image, or only the blue channel is retained, since the lesions are often more evident in this channel [18]. In vector processing, the original *RGB* image may be used directly or after conversion to other colour spaces, such as the *CIE L*a*b** [29], *CIE L*u*v** [6], and *HSV* (hue, saturation, value) spaces [31]. These colour spaces are commonly used in literature to enhance colour images, since they augment the approximate perceptual uniformity of the image colours. Several pre-processing methods were originally designed for scalar images. However, these methods may also be applied to colour images, for example, by applying the scalar method separately to each colour channel of a given colour space, and then combining the results [36], or adopting methods that deal with vector data [37].

Artefacts due to illumination variation, such as shadows and reflections, may significantly affect the skin lesion segmentation results, specifically in macroscopic images. For shading effect attenuation in macroscopic images, Cavalcanti et al. [31] proposed a method for illumination variation modelling with a quadratic function. This method converts the original *RGB* image to the *HSV* colour space, and retains the *V* channel in order to obtain a higher visibility of the shading effects. The normalized image is obtained by applying, on the *HSV* image, an estimate of the quadratic function computed from the local illumination intensity in *V* channel. Afterwards, the normalized image is converted from the *HSV* colour space back to the *RGB* colour space, but now with the shading effects significantly attenuated. Colour image segmentation is then performed on this illumination-corrected image, by using the Otsu's thresholding segmentation approach [38]. Recently, Glaister et al. [32] proposed a new multistep illumination modelling method to correct the illumination variation in macroscopic images. This method first determines a nonparametric model of the illumination by using a Monte Carlo sampling method. Then, a parametric quadratic surface model is used to

determine the final illumination estimation. Finally, the illumination-corrected image is obtained by using the reflectance component computed from the final estimated illumination.

Another factor that complicates the segmentation of skin lesions, in both macroscopic and dermoscopy images, is the low contrast of the lesions. Celebi et al. [34] presented a method to enhance the contrast in dermoscopy images. The method searches for the optimal weights to convert an original *RGB* image to the corresponding grey-level image, by maximizing an Otsu's histogram bimodality measure. Recently, Barata et al. [36] used a shades-of-grey method for colour compensation in dermoscopy images. This method only uses image information to estimate the colour of the light source. Morphological filtering [39], which is based on set theory, may also be used to enhance skin lesions in images [40]. For example, one may refer to the work of Beuren et al. [40], where colour morphological filtering is used to enhance the regions of the lesions. Moreover, morphological filtering has been applied in order to include areas with low contrast borders in the detected lesion regions [26,41], and to remove image noise [12,41].

Algorithms for hair removal, in both macroscopic and dermoscopy images, are commonly used in pre-processing steps, since this artefact may considerably affect the detection of the lesion borders. Lee et al. [42] proposed a solution for hair removal, especially thick dark hairs, which is based on one of the first widely adopted methods for hair removal in dermoscopy images, and consists of three main steps: 1) identify the hair location by applying a grey-level morphological operation to the three colour channels of the original *RGB* image separately, and build the binary hair mask image by using thresholding to divide the image into hair and non-hair regions; 2) replace the values of the detected hair pixels in the original image by the values of the corresponding nearby non-hair pixels; and 3) apply a binary morphological operation and median filter to smooth the thin lines. This method has influenced several other methods for hair detection and removal [43-46].

The presence of hairs in images may also be reduced by the application of image smoothing methods, such as the median and anisotropic diffusion filters, without losing relevant information about the lesions, and, therefore improving the accuracy of the segmentation process. The median filter [47], which is a non-linear image filtering method, has been commonly applied on noisy images showing successful results. Unlike linear filters, such as the average filter [47], this type of filter allows the smoothing of the

original image without blurring edges and thin details. The median filter has been often applied to smooth images of skin lesions, as well as to remove artefacts, maintaining the edges of the lesions, which is imperative for an adequate segmentation [6,12,48,49]. To establish the best median filtering mask for the smoothing of skin lesion images, Celebi et al. [48] 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 [50] has also been used for smoothing skin lesion images [17]. This filter is applied iteratively, such that the number of iterations is determined according to the amount of noise presented 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. [51] proposed an enhancement of the anisotropic diffusion algorithm, originally suggested by Perona and Malik [50]. 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. [52] and Nordström [53] to enhance the edges.

The results of the application of the median [47], average [47] and anisotropic diffusion [50] filters to an 256 x 256 pixel image are shown in Figure 3. A 9 x 9 convolution mask was used in the median and average filtering, since other masks did not lead to a successfully smoothed image with a reduced noise level. Regarding the anisotropic diffusion filter, the smoothing was halted after 150 iterations.

Unlike most methods proposed in literature for reducing the influence of hairs on images of skin lesions, Abbas et al. [16] suggested an effective pre-processing method for the reduction of different artefacts, in both dermoscopy and macroscopic images, and, consequently, a better detection of lesion borders. Essentially, this method consists of three steps: 1) specular reflection reduction by applying homomorphic filtering [54], Fast Fourier Transform (FFT) and high pass filtering, in order to modify the illumination and reflectance, and obtaining, therefore, high contrast skin lesions, 2) the reduction of dermoscopic-gel or air bubble artefacts, 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) [55], and the exemplar-based inpainting technique [56].

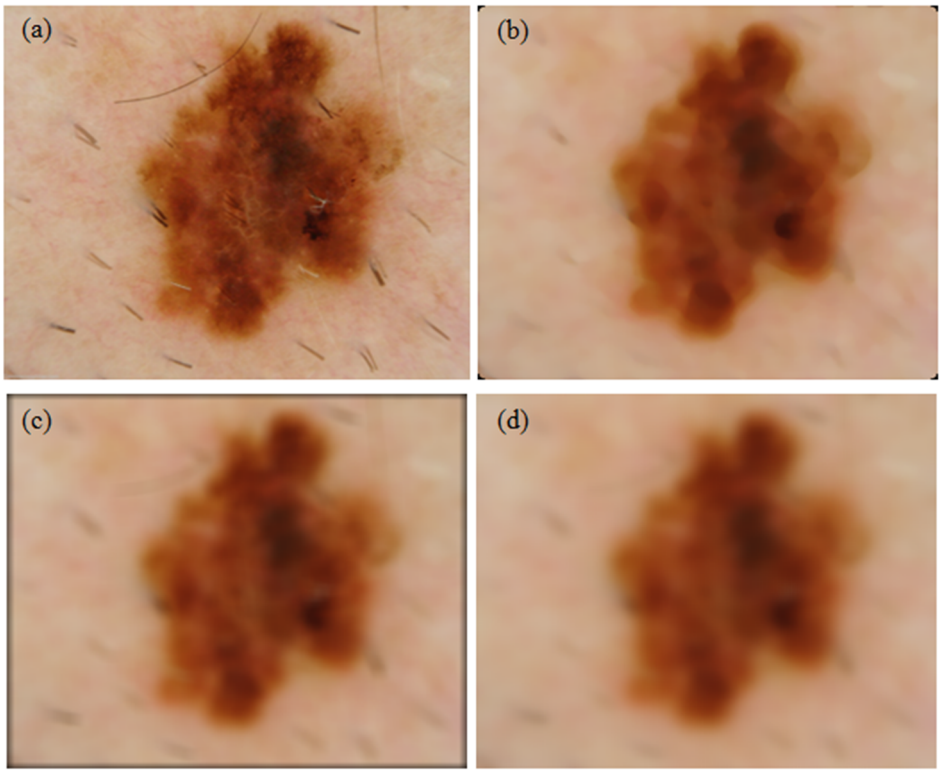

Figure 3: Application of smoothing filters: (a) original dermoscopy image of a melanoma (publicly available in [2]), and the corresponding images obtained after (b) median, (c) average, and (d) anisotropic diffusion filtering.

2.3. Image segmentation

Segmentation allows the extraction of the region of interest (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 [57]. The segmentation methods may be edge-based, i.e., the methods 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. Many segmentation methods are originally designed for scalar images. Therefore, several applications are available to convert the original colour image to scalar data [58], for example, grey-level images, pursuing the computational simplicity and convenience of scalar processing. However, in order obtain better segmentation results by using the information contained in all the colour channels of the original images, segmentation methods dedicated to process vector images have been developed [59]. However, this vector image processing is usually more computationally demanding and requires appropriate colour spaces.

In the following sections, we discuss the applicability of some methods commonly used in 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. Other methods are discussed in Section 3. The reviewed research is summarized in Table 2. Research that combines different methods [10,14,60,61], and that compares segmentation methods [12], is also include in Table 2.

Table 2: Research that has been performed related to the segmentation of skin lesions in images.

Segmentation method	Technique	References
Edge-based	Edge detectors	[17,62]
Thresholding-based	Otsu's thresholding	[6,26,29,31,34,41,44,49,63-67]
	Fuzzy logic	[14]
	Renyi's entropy	[40]
	Adaptive thresholding	[12,68]
	Iterative thresholding	[33,61,69]
	Ensemble	[24]
	Statistics	[7,23,70]
Region-based	Region growing	[6,7,10,12]
	Statistical region merging	[32,48,62,71]
	Iterative stochastic region merging	[13]
AI-based	Neural networks	[33,60,69]
	Evolutionary computation	[11]
	Fuzzy logic	[10,12,14,27,44,60,61,71]
	k-means clustering	[44,71,72]
Active contour-based	Adaptive snake	[12]
	Gradient vector flow	[12,15,65,73,74]
	Level set	[66]
	Region-based active contour algorithm	[16,28]
	Active contour without edges	[12,67]
	Expectation-maximization level set	[12]
Other methods	Hill-climbing algorithm	[29]
	Dynamic programming	[58,75]

2.3.1. Edge-based segmentation

The changes in intensity of the pixels in an image to be segmented may be determined based on the magnitude of the gradient used to detect the edges of the ROI [57]. The Prewitt, Sobel, Roberts, Laplacian [57] and Canny [76] operators are common examples of edge detectors that lead to image segmentation based on edges. According to Sonka et al. [39], 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 [39]: 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 of generating double edges, and 4) the large sensitivity to image noise.

The edge detector developed by Canny [76] has been applied to skin lesion images [17,62], due to its advantages compared with other edge detectors: 1) it provides 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. Firstly, Canny's algorithm smooths the input image $f(x,y)$, performing a convolution with a Gaussian function $G(x,y)$:

$$g(x,y) = f(x,y) * G(x,y), \tag{1}$$

where:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{2}$$

and where σ is the Gaussian function standard deviation. Then, the gradient magnitude $M(x,y)$, and the direction $\alpha(x,y)$, at each pixel in the smoothed image $g(x,y)$, are computed:

$$M(x,y) = \sqrt{g_x^2 + g_y^2}, \text{ and} \tag{3}$$

$$\alpha(x,y) = tan^{-1}\frac{g_x}{g_y}. \tag{4}$$

Subsequently, the non-maximum suppression technique is used to preserve all pixels with local maximum in the gradient image. Afterwards, double thresholding (T_1, T_2) is established to remove the weak edges. The pixels with a gradient magnitude below the T_1 are considered as weak edges, and the pixels with a gradient magnitude above T_2 are considered as strong edges. Finally, the final edges are defined by all the pixels considered as strong edges or also by the weak pixels that can be connected to any strong pixels.

Figure 4 illustrates the segmentation results from application of Canny's edge detector to two skin lesion images [76]. Usually, a median filter [47] is applied before the edge detector, in order to smooth the original image and reduce the noise. However, the edges generated by Canny's edge detector are usually not satisfactory. Although the lesions are identified by the detector, the generated edges are discontinuous; thus, the boundaries of the lesions are not fully detected. In addition, there is a large sensitivity to the noise, which generates boundaries that are not part of the lesions.

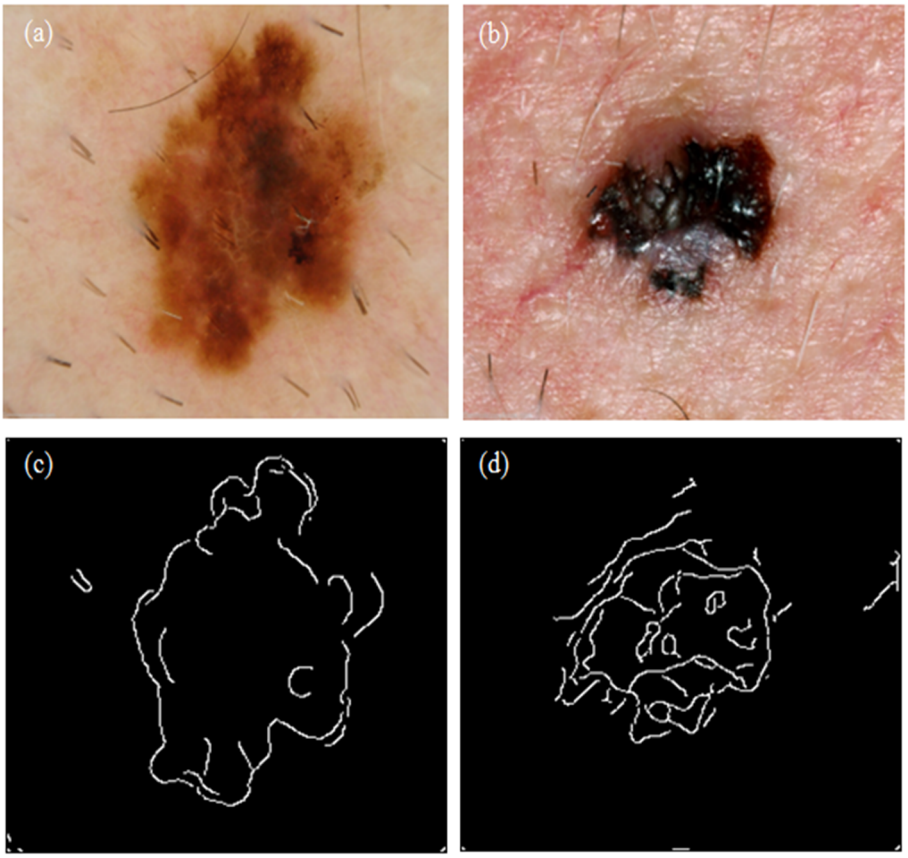

Figure 4: Segmentation results after applying Canny's edge detector to a dermoscopy image (a and c), and to a macroscopic image (b and d).

Barcelos and Pires [17] employed Canny's edge detector after the application of an anisotropic diffusion smoothing filter [51], and the results demonstrated 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.

2.3.2. Thresholding-based segmentation

The thresholding technique has been commonly used in several skin lesion segmentation methods proposed in literature. This technique is based on the histogram of the input image, which represents the distribution of the image pixels, $P_i = n_i/N$, in terms of each

possible intensity level, $i = [1, 2, \dots, L]$, where n_i is the number of pixels for a particular intensity level i, N is the total number of pixels of the image, and L is the number of intensity levels. 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 [38], which has many applications in image segmentation of skin lesion [6,26,29,41,49,63,67]. This method is based on a normalized histogram, built in order to set the optimal threshold value k, which separates the pixels of the input image into two homogeneous classes (C_0, C_1), with minimal variance (σ_B^2): one class for the ROI, $C_0 = [1,2, \dots, k]$, and the other class for the image background, $C_1 = [k + 1, k + 2, \dots, L]$:

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2, \quad (5)$$

$$\omega_0 = \sum_{i=1}^{k} P_i, \omega_1 = \sum_{i=k+1}^{L} P_i, \quad (6)$$

$$\mu_0 = \sum_{i=1}^{k} \frac{iP_i}{\omega_0}, \mu_1 = \sum_{i=k+1}^{k} \frac{iP_i}{\omega_i}, \text{ and} \quad (7)$$

$$\mu_T = \sum_{i=k+1}^{k} iP_i, \quad (8)$$

where ω_0 and ω_1 are the probabilities, and μ_0 and μ_1 the means of the classes C_0 and C_1, respectively. Thus, μ_T is the total mean of the intensities of the input image. Figure 5 presents the segmentation results after the application of Otsu's method [38] to dermoscopy and a macroscopic images. A median filter [47] 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, are not identified as part of the lesions. Furthermore, this edge detector is very sensitive to artefacts and, therefore, because of reflections, some interior regions of the lesions are wrongly identified as belonging to the lesions.

Otsu's method has revealed some problems, such as: (1) the segmented lesions tend to be smaller than they are in reality; and (2) it may lead to very irregular lesion edges. Yuksel and Borlu [14] proposed a method using the type-2 fuzzy logic technique [77] 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. Alcón et al. [23] proposed an improved thresholding

technique to overcome some issues of Otsu's method. In the proposed algorithm, the threshold is defined by finding the average value between the means of both background and lesion probability distributions. Cavalcanti et al. [31] and Gómez et al. [72] suggested building projections of the original *RGB* colour space, where they were able to properly apply Otsu's method. A thresholding method based on the Renyi's entropy [78] has also been applied to define the desired threshold value, leading to segmentations that preserve the geometry and shape of the lesions [40]. Another technique to define the threshold value is indicated by Xu et al. [70], which considers the average intensity of the strongest gradient pixels in the input image. Threshold selection by an iterative [33,61,69] or an adaptive [12,68] process has also been adopted to segment skin lesions in images. The fusion of the results provided by the ensemble of thresholding methods results in another segmentation technique based on thresholding [24].

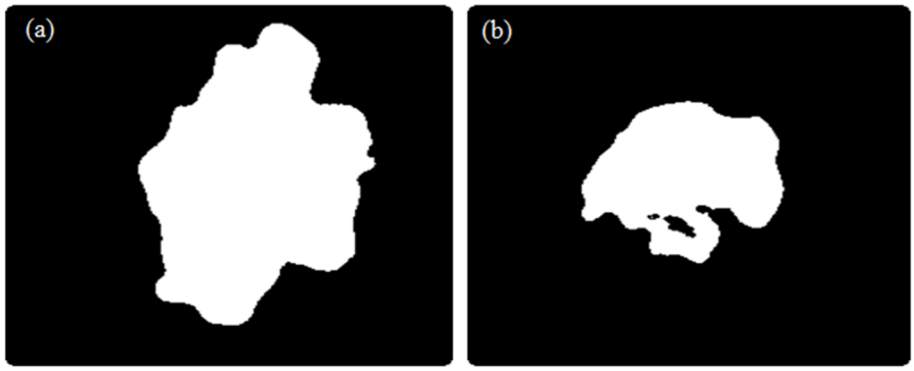

Figure 5: Segmentation results after applying Otsu's method (a) to the dermoscopy image shown in Figure 4a, and (b) to the macroscopic image shown in Figure 4b.

2.3.3. Region-based segmentation

The region growing algorithm [79], splitting and merging operations [80], and the Mumford-Shah method [81] 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 in 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 together [6,7]. The splitting and merging operations are region-based techniques applied to group similar regions [10,12]. 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 $\Omega_i = \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_n \cup k$,

where k is the boundary between them, merging the close regions by analysing their pixel intensities. This technique is based on an energy functional $E(k)$, calculated as:

$$E(k) = \sum_i \int \int_{\Omega_i} \left| |u - c_i| \right|^2 dx\, dy + \lambda\, l(k), \quad (9)$$

where u is a constant function into each image region Ω_i, $c_i = mean(u)$, dx and dy are the differentials of x and y, λ is a parameter that is incremented at each iteration, and $l(k)$ is the total length of the regions at each iteration.

The active contour model without edges [82] is based on the Mumford-Shah method and has been used in the image segmentation of skin lesions [12,16]. Examples of the results obtained by the Mumford-Shah method applied to skin lesion images are presented in Figure 6. The method was employed on two images that were previously smoothed using the median filter [47]. Observation of the resultant images, shows that the lesions are completed identified, including the lesion regions with considerable colour variation. However, some regions are erroneously identified as belonging to the lesions due to image artefacts.

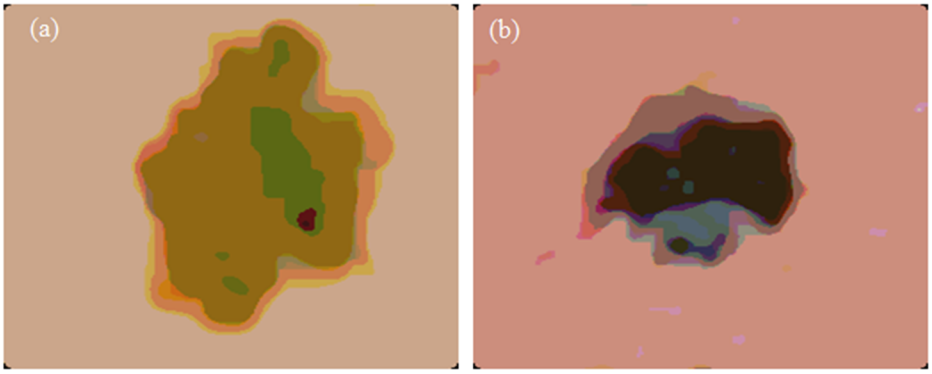

Figure 6: Segmentation results after applying the Mumford-Shah method: (a) to the dermoscopy image shown in Figure 4a, and (b) to the macroscopic image shown in Figure 4b.

Castillejos et al. [71], Celebi et al. [48] and Ganzeli et al. [62] employed the statistical region merging (SRM) algorithm [83] 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. [13], based on the SRM algorithm [83]: 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 region 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 achieves successful segmentation in cases where there is low contrast between the lesion and the skin background near the lesion boundaries.

2.3.4. Segmentation based on artificial intelligence

Techniques based on artificial intelligence (AI) have also been proposed for the image segmentation of skin lesions, in which the image 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, in order to improve segmentation performance.

Artificial neural networks (ANNs) [84], which are parallel distributed systems composed of simple processing units with the purpose of obtaining similar results to the human brain, have been applied to segment images with skin lesions [33,69]. The segmentation performance of ANNs may be improved through the application of Genetic Algorithms (GAs) [85], which are computational techniques for searching and optimization. GAs are based on natural evolution and biological genetics, with the aim of finding the best solution for a given problem; for example, GAs may be employed to optimize ANN parameters.

Roberts and Claridge [11] presented a method to segment skin lesion images through Genetic Programming (GP) [86], 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, 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, which is computationally demanding.

Fuzzy logic deals with uncertain and imprecise values. Many algorithms based on fuzzy logic have been proposed to segment skin lesions in images [10,12,14,27,60,61]. 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. [10] and Silveira et al. [12] the fuzzy method, combined with both splitting and merging techniques, was used to segment dermoscopy images. This combination, originally proposed by Maeda et al. [87, 88], generates an algorithm for the 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 [77]; the idea was to determine whether a specific intensity belongs to the ROI or to the image background [14]. Another method, named neuro-fuzzy approach [60], 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 [27,61,71]. The basic idea behind the FCM algorithm is to find the centre of each cluster, similarly to the traditional k-means algorithm. Nevertheless, this process is more flexible, since partial membership may be introduced in the clusters. For each iteration of FCM, the minimization of the objective function F is computed as:

$$F = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{k} \left\| x_i - c_j \right\|^2, \tag{10}$$

$$\mu_{ij} = \frac{1}{\sum_{m=1}^{C} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_m \right\|^{2/(k-1)}} \right)}, \text{ and} \tag{11}$$

$$c_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{k} x_i}{\sum_{i=1}^{N} \mu_{ij}^{k}}, \tag{12}$$

where N is the number of pixels in the input image, C is the number of defined clusters, c_j is the centre of each cluster j, μ_{ij} is the degree of membership for the pixels x_i in cluster j, and k is a coefficient that defines the fuzziness of the resulting clusters. The term $\|x_i - c_j\|$ is used to measure the similarity of the pixels to the centre c_j of a given cluster j.

Figure 7 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 [47]. Two clusters were defined with the initial mean intensities of 8 and 250. Using these parameters, the resultant images demonstrates that the lesions are successfully segmented. However, some lesion pixels with low contrast are not clustered into the lesion groups.

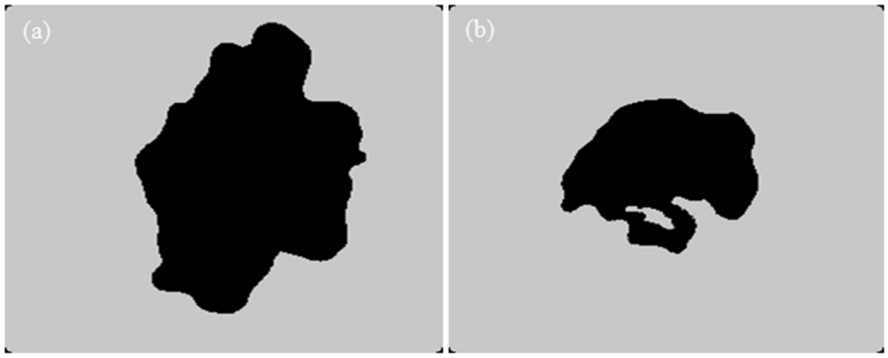

Figure 7: Segmentation results obtained after applying the fuzzy c-means method: (a) to the dermoscopy image shown in Figure 4a, and (b) to the macroscopic image shown in Figure 4b.

Zhou et al. [27] proposed a new mean shift approach, based on the FCM algorithm, called the anisotropic mean shift algorithm (AMSFCM), to segment dermoscopic images. The AMSFCM algorithm [89] 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 [90] 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. Castillejos et al. [71] proposed a cluster pre-selection algorithm based on the FCM algorithm (CPSFCM), in order to use fuzzy logic to automatically determine the optimal number of clusters, based on the input image data, such as the intensity values.

2.3.5. Segmentation based on active contours

Algorithms based on active contours have been used for segmenting skin lesion images [12,15,16,28,66]. 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 [91-93] or geometric [59,82,94-96], according to the technique used to track the curve movement.

Parametric models include the traditional active contour models, namely, snake models [92]. Typically, in these models, the curve deformation is guided by energy forces, in which an internal energy determines the smoothness level by the definition of the curve's elasticity and rigidity; in other words, it controls the degree of shrinkage or expansion of the model 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. Image-based energies may also be defined, which drive the curve toward interesting image features, such as those based on image intensity, gradient, line segments and corners. However, these models have some limitations [82,93]: 1) the curve initialization must be near the object's boundary, 2) the models have difficulty in dealing with boundaries with large curvatures, 3) the stop criterion of the curve deformation usually depends on the image gradient, which may cause bad edge localization when the gradient value is not high enough, and 4) these models have difficulty in dealing with topological changes during the curve evolution.

The gradient vector flow (GVF) [93] is another parametric model that has been used in the segmentation of skin lesions [12,15,65]. Xu and Prince [93] 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. [15,73,74] proposed a new type of dynamic energy for the segmentation of skin lesions, that combines the classical GVF model [93] and the mean shift algorithm [97]. 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, 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 method [94] and active contour model without edges, known as Chan-Vese's model [82], are such examples of geometric models. The level set method was originally proposed by Osher and Sethian [94] 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 [82] 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 [81] and Level Set [94] segmentation techniques. Essentially, Chan-Vese's model considers a "fitting" term F for the energy minimization, which is calculated by means of an energy functional based on the level set function, ϕ, to identify whether the object of interest is inside or outside the curve, C. The minimization of the energy function $F(c_1, c_2, \phi)$ allows the deformation of the curve toward the boundary of the object, where the inside and outside intensities are constant and similar:

$$F(c_1, c_2, \phi) = \mu \int_\Omega \delta\big(\phi(x,y)\big)|\nabla\phi(x,y)|dx\,dy + \nu \int_\Omega H\big(\phi(x,y)\big)dx\,dy + \lambda_1 \int_\Omega |u_0(x,y) - c_1|^2 H\big(\phi(x,y)\big)dx\,dy + \lambda_2 \int_\Omega |u_0(x,y) - c_2|^2 \Big(1 - H\big(\phi(x,y)\big)\Big)\,dx\,dy, \quad (13)$$

where u_0 is a pre-processed image, as a bounded function on $\overline{\Omega}$ and with real values. The fixed parameters μ, $\nu \geq 0$, λ_1 and $\lambda_2 > 0$ are weights for the fitting term. The terms H and δ are the Heaviside and Dirac delta functions, respectively. The constants c_1 and c_2, which are based on Mumford-Shah's segmentation model, are the average image u_0 inside and outside the curve C, respectively, and given by:

$$c_1(\phi) = \frac{\int_\Omega u_0(x,y)H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy}, \text{ and} \quad (14)$$

$$c_2(\phi) = \frac{\int_\Omega u_0(x,y)(1-H(\phi(x,y)))dxdy}{\int_\Omega (1-H(\phi(x,y)))dxdy}. \quad (15)$$

Chan-Vese's model has been used in the segmentation of skin lesions in images [12,16,67], due to its advantages when compared with other segmentation techniques based on the active contour model [82], 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.

Figure 8 presents the segmentation results obtained by applying the traditional Chan-Vese's model [82] to two images, which were previously smoothed using a median filter [47]. The segmentation process was halted 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 has provided good segmentation results, having identified low contrast boundaries and overcome the image noise.

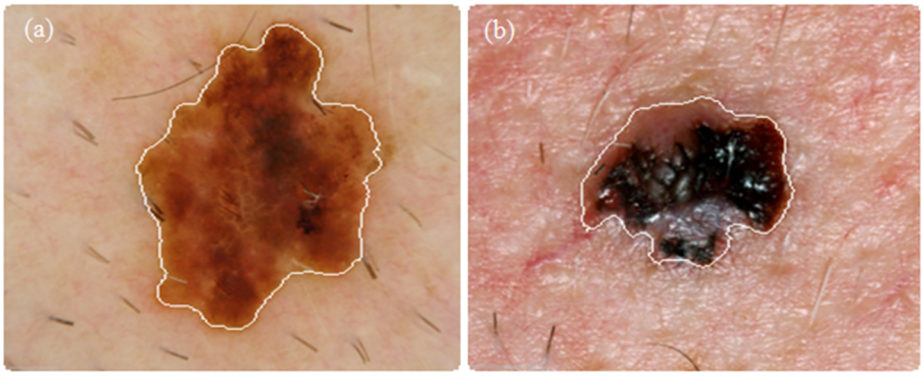

Figure 8: Segmentation results obtained after applying Chan-Vese's model: (a) to the dermoscopy image shown in Figure 4a, and (b) to the macroscopic image shown in Figure 4b.

Abbas et al. [28] proposed an improved, perceptually-oriented region-based active contour (RAC) scheme [98], where the segmentation concept is based on Chan-Vese's model [82] to determine the edges of the lesion to be segmented. The authors suggested this model due to its ability to simultaneously define multiple regions, separate

heterogeneous objects, successfully deal with image noise, and because of the automatic convergence of the modelled curve.

3. 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 [12,26,27,41,48,61]. 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 operation exclusive disjunction (XOR) [29,49,99]. The accuracy of the segmentation depends on the model and techniques used to solve the problem. Figure 9 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.

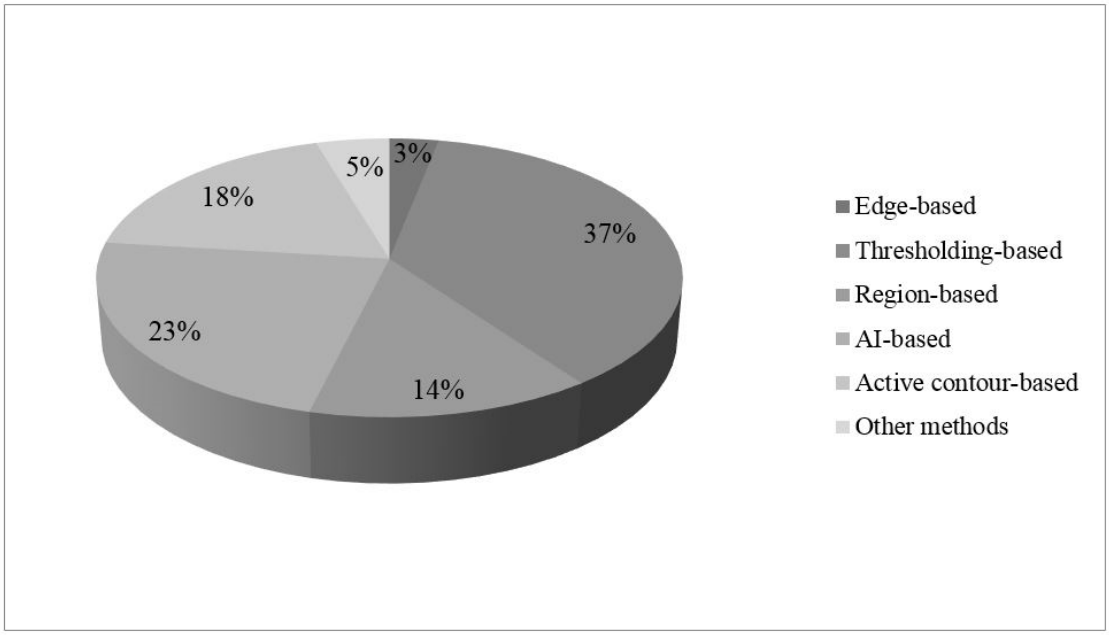

Figure 9: Distribution of the reviewed methods for the segmentation of skin lesions according to the applied principle.

Threshold-based techniques have been widely used, mainly because of their simplicity, computational efficiency and good performance. 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. Algorithms based on the active contour model have also 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. Region-based methods have also been used, since such methods have shown successful performance even in the presence of several obstacles, 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.

Clustering algorithms have also been applied to segment skin lesion images [44,71,72]. For example, the k-means clustering algorithm is used by Castillejos et al. [71]. The authors present a novel approach to segment the images based on the wavelet transform for k-means, FCM and CPSFCM algorithms. The proposed methods achieved superior results when compared with techniques that did not apply the wavelet transform. The hill-climbing algorithm (HCA) is a technique based on the clustering of points on an image, which is also applied to detect the ROIs of skin lesion images [29]. 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 relies on a simple, fast and non-parametric algorithm. In Abbas et al. [58,75], 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, with good edge-based segmentation capabilities, its ability to solve for local minima or overlapping problems, its computational efficiency, and its excellent performance in detecting lesion borders in dermoscopy images. The combination of different methods have also been adopted to improve the final result of the image segmentation process, such as finding the approximate location of lesion, and automatically defining the initial contours, mainly to be used with the active contour model [7,65,69].

Table 3 allows the performance comparison of the methods reviewed to segment both macroscopic and dermoscopy images of skin lesions, which are mostly performed

automatically. The segmentation results are compared against the ground-truth defined by one or more specialists, or their quality has been visually assessed. The table indicates the number and type of image used, the colour spaces and channels employed in the pre-processing and segmentation steps, and the values of the evaluation measures.

In order to obtain enhanced segmentation results, both from dermoscopy and macroscopic images, pre-processing methods, such as colour space transformation, illumination correction, contrast enhancement and artefact removal, have been used. The median filter [47] and anisotropic diffusion filter [51] are usually applied to smooth images, in order to reduce the noise. Nonetheless, these filters cannot deal with some obstacles, such as illumination variation and very thick dark hair. Algorithms based on hair detection and repair, for example, based on inpainting techniques, have been used for hair removal [43].These enhance the lesions, which can lead to important improvements, and, therefore, favorably affect the diagnosis.

With regards to the segmentation step, edge-based techniques are not suggested for segment skin lesions, since these techniques may produce segmentations with edges that are not completely closed. On the other hand, thresholding-, region-, and AI-based 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 [14,40], to define the threshold value, may sometimes achieve superior segmentation results. The region-based approach proposed by Wong et al. [13] has a better segmentation performance, even in the presence of boundaries with low contrast. In addition, such a method can tackle structural variations, varying illumination and colour variations. Other techniques have also been suggested to convert the FCM segmentation method into a more effective approach for segmenting skin lesions in images [27]. Using these methods, better segmentation results may be achieved, even in the presence of irregular lesions and image noise. Active contour models [82] are a good option for the segmentation of skin lesions, since these models can adequately deal with varying intensities, low contrast boundaries and noisy images. Nevertheless, these models also have disadvantages; for example, the segmentation result depends on the suitability of the initial curve.

Table 3: Comparison of the reviewed segmentation methods for skin lesions, both in macroscopic and dermoscopy images.

Ref.	Year	Image number (Type)	Pre-processing (Colour space)	Segmentation (Colour space)	Mean result (Evaluation measure)
[66]	2015	90 (DB1) / 160 (DB2) (Dermoscopy)	Smoothing (*RGB*)	Thresholding + Active contour (*CIE L*a*b** and *CIE L*u*v**)	DB1: 10.82% (XOR); DB2: 13.92% (XOR)
[73]	2013	100 (Dermoscopy)	-	Active contour (Grey-levels)	0.86 (SE), 0.99 (SP)
[67]	2013	152 (Macroscopic)	Illumination correction (*HSV*)	Thresholding + Active contour ($\bar{I}_2^N$ and $\bar{I}^L$ channels)	15.60% (XOR), 90.07% (SE), 99.11% (SP)
[24]	2013	90 (Dermoscopy)	-	Thresholding + fusion (Blue-channel - *RGB*)	8.31% (XOR)
[29]	2013	100 (Dermoscopy)	Contrast enhancement (*CIE L*a*b**)	Hill-climbing algorithm + thresholding (*CIE L*a*b**)	94.25% (TP), 3.56% (FP), 4% (EP)
[75]	2012	100 (Dermoscopy)	Artefact removal (*CIE L*a*b**)	Dynamic programming (*CIE L*a*b** and Grey-levels)	94.64% (SE), 98.14% (SP), 5.23% (EP)
[28]	2012	175 (Dermoscopy)	Illumination correction, contrast enhancement, hair removal (*JCh* and *CIECAM02*)	Active contour (*JCh* and *CIECAM02*)	Single contour initialization: 8.38% (EP); Multi-contour initialization: 4.10% (EP)
[26]	2012	426 (Dermoscopy)	Smoothing, Illumination correction (Grey-levels, *RGB*)	Thresholding (*RGB*)	NoMSLs: 84.5% (Prec.), 88.5% (Rec.); MSLs: 93.9% (Prec.), 93.8% (Rec.)
[71]	2012	50 (Dermoscopy)	-	AI-based (*RGB*)	Non-reported SE, SP, AUC
[40]	2012	100 (Macroscopic)	Mathematical morphology (*HSI*)	Thresholding (Grey-levels)	Ben.: 95.22% (FM), 4.79 (NRM); Mal.: 94.65% (FM), 5.56% (NRM)
[58]	2011	240 (Dermoscopy)	Artefact removal (*HSV*)	Dynamic programming (Grey-levels)	Ben. Mel.: 8.6% (EP); Melan.: 5.04% (EP); BCC: 9.0% (EP); MCC: 7.02% (EP); Seb. Kerat.: 2.01% (EP); Nevus: 3.24% (EP)
[16]	2011	320 (Dermoscopy)	Artefact removal (*RGB*)	Active contour (Grey-levels)	4.58% (EP)
[74]	2011	100 (Dermoscopy)	-	Active contour (Grey-levels)	0.81 (SE), 0.99 (SP)
[13]	2011	60 (Macroscopic)	-	Region-based (*RGB*)	9.16% (EP)
[33]	2011	100 (Dermoscopy)	Colour and contrast enhancement (*RGB*)	AI-based (*RGB* and Grey-levels)	RGB: 0.24, 0.16, 0.17 (XOR); Grey-levels: 0.16 (XOR)
[49]	2011	85 (Dermoscopy)	Hair removal, Smoothing, Contrast enhancement (*RGB* and Grey-levels)	Thresholding + AI-based (*XYZ*, *RGB* and Grey-levels)	89.64% (SE), 99.43% (SP)
[63]	2010	300 (Dermoscopy)	-	Thresholding (Grey-levels)	Visual
[15]	2010	100 Dermoscopy	-	Active contour (Grey-levels)	0.99 (SP), 0.81 (SE)

Table 3: Continued.

Ref.	Year	Image number (Type)	Pre-processing (Colour space)	Segmentation (Colour space)	Mean result (Evaluation measure)
[12]	2009	100 (Dermoscopy)	Mathematical morphology, smoothing (*HSV*)	Active contour (*CIEL*a*b**)	12.63% (HM), 95.47% (TDR), 36.90% (HD)
[34]	2009	367 (Dermoscopy)	Contrast enhancement	Thresholding (Grey-levels)	16.56% (XOR)
[14]	2009	Non-reported (Dermoscopy)	-	Thresholding + Fuzzy logic (Grey-levels)	Visual
[17]	2009	10 (Macroscopic)	Smoothing (*RGB*)	Edge-based (Grey-levels)	Visual
[27]	2009	100 (Dermoscopy)	-	AI-based (*RGB*)	0.78 (SE), 0.99 (SP)
[48]	2008	90 (Dermoscopy)	Black frame removal, smoothing (*HSL*)	Region-based (*RGB*)	Ground-truth 1: 11.10% (EP); Ground-truth 2: 10.27% (EP); Ground-truth 3: 10.53% (EP)
[7]	2008	319 (Dermoscopy)	Smoothing (*RGB*)	Thresholding + region-based (*RGB*)	94.1% (Prec.), 95.3% (Rec.)
[10]	2008	50 (Dermoscopy)	-	AI-based + region-based (*CIEL*a*b**)	Non-reported
[65]	2005	100 (Dermoscopy)	Smoothing (Grey-levels)	Thresholding + Active contour (Grey-levels)	Ben.: 13.77% (EP); Melan.: 19.76% (EP)
[100]	2004	Non-reported (Dermoscopy)	Mathematical morphology Smoothing (Grey-levels)	Thresholding + AI-based (*RGB*)	Non-reported
[11]	2003	100 (Dermoscopy)	-	AI-based (Non-reported)	97% (SE), 81% (SP)

Ref.: reference; Prec.: precision; Rec.: recall; Ben.: benign; Mal.: malignant; Ben. Mel.: benign melanocytic; Melan.: melanoma; Seb. Kerat.: seborrheic keratosis; DB: database; XOR: exclusive disjunction; SE: sensitivity; SP: specificity; EP: error probability; TP: true positive rate; FP: false positive rate; AUC: area under an ROC curve; ROC: receiver operating characteristics; FM: F-measure; NRM: negative rate metric; NoMSLs: non-melanocytic skin lesions; MSLs: melanocytic skin lesions; BCC: basal cell carcinoma; MCC: Merkel cell carcinoma; HM: Hammoude distance; TDR: true detection rate; HD: Hausdorff distance.

4. Conclusions

Image segmentation is an important step for the effective computational diagnosis of pigmented skin lesions in images. Skin lesion diagnosis is an area of increased interest, due both to the importance of prevention and to early diagnosis of skin cancer. Although the image segmentation of skin lesions has been addressed in several studies and successful applications, there is the potential to develop new methodologies and to improve the performance of existing methods. Here, we have presented a review about current methods that have been proposed to segment skin lesions. Additionally, we have introduced techniques used to acquire and pre-process images, with a focus on their subsequent segmentation.

From the presented review, one may conclude that 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 more adequate lesion segmentation and analysis. Nevertheless, techniques to remove or reduce the artefacts are usually necessary to obtain robust segmentation results.

The reviewed segmentation techniques were classified into: edge-, thresholding-, region-, AI- and active contour-based, and others categories. We have presented and discussed results obtained with some of these techniques applied to dermoscopy and macroscopic images of skin lesions. Active contour models can provide good results on images with colour variation and low contrast of the lesion boundaries. Therefore, such models are 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 detections.

In conclusion, the future trends regarding the image segmentation of skin lesions are 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.

Acknowledgements

The first author would like to thank the “Conselho Nacional de Desenvolvimento Científico e Tecnológico” (CNPq), in Brazil, for her PhD grant. This work is funded by the European Regional Development Funds (ERDF), through the Operational Programme ‘Thematic Factors of Competitiveness’ (COMPETE), and Portuguese Funds, through the “Fundação para a Ciência e a Tecnologia” (FCT), under the project: FCOMP-01-0124-FEDER 028160/PTDC/BBB- BMD/3088/2012.

References

[1] British Association of Dermatologists (2014). *Skin Cancer*. British Association of Dermatologists. http://www.bad.org.uk/for-the-public/skin-cancer. Accessed October 2015.

[2] Bourne, P.; Cameron, A.; Gourhant, J.-Y.; Hackett, T.; Hlaing, W.; Kittler, H.; McColl, I.; Minas, S.; Rosendahl, C. (2007). *The International Atlas of Dermoscopy and Dermatoscopy* http://www.dermoscopyatlas.com/index.cfm. Accessed October 2015.

[3] Cancer Research UK (2013). *Cancer Statistic Report: Skin Cancer*. Cancer Research UK. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/skin/?script=true. Accessed October 2015.

[4] American Cancer Society (2014). *Cancer Facts & Figures 2014*. American Cancer Society, Atlanta.

[5] INCA (2014). *Estimativa 2014: Incidência de Câncer no Brasil*. Instituto Nacional de Câncer José Alencar Gomes da Silva, Coordenação de Prevenção e Vigilância. INCA, Rio de Janeiro.

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

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

[8] Scharcanski, J.; Celebi, M. E. (2013). *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, Berlin, Heidelberg.

[9] Celebi, M. E.; Mendonca, T.; Marques, J. S. (2015). *Dermoscopy Image Analysis*, vol 10. CRC Press, Boca Raton.

[10] Maeda, J.; Kawano, A.; Yamauchi, S.; Suzuki, Y.; Marcal, A. R. S.; Mendonca, T. (2008). *Perceptual image segmentation using fuzzy-based hierarchical algorithm and its application to dermoscopy images*. In: Conference on Soft Computing in Industrial Applications, June 25-27 2008. IEEE, pp 66-71.

[11] Roberts, M. E.; Claridge, E. (2003). *An Artificially Evolved Vision System for Segmenting Skin Lesion Images*. In: Ellis, R. E., Peters, T. M. (eds) Medical Image Computing and Computer-Assisted Intervention, vol 2878. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 655-662.

[12] Silveira, M.; Nascimento, J. C.; Marques, J. S.; Marcal, A. R. S.; Mendonca, T.; Yamauchi, S.; Maeda, J.; Rozeira, J. (2009). Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):35-45.

[13] Wong, A.; Scharcanski, J.; Fieguth, P. (2011). Automatic Skin Lesion Segmentation via Iterative Stochastic Region Merging. *IEEE Transactions on Information Technology in Biomedicine*, 15 (6):929-936.

[14] Yuksel, M. E.; Borlu, M. (2009). Accurate Segmentation of Dermoscopic Images by Image Thresholding Based on Type-2 Fuzzy Logic. *IEEE Transactions on Fuzzy Systems*, 17 (4):976-982.

[15] Zhou, H.; Schaefer, G.; Celebi, M. E.; Iyatomi, H.; Norton, K.; Liu, T.; Lin, F. (2010). *Skin lesion segmentation using an improved snake model*. In: Annual

International Conference of the Engineering in Medicine and Biology Society, Buenos Aires, August 31 - September 4 2010. IEEE pp 1974-1977.

[16] Abbas, Q.; Fondón, I.; Rashid, M. (2011). Unsupervised skin lesions border detection via two-dimensional image analysis. *Computer Methods and Programs in Biomedicine*, 104 (3):e1-e15.

[17] Barcelos, C. A. Z.; Pires, V. B. (2009). An automatic based nonlinear diffusion equations scheme for skin lesion segmentation. *Applied Mathematics and Computation*, 215 (1):251-261.

[18] Celebi, M. E.; Iyatomi, H.; Schaefer, G.; Stoecker, W. V. (2009). Lesion border detection in dermoscopy images. *Computerized Medical Imaging and Graphics*, 33 (2):148-153.

[19] Ruocco, E.; Argenziano, G.; Pellacani, G.; Seidenari, S. (2004). Noninvasive Imaging of Skin Tumors. *Dermatologic Surgery*, 30:301-310.

[20] Wang, S. Q.; Rabinovitz, H.; Kopf, A. W.; Oliviero, M. (2004). Current technologies for the in vivo diagnosis of cutaneous melanomas. *Clinics in Dermatology*, 22 (3):217-222.

[21] Smith, L.; MacNeil, S. (2011). State of the art in non-invasive imaging of cutaneous melanoma. *Skin Research and Technology*, 17 (3):257-269.

[22] Cavalcanti, P. G.; Scharcanski, J. (2013). *Macroscopic Pigmented Skin Lesion Segmentation and Its Influence on Lesion Classification and Diagnosis*. In: Celebi, M. E., Schaefer, G. (eds) Color Medical Image Analysis. Springer, Dordrecht, pp 15-39.

[23] Alcón, J. F.; Ciuhu, C.; Ten Kate, W.; Heinrich, A.; Uzunbajakava, N.; Krekels, G.; Siem, D.; de Haan, G. (2009). Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):14-25.

[24] Celebi, M. E.; Wen, Q.; Hwang, S.; Iyatomi, H.; Schaefer, G. (2013). Lesion border detection in dermoscopy images using ensembles of thresholding methods. *Skin Research and Technology*, 19 (1):e252-e258.

[25] Garnavi, R.; Aldeen, M.; Bailey, J. (2012). Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 16 (6):1239-1252.

[26] Norton, K.-A.; Iyatomi, H.; Celebi, M. E.; Ishizaki, S.; Sawada, M.; Suzaki, R.; Kobayashi, K.; Tanaka, M.; Ogawa, K. (2012). Three-phase general border detection method for dermoscopy images using non-uniform illumination correction. *Skin Research and Technology*, 18 (3):290-300.

[27] Zhou, H.; Schaefer, G.; Sadka, A. H.; Celebi, M. E. (2009). Anisotropic Mean Shift Based Fuzzy C-Means Segmentation of Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):26-34.

[28] Abbas, Q.; Celebi, M. E.; Garcia, I. F. (2012). A novel perceptually-oriented approach for skin tumor segmentation. *International Journal of Innovative Computing, Information and Control*, 8 (3):1837-1848.

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

[30] Celebi, M. E.; Aslandogan, Y. A.; Stoecker, W. V.; Iyatomi, H.; Oka, H.; Chen, X. (2007). Unsupervised border detection in dermoscopy images. *Skin Research and Technology*, 13 (4):454-462.

[31] Cavalcanti, P. G.; Scharcanski, J.; Lopes, C. B. O. (2010). *Shading attenuation in human skin color images*. In: 6th International Symposium on Visual Computing, Las Vegas, November 29 - December 1 2010. Springer, pp 190-198.

[32] Glaister, J.; Amelard, R.; Wong, A.; Clausi, D. (2013). MSIM: Multistage illumination modeling of dermatological photographs for illumination-corrected skin lesion analysis. *IEEE Transactions on Biomedical Engineering*, 60 (7):1873-1883.

[33] Schaefer, G.; Rajab, M. I.; Celebi, M. E.; Iyatomi, H. (2011). Colour and contrast enhancement for improved skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 35 (2):99-104.

[34] Celebi, M. E.; Iyatomi, H.; Schaefer, G. (2009). *Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure*. In: 16th IEEE International Conference on Image Processing, Cairo, November 7-10 2009. IEEE, pp 2601-2604.

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

[36] Barata, C.; Marques, J. S.; Celebi, M. E. (2014). *Improving dermoscopy image analysis using color constancy*. In: IEEE International Conference on Image Processing (ICIP), Paris, France, October 27-30 2014. IEEE, pp 3527-3531.

[37] Celebi, M. E.; Kingravi, H. A.; Aslandogan, Y. A. (2007). Nonlinear vector filtering for impulsive noise removal from color images. *Journal of Electronic Imaging*, 16 (3):033008-033021.

[38] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9 (1):62-66.

[39] Sonka, M.; Hlavac, V.; Boyle, R. (1998). *Image processing, analysis, and machine vision*. 2 edn. PWS, USA.

[40] Beuren, A. T.; Janasieivicz, R.; Pinheiro, G.; Grando, N.; Facon, J. (2012). *Skin melanoma segmentation by morphological approach*. In: International Conference on Advances in Computing, Communications and Informatics, Chennai, August 3-5 2012. ACM, pp 972-978.

[41] Norton, K.; Iyatomi, H.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Development of a novel border detection method for melanocytic and non-melanocytic dermoscopy images*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5403-5406.

[42] Lee, T.; Ng, V.; Gallagher, R.; Coldman, A.; McLean, D. (1997). Dullrazor®: A software approach to hair removal from images. *Computers in Biology and Medicine*, 27 (6):533-543.

[43] Abbas, Q.; Celebi, M. E.; García, I. F. (2011). Hair removal methods: a comparative study for dermoscopy images. *Biomedical Signal Processing and Control*, 6 (4):395-404.

[44] Xie, F.-Y.; Qin, S.-Y.; Jiang, Z.-G.; Meng, R.-S. (2009). PDE-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma. *Computerized Medical Imaging and Graphics*, 33 (4):275-282.

[45] Zhou, H.; Chen, M.; Gass, R.; Rehg, J. M.; Ferris, L.; Ho, J.; Drogowski, L. (2008). *Feature-preserving artifact removal from dermoscopy images*. In: SPIE Medical Imaging 2008 Conference, San Diego, February 16 - 21 2008. International Society for Optics and Photonics, pp 69141B69141-69141B69149.

[46] Mirzaalian, H.; Lee, T. K.; Hamarneh, G. (2014). Hair Enhancement in Dermoscopic Images Using Dual-Channel Quaternion Tubularness Filters and MRF-Based Multilabel Optimization. *IEEE Transactions on Image Processing*, 23 (12):5486-5496.

[47] Pratt, W. K. (2001). *Digital image processing*. 3 edn. John Wiley & Sons, New York.

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

[49] Garnavi, R.; Aldeen, M.; Celebi, M. E.; Varigos, G.; Finch, S. (2011). Border detection in dermoscopy images using hybrid thresholding on optimized color channels. *Computerized Medical Imaging and Graphics*, 35 (2):105-115.

[50] Perona, P.; Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12 (7):629-639.

[51] Barcelos, C. A. Z.; Boaventura, M.; Silva Junior, E. C. (2003). A well-balanced flow equation for noise removal and edge detection. *IEEE Transactions on Image Processing*, 12 (7):751-763.

[52] Alvarez, L.; Lions, P.-L.; Morel, J.-M. (1992). Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29 (3):845-866.

[53] Nordström, N. (1990). *Biased anisotropic diffusion: A unified regularization and diffusion approach to edge detection*. In: Faugeras, O. (ed) Computer Vision, vol 427. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 18-27.

[54] Adelmann, H. G. (1998). Butterworth equations for homomorphic filtering of images. *Computers in Biology and Medicine*, 28 (2):169-181.

[55] Qin, L.; Lei, Z.; You, J.; Zhang, D.; Bhattacharya, P. (2008). *Dark line detection with line width extraction*. In: 15th International Conference on Image Processing San Diego, California, October 12-15 2008. IEEE pp 621-624.

[56] Criminisi, A.; Perez, P.; Toyama, K. (2004). Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 13 (9):1200-1212.

[57] Gonzalez, R. C.; Woods, R. E. (2002). *Digital Image Processing*. 2 edn. Prentice Hall, New Jersey.

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

[59] Chan, T. F.; Sandberg, B. Y.; Vese, L. A. (2000). Active Contours without Edges for Vector-Valued Images. *Journal of Visual Communication and Image Representation*, 11 (2):130-141.

[60] Castiello, C.; Castellano, G.; Fanelli, A. M. (2004). *Neuro-fuzzy analysis of dermatological images*. In: IEEE International Joint Conference on Neural Networks, Budapest, Hungary, July 25-29 2004. pp 3247-3252.

[61] Rahman, M. M.; Bhattacharya, P.; Desai, B. C. (2008). *A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions*. In: 8th IEEE International Conference on International Conference on BioInformatics and BioEngineering, Athens, October 8-10 2008. IEEE, pp 1-6.

[62] Ganzeli, H. S.; Bottesini, J. G.; Paz, L. O.; Ribeiro, M. F. S. (2011). SKAN: Skin Scanner - System for Skin Cancer Detection Using Adaptive Techniques. *IEEE Latin America Transactions*, 9 (2):206-212.

[63] Leo, G. D.; Paolillo, A.; Sommella, P.; Fabbrocini, G. (2010). *Automatic Diagnosis of Melanoma: A Software System Based on the 7-Point Check-List*. In: 43rd International Conference on System Sciences, Hawaii January 5-8 2010. IEEE, pp 1-10.

[64] Cavalcanti, P. G.; Scharcanski, J. (2011). Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics*, 35 (6):481-491.

[65] Erkol, B.; Moss, R. H.; Joe Stanley, R.; Stoecker, W. V.; Hvatum, E. (2005). Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. *Skin Research and Technology*, 11 (1):17-26.

[66] Ma, Z.; Tavares, J. M. R. S. (2016). A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model. *IEEE Journal of Biomedical and Health Informatics*, 20 (2):615-623.

[67] Cavalcanti, P. G.; Scharcanski, J. (2013). A coarse-to-fine approach for segmenting melanocytic skin lesions in standard camera images. *Computer Methods and Programs in Biomedicine*, 112 (3):684-693.

[68] Barata, C.; Ruela, M.; Francisco, M.; Mendonça, T.; Marques, J. S. (2013). Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features. *IEEE Systems Journal*, 8 (3):965-979.

[69] Rajab, M. I.; Woolfson, M. S.; Morgan, S. P. (2004). Application of region-based segmentation and neural network edge detection to skin lesions. *Computerized Medical Imaging and Graphics*, 28 (1):61-68.

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

[71] Castillejos, H.; Ponomaryov, V.; Nino-de-Rivera, L.; Golikov, V. (2012). Wavelet transform fuzzy algorithms for dermoscopic image segmentation. *Computational and mathematical methods in medicine*, 2012:1-11.

[72] Gómez, D. D.; Butakoff, C.; Ersbøll, B. K.; Stoecker, W. (2008). Independent histogram pursuit for segmentation of skin lesions. *IEEE Transactions on Biomedical Engineering*, 55 (1):157-161.

[73] Zhou, H.; Li, X.; Schaefer, G.; Celebi, M. E.; Miller, P. (2013). Mean shift based gradient vector flow for image segmentation. *Computer Vision and Image Understanding*, 117 (9):1004-1016.

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

[75] Abbas, Q.; Celebi, M. E.; García, I. F. (2012). Skin tumor area extraction using an improved dynamic programming approach. *Skin Research and Technology*, 18 (2):133-142.

[76] Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (6):679-698.

[77] Mendel, J. M.; John, R. I. B. (2002). Type-2 fuzzy sets made simple. *IEEE Transactions on Fuzzy Systems*, 10 (2):117-127.

[78] Sahoo, P.; Wilkins, C.; Yeager, J. (1997). Threshold selection using Renyi's entropy. *Pattern Recognition*, 30 (1):71-84.

[79] Brice, C. R.; Fennema, C. L. (1970). Scene analysis using regions. *Artificial intelligence*, 1 (3–4):205-226.

[80] Muerle, J. L.; Allen, D. C. (1968). Experimental evaluation of techniques for automatic segmentation of objects in a complex scene. *Pictorial Pattern Recognition*, 1:3-13.

[81] Mumford, D.; Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42 (5):577-685.

[82] Chan, T. F.; Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10 (2):266-277.

[83] Nock, R.; Nielsen, F. (2004). Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 (11):1452-1458.

[84] Haykin, S. S. (1999). *Neural networks: a comprehensive foundation*. Prentice Hall, Englewood Cliffs, USA.

[85] Haupt, R. L.; Haupt, S. E. (2004). *Practical Genetic Algorithms*. 2 edn. John Wiley & Sons, New Jersey.

[86] Koza, J. R. (1992). *Genetic programming: on the programming of computers by means of natural selection*. MIT Press, Cambridge.

[87] Maeda, J.; Kawano, A.; Saga, S.; Suzuki, Y. (2007). *Unsupervised perceptual segmentation of natural color images using fuzzy-based hierarchical algorithm*. In: Image Analysis. Springer, pp 462-471.

[88] Maeda, J.; Kawano, A.; Saga, S.; Suzuki, Y. (2007). *Number-driven perceptual segmentation of natural color images for easy decision of optimal result*. In: International Conference on Image Processing, San Antonio, USA, September 16 - October 19 2007. IEEE, pp II-265-II-268.

[89] Wang, J.; Thiesson, B.; Xu, Y.; Cohen, M. (2004). *Image and video segmentation by anisotropic kernel mean shift*. In: Computer Vision: ECCV 2004. Springer, pp 238-249.

[90] Comaniciu, D.; Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (5):603-619.

[91] Caselles, V.; Catté, F.; Coll, T.; Dibos, F. (1993). A geometric model for active contours in image processing. *Numerische Mathematik*, 66 (1):1-31.

[92] Kass, M.; Witkin, A.; Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1 (4):321-331.

[93] Xu, C.; Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7 (3):359-369.

[94] Osher, S.; Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79 (1):12-49.

[95] Vese, L.; Chan, T. F. (2002). A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *International Journal of Computer Vision*, 50 (3):271-293.

[96] Zhao, H.-K.; Chan, T.; Merriman, B.; Osher, S. (1996). A Variational Level Set Approach to Multiphase Motion. *Journal of Computational Physics*, 127 (1):179-195.

[97] Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 (8):790-799.

[98] Lankton, S.; Tannenbaum, A. (2008). Localizing Region-Based Active Contours. *IEEE Transactions on Image Processing*, 17 (11):2029-2039.

[99] Celebi, M. E.; Schaefer, G.; Iyatomi, H.; Stoecker, W. V.; Malters, J. M.; Grichnik, J. M. (2009). An improved objective evaluation measure for border detection in dermoscopy images. *Skin Research and Technology*, 15 (4):444-450.

[100] Rajab, M.; Woolfson, M.; Morgan, S. (2004). Application of region-based segmentation and neural network edge detection to skin lesions. *Computerized Medical Imaging and Graphics*, 28 (1):61-68.

PART B - ARTICLE 2

COMPUTATIONAL METHODS FOR PIGMENTED SKIN LESION CLASSIFICATION IN IMAGES: REVIEW AND FUTURE TRENDS

Roberta B. Oliveira, João P. Papa, Aledir S. Pereira and João Manuel R. S. Tavares

Published in: Neural Computing and Applications, 27:1-24, 2016

Abstract

Skin cancer is considered as one of the most common types of cancer in several countries and its incidence rate has increased in recent years. Melanoma cases have caused an increasing number of deaths worldwide, since this type of skin cancer is the most aggressive compared to other types. Computational methods have been developed to assist dermatologists in early diagnosis of skin cancer. An overview of the main and current computational methods that have been proposed for pattern analysis and pigmented skin lesion classification is addressed in this review. In addition, a discussion about the application of such methods, as well as future trends are also provided. Several methods for feature extraction from both macroscopic and dermoscopic images and models for feature selection are introduced and discussed. Furthermore, classification algorithms and evaluation procedures are described, and performance results for lesion classification and pattern analysis are given.

Keywords: Pattern analysis; Feature extraction and selection; Classification methods; Macroscopic and dermoscopic images.

1. Introduction

Computational methods for skin cancer diagnosis have been proposed in order to aid dermatologists in early assessment of skin cancer and in the follow-up of pigmented skin lesions [1-3]. Such lesions represent an abnormal production of melanocytes cells, which are mainly caused by excessive sun exposure. Melanocytes cells are responsible for creating the substance melanin, one of the functions of which is to provide pigmentation in the skin. Furthermore, the number of skin cancer cases has increased in the last years, and consequently, an increasing number of deaths caused by this disease has been reported, particularly due to melanoma cases (Figures 1c and d) [4-6]. Therefore, pigmented skin lesions have been a cause for global concern, since some types of benign lesions may become skin cancer, such as dysplastic nevi (Figures 1a and b).

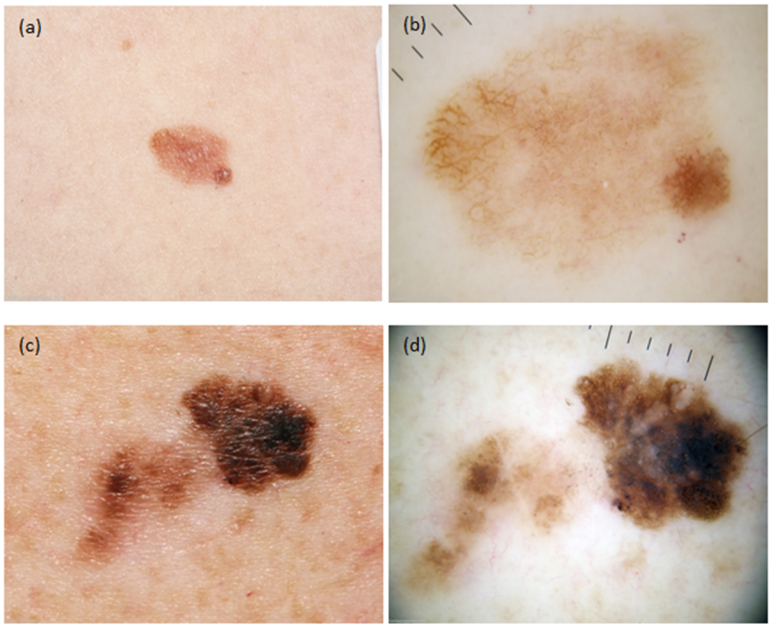

Figure 1: Two examples of macroscopic images (a and c) and dermoscopic images (b and d): (a) and (b) are images of a dysplastic nevus and, (c) and (d) are of an invasive melanoma (images publicly available from Bourne et al. [7]).

Image acquisition, pre-processing, segmentation, feature extraction, and classification are fundamental steps commonly found in computational systems for diagnosing skin lesions. Different non-invasive imaging techniques have been used to assist dermatologists [8]. Macroscopic images [9,10] and dermoscopic images [11,12] are examples of images acquired from such techniques that have been widely used in the diagnosis of pigmented skin lesions by computational methods. Macroscopic images (Figures 1a and c), commonly known as clinical images, are usually acquired from standard cameras or mobile devices. On the other hand, dermoscopic images (Figures 1b and d), may be acquired from dermatoscope devices or specific cameras in order to better visualize the pigmentation pattern on the skin surface. However, their imaging conditions are frequently inconsistent; for example, macroscopic images can be acquired from variable distances and/or under different illumination conditions. Furthermore, the images may have poor resolution, which may be challenging when the lesion under study is small. An additional problem with both macroscopy and dermoscopic images is related to the presence of artefacts, such as hair, reflections, shadows, skin lines and bubbles, which may hinder adequate analysis of the imaged skin lesions.

The identification of the regions of the lesions in such images may be performed in order to assist in the process of classification [13]. Segmentation is an important step that allows the extraction of such regions of interest (ROI) from an image [14-34]. However,

before the segmentation step, previous pre-processing methods are usually applied to reduce the effects of undesirable artefacts that may influence the outcome of the segmentation step. These methods can be based on colour space transformation [20,26,35], illumination correction [36,37], contrast enhancement [20,22,23,26,38-42], artefact removal [14,20,21,43,44] and approximate lesion localization [45]. In addition, hair removal methods are also used in pre-processing steps, since this artefact may considerably affect the detection of lesion borders [46-53]. Lee et al. [54] proposed a solution for hair removal, especially thick dark hairs, which is based on one of the first widely adopted methods for hair removal in dermoscopy images, and consists of identifying the hair location, replacing the values of the detected hair pixels in the original image by the values of the corresponding nearby non-hair pixels, and smoothing the thin lines. An overview of lesion border detection methods, including the pre-processing, segmentation and post-processing steps, is presented in Celebi et al. [55,56]. In addition, the authors also discuss performance evaluation issues and propose guidelines for future studies.

Computational methods for pigmented skin lesion classification are usually based on the features of the pixels within the segmented ROIs. Therefore, 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, 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 redundant and/or irrelevant features that should be removed from the feature set. Techniques to reduce the dimensionality of the data may be used to solve these problems according to one of the following reduction strategies: feature transformation (also known as feature extraction in literature concerning pattern recognition [57,58]), and feature selection [59].

The feature extraction strategy allows the modification of all the data of the image, in order to emphasize the most effective features, ensuring the correct separation of the classification classes [57]. 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. Contrary, new features are not created in the feature

selection strategy, meaning that a subset from the original features is defined when using this approach. Both strategies 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 algorithms can provide an automatic reduction of such excessive features. Furthermore, a larger feature space may include redundant or irrelevant data [60].

Several solutions [61-64] have been proposed for feature extraction and selection of pigmented skin lesions, in order to represent them according to a certain clinical criteria [65-67]. Such features may be used for the classification process, in order to provide dermatologists with a computer-aided diagnosis of pigmented skin lesions [2,12]. In this review, some of the most relevant solutions that have been developed to assist the skin lesion diagnosis from macroscopic and dermoscopic images are introduced, including those concerning the steps of feature extraction and selection, and image classification. Hence, this review is highly valuable for those wishing the design and/or implementation of competent expert systems for the automated classification of skin lesions in images.

This paper is organized as follows: a review of the main computational methods that have been applied to extract and select features from macroscopic and dermoscopic images of pigmented skin lesions is presented in Section 2. The main focus of that section is on the feature extraction step according to several clinical criteria. In addition, the feature selection process is addressed. The current state-of-the-art concerning the pigmented skin lesion classification, including the advantages and disadvantages of the reviewed methods, evaluation measures, and performance results for pattern and lesion classification, is presented in Section 3. Finally, conclusions and future trends about the computational methods of pigmented skin lesion classification are pointed out in the last section.

2. Image analysis of pigmented skin lesions

Computational methods regarding the feature extraction have been commonly developed based on the ABCD(E) rule, pattern analysis, seven-point checklist and Menzies' method, which are examples of clinical approaches used for the diagnosis of skin cancer from images [67-69]. The first approach can be used to extract features from both macroscopic and dermoscopic images, whereas the other approaches are usually applied to dermoscopic images in order to identify more detailed pattern features on the surfaces of

the lesions. The feature analysis based on these approaches, as well as the feature selection and extraction steps are presented with details in the following sections.

2.1. Feature analysis based on clinical approaches

The ABCD(E) rule is based on asymmetry, border, colour, diameter (or differential structures in the case of dermoscopic images), and evolution (or elevation) features, according to the criteria presented in Table 1. Such rule has been widely used for the feature extraction and automatic diagnosis of pigmented skin lesions [10,70].

Table 1: Criteria of the ABCD(E) rule for the diagnosis of skin cancer from clinical and dermoscopy analysis.

Feature	Clinical analysis		Feature	Dermoscopy analysis [a]		
	Benign lesion	Malignant lesion		Definition	Score	Weight factor
Asymmetry (A)	Shape is symmetric	Shape is asymmetric	Asymmetry (A)	Border, colours or structures are asymmetric in 0, 1, or 2 perpendicular axes	0-2	1.3
Border (B)	Border is regular or well-defined	Border is irregular or ill-defined	Border (B)	Abrupt cut-off of network at the border in 0-8 segments	0-8	0.1
Colour (C)	Colours are uniform	Colours are non-uniform	Colour (C)	Presence of six possible basic colours [b]	1-6	0.5
Diameter (D)	Size <6 mm	Size ≥ 6 mm	Differential structural (D)	Presence of five differential structural components [c]	1-5	0.5
Evolution (E)	No change	Changes in size, shape or shades of colour features				
Elevation (E)	Smooth surface	High surface				

[a] Total dermatoscopy 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.
[b] White, red, light-brown, dark-brown, blue-grey, and black.
[c] Network, structureless areas, branched streaks, dots, and globules.

The feature extraction based on pattern analysis has also been used for the pigmented skin lesion automatic diagnosis [71-74]. This approach assists in diagnosis by determining the presence of specific patterns visible in dermoscopic images, which may be divided into global and local patterns [75], as detailed in Table 2. Global patterns are represented by textured structures present in most of the lesions. Some examples of such patterns are illustrated in Figure 2. Local patterns are dermoscopic structures. Such patterns may be present or absent, as well as presenting irregular/regular or atypical/typical structures, as indicated in Table 2, which may define the type of lesion or whether it is benign or malignant. Examples of such patterns are illustrated in Figure 3.

Table 2: Pattern analysis in dermoscopic images.

Global pattern	Local pattern
Reticular	Pigmented network (present or absent/ typical or atypical)
Globular	Dots/globules (present or absent/ regular or irregular)
Cobblestone	Streaks (present or absent/ regular or irregular)
Homogeneous	Blue-whitish veil (present or absent)
Starburst	Blotches or pigmentation (present or absent/ regular or irregular)
Parallel	Hypopigmentation (present or absent)
Multicomponent (combination of three or more global patterns)	Regression structures (present or absent)
Non-specific (absent patterns)	Vascular structures (present or absent)

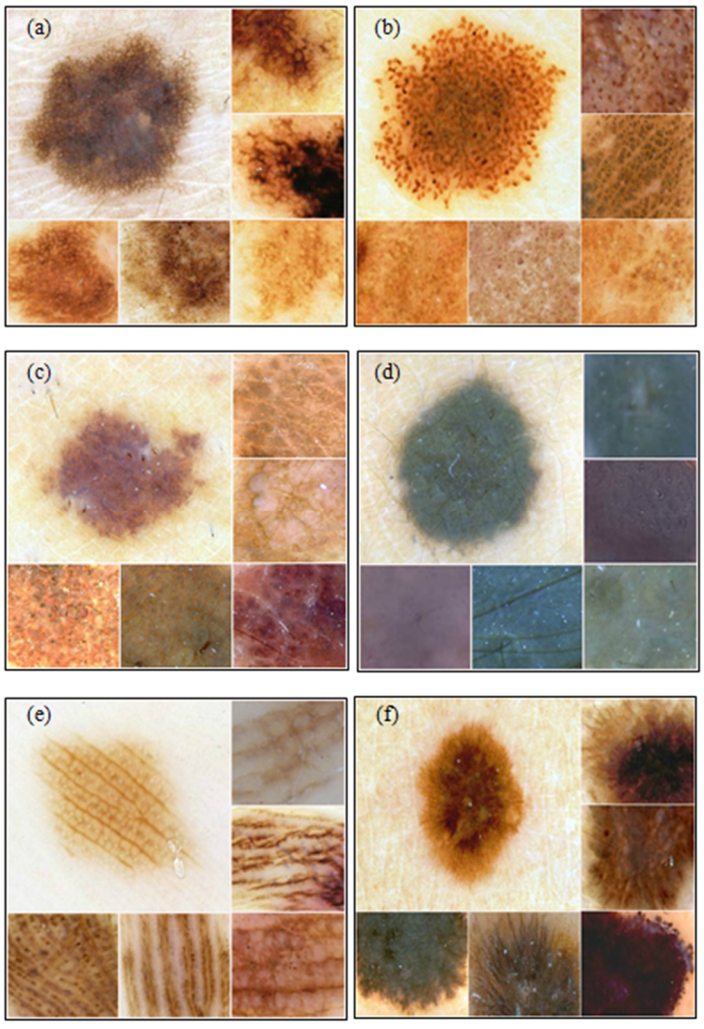

Figure 2: Examples of global patterns in dermatoscopy images: (a) reticular, (b) globular, (c) cobblestone, (d) homogeneous, (e) parallel and (f) starburst (images available in Argenziano et al. [76]).

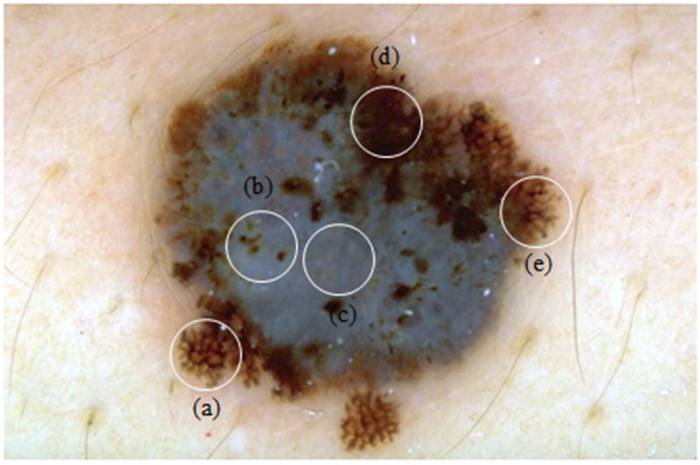

Figure 3: Examples of local patterns in a dermatoscopy image: (a) atypical pigmented network, (b) irregular dots/globules, (c) blue-whitish veil, (d) irregular pigmentation and (e) irregular streaks (adapted from Celebi et al. [77]).

The pattern analysis consists of examining the size, uniformity and distribution of the above-mentioned 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 [75]. Due to the low number of criteria to be analysed, the seven-point checklist and Menzies' method were introduced for skin lesion diagnosis from dermoscopic images in order to simplify the common pattern analysis [67]. The criteria of both clinical approaches are detailed in Table 3.

Table 3: Diagnostic criteria included the seven-point checklist and Menzies' method.

Seven-point checklist [a]		Menzies' method [d]		
Major criteria [b]	Minor criteria [c]	Colour of lesion	Symmetry of pattern	Positive feature
Atypical pigmented network	Irregular streaks	One colour	Symmetrical pattern	Blue-whitish veil
Blue-whitish veil	Irregular pigmentation	More than one colour	Asymmetrical pattern	Multiple brown dots
Atypical vascular pattern	Irregular dots/globules			Pseudopods
	Regression structures			Radial streaming
				Scarlike depigmentation
				Peripheral black dots/globules
				Multiple colours (5 or 6)
				Multiple blue/grey dots
				Broad pigment network

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

The seven-point checklist has been applied in the literature to achieve better accuracy for the computational diagnosis of dermoscopic images [24,78,79]. 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 [75]. 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 [67]. The 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 number of positive features [67]. 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 [67]. Computational methods based on the Menzies' criteria have been proposed to analyse the presence of six basic colour classes (white, red, light-brown, dark-brown, blue grey and black) for dermoscopic images [80,81].

2.2. Feature extraction

Skin lesion features can be extracted either according to a global or local manner in order to obtain information for classification. The most of works explore the global-features of the lesion, i.e., extract features from all segmented region [82]. However, some studies have used local-features, which allow the characterization of different region of the lesion. Bag-of-feature (BoF) approach is a simple strategy that has been used to compute local features [11,71,83-85]. In general, skin lesion features are categorized into shape features, colour variation and/or texture analysis [9,86]. These features can be extracted to detect patterns [73], or diagnose skin lesions [82] from both macroscopic and dermoscopic images. Extracted features of pigmented skin lesions from both of images are summarized in Table 4 and discussed in the following sections.

2.2.1. Shape features

Shape features allow the assessment of lesion's asymmetry or border's irregularity. The asymmetry features 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. 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 [87]. In

some studies, the axis of symmetry is defined based on the principal axis of inertia [87], major and minor axis orientation [12,88] and longest or shortest diameter [89].

Table 4: Extracted features of pigmented skin lesions from both macroscopic and dermoscopic images.

Feature	References
Shape	
Asymmetry index	[87,90] [a]; [12,84,88,91] [b]
Statistical geometrical measures	[9,10,61,87,92] [a]; [12,63,84,88,89,93-103] [b]
Statistical measures based on border's gradient or periphery regions	[9,10,61,70,104] [a]; [74,78,84,88,93,94,96,97,105] [b]
Border features (irregularity index)	[61,90,92,104] [a]; [12,89,103,106,107] [b]
Colour	
Statistical measures based on colour models	[9,10,70,82,87,90] [a]; [63,71,73,84,88,93,94,96,101,105,108-113] [b]
Colour occurrence or percentage	[74,81,102,103,113-116] [b]
Absolute or relative colour features	[87] [a]; [77,88,93,94,96,99,113,117-119] [b]
Colour asymmetry	[73,88,105,110] [b]
Histogram-based features (colour distribution)	[24,62,78,79,83,84,88,97,99,102,110,112] [b]
Colour features based on cluster analysis	[102,112,120] [b]
Border's gradient-based features	[96,113] [b]
Texture	
Statistical	[9,10,70] [a]; [11,63,64,77,83,84,88,93,94,96,101,102,105,108-112,120-122] [b]
Model-based	[90] [a]; [12,123,124] [b]
Filter-based	[12,24,62,73,78,79,83,84,111,112,121,123] [b]
Other features	
Colour-texture features	[82] [a]; [72,124] [b]
High-level intuitive features	[125] [a]
Manual information	[10,82] [a]
Diameter	[102] [b]
Differential structures	[103,126] [b]
Evolution measures	[127] [a]; [128] [b]

[a] Macroscopic images
[b] Dermoscopic images

Geometrical measures from the segmented lesion area have been commonly computed for assessing the lesion's asymmetry and border's irregularity [12,63,70,84,129]. Such measures include the area of the lesion (computed as the number of pixels inside the lesion region [10] or by applying the bit quads method [88]), aspect ratio, compactness, perimeter, greatest diameter, shortest diameter, equivalent, convex hull, eccentricity, solidity, rectangularity, entropy measures, circularity index (namely thinness ratio), and irregularity index. Shape features based on wavelet transform [12,61,106], Fourier transform [104] and fractal dimension [92,103] have also been used for assessing the border's irregularity.

Shape features of differential structures inside the lesion in dermoscopic images may also be considered [66], such as solid pigments of the lesions computed according to Chang et al. [87]. In other studies [63,93,96], the asymmetry is assessed, according to pre-defined regions inside the lesion under analysis.

In order to identify the sharp transition between inside and outside regions of a lesion concerning its border, Iyatomi et al. [94,96] 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. [88] 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.

2.2.2. Colour variation

The *RGB* colour space is commonly used to represent the colours of skin lesions [63,110]. Other colour spaces have also been applied in order to obtain more specific information about a lesion's colours, such as: normalized *RGB* [110,111], *HSV* [11,84,110], *HVC* [109], *CMY* [108], *YUV* [108], *I1/2/3* [110], *Opp* [83,84], $\bar{I}_i^N$ [70], *JCh* [73], *L*C*H* [87], *CIEXYZ* [111], *CIELAB* [11,83,84] and *CIELUV* [11,110].

Statistical measures are widely applied to the feature extraction from skin lesion images [10,63,70,93]. 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 as part of the lesion. Peripheral regions may be defined by a recursive erosion process [93,110], a fast Euclidean distance transform algorithm [88], or a circular region with centre point upon the lesion's centroid [87]. In addition, such regions may reduce the effects of peripheral inflammation and errors caused by automatic border detection, as proposed by Celebi et al. [88].

Skin lesion features based on relative colours have been proposed [77,88,93], 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 [77].

The occurrence of the possible basic colours present in the skin lesions has also been analysed [74,81], as well as the number or percentage of pixels within the segmented area for each of the basic colours [9,73].

2.2.3. 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. Texture descriptors with statistical-, model- and filter-based approaches [130], have been used for texture quantification of skin lesions. Among the various statistical-based texture descriptors applied in the literature, the grey-level co-occurrence matrix (GLCM) proposed by Haralick et al. [131] has been one of the most commonly used [63,84,101,105,110,111]. The GLCM 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 GLCM, such as variance, entropy, dissimilarity, correlation, contrast, energy, maximum probability, inverse difference, angular second moment (ASM), mean, standard deviation and homogeneity. In Schaefer et al. [110], the authors computed the ratio and difference of the same co-occurrence features between different image regions.

Skin lesion features from histograms, which are also statistical-based descriptors, are extracted by some researchers to represent texture features [11,84]. Tanaka et al. [121] computed some aforementioned statistical measures based on the intensity histogram, whereas Barata et al. [83] applied gradient histograms, such as the gradient amplitude and 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. Local binary pattern (LBP) that is a discriminative rotation invariant feature descriptor [84,102,112], statistical measures based on pixel intensities [9,70], run-length matrix [121], and entropy features [120], have also been applied to texture extraction based on statistical approaches.

Model-based texture descriptors have also been proposed to assess the skin lesion's texture, such as fractal dimensional [12], auto-regression [123], and Markov random fields (MRF) [124]. Among these, fractal dimension have been applied with the box-counting method (BCM), being one of the most commonly used methods, since it is simple and effective [132]. Image-based fractal dimension [132] is a procedure for

splitting the image in several quadrants to quantify the irregularity level or self-similarity of the image's fractals.

Wavelet transform [12,71,112], Fourier transform [24,78,79,112,121], Gabor filtering [83,84,123], scale-invariant feature transform (SIFT) [84] and steerable pyramid transforms [73], which are filter-based texture descriptors, have also been proposed for feature extraction of skin lesion images. Such descriptors allow the decomposition of the input image into component parts in order to extract features from the structures of interest. Sobel, Hessian, Gaussian and difference of Gaussians (DoG) features have also been extracted based on bank of Gaussian filters [111]. Further details regarding texture analysis techniques for image feature extraction are presented in Xie [130].

2.2.4. Other features

Skin lesion features based on shape, colour and texture properties have been commonly used for skin lesion recognition. However, other features have also been considered, such as information regarding the part of the body, size and gender, since they can assist in skin lesion diagnosis [10,82]. Colour-texture descriptors have also been recently used to assess skin lesion features; e.g., colour image analysis learning vector quantization (CIA-LVQ) in the *RGB* colour space [82], and joint distribution of colour (JDC) in the *L*a*b** colour space [72]. Further details regarding colour-texture descriptors are presented in Xie [130].

The lesion's diameter is another feature that can be used for skin lesion diagnosis. This feature 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 [65]. This feature is not commonly applied to skin lesion classification due to its great dependence on the image resolution [88], since the image size affects the number of pixels for each segmented lesion's region. An application of this feature is presented in Møllersen et al. [102], in which the diameter of a lesion is defined as the length of the major axis of the best-fit ellipse. The differential structures of skin lesions may also be assessed, more specifically in dermoscopic images. For example, in Torre et al. [126] 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.

Elevation and evolution features can be assessed to assist in skin lesion classification process [66,133]. The former is a morphological feature that may be measured considering its surface. The latter may represent the historical evolution of the lesion in

order to diagnose it, including changes in its shape, size, shades of colour, or surface features. To the best of our knowledge, few previous image analysis systems of skin lesions surveyed in the literature have used such features [127,128]. 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. [134] and Fadzil et al. [135] 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 methods may be applied to track skin lesions in images [136], or to detect changes in their structure over time, as the algorithm introduced by Huang and Bergstresser [127]. The authors proposed a new method 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 transform the point registration problem in images to a bipartite graph-matching problem.

2.3. Feature selection

A feature selection step [137] has been used for pattern analysis and 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 [93,98,108,121]. Moreover, such features may influence the performance of the classification process, i.e., render it a slower process [138]. Several benefits are associated with the application of feature selection schemes, such as [88]: 1) to reduce the feature extraction time, 2) to decrease the classification complexity, 3) to improve the classification accuracy rate, 4) to decrease training and testing time, and 5) to simplify the understanding and visualizing the data.

Essentially, the feature selection process has the following steps: 1) feature subset selection, 2) feature subset evaluation, 3) stopping criterion, and 4) validation procedure [139]. 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.

Feature subset selection step 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 [108], ranker [12,108], incremental stepwise [93,98] and random [87,88], have been used for the feature subset selection process. Exhaustive and genetic searches are other examples of such algorithms that may be applied [140]. These algorithms influence the search direction and execution time of the selection process depending on the adopted search strategy, which may be complete, sequential, or random [139,141]. 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 [138].

In evaluation step, the selected feature subset is then evaluated according to the type of search algorithm applied before. The filter model [141] has been commonly used for the evaluation process of skin lesion feature selection. This model allows for evaluating the goodness of selected features without using any classification 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. Examples of filter methods applied in the literature based on the aforementioned independent measures are: gain ratio feature selection (GRFS) [12], information gain measure [12], chi-squared [12], correlation-based feature selection (CFS) [10,12,61], ReliefF [12,88], mutual information-based feature selection (MIFS) [88], sequential feature selection (SFS) [80], generalized sequential feature selection (GSFS) [108], and fast correlation-based feature filter (FCBF) [110].

Wrapper [142], hybrid [141] and embedded [138] models can also be used to evaluate the selected feature subset by a search strategy. The evaluation of feature subsets based

[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 of finding a minimum number of features that may separate classes as consistently as the full set of features may.

on the wrapper model is similar to the filter model. The main difference between these two models is the use of classification algorithms to evaluate each candidate subset in order to determine the most relevant subset, for which the classification algorithm tends to perform better when searching for such a subset [141]. 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. 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 algorithms, such as classification and regression tree (CART), are examples of such a model [137].

Feature selection methods based on a filter model [141] 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 [10,12,98]. Nevertheless, the features selected by using a filter model may not be the most relevant for the application, whereas the wrapper model [142] may be applied to search for the most relevant features based on classification 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. [88]. 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 [137].

The stopping criterion determines the situation in which the feature selection process must stop. Some examples of such criteria occur when: 1) the search is complete, 2) the predefined minimum number of features is achieved, 3) the predefined 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 [141]. The validation procedure 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.

Principal component analysis (PCA) [143] and linear discriminant analysis (LDA) [62], which are methods for space dimensionality reduction, have also been applied to feature selection [24,70,73]. Maglogiannis and Doukas [108] applied several classification methods to evaluate the obtained subsets by using feature selection algorithms such as the CFS, PCA and GSFS. Furthermore, the achieved results are compared to the ones 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. On the other hand, Ma and Staunton [61] used a feature selection scheme based on correlation analysis for skin lesion classification based on a neural network, since it achieved better result than original feature-based classification. Arroyo and Zapirain [111] analysed the relevant features based on the minimum number of samples per leaf by using decision tree classifier. Several other study have achieved good classification results by using a feature selection scheme [80,84,110].

Another means of determining the most discerning features based on colour and texture was addressed by Barata et al. [11], 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. On the other hand, Rastgoo et al. [84] evaluated the most discerning features between shape, colour and texture features and the evaluation revealed the potential of texture features for skin lesion classification.

3. 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 methods may be applied to overcome these problems.

Feature normalization is a pre-processing step, in which methods may be applied in order to solve the problem of different ranges. The z-score transformation is a common method used for data normalization, which allows transforming all numeric features in values within the same range, as discussed by Celebi et al. [88] and by Cavalcanti and Scharcanski [70]. 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.

Unbalanced datasets concerning the number of samples in each class is also 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. Sampling techniques, such as over- and under-sampling [140,144], have been used to solve this problem [88,110]. Nevertheless, random under-sampling may remove important samples, and random over-sampling may lead to over-fitting. Synthetic minority oversampling technique (SMOTE) [145] is an over-sampling techniques for overcoming the over-fitting and expand the decision region of minority class samples. Such techniques can also be combined with ensemble methods for addressing unbalanced classes [110]. Another method to solve the unbalanced dataset problem was used by Barata et al. [11], 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.

As mentioned previously, feature selection [137], which is a pre-processing step in machine learning, can be addressed to deal with datasets contain a large number of features for skin lesion classification (Section 3.3). The classification methods used for skin lesion diagnosis, as well as its evaluation procedures, are presented with details in following sections. Furthermore, some results of recent studies for classification of skin lesion and its patterns are also provided.

3.1. Methods for classification

Classification methods based on instance-based learning [140], decision trees [138], Bayesian learning [146], artificial neural networks (ANNs) [147], support vector machines (SVMs) [148], and ensemble methods [140], have been commonly applied to discriminate skin lesions in images. A description and the main advantages and disadvantages of such methods are summarized in the following, while their algorithms applied to the learning objective are presented in Table 5.

Table 5: Classification methods applied to discriminate skin lesions from images.

Classification Method	References
Instance-based learning	
KNN	[11,63,70,73,83,109,149]
KStar	[108]
LWL	[108]
Decision tree	
NBTree	[64,108]
AD-Tree	[64]
CART	[70,80,102,108]
J48/ C4.5/ C5.0	[10,64,77,87,106,111,150]
C&R	[106]
LMT	[10,12,24,79,151]
Decision Stump	[10,64]
Bayesian network	
BayesNet	[10,62,82,108]
NBL	[108]
HNB	[12]
ANN	
MLP architecture	[61,63,81,94,97,99,106,117]
RBF network	[108]
SVM	
Linear kernel	[125]
RBF kernel	[11,12,73,88,90,108,109,152]
Polynomial kernel	[63,110,123]
PUK kernel	[63]
Ensemble of classifiers	
Homogeneous ensemble	[110,112,153]
Heterogeneous ensemble	[82,109]
Bagging	[110]
Random forest	[12,63,64,83,84]
Boosting	[84,100,101,154]
AdaBoost	[73,83]
Other methods	
Linear classifier	[93,96,105]
Regression analysis	[104,108]
Prototype-based	[82]
Discriminant analysis	[80,102,155]
Maximum likelihood	[62,109,149]

KNN: k-nearest neighbour; LWL: locally weighted learning; NBTree: naïve Bayes/decision tree; AD-Tree: alternative decision tree; CART: classification and regression trees; C&R: classification and regression; LMT: logistic model tree; RF: random forest; NBL: naïve Bayes multinomial; HNB: hidden naïve Bayes; ANN: artificial neural network; MLP: multilayer perceptron; RBF: radial basis function; SVM: support vector machine; PUK: Pearson VII function-based universal kernel.

In instance-based classifiers [140], a distance function is used to assess which sample of the training set is closest to an unknown sample and then assigning the unknown sample to the class with the majority of the nearest neighbours. These classifiers have

been applied due to their simplicity of implementation and their facility to deal with the existence of correlated features. In addition, new samples can be added to the training set at any time. However, they are sensitive to the existence of irrelevant features, and they require a great deal of time for classifying large datasets. Barata et al. [11] used the k-nearest neighbours (KNN) algorithm to classify the lesions and compared several distance functions, such as Euclidean, Kolmogorov and Kullback–Leibler, in order to measure the distance of k-nearest neighbours from different k values. The authors concluded that it is not clear which of these three used 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. [109] 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 [138] 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. Several methods based on decision trees have been frequently applied to classify skin lesions [10,12,24,77,106,108]. Understanding such a structure, as well as ease of rule generation, is quite straightforward. However, the excess of adjustments (over-fitting) and the difficulties in dealing with correlated features are the major drawbacks of decision trees.

Bayesian learning-based methods [146] compute the probability of a given set of features to belong to each class, assuming that the features are independent. These methods have been applied to classify skin lesions particularly because of their fast training [10,12,108]. Although Bayesian methods provide fast training and no sensitivity to irrelevant features, they assume that the features need to be independent, which can be a disadvantage of these methods.

ANNs [147] 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 [81,106], 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 [61], which consists of forward and backward processes applied to adjust the weight values of the connections. 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 [148] involve a method based on statistical learning applied to building a hyper-plane to separate the data according to the defined classes. This kind of 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. However, these classifiers are sensitive to noise and the classification process is based on a binary class. The radial basis function (RBF) kernel have been commonly adopted in several studies [11,12,108] 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 [88].

The ensemble methods [140] have been recently adopted to diagnoses skin lesions [82-84,110,156]. Ensemble models may be constructed with either several classification algorithms, classified as heterogeneous, or only with one classification algorithm, classified as homogeneous, which can be developed through data manipulation [157]. 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 used for this same purpose. The common algorithms applied to manipulate the training samples are the Bagging and Boosting algorithms [157]. Random Forest [158] and AdaBoost [159] are also popular ensemble methods. Random Forest is a variation of the Bagging algorithm that is used to create individual decision trees, whereas AdaBoost is a popular boosting algorithm that maintains a set of weighting systems over the training samples. Ensemble methods consist 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. However, such methods can present a high computational complexity.

3.2. Evaluating the classification

The main objective of the classification process of skin lesions is to achieve good results for distinguishing between different lesion 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 [140] is the most commonly used in the literature to evaluate the results of skin lesion classification, since it avoids over-fitting while testing the capacity of the classifier to generalize. The k-fold cross-validation [12,108] and leave-one-out [11,93] are examples of cross-validation procedures proposed for classifying skin lesions in images. The half-and-half test is another evaluation procedure, which was applied by Iyatomi et al. [96]. In addition, 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.

Statistical measures based on performance metrics [160] 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: 1) true positive (TP), 2) true negative (TN), 3) false positive (FP), and 4) false negative (FN). These outcomes represent the number 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 number of correctly classified positive samples, the TN rate is the number of correctly classified negative samples, the FP rate is the number of incorrectly classified negative samples, and the FN rate is the number of incorrectly classified positive samples.

The aforementioned 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 [10,12,81], 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 or sensitivity, which is the percentage of correctly classified positive samples with respect to all positive samples, 3) the specificity, which is the percentage of correctly classified negative samples with respect to all negative samples, and 4) the accuracy that is the percentage of correctly classified positive and negative samples based on all samples. Area under the ROC curve (AUC) is an additional term associated with the receiver operating characteristics (ROC) graph [160], which is also used to compare the performance of the classification, since it is a very useful tool for visualizing and

evaluating classifiers [10,11,81]. Currently, such measure is commonly used and are able to provide a more robust classification performance measure than other evaluation measures [160].

3.3. Skin lesion classification performance

For the skin lesion classification process, one or several techniques 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 classification process may be binary or multi-class, and includes different classes according to the classification goal, such as: 1) malignancy of the lesions (benign versus malignant) [12,106], and 2) distinct types of skin lesions (melanoma versus nevus [94,108], melanocytic versus non-melanocytic [93], and dysplastic versus non-dysplastic versus melanotic [108]). Furthermore, skin lesion features are also classified in terms of: 1) border features (regular versus irregular [90,106] and irregularity level [95]), 2) presence of main colours existing in malignant lesions [81,112], 3) presence of features of the seven-point checklist [24,79,161], 4) presence of global patterns [72,73,162] and 5) presence of local patterns [71,128].

Table 6 summarizes the best results of recent studies concerning skin lesion classification. The table indicates the number and type of image used, the techniques employed in the segmentation step and feature selection, the number of extracted and selected features, the classification algorithms and the values of the evaluation measures used. The performance of several classifiers has been compared, e.g. in terms of the discrimination between benign lesions and melanomas, by several authors.

Zortea et al. [80] compared the classification performance of quadratic discriminant analysis (QDA), linear discriminant analysis (LDA) and classification and regression trees (CART), and obtained the best results with QDA. In the study of Rastgoo et al. [84], better results were achieved using a random forest than a gradient boosting and SVM classifier. Likewise, Barata et al. [83] have also obtained the best results by applying a random forest than using AdaBoost, SVM and KNN. Schaefer et al. [110] proposed an ensemble method based on a SVM (polynomial kernel), non-pairwise measure of diversity (fuzzy Shannon), and neural network based on classifier fusion, which obtained the best results when compared with other ensemble methods, as well as with individual SVM classifier.

Table 6: Results of recent studies focused on the skin lesion classification.

Ref.	Number of image (Type)	Segmentation	Feature selection (EF/ SF)	Classifier	Classification	Mean results (Evaluation measures)
[112]	200 (Derm.)	-	- (NM/ -)	Ensemble method (SVM)	Malignant/ benign	91% (ACC), 97% (SE), 65% (SP), 92% (Prec.), 94% (FM), 95% (AUC).
[82]	170 (Macro.)	K-means Clustering	- (NM/ -)	Ensemble method (CLAM, CIA-LVQ, Naïve Bayes)	Melanoma/ nevus	81% (ACC), 0.741 (PPV), 0.859 (NPV).
[105]	968 (Derm.)	Thresholding	Incremental Stepwise (828/ 25)	Linear classifier	Melanoma/ nevus/ BCC/ SK	Melanoma: 90.48% (DR); Nevus: 82.51% (DR); BCC: 82.61% (DR); SK: 80.61% (DR).
[84]	180 (Derm.)	Thresholding	PCA (NM/ NM)	Ensemble method (Random forest)	Melanoma/ dysplastic nevus	98% (SE), 70% (SP).
[83]	DB1: 200 DB2: 482 (Derm.)	NM	Fusion strategies (NM/ NM)	Ensemble method (Random forest)	Melanoma/ nevus	DB1: 98% (SE), 90% (SP); DB2: 83% (SE), 76% (SP).
[102]	210 (Derm.)	NM	Feature analysis/ wrapper + filter (59/ 19)	Discriminant analysis	Not-cut/ cut (benign lesion/ suspicious lesion and melanoma)	81% (CR), 83% (SE), 80% (SP).
[125]	206 (Macro.)	-	- (62/ -)	SVM	Melanoma/ non-melanoma	83.59% (ACC), 91.01% (SE), 73.45% (SP).
[80]	206 (Derm.)	NM	SFS (53/ 7.6)	Discriminant analysis	Melanoma/ benign	86% (SE), 52% (SP), 63.3% (CR).
[110]	564 (Derm.)	Thresholding, region-growing and merging	FCBF (437/ 74)	Ensemble method (SVM)	Melanoma/ benign	93.83% (ACC), 93.76% (SE), 93.84% (SP).
[61]	134 (Macro.)	-	Correlation analysis (25/ 13)	ANN	Melanoma/ benign	0.83 (SE), 0.90 (SP), 0.89 (AUC).
[149]	152 (Macro.)	Thresholding	- (Stage one: 52; stage two: 12/ -)	Stage one: KNN; Stage two: maximum likelihood	Malignant/ benign	99.34% (ACC), 100% (SE), 97.78% (SP).
[11]	176 (Derm.)	Thresholding	Individual and combined feature analysis (NM/ -)	Ensemble method (AdaBoost)	Melanoma/ nevus	96% (SE), 80% (SP).
[103]	120 (Derm.)	Dynamic programming	SFFS (NM/ NM)	SVM	Melanoma/ nevus	Melanoma: 88.2% (SE), 91.30% (SP), 0.880 (AUC); Nevus: 86.5% (SE), 88.2% (SP), 0.824 (AUC).
[12]	289 (Derm.)	Thresholding	GRFS (35,455/ 23)	Ensemble method (random forest)	Malignant/ benign	91.26% (ACC), 0.937 (AUC).
[70]	152 (Macro.)	Thresholding	- (52/ -)	KNN/ KNN-DT	Malignant/ benign	96.71% (ACC), 96.26% (SE), 97.78% (SP).
[104]	167 (Macro.)	Ncut	- (NM/ -)	Regression analysis	Melanoma/ benign	70.5% (ACC), 71.8% (SE), 69.8% (SP).
[93]	655 (Derm.)	Thresholding, morphological operations	Incremental Stepwise (428/ 2)	Linear classifier	Melanocytic/ non-melanocytic	97.99% (SE), 86.64% (SP).
[10]	152 (Macro.)	Thresholding	CFS (45/ 5)	LMT	Melanoma/ nevus	86% (ACC), 94% (SE), 68% (SP), 0.890 (AUC).
[106]	30 (Derm.)	-	- (NM/ -)	C5.0	Malignant/ benign	93.30% (ACC), 80% (SE), 96% (SP).
[108]	3639 (Derm.)	NM	- (31/ -)	MLR/ SVM/ LWL/ CART	Melanoma/ nevus	100% (ACC), 1.0 (AUC).

Table 6: Continued.

Ref.	Number of image (Type)	Segmentation	Feature selection (EF/ SF)	Classifier	Classification	Mean results (Evaluation measures)
[108]	3639 (Derm.)	NM	- (31/ -)	ANN/ SVM/ Bayes networks	Dysplastic/ non-dysplastic	73.29% (ACC), 0.688 (AUC)/ 76.08% (ACC), 0.607 (AUC)/ 68.94% (ACC), 0.663 (AUC).
[108]	3639 (Derm.)	NM	- (31/ -)	SVM	Melanotic/ dysplastic/ non-dysplastic	77.06% (ACC), 1.0 (AUC).
[94]	1258 (Derm.)	Thresholding, region-growing	Incremental Stepwise (428/ 72)	ANN	Melanoma/ nevus	94.10% (ACC), 85.90% (SE), 86.0% (SP), 0.928 (AUC).
[96]	199 (Derm.)	NM	Incremental Stepwise (482/ 10)	Linear classifier	Melanoma/ nevus	100% (SE), 95.9% (SP), 0.993 (AUC).
[109]	358 (Derm.)	FCM, thresholding	PCA (128/ 10)	Ensemble method (KNN, SVM, GML)	Malignant/ benign/ dysplastic	75.69% (ACC).
[88]	564 (Derm.)	Region-growing and merging	CFS (473/ 18)	SVM	Melanoma/ benign	92.34% (SE), 93.33% (SP), 0.966 (AUC).

Ref.: reference; Macro.: macroscopic; Derm.: dermoscopic; BCC: basal cell carcinoma; SK: seborrheic keratosis; EF: extracted features; SF: selected features; NM: non-mentioned; ACC: accuracy; SE: sensitivity; SP: specificity; Prec.: precision; FM: F-Measure; AUC: area under the ROC curve; PPV: positive predictive value; NPV: negative predictive value; DR: detection rate; CR: correct rate; DB: database; Ncut: normalized cut; FCM: fuzzy c-means; GRFS: gain ratio feature selection; CFS: correlation-based feature selection; PCA: principal component analysis; SFS: sequential feature selection; SFFS: sequential floating feature selection; FCBF: fast correlation-based feature filter; LMT: logistic model tree; MLR: multinomial logistic regression; SVM: support vector machine; LWL: locally weighted learning; CART: classification and regression trees; ANN: artificial neural network; GML: Gaussian maximum likelihood; KNN: k-nearest neighbours; KNN-DT: k-nearest neighbours-decision tree; FKNN: fuzzy k-nearest neighbours; CLAM: cluster-based adaptive metric; CIA-LVQ: colour image analysis-leaning vector quantization.

Ensemble methods have performed better than individual classifiers in several studies [11,109], whereas Alcón et al. [10] obtained the best results in both the individual logistic model tree (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 discriminative features, as previously discussed.

The features extracted from the lesion have also been used for pattern detection or classification in order to assist in skin lesion diagnosis. Table 7 summarizes the best results of recent studies concerning global and local pattern classification in dermoscopic images. The table indicates the number of image used, the target of the detection or classification, and the values of the evaluation measures.

Table 7: Results of recent studies focused on the global and local pattern analysis in dermoscopic images.

References	Year	Number of images	Detection/classification	Mean results (Evaluation measures)
Global pattern				
[73]	2013	350	Reti./ Glob./ Cob./ Homo./ Paral./ Starb./ Mult.	89.28% (SE), 93.75% (SP), 0.986 (AUC).
[162]	2012	180	Reti./ Glob./ Cob./ Homo./ Paral./ Starb./	93.08% (SE), 91.45% (SP), 0.948 (AUC).
[72]	2012	325	Reti./ Glob./ Cob./ Homo./ Paral./	86.8% (ACC).
[74]	2011	160	Reti.; Glob./	89% (ACC); 95% (ACC).
[71]	2011	360	Mult.	NM
[98]	2009	100	Reti./ Glob./ Cob./ Homo./ Paral./	94% (ACC)
[124]	2009	100	Reti./ Glob./ Cob./ Homo./ Paral./	86% (ACC)
[96]	2008	213	Paral. ridge; paral. furrow; fibrillar	0.985 (AUC); 0.931 (AUC); 0.890 (AUC).
[121]	2008	44	Reti./ Glob./ Homo./	94% (ACC)
Pigmented network				
[112]	2015	NM	Typical	74% (ACC), 0.82 (AUC), 79% (Prec.).
[163]	2014	122	Present/ absent; typical/ atypical	85% (ACC), 0.821 (AUC); 100% (ACC).
[111]	2014	220	Present/ absent	86% (SE), 81.67% (SP).
[100]	2012	200	Present/ absent	86.2% (ACC), 91.1% (SE), 82.1% (SP).
[62]	2011	734	Present/ absent	0.922 (AUC)
[71]	2011	360	Melanoma/ benign	NM
[24]	2010	115	Atypical/ absent	80% (SE), 82% (SP).
[128]	2010	NM	Present/ absent	NM
[101]	2010	436	Present/ absent; absent/ typical/ atypical	93% (ACC), 0.935 (Prec.), 0.933 (Rec.); 82% (ACC), 0.820 (Prec.), 0.823 (Rec.).
[64]	2010	106	Melanoma/ benign	95.4% (ACC)
[164]	2008	173	Typical/ atypical	85% (ACC)
[165]	2006	60	No network/ partial/ complete	88.3% (ACC)
[166]	2006	30	Typical/ atypical	NM
[122]	2004	155	Present/ absent	80% (ACC)
[155]	1998	NM	Present/ absent	NM
Dots/globules				
[112]	2015	NM	Absent; typical; atypical	47% (ACC), 0.53 (AUC), 47% (Prec.); 70% (ACC), 0.55 (AUC), 39% (Prec.); 61% (ACC), 0.51 (AUC), 29% (Prec.).
[63]	2015	108	Malignant/ non-malignant	0.903 (ACC), 0.884 (SE), 0.923 (SP).
[128]	2010	NM	Present/ absent	NM
[71]	2011	360	Melanoma/ benign	NM
[155]	1998	NM	Present/ absent	NM
Streaks				
[112]	2015	NM	Absent	85% (ACC), 0.79 (AUC), 95% (Prec.).
[154]	2013	945	Present/ absent; regular/ irregular; absent/ regular/ irregular	78.3% (ACC), 83.2% (AUC); 83.6% (ACC), 88.9% (AUC); 76.1% (ACC), 85% (AUC).
[152]	2012	99	absent/ regular/ irregular	91% (ACC)
[151]	2010	53	Present/ absent	86% (SE), 88% (SP).
[24]	2010	200	Irregular/ absent	86% (SE), 88% (SP).
[78]	2005	10	Present/ absent	NM
Blue-whitish veil				
[112]	2015	NM	Absent	90% (ACC), 0.96 (AUC), 99% (Prec.).
[118]	2013	200; 100	Present/ absent	87% (ACC); 67% (ACC).
[150]	2011	887	Present/ absent	80.50% (SE), 90.93% (SP).
[74]	2011	160	Present/ absent	86% (ACC)
[71]	2011	360	Melanoma/ benign	NM
[24]	2010	110	Present/ absent	90% (SE), 93% (SP).
[151]	2010	135	Present/ absent	87% (SE), 85% (SP).
[77]	2008	100; 545	Present/ absent; Melanoma/ benign	84.33% (SE), 96.19% (SP); 69.35% (SE), 89.97% (SP).

Table 7: Continued.

References	Year	Number of images	Detection/classification	Mean results (Evaluation measures)
Blotches				
[24]	2010	110	Irregular/ absent	87% (SE), 90% (SP).
[99]	2009	424	Melanoma/ benign	81.2% (ACC)
[120]	2009	50	Present/ absent	NM
[117]	2005	512	Melanoma/ benign	77% (ACC)
Hypopigmentation				
[97]	2011	244	Melanoma/ nevus	0.952 (AUC)
Regression structures				
[112]	2015	NM	Absent	89% (ACC), 0.86 (AUC), 98% (Prec.).
[24]	2010	110	Present/ absent	80% (SE), 83% (SP).
[151]	2010	80	Present/ absent	80% (SE), 83% (SP).
Vascular structures				
[166]	2006	NM	Present/ absent	NM

The references of research about local features also include the works focused on the seven-point checklist method.
NM: non-mentioned; ACC: accuracy; SE: sensitivity; SP: specificity; AUC: area under the ROC curve; Prec.: precision; Rec.: recall; Reti.: reticular; Glob.: globular; Cob.: cobblestone; Homo.: homogeneous; Paral.: parallel; Starb.: starburst; Multi.: multicomponent.

Several methods have been proposed for the pattern analysis task in skin lesion diagnosis. Some of these methods have also used feature selection techniques, and the performance of several classifiers has also been taken into account [62,63,73]. 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. To the best of our knowledge, only one study dealing with the classification of all global patterns of skin lesions has been proposed [73], and no previous study has addressed the issue to identify the absence of such patterns. Indeed, it should be noted that the multicomponent pattern and the absence of patterns can indicate a higher probability of being a malignant lesion.

Abbas et al. [73] proposed the classification of skin lesion global patterns by using AdaBoost algorithm based on colour and texture properties from a perceptually uniform colour space. Furthermore, the authors developed a multi-label learning algorithm (AdaBoost.MC) to solve the problem of multicomponent pattern. This pattern is determined by fusing the results produced by AdaBoost.MC based on maximum a posteriori (MAP) and robust ranking principles. The method achieved superior results compared with the multi-label SVM and KNN.

Local pattern detection of dermoscopy images is a challenging task to assist in discriminating between benign and malignant skin lesions. 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 malignant lesion. To the best of our knowledge, no previous study has dealt with all skin lesion local patterns. Leo et al. [24]

proposed a method based on LMT to classify five local patterns based on the seven-point checklist method. The authors segmented the lesion colour by using 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 methods to extract texture features, such as median filter, close-opening operation, fast Fourier transform (FFT), high-pass filtering, inverse fast Fourier transform (IFFT) and suitable thresholding, in order to detect the atypical pigment network, and irregular streaks. The authors achieved good results in the detection of such local patterns.

Most studies have been proposed for the pigmented network detection [100,111,163]. In addition, other studies have considered feature extracted from patterns for discriminating between benign and malignant skin lesions [63,64,77]. Maglogiannis and Delibasis [63] classified the skin lesion into malignant and non-malignant and achieved superior results with inclusion of the dot-related features to the lesion-related features. The SVM classifier (polynomial kernel) yielded better results than MLP, KNN, random forest and SVM (PUC kernel) based on dot-related features. The dots were segmented using a circularity function and definition of diffusivity after enhancing dark circular structures using inverse non-linear diffusion.

4. Discussion

Dermoscopic images have been widely used for diagnosis of pigmented skin lesions [167,168], 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 dermoscopic compared to diagnosis by macroscopic image [169]. Among the several skin lesion diagnostic methods using dermoscopic images [67], the ABCD rule has been commonly applied to extract features for computational analysis [84,149]. This rule allows for easy understanding and provides simplicity of application while showing reliable results for the melanoma diagnosis. On other hand, previous clinical studies [69] 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 shape, 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 [73,81]. Figure 4 illustrates the distribution of the methods that have been proposed for skin lesion classification reviewed in this article according to the main feature used.

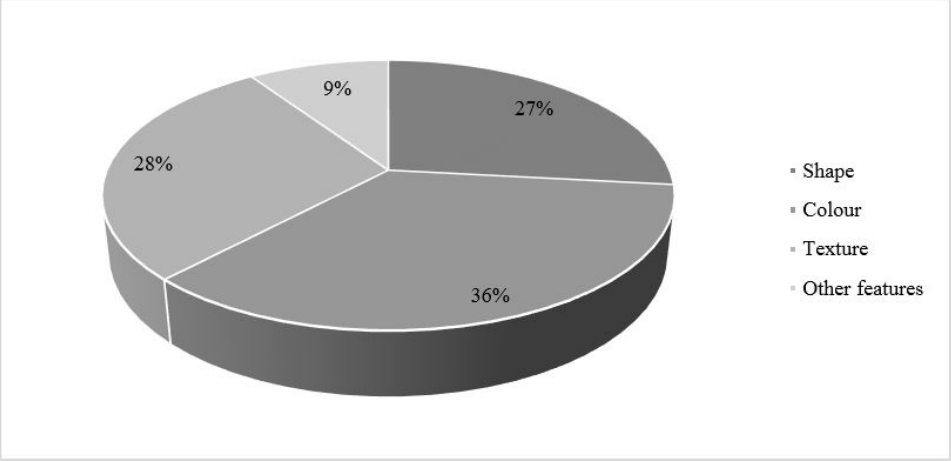

Figure 4: Distribution of the reviewed skin lesion classification methods according to the main feature used.

The classification process of skin lesions in images must be effective, since it is crucial to assist dermatologists in the diagnosis of these lesions by means of CAD systems. In addition, the evaluation and improvement of the performance of classifiers are essential for the pattern recognition research field [58]. A relevant problem that affects the performance of classifiers is the definition of the meaningful features for representing the classes. Consequently, the feature extraction and selection steps are very important to achieve better performance for the computational diagnosis of skin lesions in images. The application of several descriptions may be required considering the large number of features extracted from images. For dealing with this issue, feature selection methods have been applied to establish the most relevant features [61,80,84,110], 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. 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 algorithms should be chosen based on the classification problem and available data regarding

advantages and disadvantages of each algorithm. Figure 5 illustrates the distribution of the classification algorithms used in the methods reviewed in this article for skin lesion classification.

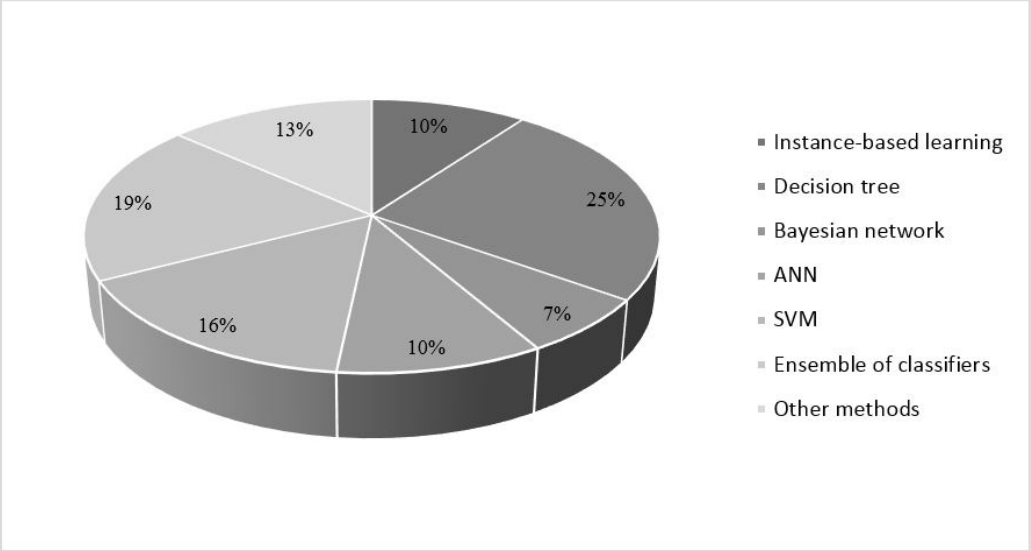

Figure 5: Distribution of the classification algorithms used by the reviewed methods for skin lesion classification.

Classification methods based on a decision tree have been used by many authors for the skin lesion classification [10,12,24]. 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. Ensemble methods [140], which aim to combine the strengths of different classifiers, have also been commonly proposed to improve the performance of the classification of skin lesions. These methods have performed better than individual classifiers [11,110]. The SVM classifier [148] 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 [63,73]. Despite the long training time, ANNs have been proposed in various studies [61,81,106] to deal with complex pattern recognition problems. Recently, the linear classifier [93], regression analysis [104], prototype-based classifier [82], discriminant analysis [80] and maximum likelihood [149], have also been proposed to solve problems of skin lesion classification.

5. Conclusion and future trends

Pigmented skin lesion classification is an area of great research interest due to its importance in skin cancer prevention, as well as in the early diagnosis. This review

provides an overview of current developments of computational methods for skin lesion image classification. Studies specifically addressing automatic methods applied to the feature selection and extraction steps, based on several clinical approaches, were presented in this review. In addition, the skin lesion classification step was addressed by including classifiers and evaluation procedures, as well as some performance results for pattern and lesion classification.

From this review, one may conclude that several studies focused on skin lesion classification have been proposed for use in CAD systems. Such systems aim at an effective computational diagnosis of pigmented skin lesions to assist dermatologists in their diagnosis. Although this research topic has been addressed in several studies, resulting in successful systems, new methodologies may be proposed to fill gaps that still have not been fully addressed, as well as to improve the performance of existing methods. Most studies involve extraction of several features from dermoscopic 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 patterns 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 expert systems for the computational diagnosis based on macroscopic and dermoscopic images. Hence, several issues may be addressed to achieve this goal, in particular: 1) the evolution features 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 features may complement the diagnosis; 2) the development and evaluation of new computational methods to identify the presence of global patterns, mainly the starburst and multicomponent patterns, since few studies have explored such patterns; 3) the lack of computational methods to detect some skin lesion local patterns and access their irregularity that can also be important to assist in diagnosis of specific lesions; 4) the development of new approaches for colour and asymmetry patterns, and positive feature analysis based on Menzies's method is important for future applications of this method for computational diagnosis of skin lesions; 5) in order to find more relevant features for the given problem, different feature

selection models may be compared; and 6) 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 the issues aforementioned may perform better and more effectively in diagnosing skin lesions in images. In addition, such methods may cover several problems regarding skin lesion classification, which convert CAD systems into more complete expert systems for diagnosing such lesions based on macroscopic and dermoscopic images.

Acknowledgements

The first author would like to thank the CNPq ("Conselho Nacional de Desenvolvimento Científico e Tecnológico"), in Brazil, for her PhD grant. This work is funded by European Regional Development Funds (ERDF), through the Operational Programme 'Thematic Factors of Competitiveness' (COMPETE), and Portuguese Funds, through "Fundação para a Ciência e a Tecnologia (FCT)", under the project: FCOMP-01-0124-FEDER-028160/PTDC/BBB- BMD/3088/2012. Authors gratefully acknowledge the funding of Project NORTE-01-0145-FEDER-000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, cofinanced by "Programa Operacional Regional do Norte" (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional" (FEDER).

References

[1] Razmjooy, N.; Mousavi, B. S.; Soleymani, F.; Khotbesara, M. H. (2013). A computer-aided diagnosis system for malignant melanomas. *Neural Computing and Applications*, 23 (7):2059-2071.

[2] Ruela, M.; Barata, C.; Marques, J. S.; Rozeira, J. (2015). A system for the detection of melanomas in dermoscopy images using shape and symmetry features. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*:1-11.

[3] Scharcanski, J.; Celebi, M. E. (2013). *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, Berlin, Heidelberg.

[4] INCA (2014). *Estimativa 2014: Incidência de Câncer no Brasil*. Instituto Nacional de Câncer José Alencar Gomes da Silva, Coordenação de Prevenção e Vigilância. INCA, Rio de Janeiro.

[5] American Cancer Society (2014). *Cancer Facts & Figures 2014*. American Cancer Society, Atlanta.

[6] Cancer Research UK (2013). *Cancer Statistic Report: Skin Cancer*. Cancer Research UK. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/skin/?script=true. Accessed October 2015.

[7] Bourne, P.; Cameron, A.; Gourhant, J.-Y.; Hackett, T.; Hlaing, W.; Kittler, H.; McColl, I.; Minas, S.; Rosendahl, C. (2007). *The International Atlas of Dermoscopy and Dermatoscopy* http://www.dermoscopyatlas.com/index.cfm. Accessed October 2015.

[8] Smith, L.; MacNeil, S. (2011). State of the art in non-invasive imaging of cutaneous melanoma. *Skin Research and Technology*, 17 (3):257-269.

[9] Cavalcanti, P. G.; Scharcanski, J. (2013). *Macroscopic Pigmented Skin Lesion Segmentation and Its Influence on Lesion Classification and Diagnosis*. In: Celebi, M. E., Schaefer, G. (eds) Color Medical Image Analysis. Springer, Dordrecht, pp 15-39.

[10] Alcón, J. F.; Ciuhu, C.; Ten Kate, W.; Heinrich, A.; Uzunbajakava, N.; Krekels, G.; Siem, D.; de Haan, G. (2009). Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):14-25.

[11] Barata, C.; Ruela, M.; Francisco, M.; Mendonça, T.; Marques, J. S. (2013). Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features. *IEEE Systems Journal*, 8 (3):965-979.

[12] Garnavi, R.; Aldeen, M.; Bailey, J. (2012). Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 16 (6):1239-1252.

[13] Oliveira, R. B.; Filho, M. E.; Ma, Z.; Papa, J. P.; Pereira, A. S.; Tavares, J. M. R. S. (2016). Computational methods for the image segmentation of pigmented skin lesions: a review. *Computer Methods and Programs in Biomedicine*, 131:127-141.

[14] Silveira, M.; Nascimento, J. C.; Marques, J. S.; Marcal, A. R. S.; Mendonca, T.; Yamauchi, S.; Maeda, J.; Rozeira, J. (2009). Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):35-45.

[15] Wong, A.; Scharcanski, J.; Fieguth, P. (2011). Automatic Skin Lesion Segmentation via Iterative Stochastic Region Merging. *IEEE Transactions on Information Technology in Biomedicine*, 15 (6):929-936.

[16] Yuksel, M. E.; Borlu, M. (2009). Accurate Segmentation of Dermoscopic Images by Image Thresholding Based on Type-2 Fuzzy Logic. *IEEE Transactions on Fuzzy Systems*, 17 (4):976-982.

[17] Zhou, H.; Schaefer, G.; Celebi, M. E.; Iyatomi, H.; Norton, K.; Liu, T.; Lin, F. (2010). *Skin lesion segmentation using an improved snake model*. In: Annual International Conference of the Engineering in Medicine and Biology Society, Buenos Aires, August 31 - September 4 2010. IEEE pp 1974-1977.

[18] Zhou, H.; Li, X.; Schaefer, G.; Celebi, M. E.; Miller, P. (2013). Mean shift based gradient vector flow for image segmentation. *Computer Vision and Image Understanding*, 117 (9):1004-1016.

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

[20] Abbas, Q.; Celebi, M. E.; Garcia, I. F. (2012). A novel perceptually-oriented approach for skin tumor segmentation. *International Journal of Innovative Computing, Information and Control*, 8 (3):1837-1848.

[21] Abbas, Q.; Fondón, I.; Rashid, M. (2011). Unsupervised skin lesions border detection via two-dimensional image analysis. *Computer Methods and Programs in Biomedicine*, 104 (3):e1-e15.

[22] Norton, K.; Iyatomi, H.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Development of a novel border detection method for melanocytic and non-melanocytic dermoscopy images*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5403-5406.

[23] Norton, K.-A.; Iyatomi, H.; Celebi, M. E.; Ishizaki, S.; Sawada, M.; Suzaki, R.; Kobayashi, K.; Tanaka, M.; Ogawa, K. (2012). Three-phase general border detection method for dermoscopy images using non-uniform illumination correction. *Skin Research and Technology*, 18 (3):290-300.

[24] Leo, G. D.; Paolillo, A.; Sommella, P.; Fabbrocini, G. (2010). *Automatic Diagnosis of Melanoma: A Software System Based on the 7-Point Check-List*. In: 43rd International Conference on System Sciences, Hawaii January 5-8 2010. IEEE, pp 1-10.

[25] Garnavi, R.; Aldeen, M.; Celebi, M. E.; Varigos, G.; Finch, S. (2011). Border detection in dermoscopy images using hybrid thresholding on optimized color channels. *Computerized Medical Imaging and Graphics*, 35 (2):105-115.

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

[27] Flores, E.; Scharcanski, J. (2016). Segmentation of melanocytic skin lesions using feature learning and dictionaries. *Expert Systems with Applications*, 56:300-309.

[28] Ma, Z.; Tavares, J. M. R. S. (2016). A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model. *IEEE Journal of Biomedical and Health Informatics*, 20 (2):615-623.

[29] Celebi, M. E.; Wen, Q.; Hwang, S.; Iyatomi, H.; Schaefer, G. (2013). Lesion border detection in dermoscopy images using ensembles of thresholding methods. *Skin Research and Technology*, 19 (1):e252-e258.

[30] Abbas, Q.; Celebi, M. E.; García, I. F. (2012). Skin tumor area extraction using an improved dynamic programming approach. *Skin Research and Technology*, 18 (2):133-142.

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

[32] Garnavi, R.; Aldeen, M.; Celebi, M. E. (2011). Weighted performance index for objective evaluation of border detection methods in dermoscopy images. *Skin Research and Technology*, 17 (1):35-44.

[33] Celebi, M. E.; Schaefer, G.; Iyatomi, H.; Stoecker, W. V.; Malters, J. M.; Grichnik, J. M. (2009). An improved objective evaluation measure for border detection in dermoscopy images. *Skin Research and Technology*, 15 (4):444-450.

[34] Zhou, H.; Schaefer, G.; Sadka, A. H.; Celebi, M. E. (2009). Anisotropic Mean Shift Based Fuzzy C-Means Segmentation of Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):26-34.

[35] Celebi, M. E.; Aslandogan, Y. A.; Stoecker, W. V.; Iyatomi, H.; Oka, H.; Chen, X. (2007). Unsupervised border detection in dermoscopy images. *Skin Research and Technology*, 13 (4):454-462.

[36] Cavalcanti, P. G.; Scharcanski, J.; Lopes, C. B. O. (2010). *Shading attenuation in human skin color images*. In: 6th International Symposium on Visual Computing, Las Vegas, November 29 - December 1 2010. Springer, pp 190-198.

[37] Glaister, J.; Amelard, R.; Wong, A.; Clausi, D. (2013). MSIM: Multistage illumination modeling of dermatological photographs for illumination-corrected skin lesion analysis. *IEEE Transactions on Biomedical Engineering*, 60 (7):1873-1883.

[38] Schaefer, G.; Rajab, M. I.; Celebi, M. E.; Iyatomi, H. (2011). Colour and contrast enhancement for improved skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 35 (2):99-104.

[39] Celebi, M. E.; Iyatomi, H.; Schaefer, G. (2009). *Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure*. In: 16th IEEE International Conference on Image Processing, Cairo, November 7-10 2009. IEEE, pp 2601-2604.

[40] Beuren, A. T.; Janasieivicz, R.; Pinheiro, G.; Grando, N.; Facon, J. (2012). *Skin melanoma segmentation by morphological approach*. In: International Conference on Advances in Computing, Communications and Informatics, Chennai, August 3-5 2012. ACM, pp 972-978.

[41] Barata, C.; Celebi, M. E.; Marques, J. S. (2015). Improving dermoscopy image classification using color constancy. *IEEE Journal of Biomedical and Health Informatics*, 19 (3):1146-1152.

[42] Barata, C.; Celebi, M. E.; Marques, J. S. (2015). *Towards a Robust Analysis of Dermoscopy Images Acquired under Different Conditions*. In: Celebi, M. E., Mendonca, T., Marques, J. S. (eds) Dermoscopy Image Analysis. CRC Press, Boca Raton, pp 1-22.

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

[44] Barcelos, C. A. Z.; Pires, V. B. (2009). An automatic based nonlinear diffusion equations scheme for skin lesion segmentation. *Applied Mathematics and Computation*, 215 (1):251-261.

[45] Celebi, M. E.; Iyatomi, H.; Schaefer, G.; Stoecker, W. V. (2009). Approximate lesion localization in dermoscopy images. *Skin Research and Technology*, 15 (3):314-322.

[46] Zhou, H.; Chen, M.; Gass, R.; Rehg, J. M.; Ferris, L.; Ho, J.; Drogowski, L. (2008). *Feature-preserving artifact removal from dermoscopy images*. In: SPIE Medical Imaging 2008 Conference, San Diego, February 16 - 21 2008. International Society for Optics and Photonics, pp 69141B69141-69141B69149.

[47] Wighton, P.; Lee, T. K.; Atkins, M. S. (2008). *Dermascopic hair disocclusion using inpainting*. In: SPIE Medical Imaging 2008 Conference, San Diego, February 16 - 21 2008. International Society for Optics and Photonics, pp 6914271-6914278.

[48] Xie, F.-Y.; Qin, S.-Y.; Jiang, Z.-G.; Meng, R.-S. (2009). PDE-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma. *Computerized Medical Imaging and Graphics*, 33 (4):275-282.

[49] Kiani, K.; Sharafat, A. R. (2011). E-shaver: An improved DullRazor® for digitally removing dark and light-colored hairs in dermoscopic images. *Computers in Biology and Medicine*, 41 (3):139-145.

[50] Abbas, Q.; Celebi, M. E.; García, I. F. (2011). Hair removal methods: a comparative study for dermoscopy images. *Biomedical Signal Processing and Control*, 6 (4):395-404.

[51] Abbas, Q.; Garcia, I. F.; Emre Celebi, M.; Ahmad, W. (2013). A Feature-Preserving Hair Removal Algorithm for Dermoscopy Images. *Skin Research and Technology*, 19 (1):e27-e36.

[52] Toossi, M. T. B.; Pourreza, H. R.; Zare, H.; Sigari, M. H.; Layegh, P.; Azimi, A. (2013). An effective hair removal algorithm for dermoscopy images. *Skin Research and Technology*, 19 (3):230-235.

[53] Mirzaalian, H.; Lee, T. K.; Hamarneh, G. (2014). Hair Enhancement in Dermoscopic Images Using Dual-Channel Quaternion Tubularness Filters and MRF-Based Multilabel Optimization. *IEEE Transactions on Image Processing*, 23 (12):5486-5496.

[54] Lee, T.; Ng, V.; Gallagher, R.; Coldman, A.; McLean, D. (1997). Dullrazor®: A software approach to hair removal from images. *Computers in Biology and Medicine*, 27 (6):533-543.

[55] Celebi, M. E.; Iyatomi, H.; Schaefer, G.; Stoecker, W. V. (2009). Lesion border detection in dermoscopy images. *Computerized Medical Imaging and Graphics*, 33 (2):148-153.

[56] Celebi, M. E.; Wen, Q.; Iyatomi, H.; Shimizu, K.; Zhou, H.; Schaefer, G. (2015). *A State-of-the-Art Survey on Lesion Border Detection in Dermoscopy Images*. In: Celebi, M. E., Mendonca, T., Marques, J. S. (eds) Dermoscopy Image Analysis. CRC Press, Boca Raton, pp 97-129.

[57] Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. 2 edn. Academic press, San Diego.

[58] Webb, A. R. (2003). *Statistical pattern recognition*. 2 edn. John Wiley & Sons, England.

[59] Guyon, I.; Gunn, S.; Nikravesh, M.; Zadeh, L. (2006). *Feature extraction: foundations and applications*, vol 207. Studies in Fuzziness and Soft Computing. Springer-Verlag, Berlin, Heidelberg.

[60] Liu, H.; Motoda, H. (1998). *Feature extraction, construction and selection: A data mining perspective*. Springer, Norwell.

[61] Ma, L.; Staunton, R. C. (2013). Analysis of the contour structural irregularity of skin lesions using wavelet decomposition. *Pattern Recognition*, 46 (1):98-106.

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

[63] Maglogiannis, I.; Delibasis, K. K. (2015). Enhancing classification accuracy utilizing globules and dots features in digital dermoscopy. *Computer Methods and Programs in Biomedicine*, 118 (2):124-133.

[64] Shrestha, B.; Bishop, J.; Kam, K.; Chen, X.; Moss, R. H.; Stoecker, W. V.; Umbaugh, S.; Stanley, R. J.; Celebi, M. E.; Marghoob, A. A. (2010). Detection of atypical texture features in early malignant melanoma. *Skin Research and Technology*, 16 (1):60-65.

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

[66] Blum, A.; Rassner, G.; Garbe, C. (2003). Modified ABC-point list of dermoscopy: A simplified and highly accurate dermoscopic algorithm for the diagnosis of cutaneous melanocytic lesions. *Journal of the American Academy of Dermatology*, 48 (5):672-678.

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

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

[69] Argenziano, G.; Fabbrocini, G.; Carli, P.; De Giorgi, V.; Sammarco, E.; Delfino, M. (1998). Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: Comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Archives of Dermatology*, 134 (12):1563-1570.

[70] Cavalcanti, P. G.; Scharcanski, J. (2011). Automated prescreening of pigmented skin lesions using standard cameras. *Computerized Medical Imaging and Graphics*, 35 (6):481-491.

[71] Situ, N.; Yuan, X.; Zouridakis, G. (2011). Assisting Main Task Learning by Heterogeneous Auxiliary Tasks with Applications to Skin Cancer Screening. *Journal of machine learning research*, 15:688-697.

[72] Sadeghi, M.; Lee, T. K.; McLean, D.; Lui, H.; Atkins, M. S. (2012). *Global pattern analysis and classification of dermoscopic images using textons*. In: SPIE 8314, Medical Imaging 2012: Image Processing, San Diego, February 4-9 2012. International Society for Optics and Photonics, pp 83144X83141-83144X83146.

[73] Abbas, Q.; Celebi, M. E.; Serrano, C.; Fondón García, I.; Ma, G. (2013). Pattern classification of dermoscopy images: A perceptually uniform model. *Pattern Recognition*, 46 (1):86-97.

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

[75] Argenziano, G.; Soyer, H. P.; Chimenti, S.; Talamini, R.; Corona, R.; Sera, F.; Binder, M.; Cerroni, L.; De Rosa, G.; Ferrara, G.; Hofmann-Wellenhof, R.; Landthaler, M.; Menzies, S. W.; Pehamberger, H.; Piccolo, D.; Rabinovitz, H. S.; Schiffner, R.; Staibano, S.; Stolz, W.; Bartenjev, I.; Blum, A.; Braun, R.; Cabo, H.; Carli, P.; De Giorgi, V.; Fleming, M. G.; Grichnik, J. M.; Grin, C. M.; Halpern, A. C.; Johr, R.; Katz, B.; Kenet, R. O.; Kittler, H.; Kreusch, J.; Malvehy, J.; Mazzocchetti, G.; Oliviero, M.; Özdemir, F.; Peris, K.; Perotti, R.; Perusquia, A.; Pizzichetta, M. A.; Puig, S.; Rao, B.; Rubegni, P.; Saida, T.; Scalvenzi, M.; Seidenari, S.; Stanganelli, I.; Tanaka, M.; Westerhoff, K.; Wolf, I. H.; Braun-Falco, O.; Kerl, H.; Nishikawa, T.; Wolff, K.; Kopf, A. W. (2003). Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet. *Journal of the American Academy of Dermatology*, 48 (5):679-693.

[76] Argenziano, G.; Soyer, H.; De Giorgi, V.; Piccolo, D.; Carli, P.; Delfino, M.; al., e. (2002). Dermoscopy: A tutorial. *EDRA Medical Publishing & NewMedia*.

[77] Celebi, M. E.; Iyatomi, H.; Stoecker, W. V.; Moss, R. H.; Rabinovitz, H. S.; Argenziano, G.; Soyer, H. P. (2008). Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, 32 (8):670-677.

[78] Betta, G.; Di Leo, G.; Fabbrocini, G.; Paolillo, A.; Scalvenzi, M. (2005). *Automated Application of the "7-point checklist" Diagnosis Method for Skin Lesions: Estimation of Chromatic and Shape Parameters*. In: Instrumentation and Measurement Technology Conference, Ottawa, May 16-19 2005. IEEE, pp 1818-1822.

[79] Leo, G. D.; Fabbrocini, G.; Paolillo, A.; Rescigno, O.; Sommella, P. (2009). *Towards an automatic diagnosis system for skin lesions: estimation of blue-whitish veil and regression structures*. In: International Multi-Conference on Systems, Signals and Devices Djerba, March 23-26 2009. IEEE, pp 1-6.

[80] Zortea, M.; Schopf, T. R.; Thon, K.; Geilhufe, M.; Hindberg, K.; Kirchesch, H.; Møllersen, K.; Schulz, J.; Skrøvseth, S. O.; Godtliebsen, F. (2014). Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists. *Artificial intelligence in medicine*, 60 (1):13-26.

[81] Silva, C. S.; Marcal, A. R. (2013). Colour-based dermoscopy classification of cutaneous lesions: an alternative approach. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1 (4):211-224.

[82] Giotis, I.; Molders, N.; Land, S.; Biehl, M.; Jonkman, M. F.; Petkov, N. (2015). MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert Systems with Applications*, 42 (19):6578-6585.

[83] Barata, C.; Emre Celebi, M.; Marques, J. S. (2015). *Melanoma detection algorithm based on feature fusion*. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society Milan, August 25-29 2015. IEEE, pp 2653-2656.

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

[85] Barata, C.; Marques, J. S.; Celebi, M. E. (2013). *Towards an automatic bag-of-features model for the classification of dermoscopy images: The influence of segmentation*. In: 8th International Symposium on Image and Signal Processing and Analysis Trieste, September 4-6 2013. IEEE, pp 274-279.

[86] Sáez, A.; Acha, B.; Serrano, C. (2014). *Pattern Analysis in Dermoscopic Images*. In: Scharcanski, J., Celebi, M. E. (eds) Computer Vision Techniques for the Diagnosis of Skin Cancer. Series in BioEngineering. Springer, Berlin, Heidelberg, pp 23-48.

[87] Chang, Y.; Stanley, R. J.; Moss, R. H.; Van Stoecker, W. (2005). A systematic heuristic approach for feature selection for melanoma discrimination using clinical images. *Skin Research and Technology*, 11 (3):165-178.

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

[89] Ng, V. T. Y.; Fung, B. Y. M.; Lee, T. K. (2005). Determining the asymmetry of skin lesion with fuzzy borders. *Computers in Biology and Medicine*, 35 (2):103-120.

[90] Oliveira, R. B.; Marranghello, N.; Pereira, A. S.; Tavares, J. M. R. S. (2016). A computational approach for detecting pigmented skin lesions in macroscopic images. *Expert Systems with Applications*, 61:53-63.

[91] D'Amico, M.; Ferri, M.; Stanganelli, I. (2004). *Qualitative Asymmetry Measure for Melanoma Detection*. In: IEEE International Symposium on Biomedical Imaging: Nano to Macro, Arlington, April 15-18, 2004. IEEE, pp 1155-1158.

[92] Lee, T. K.; McLean, D. I.; Atkins, M. S. (2003). Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. *Medical image analysis*, 7 (1):47-64.

[93] Iyatomi, H.; Norton, K.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Classification of melanocytic skin lesions from non-melanocytic lesions*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5407-5410.

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

[95] Jaworek-Korjakowska, J. (2015). Novel method for border irregularity assessment in dermoscopic color images. *Computational and mathematical methods in medicine*, 2015 (Article ID 496202):1-11.

[96] Iyatomi, H.; Oka, H.; Celebi, M. E.; Ogawa, K.; Argenziano, G.; Soyer, H. P.; Koga, H.; Saida, T.; Ohara, K.; Tanaka, M. (2008). Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. *Journal of Investigative Dermatology*, 128 (8):2049-2054.

[97] Dalal, A.; Moss, R. H.; Stanley, R. J.; Stoecker, W. V.; Gupta, K.; Calcara, D. A.; Xu, J.; Shrestha, B.; Drugge, R.; Malters, J. M.; Perry, L. A. (2011). Concentric decile segmentation of white and hypopigmented areas in dermoscopy images of skin lesions allows discrimination of malignant melanoma. *Computerized Medical Imaging and Graphics*, 35 (2):148-154.

[98] Mendoza, C. S.; Serrano, C.; Acha, B. (2009). *Scale invariant descriptors in pattern analysis of melanocytic lesions*. In: 16th IEEE International Conference on Image Processing, Cairo, November 7-10 2009. IEEE, pp 4193-4196.

[99] Khan, A.; Gupta, K.; Stanley, R. J.; Stoecker, W. V.; Moss, R. H.; Argenziano, G.; Soyer, H. P.; Rabinovitz, H. S.; Cognetta, A. B. (2009). Fuzzy logic techniques for blotch feature evaluation in dermoscopy images. *Computerized Medical Imaging and Graphics*, 33 (1):50-57.

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

[101] Sadeghi, M.; Razmara, M.; Wighton, P.; Lee, T. K.; Atkins, M. S. (2010). *Modeling the dermoscopic structure pigment network using a clinically inspired feature set*. In: Medical Imaging and Augmented Reality. Springer, pp 467-474.

[102] Møllersen, K.; Zortea, M.; Hindberg, K.; Schopf, T. R.; Skrøvseth, S. O.; Godtliebsen, F. (2015). *Improved skin lesion diagnostics for general practice by*

computer-aided diagnostics. In: Celebi, M. E., Mendonca, T., Marques, J. S. (eds) Dermoscopy Image Analysis. CRC Press, Boca Raton, pp 247-292.

[103] Abbas, Q.; Celebi, M. E.; Garcia, I. F.; Ahmad, W. (2013). Melanoma recognition framework based on expert definition of ABCD for dermoscopic images. *Skin Research and Technology*, 19 (1):e93-e102.

[104] Zhou, Y.; Smith, M.; Smith, L.; Warr, R. (2010). A new method describing border irregularity of pigmented lesions. *Skin Research and Technology*, 16 (1):66-76.

[105] Shimizu, K.; Iyatomi, H.; Celebi, M. E.; Norton, K.-A.; Tanaka, M. (2015). Four-class classification of skin lesions with task decomposition strategy. *IEEE Transactions on Biomedical Engineering*, 62 (1):274-283.

[106] Clawson, K. M.; Morrow, P.; Scotney, B.; McKenna, J.; Dolan, O. (2009). *Analysis of pigmented skin lesion border irregularity using the harmonic wavelet transform*. In: 13th International Machine Vision and Image Processing Conference Dublin, September 2-4 2009. IEEE, pp 18-23.

[107] Schmid-Saugeon, P. (2000). Symmetry axis computation for almost-symmetrical and asymmetrical objects: application to pigmented skin lesions. *Medical image analysis*, 4 (3):269-282.

[108] Maglogiannis, I.; Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine*, 13 (5):721-733.

[109] Rahman, M. M.; Bhattacharya, P.; Desai, B. C. (2008). *A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions*. In: 8th IEEE International Conference on International Conference on BioInformatics and BioEngineering, Athens, October 8-10 2008. IEEE, pp 1-6.

[110] Schaefer, G.; Krawczyk, B.; Celebi, M. E.; Iyatomi, H. (2014). An ensemble classification approach for melanoma diagnosis. *Memetic Computing*, 6 (4):233-240.

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

[112] Abedini, M.; Chen, Q.; Codella, N. C. F.; Garnavi, R.; Sun, X. (2015). *Accurate and Scalable System for Automatic Detection of Malignant Melanoma*. In: Celebi, M. E., Mendonca, T., Marques, J. S. (eds) Dermoscopy Image Analysis. CRC Press, Boca Raton, pp 293-343.

[113] Iyatomi, H.; Celebi, M. E.; Schaefer, G.; Tanaka, M. (2011). Automated color calibration method for dermoscopy images. *Computerized Medical Imaging and Graphics*, 35 (2):89-98.

[114] Celebi, M. E.; Zornberg, A. (2014). Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification. *IEEE Systems Journal*, 8 (3):980-984.

[115] Barata, C.; Figueiredo, M. A.; Celebi, M. E.; Marques, J. S. (2014). *Color identification in dermoscopy images using gaussian mixture models*. In: IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, May 4-9 2014. IEEE, pp 3611-3615.

[116] Barata, C.; Celebi, M. E.; Marques, J. S. (2015). *Color Detection in Dermoscopy Images Based on Scarce Annotations*. In: 7th Iberian Conference on Pattern Recognition and Image Analysis, Santiago de Compostela, June 17-19 2015. Springer, pp 309-316.

[117] Stoecker, W. V.; Gupta, K.; Stanley, R. J.; Moss, R. H.; Shrestha, B. (2005). Detection of asymmetric blotches (asymmetric structureless areas) in dermoscopy images of malignant melanoma using relative color. *Skin Research and Technology*, 11 (3):179-184.

[118] Madooei, A.; Drew, M. S. (2013). *A Colour Palette for Automatic Detection of Blue-White Veil*. In: 21st Color and Imaging Conference Final Program and Proceedings, Albuquerque, November 4-8 2013. Society for Imaging Science and Technology, pp 200-205.

[119] Madooei, A.; Drew, M. S.; Sadeghi, M.; Atkins, M. S. (2013). *Automatic detection of blue-white veil by discrete colour matching in dermoscopy images*. In: 16th International Conference on Medical Image Computing and Computer-Assisted Intervention, Nagoya, September 22-26 2013. Springer, pp 453-460.

[120] Madasu, V. K.; Lovell, B. C. (2009). *Blotch Detection in Pigmented Skin Lesions Using Fuzzy Co-clustering and Texture Segmentation*. In: Digital Image Computing: Techniques and Applications, Melbourne, December 1-3 2009. IEEE, pp 25-31.

[121] Tanaka, T.; Torii, S.; Kabuta, I.; Shimizu, K.; Tanaka, M. (2008). Pattern classification of nevus with texture analysis. *IEEJ Transactions on Electrical and Electronic Engineering*, 3 (1):143-150.

[122] Anantha, M.; Moss, R. H.; Stoecker, W. V. (2004). Detection of pigment network in dermatoscopy images using texture analysis. *Computerized Medical Imaging and Graphics*, 28 (5):225-234.

[123] Yuan, X.; Yang, Z.; Zouridakis, G.; Mullani, N. (2006). *SVM-based texture classification and application to early melanoma detection*. In: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, New York, August 30 - September 3 2006. IEEE, pp 4775-4778.

[124] Serrano, C.; Acha, B. (2009). Pattern analysis of dermoscopic images based on Markov random fields. *Pattern Recognition*, 42 (6):1052-1057.

[125] Amelard, R.; Glaister, J.; Wong, A.; Clausi, D. A. (2015). High-level intuitive features (HLIFs) for intuitive skin lesion description. *IEEE Transactions on Biomedical Engineering*, 62 (3):820-831.

[126] Torre, E. L.; Caputo, B.; Tommasi, T. (2010). Learning methods for melanoma recognition. *International Journal of Imaging Systems and Technology*, 20 (4):316-322.

[127] Huang, H.; Bergstresser, P. (2007). *A new hybrid technique for dermatological image registration*. In: 7th IEEE International Conference on Bioinformatics and Bioengineering Boston, October 14-17 2007. IEEE, pp 1163-1167.

[128] Skrovseth, S. O.; Schopf, T. R.; Thon, K.; Zortea, M.; Geilhufe, M.; Mollersen, K.; Kirchesch, H. M.; Godtliebsen, F. (2010). *A computer aided diagnostic system for malignant melanomas*. In: 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies Rome, November 7-10 2010. IEEE, pp 1-5.

[129] Jaworek-Korjakowska, J.; Tadeusiewicz, R. (2014). *Determination of border irregularity in dermoscopic color images of pigmented skin lesions*. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, August 26-30 2014. IEEE, pp 6459-6462.

[130] Xie, X. (2008). A review of recent advances in surface defect detection using texture analysis techniques. *Electronic Letters on Computer Vision and Image Analysis*, 7 (3):1-22.

[131] Haralick, R. M.; Shanmugam, K.; Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3 (6):610-621.

[132] Al-Akaidi, M. (2004). *Fractal speech processing*. Cambridge university press, New York.

[133] Strayer, S. M.; Reynolds, P. (2003). Diagnosing skin malignancy: assessment of predictive clinical criteria and risk factors. *The Journal of family practice*, 52 (3):210-218.

[134] Hani, A. F. M.; Fitriyah, H.; Prakasa, E.; Asirvadam, V. S.; Hussein, S. H.; Azura, M. A. (2010). *In vivo 3D thickness measurement of skin lesion*. In: IEEE Conference on Biomedical Engineering and Sciences, Kuala Lumpur, November 30 - December 2 2010. IEEE, pp 155-160.

[135] Fadzil, M. A.; Fitriyah, H.; Prakasa, E.; Nugroho, H.; Hussein, S. H.; Affandi, A. M. (2009). *Thickness Characterization of 3D Skin Surface Images Using Reference Line Construction Approach*. In: International Visual Informatics Conference, Kuala Lumpur, November 11-13 2009. Springer, pp 448-454.

[136] Mirzaalian, H.; Lee, T. K.; Hamarneh, G. (2016). Skin lesion tracking using structured graphical models. *Medical image analysis*, 27:84-92.

[137] Guyon, I.; Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157-1182.

[138] Han, J.; Kamber, M. (2006). *Data Mining: concepts and techniques*. Elsevier, San Francisco.

[139] Dash, M.; Liu, H. (1997). Feature selection for classification. *Intelligent data analysis*, 1 (3):131-156.

[140] Witten, I. H.; Frank, E.; Hall, M. A. (2011). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco.

[141] Liu, H.; Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, 17 (4):491-502.

[142] Kohavi, R.; John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, 97 (1):273-324.

[143] Hand, D.; Mannila, H.; Smyth, P. (2001). *Principles of Data Mining*. The MIT Press, London.

[144] Chawla, N. V. (2005). *Data mining for imbalanced datasets: An overview*. In: Maimon, O., Rokach, L. (eds) Data mining and knowledge discovery handbook. Springer, New York, pp 853-867.

[145] Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321-357.

[146] Congdon, P. (2007). *Bayesian statistical modelling*, vol 704. 2 edn. John Wiley & Sons, Chichester.

[147] Haykin, S. S. (1999). *Neural networks: a comprehensive foundation*. Prentice Hall, Englewood Cliffs, USA.

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

[149] Cavalcanti, P. G.; Scharcanski, J.; Baranoski, G. V. (2013). A two-stage approach for discriminating melanocytic skin lesions using standard cameras. *Expert Systems with Applications*, 40 (10):4054-4064.

[150] Arroyo, J. L. G.; Zapirain, B. G.; Zorrilla, A. M. (2011). *Blue-white veil and dark-red patch of pigment pattern recognition in dermoscopic images using machine-learning techniques*. In: IEEE International Symposium on Signal Processing and Information Technology, Bilbao, December 14-17 2011. IEEE, pp 196-201.

[151] Fabbrocini, G.; Betta, G.; Di Leo, G.; Liguori, C.; Paolillo, A.; Pietrosanto, A.; Sommella, P.; Rescigno, O.; Cacciapuoti, S.; Pastore, F. (2010). Epiluminescence image processing for melanocytic skin lesion diagnosis based on 7-point check-list: a preliminary discussion on three parameters. *The open dermatology journal*, 4:110-115.

[152] Mirzaalian, H.; Lee, T. K.; Hamarneh, G. (2012). *Learning features for streak detection in dermoscopic color images using localized radial flux of principal intensity curvature*. In: Workshop on Mathematical Methods in Biomedical Image Analysis, Breckenridge, January 9-10 2012. IEEE, pp 97-101.

[153] Schaefer, G.; Krawczyk, B.; Celebi, M. E.; Iyatomi, H.; Hassanien, A. E. (2014). *Melanoma classification based on ensemble classification of dermoscopy image features*.

In: International Conference on Advanced Machine Learning Technologies and Applications, Cairo, November 28-30 2014. Springer, pp 291-298.

[154] Sadeghi, M.; Lee, T. K.; McLean, D.; Lui, H.; Atkins, M. S. (2013). Detection and analysis of irregular streaks in dermoscopic images of skin lesions. *IEEE Transactions on Medical Imaging*, 32 (5):849-861.

[155] Fleming, M. G.; Steger, C.; Zhang, J.; Gao, J.; Cognetta, A. B.; Dyer, C. R. (1998). Techniques for a structural analysis of dermatoscopic imagery. *Computerized Medical Imaging and Graphics*, 22 (5):375-389.

[156] Abedini, M.; Codella, N. C. F.; Connell, J. H.; Garnavi, R.; Merler, M.; Pankanti, S.; Smith, J. R.; Syeda-Mahmood, T. (2015). A generalized framework for medical image classification and recognition. *IBM Journal of Research and Development*, 59 (2-3):1-18.

[157] Dietterich, T. G. (2000). *Ensemble methods in machine learning*. In: Multiple Classifier Systems, vol 1857. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 1-15.

[158] Breiman, L. (2001). Random forests. *Machine learning*, 45 (1):5-32.

[159] Zhu, J.; Zou, H.; Rosset, S.; Hastie, T. (2009). Multi-class adaboost. *Statistics and its Interface*, 2 (3):349-360.

[160] Fawcett, T. (2004). ROC graphs: Notes and practical considerations for researchers. *Machine learning*, 31 (1):1-38.

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

[162] Abbas, Q.; Celebi, M. E.; Fondón, I. (2012). Computer-aided pattern classification system for dermoscopy images. *Skin Research and Technology*, 18 (3):278-289.

[163] Barhoumi, W.; Baâzaoui, A. (2014). Pigment network detection in dermatoscopic images for melanoma diagnosis. *IRBM*, 35 (3):128-138.

[164] Leo, G. D.; Liguori, C.; Paolillo, A.; Sommella, P. (2008). *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, Istanbul, July 14-16 2008. IEEE, pp 190-194.

[165] Grana, C.; Cucchiara, R.; Pellacani, G.; Seidenari, S. (2006). *Line detection and texture characterization of network patterns*. In: 18th International Conference on Pattern Recognition, Hong Kong, August 20-24 2006. IEEE, pp 275 - 278.

[166] Betta, G.; Di Leo, G.; Fabbrocini, G.; Paolillo, A.; Sommella, P. (2006). *Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern*. In: IEEE International Workshop on Medical Measurement and Applications, Benevento, April 20-21 2006. IEEE, pp 63-67.

[167] Celebi, M. E.; Schaefer, G. (2012). *Color medical image analysis*, vol 6. Springer, Dordrecht.

[168] Celebi, M. E.; Mendonca, T.; Marques, J. S. (2015). *Dermoscopy Image Analysis*, vol 10. CRC Press, Boca Raton.

[169] Mayer, J. (1997). Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma. *The Medical Journal of Australia*, 167 (4):206-210.

PART B - ARTICLE 3

A COMPUTATIONAL APPROACH FOR DETECTING PIGMENTED SKIN LESIONS IN MACROSCOPIC IMAGES

Roberta B. Oliveira, Norian Marranghello, Aledir S. Pereira and João Manuel R. S. Tavares

Published in: Expert Systems with Applications, 61:53-63, 2016

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. Computational methods have been developed to assist dermatologists in early diagnosis of skin cancer. Computational analysis of skin lesion images has become a challenging research area due to the difficulty in discerning some types of skin lesions. A novel computational approach is presented for extracting skin lesion features from images based on asymmetry, border, colour and texture analysis, in order to diagnose skin lesion types. The approach is based on an anisotropic diffusion filter, an active contour model without edges and a support vector machine. Experiments were performed regarding the segmentation and classification of pigmented skin lesions in macroscopic images, with the results obtained being very promising.

Keywords: Image pre-processing; Image segmentation; Image classification; Anisotropic diffusion filter; Active contour model without edges; Support vector machine.

1. Introduction

Computational analysis of skin lesion images is an area of great research interest due to its importance in skin cancer prevention, particularly in achieving a successful early diagnosis [1-3]. Such lesions, which can be classified as benign or malignant, are mainly due to abnormal production of melanocyte cells originating from factors such as excessive sun exposure. Melanocyte cells are responsible for creating the substance melanin, whose main function is to provide skin pigmentation. In the case of malignant cells, i.e. melanoma (Figure 1a), such cells divide quickly and may invade other parts of the body. An increasing number of deaths due to melanoma have been observed worldwide, since this type of malignant lesion is the most aggressive compared to other lesion types due to its high level of metastasis [4]. Benign lesions display a more organized structure than malignant lesions, since the former are unable to proliferate into other tissues. Seborrheic keratosis (Figure 1b) and melanocytic nevus (Figure 1c) are examples of benign lesions. However, these skin lesions have also been of global concern, since some types of nevi may become melanoma; moreover a melanoma may resemble a seborrheic keratosis or a nevus in its initial state.

Figure 1: Three examples of pigmented skin lesions: (a) melanoma, (b) seborrheic keratosis and (c) melanocytic nevus.

Different non-invasive imaging techniques have been employed to assist dermatologists in diagnosing skin lesions [5]. Macroscopic images, commonly known as clinical images, are normally used in computational analysis of skin lesions [6,7], since such images may be obtained using common digital video or image cameras. Figure 1 presents examples of macroscopic images. However, their imaging conditions are frequently inconsistent; for example, images are acquired from variable distances and/or under different illumination conditions. Furthermore, the images may have poor resolution, which may be challenging when the lesion under study 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 adequate analysis of the imaged skin lesions.

Pre-processing, segmentation, feature extraction, and classification are fundamental steps commonly found in computational systems of image analysis. In terms of skin lesions, the image pre-processing step is an important aspect for good segmentation, i.e. identification, of the image's pigmented skin lesions. Effective approaches based on colour space transformation [8], contrast enhancement [9] and artefact removal [10] have been proposed for this step in order to improve the accuracy of the following segmentation step. Segmentation allows for extracting the region of interest (ROI) from the macroscopic image under analysis. Previous studies [6,11,12] have shown that computational methods for image segmentation may provide suitable results for the identification of skin lesions in images.

The feature extraction of skin lesion images is usually based on methods used by dermatologists in their clinical routine diagnosis. Of these methods, the ABCD rule is mostly used, being a criteria based on the Asymmetry, Border, Colour and Diameter characteristics of the lesion under study [13]. The asymmetry criterion may be examined

by dividing the region of the lesion into two sub-regions by an axis of symmetry, in order to analyse the similarity of the area by overlapping the two sub-regions along the axis. The lesion is 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. The border criterion corresponds to the measure of the regularity of the lesion's shape. According to this criterion, a border of regular shape is associated with benign lesions while a border of irregular shape is associated with malignant lesions instead. The colour criterion consists of analysing the tonality variation of the pigmented skin lesions in order to identify the malignant lesions, which usually present non-uniform colours. The diameter criterion is associated with the size of the lesion and is defined by the greatest distance between any two points of the lesion's border. As such, a diameter equal to or greater than 6 (six) millimetres may indicate malignancy. Texture analysis may also be performed for image-based examination of skin lesions, since it assists in discriminating benign from malignant lesions by assessing the roughness of their structure [7].

Several computational solutions [1,14] have been proposed for extracting features from pigmented skin lesions in images in order to represent them according to certain criteria. Then, the classification step consists of recognizing and interpreting the information about the pigmented skin lesions based on these features. Hence, computational classifiers are important tools to assist the computational diagnosis of skin lesions in macroscopic images [15-17].

The objective of this work was to develop a novel computational approach based on the ABCD rule and texture analysis for the identification and classification of pigmented skin lesions in macroscopic images, in order to provide information that may assist dermatologists in their diagnosis. In this approach, an anisotropic diffusion filter [18] is applied to reduce the noise present in the image under study. Then, the active contour model without edges [19] is employed in the segmentation of the lesion in the pre-processed image. Afterwards, features related to the asymmetry, border, colour and texture of the lesion are extracted from the segmented image. Finally, the features are used as input to a support vector machine (SVM) classifier [20] to classify the skin lesion.

This paper is organized as follows: a review of the computational methods that have been applied to classify pigmented skin cancers and other skin lesions is provided in Section 2. A novel approach for detecting and classifying skin lesions in macroscopic

images is proposed in Section 3. The results and their discussion are provided in Section 4. Finally, conclusions drawn and proposal for future studies are in the last section.

2. Related studies

Computer-aided diagnosis (CAD) systems for skin lesions in images have been proposed in order to assist dermatologists, predominantly in the early assessment of skin cancer. In these systems, image filters are commonly applied to pre-process the input images in order to increase the accuracy of the segmentation step. A median filter, which is a non-linear image filtering algorithm, has been applied often to smooth images of skin lesions as well as to remove artefacts, preserving the border of the lesion, which is imperative to assure adequate segmentation [1,11,21]. An anisotropic diffusion filter has also been regularly used for smoothing skin lesion images, particularly to remove artefacts with good results and without losing relevant information about lesions [22]. Based on set theory, morphological filtering [23] enables removing image noise [11,24], and may also be used to enhance skin lesions in images [25], as well as to include areas with borders of low contrast in previously detected lesion regions [24,26].

Algorithms of image segmentation have been developed based on several techniques to assist the diagnosis of skin lesions from images [27]. From these, threshold-based algorithms have been widely used, mainly because of their simplicity. Thus, thresholding algorithms, such as the Otsu [1,24,26,28], type-2 fuzzy logic [29] and the Renyi entropy method [25], aim to establish the threshold values in order to separate the regions of interest (ROIs) in the input images. However, these techniques may reveal some problems; for example the segmented lesions tend to be smaller than their real size, and the segmentation process may lead to highly irregular lesion borders [29].

Algorithms based on active contour models (ACM) have been frequently proposed for the segmentation of skin lesions in images [8,11,12]. In these algorithms, initial curves move toward the boundaries of the objects of interest through appropriate deformation. The algorithms of active contour may be classified as edge- or region-based models [30] according to the technique used to track the curves movement. Additionally, mixed models have been also adopted, see, for example, Li et al. [31]. The edge-based models include classic parametric models [32], gradient vector flow (GVF) [33] and geometric (or geodesic) active contours (GAC) [34]. However, classic parametric models and GVF have difficulty in dealing with topological changes and large curvatures. On the other

hand, GAC models, such as level-set-based algorithms, do not present such problems. The region-based active contour model proposed by Chan and Vese [19] has been used in the segmentation of skin lesions [11] due to its advantages relatively to other segmentation algorithms based on ACM [19], such as: 1) the initial curve may be defined more freely in the input image, 2) the inner contours are automatically detected without the need to define additional curves in the image, and 3) the segmentation is successfully carried out even in the presence of intensity variations, very smooth boundaries and boundaries not successfully detected by gradient operators. Region-based algorithms, like the region growing, splitting and merging methods have also been used to segment skin lesions in images [1,11,17,35]. These methods consist of grouping similar neighbouring pixels, or sub-regions, into larger homogeneous regions according to a growing criterion. Such methods have shown successful performance even under complex conditions such as great variations of illumination and colour. However, some of these methods may not adequately identify lesion regions that present low contrast relatively to the skin background.

The wide use of algorithms based on artificial intelligence (AI) is justified by the advantages they offer [11], such as the possibility of learning from sample cases provided by artificial neural networks (ANNs) [9], the search and optimization for the best segmentation results provided by techniques based on genetic algorithms (GAs) [36], and the capability of dealing with imprecise values provided by fuzzy logic, e.g., by applying the type-2 fuzzy logic technique [29]. In addition, fuzzy logic combined with clustering techniques have been employed in the image segmentation of skin lesions, such as the fuzzy c-means (FCM) algorithm [37] and the anisotropic mean shift approach based on the FCM algorithm (AMSFCM) [38]. The hill-climbing algorithm (HCA) is a technique based on the clustering of points on an image, which is also applied to detect ROIs in skin lesion images [39]. In Abbas et al. [40], a new segmentation method based on dynamic programming was proposed to overcome the limitation of thresholding, region-growing and clustering, as well as level-set-based segmentation methods. However, some algorithms based on AI may also present disadvantages regarding the complexity of their implementation and the presence of unnecessary steps, which requires high computational efforts [36].

The ABCD rule and texture analysis are examples of approaches employed in the literature for the computational analysis of skin lesions in macroscopic images. However,

other descriptors have also been extracted for the characterization of skin lesions in images:

- Asymmetry (A): asymmetry index descriptors based on axis of symmetry [41,42], and geometrical descriptors [7];
- Border (B): geometrical descriptors based on the best-fit of ellipse axes [41,42], and statistical descriptors based on border gradient and edge regions [7];
- Colour (C): statistical descriptors based on colour models [7,41,42], amount of colour pixels [7], and relative colour descriptors [41];
- Diameter (D): semi-major axis of the best-fit ellipse [42]; and
- Texture (T): statistical descriptor based on the intensity of the pixels inside the lesion regions [7].

For the classification process, one or several methods have been evaluated to achieve the best results. The performance of this process depends on several issues, such as the quality of the segmented image and extracted features, as well as on the classification method used. The output of the skin lesion classification process may be binary or multi-class, and concern different classes according to the classification goal, e.g., malignancy of the lesions (benign versus malignant) [43], and distinct types of skin lesions (melanoma versus nevus [16,17], melanocytic versus non-melanocytic [14], and dysplastic versus non-dysplastic versus melanotic [16]). Furthermore, skin lesion features, such as border features (regular versus irregular) [44] can also be classified.

Classification methods based on a decision tree have been used in the classification of skin lesions by many authors [15,16,41]. The simplicity of the classification structure in terms of ease of understanding and visualization, as well as the easy rule generation, is one of the important advantages of this approach. However, the difficulties in dealing with correlated features and the possibility of excessive adjustments (over-fitting) are its major drawbacks. Bayesian learning-based methods have also been applied to classify skin lesions [16,43]. Although Bayesian methods provide fast training and no sensitivity to irrelevant features, they assume that the features are independent. Despite the long training time, artificial neural networks have been proposed in various studies [16,17] to cope with many complex pattern recognition problems, since such classifiers present good capability and flexibility to solve several non-separable problems. The SVM

classifier [20] 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 [1,16]. The SVM performed better than other computer classifiers in several studies [16]. However, this classifier may be sensitive to noise and the classification process is binary.

3. Proposed approach

In this section, a computational approach for identification and classification of pigmented skin lesions in macroscopic images is presented, in order to provide information that may assist dermatologists in their diagnosis. Figure 2 illustrates the pipeline of the proposed approach, which involves the following steps: 1) image pre-processing, 2) image segmentation, 3) image post-processing, 4) feature extraction, and 5) lesion classification. The first step is mainly applied to deal with noisy images based on an anisotropic diffusion filter [18]. The second step is responsible for identifying the lesion presented in the image being studied by using an active contour model without edges, known as Chan-Vese's model [19]. The third step consists of the post-processing of the segmented image based on morphological filtering [23] in order to improve the quality of the segmentation result. In the fourth step, features are extracted from the identified lesion, including the lesion's asymmetry, border, colour and texture properties. Finally, the last step concerns the lesion classification based on the extracted features that are inputted into an SVM classifier [20]. In the next sections, each step of the proposed approach is detailed.

3.1. Image pre-processing

As mentioned previously, the image under analysis may contain several artefacts that can affect the accuracy of the image segmentation step. Hence, an anisotropic diffusion filter [18] is applied to smooth the input image, mainly in order to reduce the presence of hairs. Hence, initially, the original *RGB* (red, green, blue) image is converted into a grey-level image, since the segmentation method used is applied to grey-level images. Afterwards, the anisotropic diffusion filter is applied to the converted image according to the solution proposed by Barcelos et al. [18], which aims at smoothing very noisy images without removing relevant borders.

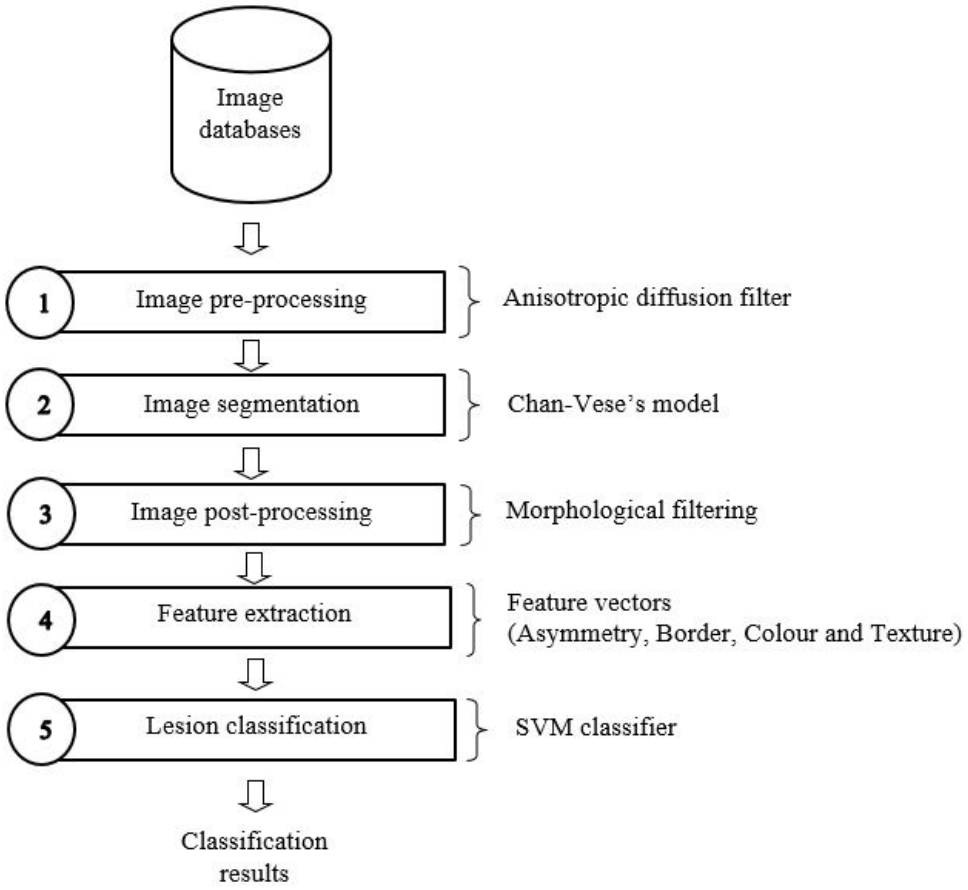

Figure 2: Pipeline of the proposed approach for detecting and classifying pigmented skin lesions in images.

The implementation of this filter is based on the following equations:

$$u_t = g(|(G_\sigma * \nabla u|)\,|\nabla \mathrm{u}|\; div\left(\frac{\nabla u}{|\nabla \mathrm{u}|}\right) - \lambda\,(1-g)(u-I)\,, \tag{1}$$

$$g(|(G_\sigma * \nabla u|) = \frac{1}{1+k|(G_\sigma * \nabla u)|^2}\,, \text{ and} \tag{2}$$

$$G_t(x,y) = \frac{1}{2\pi\sigma t^2} e^{\frac{-|x^2+y^2|}{2\sigma t^2}}\,, \tag{3}$$

where $u(x,y,0) = I(x,y)$, $x \in \Omega$, $t > 0$, $I(x,y)$ is the original image to be processed, u the smoothed image at scale t, div the divergence operator, ∇u the gradient of u, and λ a parameter related to the diffusion speed. The term $g(|(G_\sigma * \nabla u|)$ is used for border detection, where k is a parameter, G_σ the Gaussian function, and σ the standard deviation of G_σ. The convolution $G_\sigma * \nabla u$ is a Gaussian scale space of g given by: $T_g(x,y,t) = g * G_t(x,y)$ where G_t is given by Eq. (3) and t is the scale. Considering a neighbourhood of a pixel x, when the gradient ∇ has a low average value; i.e., there are few noisy pixels in the input image, x is considered an inner pixel (homogeneous region), resulting in $g \cong$

1. Otherwise, x will be a pixel of a contour, $g \cong 0$. The moderation selector $(1-g)$ [18] allows a balanced diffusion of the input image, i.e., the homogeneous regions are smoothed even more with respect to the borders of the regions. This filter is iteratively applied to the image, such that the number of iterations (NI) is determined according to the amount of noise presented in the input image. However, relevant borders may be removed when the number of iterations is too large.

3.2. Image segmentation

The segmentation process should be effective, so information of the lesion may be extracted with high confidence. In addition, the accuracy of this process directly influences the feature extraction step, which is required to suitably represent the lesion for its classification process. Therefore, an appropriate segmentation technique is crucial to obtain good classification results for the problem in question. The Chan-Vese model [19] is based on the average of the intensities of the image's pixels, and not on the image's gradient. This model uses the concepts of the Mumford-Shah and level-set segmentation models. Essentially, Chan-Vese's model considers a "fitting" term F for the energy minimization, which allows the deformation of the curve toward the boundary of the object to be segmented, in which the inside and outside intensities are constant and similar. In order to identify whether the object of interest is inside or outside the curve, the energy minimization $F(c_1, c_2, \phi)$ is calculated as:

$$F(c_1, c_2, \phi) = \mu \int_\Omega \delta\big(\phi(x,y)\big)|\nabla\phi(x,y)|dxdy + \nu \int_\Omega H\big(\phi(x,y)\big)dxdy + \lambda_1 \int_\Omega |u_0(x,y) - c_1|^2 H\big(\phi(x,y)\big)dxdy + \lambda_2 \int_\Omega |u_0(x,y) - c_2|^2 (1 - H(\phi(x,y)))\, dxdy\ , \quad (4)$$

where u_0 is a pre-processed image, as a bounded function on $\overline{\Omega}$ and with real values. The fixed parameters μ, $\nu \geq 0$, and λ_1 and $\lambda_2 > 0$ are weights for the fitting term. The terms H and δ are the Heaviside and Dirac delta functions, respectively, used in order to obtain the level-set energy function $F(c_1, c_2, \phi)$. The constants c_1 and c_2, which are based on the Mumford-Shah segmentation model, are the average image u_0 inside and outside curve C, respectively. Such constants are given by:

$$c_1(\phi) = \frac{\int_\Omega u_0(x,y) H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy}, \quad (5)$$

$$c_2(\phi) = \frac{\int_\Omega u_0(x,y)(1-H(\phi(x,y)))dxdy}{\int_\Omega (1-H(\phi(x,y)))dxdy}. \quad (6)$$

3.3. Image post-processing

Frequently, the segmentation results are post-processed to improve the accuracy of the obtained lesion region. In many cases, morphological operations [23] are employed for this purpose [21,26,38]. Here, a morphological filtering, presented in Eq. (7), is applied to the segmented image I by using a structuring element E. This process allows the smoothing of borders, the removing of isolated regions, and/or even filling the segmented lesion region. This filter consists of the opening operation $I \circ E$, defined by Eq. (8), followed by the closing operation of the result by E, defined by Eq. (9), respectively:

$$(I \circ E) \cdot E \,, \tag{7}$$

$$I \circ E = (I \ominus E) \oplus E \,, \tag{8}$$

$$I \cdot E = (I \oplus E) \ominus E \,, \tag{9}$$

where $I \oplus E$ is the dilation operation given by Eq. (10) and $I \ominus E$ is the erosion operation given by Eq. (11). Therefore, the opening of set I by E is the erosion of I by E, followed by the dilation of the result by E. The closing of the set I by E is the dilation of I by E, followed by the erosion of the result by E:

$$I \oplus E = \left\{x \middle| \left[\left(\hat{E}\right)_x \cap \quad I\right] \subseteq I\right\}, \tag{10}$$

$$I \ominus E = \{x|(E)_x \subseteq I\} \,, \tag{11}$$

where $\hat{E}$ is the reflection of set E (structuring element), $\left(\hat{E}\right)_x$ is the translation of set $\hat{E}$ by pixel x, and $(E)_x$ is the translation of set E by pixel x.

3.4. Feature extraction

After the ROI identification, the next step is to extract a lesion's features based on the ABCD rule in order to numerically describe its properties. The clinical assessment is usually based on all of the rule's criteria to diagnose the malignancy of lesions in images. However, the diameter criterion was not applied here due to its great dependence on the image resolution [1], since the size of the image highly affects the number of pixels of each segmented lesion regions. Instead, a texture analysis is performed to assess the surface roughness of the lesion. Therefore, asymmetry, border, colour and texture properties are extracted from the original *RGB* image using the segmented image after post-processing as a feature extraction mask.

3.4.1. Asymmetry

In order to extract features based on the asymmetry criterion, the region of the lesion under analysis is dividing into two sub-regions (R_1, R_2) by an axis according to the longest diagonal d defined by Euclidian distance [23]:

$$D_{(p,q)} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\,, \tag{12}$$

where (x_1, y_1) and (x_2, y_2) are the coordinates of the border's pixels p and q. All the border's pixels are analysed in order to find which pair has the greatest distance $D_{(p,q)}$.

Perpendicular lines from the pixels of the longest diagonal d are computed to analyse the similarity between two sub-regions of the lesion. The number of perpendicular lines may be different for each image, since it depends on the size of the diagonal d of the lesion. Therefore, $N = T/P$ is computed to determine the number of perpendicular lines for all images to be classified; i.e. it determines a set of perpendicular lines S, where T is the total number of perpendicular lines along the diagonal d, and P is a pre-defined fixed number of expected perpendicular lines. In order to determine an adequate set S, the following values for P have been experimentally established, $P = \{10, 20, 30, 40, 50\}$. Ten perpendicular lines $P = 10$ obtained the best results in experimental tests to represent the size of the set of perpendicular lines for each image. Afterwards, two semi-lines were determined from each perpendicular line of the set S, one semi-line represents the sub-region R_1, and the other represents the sub-region R_2. For each perpendicular, the distance $D_{(p,q)}$ of the semi-line for both sub-regions (R_1, R_2) is computed, where p is a pixel of the diagonal d and q is a pixel of the border.

Eleven features are extracted to represent the asymmetry criterion:

- The ratio between the shortest and longest distances based on the semi-lines (R_1, R_2) from each perpendicular line of set S (10 features);
- The standard deviation from ratios based on all perpendicular lines of set S (1 feature).

The ratio between the two semi-lines allows for determining whether the lesion area may be more symmetric or more asymmetric to a particular pixel of the longest diagonal, i.e., the area is either more asymmetric when its coefficient is closer to zero, or more symmetrical when its coefficient is closer to one.

3.4.2. Border

A border is represented by pixels comprising the lesion's boundary, obtained as a result of the lesion segmentation process. A one-dimensional border [23] of the lesion under analysis is defined to extract features based on this criterion. The number of peaks, valleys and straight lines of the border is extracted by vector product and inflexion point descriptors by means of the one-dimensional border. An inflexion point descriptor is applied to measure small irregularities in the border, whereas a vector product descriptor is applied to measure substantial irregularities in the border [45].

The inflexion point descriptor aims to analyse border's pixels P_i to define which pixels show a change of direction. Therefore, a four-point neighbourhood N_j for both left and right directions is considered for each border's pixel P_i. In order to detect if a given pixel P_i is an inflexion, weights w_j are assigned to its neighbour pixels. From the analysis of the y axis of a system of coordinates, each neighbour pixel N_j that is below the pixel under analysis P_i receives $w_j = 1$. Otherwise, each neighbour pixel receives $w_j = -1$. Afterwards, the weights w_j corresponding to each direction (left, D_l, and right, D_r) are added separately, $D_l, D_r = \sum_j w_j$. Pre-defined thresholds $T_1 = 2$ and $T_2 = -2$ [45] are considered to analyse small irregularities in the border, based on the sum of the weights D_l, D_d. An inflexion pixel P_i is achieved when D_l and $D_r \geq T_1$ or D_l and $D_r \leq T_2$. The sum of the weights for both left and right neighbour pixels $S_i = D_l + D_r$ identifies the inflexion pixel P_i as a peak when $S_i > 0$, as a valley when $S_i < 0$, or as a straight line when $S_i = 0$.

On the other hand, the vector product descriptor aims to analyse a border's pixels to identify peaks and valleys with substantial irregularities. The vector product V_i is based on three border pixels $p_1, p_2,$ and p_3 established according to a difference of fifteen pixels between them, totalling a difference of thirty pixels between p_1 and p_3 [45]. The vector product V_i is computed for each border's pixels as:

$$V_i = (x_2 - x_1)(y_3 - y_1) - (y_2 - y_1)(x_3 - x_1)\,, \tag{13}$$

where (x_1, y_1), (x_2, y_2) and (x_3, y_3) are the three aforementioned pixels $p_1, p_2,$ and p_3. Such points determine whether a segment belongs to a peak, valley or straight line. Therefore, a border's pixel P_i is identified as a peak when $V_i > 0$, as a valley when $V_i < 0$, or as a straight line when $V_i = 0$.

Six features are extracted to represent the border criterion:

- The number of peaks, valleys and straight lines based on small irregularities of the border by using the inflexion point descriptor (3 features);
- The number of peaks, valleys and straight lines based on large irregularities of the border by using the vector product descriptor (3 features).

The peak, valley and straight-line values may be relatively different for each image, since they depend on the size of the lesion's border. In order to solve the problem of different ranges that may influence the classification results, such values are adjusted into an interval between 0 (zero) and 1 (one). Therefore, the values obtained by the inflexion point descriptor are divided by the total number of pixels obtained, and the values obtained by the vector product descriptor are divided by the total number of border's pixels. These features allow the assessment of the regularity or irregularity of the lesion's border.

3.4.3. Colour

The *RGB* colour model is commonly employed to represent the colours of skin lesions in images [1,14,17,41,42]. Therefore, statistical measures based on this model are applied to represent the colour criterion. The mean, variance and standard deviation values for each *RGB* channel were extracted (nine features). These features allow for analysing tonality changes of pigmented skin lesions in order to identify malignant lesions.

3.4.4. Texture

In order to extract texture properties of the skin lesions, fractal dimensions are computed from the image under study by using a box-counting method (BCM), since it is simple and effective [43,46]. A fractal dimension [47] is a procedure for splitting the input image into several quadrants to quantify the irregularity level or self-similarity of the image's fractals according to:

$$D = \frac{\log(\mathrm{N})}{\log(1/\mathrm{T})}, \tag{14}$$

where N represent the number of elements of the self-similar parts that reconstruct the original image, and T is the amount of quadrants corresponding to a fraction of its previous size.

The BCM method demarcates a grid over the image; i.e., it divides the image into several squares. The process is iterative, in which the size of each square decreases and

the amount of squares that covered the fractal is counted at each iteration. The box-counting algorithm uses a least squares error to compute the fractal dimension:

$$e = \sum_i \left(f_i - \widehat{f_i}\right)^2 , \text{with } i = 1,2, \dots, N, \tag{15}$$

where N is the number of elements and the term $\widehat{f_i}$, which is an approximation to function f_i, is defined as:

$$\widehat{f_i} = \beta x_i + c \, , \tag{16}$$

where the slope β and intercept c of the line $\widehat{f_i}$ are computed as:

$$\beta = \frac{N \sum_i f_i x_i - (\sum_i f_i)(\sum_i x_i)}{N \sum \ x_i^2 - (\sum_i x_i)^2} \, , \tag{17}$$

$$c = \frac{\sum_i f_i - \beta \ \sum_i x_i}{N} \, . \tag{18}$$

The image-based fractal dimension D_2 is computed individually for each row and column of the image. Afterwards, the final fractal dimension is defined as:

$$D_2 = \left(\frac{\sum_i D_i}{t}\right) + 1 \, , \tag{19}$$

where D_i is the fractal dimension obtained at each iteration and t is the total amount of fractal dimensions.

Eighteen features are extracted to represent the texture properties of the lesion under analysis:

- The fractal dimension of the lesion's area (1 feature);
- The fractal dimension of the original image (1 feature);
- The fractal dimension of sixteen parts of the image, with the original image divided into parts of the same size to measure their fractal dimension (16 features).

The fractal dimension is a value between two and three, which allows for measuring the irregularity level or self-similarity of the image surface.

Overall, the number of features m extracted from each image under study is 44 (11 asymmetry, 6 border, 9 colour and 18 texture features). From this set of features, datasets are constructed with a set of samples (x_i), according to the number of images n for a given classification problem, $i = 1, \dots, n$. Each sample (x_i) is composed of m features

(x_{im}) and the class to which it belongs (y_i). Such datasets are used for the classification process.

3.5. Lesion classification

After building the set of features, the next step is the lesion classification based on the extracted features. The classification process occurs by randomly dividing the available image samples into training and test sets. The training step consists of developing a classification model based on the training samples, 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 in the training step over the set of tests. The classification process should have high performance and robustness, since its results are often used to assist dermatologists in their diagnosis. Therefore, the SVM classifier [20] was used mainly due to its good generalization properties.

The SVM classifier involves an algorithm based on statistical learning applied to build a hyperplane that separates the data according to the defined classes. Such data may be linearly separable or linearly non-separable. Let us consider the training data $\{x_i, y_i\}$, with $x_i \in$ X and $y_i \in$ Y, where X is the set of samples and Y is the class to which they belong $\{-1, +1\}$. A separating hyperplane may be defined as $f(x) = w \cdot x + b$. Then, the points x that lie on the hyperplane satisfy $f(x) = 0$, where w is the normal distance to the hyperplane, and $|b|/\|w\|$ is the perpendicular distance from the hyperplane to the origin, with $b \in \Re$ and $\|w\|$ being the Euclidian norm of w. Therefore, the $f(x)$ divides X into two regions: positive samples if $f(x) > 0$, and negative samples if $f(x) < 0$. For the linearly separable case, the algorithm is used to search the data with largest distance ("named as largest margin") from the hyperplane based on the following constraints:

$$y_i(w \cdot x_i + b) - 1 \geq 0, \text{ with } \forall_i = 1, \ldots, n, \tag{20}$$

where $w \cdot x_i + b \geq +1$ for $y_i = +1$, and $w \cdot x_i + b \leq -1$ for $y_i = -1$.

The largest border is represented by a pair of parallel hyperplanes, H_1 and H_2. The points defined for these hyperplanes are the training points used for classification, called support vectors. The pair of hyperplanes is obtained by minimization of $\|w\|^2$ based on the constraints defined in Eq. (20). Such minimization is given by the Lagrangian function subject to the conditions $w = \sum_{i=1}^{n} \alpha_i y_i x_i$ and $\sum_{i=1}^{n} \alpha_i y_i = 0$, where α_i are

positive Lagrange multipliers for each of the constraints (Eq. (20)). The Lagrangian function is defined as:

$$L_1 = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j \text{ , with } \alpha_i \geq 0 \text{ .} \quad (21)$$

For the linearly non-separable case, positive slack variables ξ_i, $i = 1, \ldots, n$, are introduced in the constraints:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \text{ , with } \xi_i \geq 0, \text{and } \forall_i = 1, \ldots, n \text{ ,} \quad (22)$$

where $w \cdot x_i + b \geq +1 - \xi_i$ for $y_i = +1$, and $w \cdot x_i + b \leq -1 + \xi_i$ for $y_i = -1$. In order to deal with noise and outliers, parameter C is introduced for assigning a penalty to errors, which becomes:

$$L_2 = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j \text{ ,} \quad (23)$$

subject to $0 \leq \alpha_i \leq C$, $\forall_i = 1, \ldots, n$, and $\sum_{i=1}^{n} \alpha_i y_i = 0$.

In order to simplify the process of separating the non-linear data, a kernel function may be applied to map the set of samples of the original space X to a new space with infinite dimensional $\Im$, defined as $\Phi: \mathrm{X} \rightarrow \Im$. The kernel function receives two points of the original space (x_i, x_j), and computes the scalar product in the new space, defined as $K(\mathrm{x}_i, \mathrm{x}_j) = \Phi(\mathrm{x}_i) \cdot \Phi(\mathrm{x}_j)$. The mapping, by using kernel function based on a dual problem presented in Eq. (23), is defined as:

$$L_3 = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(\mathrm{x}_i, \mathrm{x}_j) \text{ ,} \quad (24)$$

subject to $0 \leq \alpha_i \leq C$, and $\sum_{i=1}^{n} \alpha_i y_i = 0$. The application of kernel functions for non-linear data makes the algorithm efficient, so that simple hyperplanes are constructed in a space with high dimensions.

In this study, the histogram intersection kernel [48] is adopted, as defined by Eq. (25), since such a kernel is proposed especially for image classification and it has achieved superior results compared to other kernels. The histogram intersection kernel has been proposed for colour-based image recognition [48], whereas in this study it is based on all extracted lesion features, i.e., asymmetry, border, colour and texture:

$$K(\mathrm{x}_i, \mathrm{x}_j) = \sum_{i=1}^{n} \min(x_i, x_j) \text{ .} \quad (25)$$

Here, the classification algorithm is based on supervised learning and the classification process is binary, since the SVM classifier is originally binary. The image classification

is divided into two steps: feature classification and skin lesion classification. The feature classification step consists of analysing the following classification processes: 1) region asymmetry (symmetric or asymmetric), 2) border irregularity (regular or irregular), 3) colour uniformity (uniform or non-uniform), and 4) texture irregularity (regular or irregular). Each process takes into account only the features related to the classification goal, i.e., a subset of features. The skin lesion classification step consists of distinguishing the following types of skin lesions: 1) nevus and seborrheic keratosis, 2) nevus and melanoma, and 3) seborrheic keratosis and melanoma. In this case, each classification process considers the entire set of features.

4. Experimental Results and Discussion

In this section, segmentation and classification results are described and discussed. First, the image databases used to evaluate the results are described. Second, the experiments for border detection, regarding the pre-processing, segmentation and post-processing steps are presented. Finally, the experiments on the feature extraction of skin lesions, which correspond to the lesion's asymmetry, border, colour, and texture, are presented as well as those for lesion classification.

4.1. Image databases

The databases used to evaluate the proposed approach are composed of macroscopic images of pigmented skin lesions. Examples of such images are shown in Figure 1. A great deal of information concerning the diagnosis of the imaged lesions provided by expert dermatologists is available in these databases, including among them, diagnostics on the lesions and their features (i.e., asymmetry, border, colour and texture). All the information contained in the datasets has been used for the development and evaluation of this work.

The used databases have a total of 408 images, which were collected from Loyola University Chicago [49], YSP Dermatology Image Database [50], DermAtlas [51], DermIS [52], *Saúde Total* [53], Skin Cancer Guide [54], and Dermnet - Skin Disease Atlas [55,56]. Of these, 62 images were melanocytic nevi, 86 images were seborrheic keratosis, and 260 images were melanoma. In regard to the asymmetry criterion, the lesions were symmetric in 137 images and in 271 images were asymmetric. In regard to the border criterion, the lesions have regular borders in 77 images and irregular borders

in 331 images. In regard to the colour criterion, the lesions present uniform colours in 32 images and non-uniform colours in 376 images. In regard to the texture criterion, the lesions present regular texture in 224 images and in 184 images they present irregular texture. The images of the databases have been resized to 200×200 pixels to simplify their processing.

4.2. Border detection

In order to remove noise and enhance the lesions, an anisotropic diffusion filter was applied to the input images according to the discretization of Eq. (1). The parameters were defined by experimental tests, based on parameters suggested by Barcelos and Pires [22], with the following values: $\Delta t = 0.1$, $\sigma = 1$, $\lambda = 1$, $k = 0.0008$, and $NI = 100$. The smoothing results obtained by applying the anisotropic diffusion filter to grey-level images are shown in Figure 3(a-c). The resultant images in (d-f) indicate that the filter has successfully reduced the presence of hairs. However, this filter may not remove other artefacts, such as, reflections and shadows.

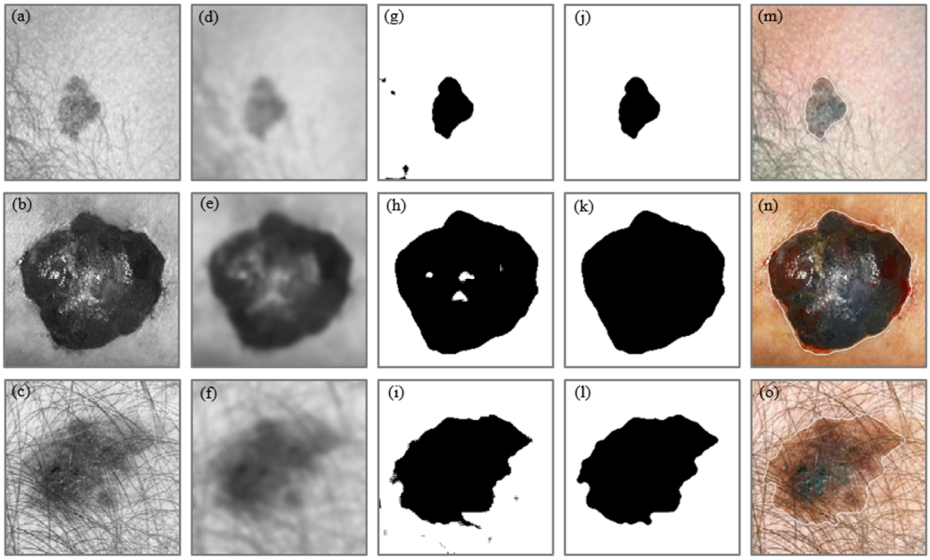

Figure 3: Image processing results for each step of the proposed approach: (a-c) grey-level images, (d-f) smoothed images, (g-i) segmented images, (j-l) post-processed images, and (m-o) original images with the detected borders (white contours) overlapped.

Afterwards, Chan-Vese's model was applied to segment the smoothed image according to Eq. (4). The parameters were defined by experimental tests, based on the parameters proposed by Chan and Vese [19]: $\mu = 0.2$, $\nu = 0$, λ_1 and $\lambda_2 = 1$, $h = 1$, $\Delta t = 0.1$, and 500 iterations were established for the evolution of the curve. In order to define an appropriate curve C, several initial shapes and sizes were tried and visually assessed.

A square-shaped curve was defined and positioned close to the image's centre. However, the imaging conditions are usually inconsistent, and the clinical images are acquired from variable distances, implying that the size of the lesions may be different as they are dependent on the distance adopted in the image acquisition. Therefore, two curves, C_s and C_b, with different sizes were established: for small lesions, $C_s = 40 \times 40$ pixels, and for large lesions, $C_b = 140 \times 140$ pixels. Examples of the segmentation results obtained by applying the Chan-Vese model to the smoothed images (d-f) are shown in Figure 3. Although the resultant binary images (g-i) are of good quality, some binary images presented holes in the interior of the segmented lesion region and/or split regions, which were mainly caused by reflections and shadows.

A morphological filter [23] was applied to the segmented binary images to achieve better segmentation results. In order to define an appropriate structuring element E, several shapes and sizes were tested. Ellipse-shaped structuring elements with radii $r_1, r_2 = 4$, presented the best results according to a visual assessment. The post-processing results obtained by applying the morphological filter to the binary images (g-i) are shown in Figure 3. The resultant images (j-l) confirm the removing of isolated regions and the filling of hole regions, as well as the smoothing of the borders without losing their important characteristics. Afterwards, the borders found were overlapped on the original images (m-o) based on the post-processing image results (j-l).

A subjective evaluation [57] was applied to evaluate the proposed approach, which included a visual assessment of the border detection results by a specialist. The first evaluation analysed whether the lesions were correctly or incorrectly segmented; Figure 4 includes some example results. The evaluation of the results obtained revealed that the proposed approach is effective in detecting skin lesions and extracting their contours from clinical images. The proposed approach adequately tackled the noisy images. However, some images with low contrast boundaries, shadows and reflections were incorrectly segmented.

The second evaluation compared the segmentation results obtained by the proposed approach against the threshold-based segmentation results achieved by using Otsu's method [58], since this method has been widely applied in this domain [1,24,26,39]. Figure 5 presents examples of the segmentation results obtained by applying both segmentation methods to the original images (a-e). The evaluation performed on the results obtained revealed that the proposed approach defined the border of the lesion in a

more effective way than Otsu's method in several cases. Furthermore, the proposed approach also achieved better results when dealing with images of low contrast, and with shadows and reflections. The percentages of correctly segmented images for both segmentation methods, based on the visual assessment of the resultant borders by a specialist, are shown in Table 1. It may be seen that the proposed approach obtained significantly superior results compared to the threshold-based method. The quality of the detected borders of the 385 images correctly segmented by the proposed approach was also visually evaluated by the specialist, with 91.43% of these considered having good quality and the remaining ones having acceptable quality.

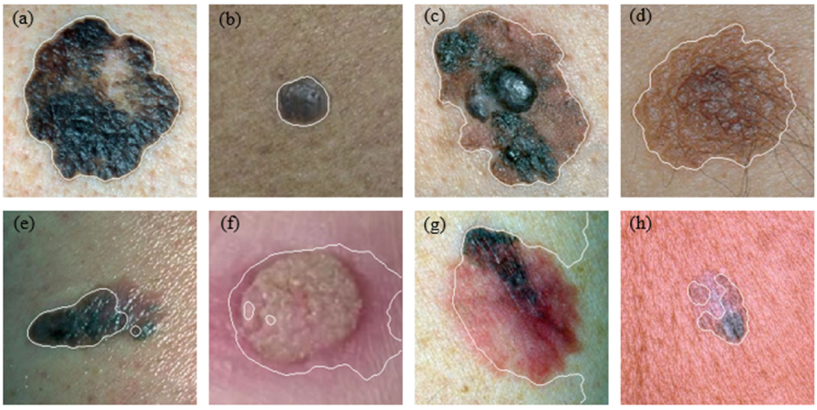

Figure 4: Example of border detection results obtained by applying the proposed approach: (a-d) examples of correctly segmented images and (e-h) examples of incorrectly segmented images.

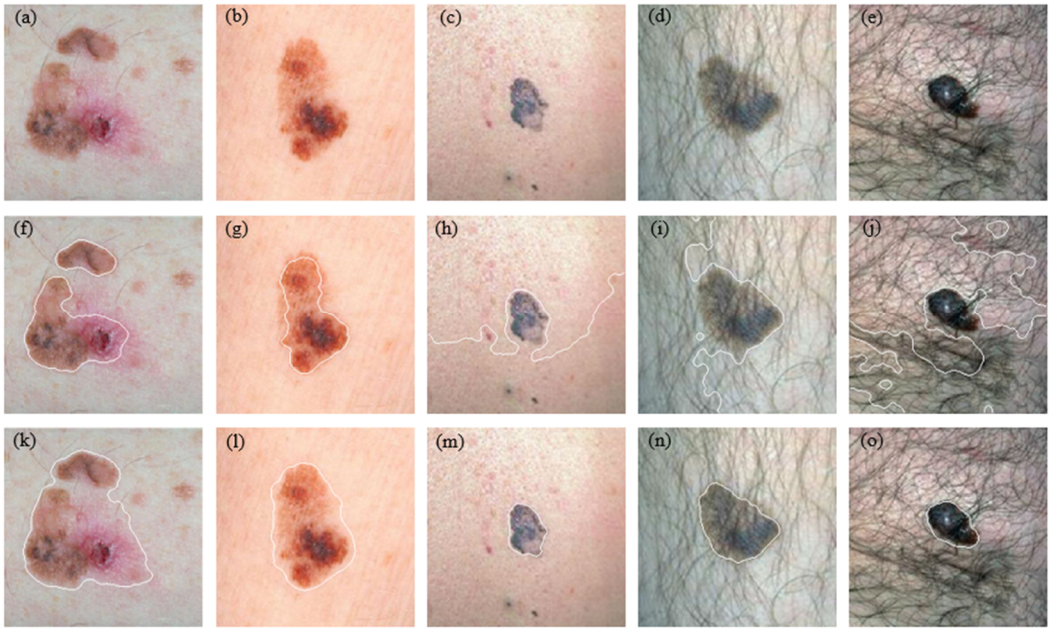

Figure 5: Comparison of the two segmentation methods: (a-e) original images, (f-j) borders detected by Otsu's method, and (k-o) detected by the proposed approach.

Table 1: Skin lesion segmentation results.

Segmentation method	Melanocytic nevus (%)	Seborrheic keratosis (%)	Melanoma (%)	All (%)
Thesholding	80.65	81.40	80	80.39
Proposed approach	96.77	93.02	94.23	94.36

4.3. Skin lesion classification

In order to differentiate types of skin lesions and to detect their features, several classification experiments were performed. The sets of training and test for the classification process were randomly defined from the available image samples, i.e., from the 385 correctly segmented images. In order to define adequate training sets and test for each classification problem, several sizes for the training set were assessed, with the remaining ones employed as test sets. The size values considered for the training set were $T_S = \{10, 20, 30, 40, 50\}$ (in percentage). Each classification model was obtained by applying the SVM classifier [20] by using a histogram intersection kernel [48] based on the set or subset of features and on the samples of the training set. Afterwards, the samples of the test set were classified based on the classification model and the predicted classes were compared to the known classes. Classification performance metrics, such as the precision for each class and the accuracy for each model, were measured to assess the quality of the results obtained.

The following experiments for feature classification were performed: 1) the first experiment involved asymmetry classification, in which $T_S = 10$ was considered the best training set, 2) the second experiment comprised the border classification, in which $T_S = 50$ was considered the best training set, 3) the third experiment comprised the colour classification, in which $T_S = 50$ was considered the best training set, and 4) the last experiment was the texture classification, in which $T_S = 30$ was considered the best training set. The feature classification results are shown in Table 2. The asymmetry classification obtained good results for both classes. In contrast, the texture and colour feature classifications have not led to good generalization between the classes, whereas border feature classification has resulted in an average distinction between the two classes.

Table 2: Feature classification results of the proposed approach.

Classification	Radial basis function			Histogram intersection		
	Class 1 Precision (%)	Class 2 Precision (%)	Accuracy (%)	Class 1 Precision (%)	Class 2 Precision (%)	Accuracy (%)
Asymmetry	Symmetric	Asymmetric		Symmetric	Asymmetric	
	60.71	73.62	69.45	89.29	100	96.54
Border	Regular	Irregular		Regular	Irregular	
	22.86	94.30	81.35	71.43	74.68	74.09
Colour	Uniform	Non-uniform		Uniform	Non-uniform	
	43.75	75.71	73.06	56.25	75.14	73.58
Texture	Regular	Irregular		Regular	Irregular	
	61.69	64.10	62.73	60.39	69.23	64.21

The results obtained for the skin lesion classification are shown in Table 3. The following experiments for skin lesion classification were performed: 1) the first experiment involved classification between nevus and seborrheic keratosis, in which $T_S = 40$ was considered the best training set. Although these two types of lesions are benign, the classification model had an average separation between the two classes, 2) the second experiment was determined by the classification between nevus and melanoma, in which $T_S = 50$ was considered the best training set. The classification result between these two types of lesion has not been quite expressive, since several samples of the database are composed of skin lesions that do not exactly follow the rule that distinguishes these lesions, and 3) the last experiment was based on the classification between seborrheic keratosis and melanoma, in which $T_S = 20$ was considered the best training set. In this case, such lesions are usually too similar, with texture being the main feature used to differentiate them. Therefore, the outcome of the texture classification properly explains why the classification results between seborrheic keratosis and melanoma were not so expressive.

Table 3: Skin lesion classification results of the proposed approach.

Classification	Radial basis function			Histogram intersection		
	Class 1 Precision (%)	Class 2 Precision (%)	Accuracy (%)	Class 1 Precision (%)	Class 2 Precision (%)	Accuracy (%)
Nevus – Keratosis (Class 1 – Class 2)	72.22	73.33	72.84	77.78	80	79.01
Nevus – Melanoma (Class 1 – Class 2)	56.67	73.02	69.87	76.67	73.81	74.36
Keratosis – Melanoma (Class 1 – Class 2)	60	72.64	69.73	80	72.64	74.33

The classification results obtained by applying the histogram intersection kernel for the SVM classifier were compared with the results obtained by applying the radial basis function (RBF) kernel [1,16,37]. The comparison results between the two kernels for both

feature and skin lesion classifications are shown in Table 2 and Table 3, respectively. The application of a histogram intersection kernel showed superior performances for the image classifications. Although the border classification by using an RBF kernel had better accuracy than the classification by using a histogram intersection kernel, the precision of the regular border classification was somewhat low (22.86%). On the other hand, the border classification by using a histogram intersection kernel achieved a more balanced classification result between the regular and irregular classes. In regard to the colour and texture classification, the results were similar for both kernels. In contrast, the asymmetry classification presented significantly superior results. Moreover, the application of a histogram intersection kernel presented much better results for all skin lesion classifications than the RBF kernel.

The proposed approach has been developed using: 1) Matlab 8.4.0.150421 environment for the algorithms of pre-processing, segmentation, post-processing and feature extraction; and 2) Dev-C++ 5.11 environment for the algorithms of texture extraction and classification. The pre-processing step took 63.76 s in smoothing the 385 images. As to the segmentation step, the algorithm took around 49.12 min to segment the images. The post-processing step required 5.09 s to enhance the segmented images. The extraction of the image features from the enhanced images required 1.54 min: asymmetry, 48.65 s; border, 7.35 s; colour, 6.53 s; and texture, 29.44 s. Finally, the classifier required a total of 4.48 s for the training and testing steps. From these values, which are the average times over 10 runs, one can note that the segmentation step was the most time-consuming; however, the computation time required by this step can be considerably decreased by using optimized C/C++ implementations. All algorithms were performed on an Intel(R) Core(TM) i5 CPU 650 @ 3.20 GHz 3.33 GHz with 8 GB of RAM, running Microsoft Windows 7 Professional 64-bits.

5. Conclusion and future works

There are several approaches in the literature for pigmented skin lesion classification. Nevertheless, most of the studies involve only dermoscopy images, in which these images may be more difficult to obtain, since they require a dermoscopy device. In contrast, macroscopic images may be obtained using common digital video or image cameras, so that many computational methods to process them become accessible to dermatologists in several regions of the world. Furthermore, the feature classifications in macroscopic

images are still little explored in research on automated diagnosis, and most studies do not deal with the classification of all features considered in this paper.

An approach was presented for the segmentation and classification of pigmented skin lesions in macroscopic images. This approach is based on an anisotropic diffusion filter, Chan-Vese's model and a SVM classifier to allow for extracting lesion features and the distinguishing between some types of skin lesions, in order to assist dermatologists in their diagnosis. Asymmetry, border, colour and texture properties were considered for the classification process. Although the proposed approach achieved good segmentation results, mainly with noisy images, it may not perform well on images with too low contrast boundaries, shadows and reflections. Both feature and skin lesion classification presented significant results. However, some classification results were not expressive, e.g., the colour and texture based classifications. Whereas these features were extracted from the original *RGB* images of the databases, in which some images contain too much hair and too many reflections and shadows. Therefore, such artefacts may harm the assessement of the colour and texture properties of the lesions. In addition, the features of some images of the databases are too heterogeneous for both classes, which can adversely affect the classification results. Unbalanced databases regarding the number of samples for each class may have decreased the accuracy of the classification results, since the classifier tends to be based on classes with the highest occurrence.

In conclusion, future studies regarding the segmentation and classification of pigmented skin lesion images should involve searching for new methods aiming to develop more efficient and effective systems for better computational diagnosis based on macroscopic images. For example, the development of methods for dealing with reflections and shadows may be considered, in order to solve the previously discussed problems concerning the image segmentation step. Other features and types of pigmented skin lesions may also be approached for the purpose of lesions classification from macroscopic images. The skin lesion classification results can be improved using deep learning architectures, since these architectures have presented excellent performances in different applications, including of Computational Vision. From the advantages that these architectures have revealed, one can stress the capacity of learning from large amount of data in an unsupervised way [59]. Therefore, deep learning architectures should be taken into account in future works related to the classification of skin lesions in images.

Acknowledgments

The first author would like to thank CNPq ("Conselho Nacional de Desenvolvimento Científico e Tecnológico"), in Brazil, for her PhD grant. The authors would also like to thank CAPES ("Coordenação de Aperfeiçoamento de Pessoal de Nível Superior"), in Brazil, for the financial support. Authors gratefully acknowledge the funding of Project NORTE-01-0145-FEDER-000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, cofinanced by "Programa Operacional Regional do Norte" (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional" (FEDER). Furthermore, the authors thank Dr. Ricardo Baccaro Rossetti, from Derm Clínica's Dermatologist of São José do Rio Preto, in Brazil, for his suggestions and for evaluating the results obtained.

References

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

[2] Scharcanski, J.; Celebi, M. E. (2013). *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, Berlin, Heidelberg.

[3] Filho, M.; Ma, Z.; Tavares, J. M. R. S. (2015). A Review of the Quantification and Classification of Pigmented Skin Lesions: From Dedicated to Hand-Held Devices. *Journal of Medical Systems*, 39 (11):1-12.

[4] INCA (2014). *Estimativa 2014: Incidência de Câncer no Brasil*. Instituto Nacional de Câncer José Alencar Gomes da Silva, Coordenação de Prevenção e Vigilância. INCA, Rio de Janeiro.

[5] Smith, L.; MacNeil, S. (2011). State of the art in non-invasive imaging of cutaneous melanoma. *Skin Research and Technology*, 17 (3):257-269.

[6] Wong, A.; Scharcanski, J.; Fieguth, P. (2011). Automatic Skin Lesion Segmentation via Iterative Stochastic Region Merging. *IEEE Transactions on Information Technology in Biomedicine*, 15 (6):929-936.

[7] Cavalcanti, P. G.; Scharcanski, J. (2013). *Macroscopic Pigmented Skin Lesion Segmentation and Its Influence on Lesion Classification and Diagnosis*. In: Celebi, M. E., Schaefer, G. (eds) Color Medical Image Analysis. Springer, Dordrecht, pp 15-39.

[8] Abbas, Q.; Celebi, M. E.; Garcia, I. F. (2012). A novel perceptually-oriented approach for skin tumor segmentation. *International Journal of Innovative Computing, Information and Control*, 8 (3):1837-1848.

[9] Schaefer, G.; Rajab, M. I.; Celebi, M. E.; Iyatomi, H. (2011). Colour and contrast enhancement for improved skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 35 (2):99-104.

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

[11] Silveira, M.; Nascimento, J. C.; Marques, J. S.; Marcal, A. R. S.; Mendonca, T.; Yamauchi, S.; Maeda, J.; Rozeira, J. (2009). Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):35-45.

[12] Zhou, H.; Schaefer, G.; Celebi, M. E.; Iyatomi, H.; Norton, K.; Liu, T.; Lin, F. (2010). *Skin lesion segmentation using an improved snake model*. In: Annual International Conference of the Engineering in Medicine and Biology Society, Buenos Aires, August 31 - September 4 2010. IEEE pp 1974-1977.

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

[14] Iyatomi, H.; Norton, K.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Classification of melanocytic skin lesions from non-melanocytic lesions*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5407-5410.

[15] Celebi, M. E.; Iyatomi, H.; Stoecker, W. V.; Moss, R. H.; Rabinovitz, H. S.; Argenziano, G.; Soyer, H. P. (2008). Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, 32 (8):670-677.

[16] Maglogiannis, I.; Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine*, 13 (5):721-733.

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

[18] Barcelos, C. A. Z.; Boaventura, M.; Silva Junior, E. C. (2003). A well-balanced flow equation for noise removal and edge detection. *IEEE Transactions on Image Processing*, 12 (7):751-763.

[19] Chan, T. F.; Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10 (2):266-277.

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

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

[22] Barcelos, C. A. Z.; Pires, V. B. (2009). An automatic based nonlinear diffusion equations scheme for skin lesion segmentation. *Applied Mathematics and Computation*, 215 (1):251-261.

[23] Gonzalez, R. C.; Woods, R. E. (2002). *Digital Image Processing*. 2 edn. Prentice Hall, New Jersey.

[24] Norton, K.; Iyatomi, H.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Development of a novel border detection method for melanocytic and non-melanocytic dermoscopy images*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5403-5406.

[25] Beuren, A. T.; Janasieivicz, R.; Pinheiro, G.; Grando, N.; Facon, J. (2012). *Skin melanoma segmentation by morphological approach*. In: International Conference on Advances in Computing, Communications and Informatics, Chennai, August 3-5 2012. ACM, pp 972-978.

[26] Norton, K.-A.; Iyatomi, H.; Celebi, M. E.; Ishizaki, S.; Sawada, M.; Suzaki, R.; Kobayashi, K.; Tanaka, M.; Ogawa, K. (2012). Three-phase general border detection method for dermoscopy images using non-uniform illumination correction. *Skin Research and Technology*, 18 (3):290-300.

[27] Oliveira, R. B.; Filho, M. E.; Ma, Z.; Papa, J. P.; Pereira, A. S.; Tavares, J. M. R. S. (2016). Computational methods for the image segmentation of pigmented skin lesions: a review. *Computer Methods and Programs in Biomedicine*, 131:127-141.

[28] Celebi, M. E.; Wen, Q.; Hwang, S.; Iyatomi, H.; Schaefer, G. (2013). Lesion border detection in dermoscopy images using ensembles of thresholding methods. *Skin Research and Technology*, 19 (1):e252-e258.

[29] Yuksel, M. E.; Borlu, M. (2009). Accurate Segmentation of Dermoscopic Images by Image Thresholding Based on Type-2 Fuzzy Logic. *IEEE Transactions on Fuzzy Systems*, 17 (4):976-982.

[30] Zhang, K.; Song, H.; Zhang, L. (2010). Active contours driven by local image fitting energy. *Pattern Recognition*, 43 (4):1199-1206.

[31] Li, C.; Xu, C.; Gui, C.; Fox, M. D. (2010). Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, 19 (12):3243-3254.

[32] Kass, M.; Witkin, A.; Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1 (4):321-331.

[33] Xu, C.; Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7 (3):359-369.

[34] Paragios, N.; Deriche, R. (2002). Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46 (3):223-247.

[35] Celebi, M. E.; Aslandogan, Y. A.; Stoecker, W. V.; Iyatomi, H.; Oka, H.; Chen, X. (2007). Unsupervised border detection in dermoscopy images. *Skin Research and Technology*, 13 (4):454-462.

[36] Roberts, M. E.; Claridge, E. (2003). *An Artificially Evolved Vision System for Segmenting Skin Lesion Images*. In: Ellis, R. E., Peters, T. M. (eds) Medical Image Computing and Computer-Assisted Intervention, vol 2878. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 655-662.

[37] Rahman, M. M.; Bhattacharya, P.; Desai, B. C. (2008). *A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions*. In: 8th IEEE International Conference on International Conference on BioInformatics and BioEngineering, Athens, October 8-10 2008. IEEE, pp 1-6.

[38] Zhou, H.; Schaefer, G.; Sadka, A. H.; Celebi, M. E. (2009). Anisotropic Mean Shift Based Fuzzy C-Means Segmentation of Dermoscopy Images. *IEEE Journal of Selected Topics in Signal Processing*, 3 (1):26-34.

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

[40] Abbas, Q.; Celebi, M. E.; García, I. F. (2012). Skin tumor area extraction using an improved dynamic programming approach. *Skin Research and Technology*, 18 (2):133-142.

[41] Chang, Y.; Stanley, R. J.; Moss, R. H.; Van Stoecker, W. (2005). A systematic heuristic approach for feature selection for melanoma discrimination using clinical images. *Skin Research and Technology*, 11 (3):165-178.

[42] She, Z.; Liu, Y.; Damatoa, A. (2007). Combination of features from skin pattern and ABCD analysis for lesion classification. *Skin Research and Technology*, 13 (1):25-33.

[43] Garnavi, R.; Aldeen, M.; Bailey, J. (2012). Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 16 (6):1239-1252.

[44] Clawson, K. M.; Morrow, P.; Scotney, B.; McKenna, J.; Dolan, O. (2009). *Analysis of pigmented skin lesion border irregularity using the harmonic wavelet transform*. In: 13th International Machine Vision and Image Processing Conference Dublin, September 2-4 2009. IEEE, pp 18-23.

[45] Araujo, A. F. d. (2010). *Método para extração e caracterização de lesões de pele usando difusão anisotrópica, crescimento de regiões, watersheds e contornos ativos*. Unpublished master's thesis, Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas, Brasil (in Portuguese), São José do Rio Preto.

[46] Dobrescu, R.; Dobrescu, M.; Mocanu, S.; Popescu, D. (2010). Medical images classification for skin cancer diagnosis based on combined texture and fractal analysis. *WISEAS Transactions on Biology and Biomedicine*, 7 (3):223-232.

[47] Al-Akaidi, M. (2004). *Fractal speech processing*. Cambridge university press, New York.

[48] Barla, A.; Odone, F.; Verri, A. (2003). *Histogram intersection kernel for image classification*. In: International Conference on Image Processing, Italy, September 14-17 2003. IEEE, pp 513-516.

[49] Melton, J. L.; Swanson, J. R. (2012). *Skin Cancer and Benign Tumor Image Atlas*. Loyola University Dermatology Medical Education. http://www.meddean.luc.edu/lumen/MedEd/medicine/dermatology/melton/content1.htm. Accessed August 2012.

[50] Suzumura, Y. (2012). *YSP Dermatology Image Database*. http://homepage1.nifty.com/ysh/soft_e_ysp.htm. Accessed August 2012.

[51] Cohen, B. A.; Lehmann, C. U. (2012). *Dermatology Image Atlas*. Johns Hopkins University - DermAtlas. http://dermatlas.med.jhmi.edu/. Accessed August 2012.

[52] Diepgen, T. L.; Yihune, G. (2012). *Dermatology Online Atlas*. Dermatology Information System - DermIS. http://www.dermis.net/dermisroot/en/home/index.htm. Accessed August 2012.

[53] Saúde Total (2012). *Câncer da Pele: fotoproteção, Vida saudável com o sol*. http://www.saudetotal.com.br/prevencao/topicos/default.asp. Accessed August 2012.

[54] Skin Cancer Guide (2004). *Melanoma*. http://www.skincancerguide.ca/melanoma/images/melanoma_images.html. Accessed August 2012.

[55] Campbell Jr., J. L. (2012). *Nevi, melanoma*. Dermnet, Skin Disease Atlas. http://www.dermnet.com/videos/nevi-melanoma/. Accessed August 2012.

[56] Chapman, S. (2012). *Benign tumors*. Dermnet, Skin Disease Atlas. http://www.dermnet.com/videos/benign-tumors/. Accessed August 2012.

[57] Celebi, M. E.; Iyatomi, H.; Schaefer, G.; Stoecker, W. V. (2009). Lesion border detection in dermoscopy images. *Computerized Medical Imaging and Graphics*, 33 (2):148-153.

[58] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9 (1):62-66.

[59] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2 (1):1-127.

PART B - ARTICLE 4

COMPUTATIONAL DIAGNOSIS OF SKIN LESIONS FROM DERMOSCOPIC IMAGES USING A COMBINATION OF FEATURES

Roberta B. Oliveira, Aledir S. Pereira and João Manuel R. S. Tavares

Submitted to an international journal (under review), 2017

Abstract

There has been an alarming increase in the number of skin cancer cases worldwide in recent years, which has raised interest in computational systems for automatic diagnosis to assist early diagnosis and prevention. Feature extraction to describe skin lesions is a challenging research area due to the difficulty in selecting meaningful features. The main objective of this work is the find the best combination of features, based on shape properties, colour variation and texture analysis, to be extracted using various feature extraction methods. Several colour spaces are used for the extraction of both colour- and texture-related features. Different categories of classifiers were adopted to evaluate the proposed feature extraction step and different feature selection algorithms were compared for the classification of skin lesions. The developed skin lesion computational diagnosis system was applied to a set of 1104 dermoscopic images using a cross-validation procedure. The best results were obtained by an optimum-path forest classifier with very promising results. The proposed system achieved an accuracy of 92.3%, sensitivity of 87.5% and specificity of 97.1% when the full set of features was used. Furthermore, it achieved an accuracy of 91.6%, sensitivity of 87% and specificity of 96.2%, when 50 features were selected using a correlation-based feature selection algorithm.

Keywords: Feature extraction and selection; Fractal dimension analysis; Discrete wavelet transform; Co-occurrence matrix; Skin lesion image classification.

1. Introduction

Dermoscopic images are widely applied for automated diagnosis of pigmented skin lesions. Such images can be acquired from dermatoscope devices or specific cameras to provide a better visualization of the pigmentation pattern on the skin surface. Several computational systems have been proposed to assist dermatologists in obtaining an effective diagnosis [1]. These systems can be used to monitor benign skin lesions, and malignant lesions may be diagnosed at an early stage, so that the patient has a higher probability of being cured with less aggressive therapies.

The features extracted from skin lesion images must represent their class, e.g., benign or malignant. Several methods to extract shape-, colour- and texture-related features for the automated diagnosis have been proposed in the literature [2,3]. Such features can

represent skin lesion properties adequately. Nevertheless, few of them combine different methods to extract features in a similar category, e.g., texture analysis. The assessment of classifiers is an important issue for pattern recognition processes [4]. Other difficulties involve defining which features are meaningful to describe the skin lesions, including the presence of highly correlated, redundant and irrelevant features. Some studies have proposed feature selection methods [5] to overcome these problems [2,3]. An overview of the computational methods for pigmented skin lesion classification in images, which addresses the feature extraction and selection, and the classification steps, is presented in Oliveira et al. [6].

The aim of the present study is to find and extract the most relevant features for skin lesion computational diagnosis based on shape properties, colour variation and texture analysis using different techniques. Figure 1 provides an overview of the skin lesion classification approach proposed in this study. The main contribution of this study is the texture analysis based on twelve colour channels, since texture features are usually extracted using grey-level images or few colour channels [3,7]. In addition, these features are combined with shape and colour features to construct a set of features. A dataset is constructed from this set of features with a number of samples (x_i). According to the number of images n for a given classification problem, $i = 1, \dots, n$. Each sample (x_i) is composed of m features (x_{im}) and the class to which it belongs (y_i). Such a dataset is used in the image classification process of benign or malignant lesions using different classifiers and feature selection algorithms.

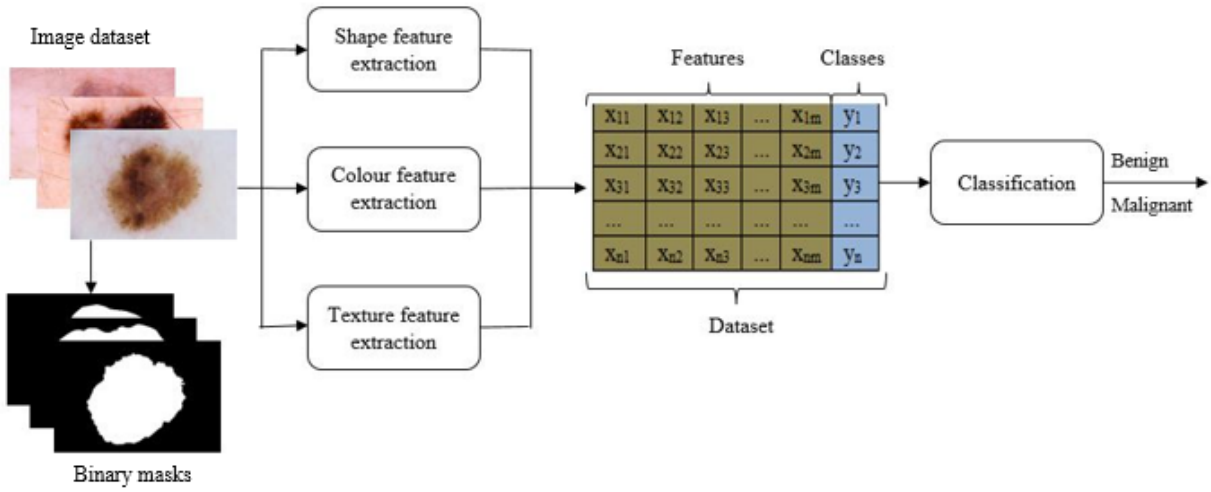

Figure 1: Overview of the proposed approach for the skin lesion classification.

This paper is organized as follows: The proposed feature extraction system, based on shape, colour and texture properties, is explained in Section 2. The algorithms used for selecting features and classifying skin lesions in dermoscopic images are detailed in Section 3. The experimental results and their discussion, are presented in Section 4. Finally, the conclusions drawn and the proposals for future studies are expressed in Section 5.

2. Proposed feature extraction

In this section, a combination of features to represent the skin lesion images is proposed. These features are based on clinical approaches commonly used by dermatologists when diagnosing skin lesions. Various features have been proposed in the literature for skin lesion diagnosis in dermoscopic images [6]. The feature extraction step is based on the intensities of the pixels in the regions of interest (ROIs) defined by specialists, i.e., binary images, where the non-zero pixels belong to the lesion, and the others to the background skin. The features described are categorized in shape properties, colour variation and texture analysis in Table 1.

Table 1: Features extracted from skin lesion images based on shape properties, colour variation and texture analysis.

Skin lesion features	Denotation	Number of features (channels)
Shape properties		
Geometrical properties	$A, P, ED, CO, CI, S, R, AR, e$	9
Lesion asymmetry	μ_s, s_s^2, s_s	3
Border irregularity	$p_S, v_S, l_S, p_L, v_L, l_L$	6
Colour variation		
Colour average, variance and standard deviation	μ_c, s_c^2, s_c	3 (× 12)
Minimum and maximum colours	min_c, max_c	2 (× 12)
Colour skewness	SK_c	1 (× 12)
Texture analysis		
Fractal dimension analysis	D_c^2	1 (× 12)
Discrete wavelet transform	$E(Sb)_c, H(Sb)_c; Sb = 10$	20 (× 12)
Co-occurrence matrix	$ASM_c, C_c, CRL_c, VAR_c, IDM_c, SA_c, SV_c, SH_c, H_c, DV_c, DH_c, CRL1_c, CRL2_c, MCC_c$	14 (× 12)

2.1. Shape properties

Shape properties provide measures of the lesions based on their geometrical properties, their asymmetry or irregularity of their borders. These features are important for skin lesion diagnosis, as an asymmetric shape, border irregularity or ill-defined structure can

characterize malignant lesions. Other geometrical properties of the lesion area which are commonly computed include the number of pixels inside the lesion region, aspect ratio, compactness, perimeter, greatest and shortest diameters, equivalent diameter, eccentricity, solidity, rectangularity, and circularity [2,3,8].

The lesion asymmetry can be evaluated by dividing the region of the lesion under analysis into two sub-regions using an axis of symmetry, and thereby analyse the similarity of the area by overlapping the two sub-regions of the lesion along the axis. In some studies, the axis of symmetry is based both major and minor axes [2,3]. Features extracted from a wavelet transform [3,9], Fourier transform [10], fractal dimension [11], and irregularity index [3] have also been used to assess border irregularity. In this study, 18 lesion shape features were extracted from each image under analysis. These features are based on some of the standard features previously mentioned and some new features presented in a previous study [12].

2.1.1. Geometrical property measures

These measures can provide the geometrical properties of a lesion by comparing the shape of the lesion with geometrical objects, e.g., a circle or a rectangle. However, some of these features depend on the image resolution and frequently the properties of the images are different as they may have been acquired from different distances and therefore, have different resolutions. Consequently, a normalization procedure is required. This will be considered in the following sections.

a) Lesion area and border perimeter: The lesion area A is the number of pixels within the lesion border, and the border perimeter P is the number of pixels along the lesion border.

b) Equivalent diameter, compactness and circularity: These measures are based on a circle. The equivalent diameter ED is the diameter of a circle whose area is same as the lesion area A, which is given by $ED = \sqrt{4\,A/\pi}$. The compactness CO measures the ratio of the lesion area to a circle with the same perimeter. Nonetheless, an alternative version based on the perimeter can be calculated as the ratio between the equivalent diameter ED and maximum diameter MD of the lesion [2], $CO = ED/MD$. The circularity CI is the measure of how closely the lesion area approaches that of a circle, $CI = 4\,A\,\pi/P^2$.

c) Solidity and rectangularity: These measures are based on a convex hull (it checks a curve for convexity defects and corrects them) and a bounding rectangle from the

lesion area. The solidity S is computed by the ratio of lesion area A to its convex hull area CH, $S = A/CH$. Rectangularity R is the ratio of the lesion area to the bounding rectangle area BA, i.e., a bounding-box, $R = A/BA$, where $BA = width * height$.

d) Aspect ratio and eccentricity: These measures can be based on the structure of moments, up to the 3rd order of a lesion shape [2]. The aspect ratio AR is determined by the ratio of the length of the major axis A_1 to the length of the minor axis A_2, $AR = A_1/A_2$, where A_1 and A_2 are given by:

$$A_1, A_2 = \left\{8\left\{mu_{02} + mu_{20} \pm [(mu_{02} - mu_{20})^2 + 4mu_{11}]^{1/2}\right\}\right\}^{1/2}, \quad (1)$$

where mu_{ij}, defined in Eq. (2), is the (i,j)th order of central moments of the lesion shape. The relation (c_x, c_y) denotes the lesion shape centroid given by: $c_x = m_{10}/m_{00}$ and $c_y = m_{10}/m_{00}$, which is computed from the geometric moments, m_{ij}, given by Eq. (3).

$$mu_{ij} = \sum_{x=1}^{rows} \sum_{y=1}^{cols} (x - c_x)^i \cdot \left(y - c_y\right)^j, \quad (2)$$

$$m_{ij} = \sum_{x=1}^{rows} \sum_{y=1}^{cols} x^i \cdot y^j \ . \quad (3)$$

The eccentricity e is a measure of the shape elongation of the lesion region, which can be computed as:

$$e = [(mu_{02} - mu_{20})^2 4mu_{11}]/(mu_{02} + mu_{20})^2, \quad (4)$$

where mu_{ij} is the central moments defined in Eq. (2).

2.1.2. Lesion asymmetry

In order to extract features based on the asymmetry properties, adapted from Oliveira et al. [12], the region of the lesion under analysis is divided into two sub-regions (R_1, R_2) by an axis, according to the longest diagonal, d, defined by the Euclidian distance: $D_{(p,q)} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, where (x_1, y_1) and (x_2, y_2) are the coordinates of the border pixels p and q. All the border pixels are analysed in order to find which pair has the greatest distance $D_{(p,q)}$. Perpendicular lines S_i from the pixels of the longest diagonal d are computed to analyse the similarity between two sub-regions of the lesion. Afterwards, two semi-lines are determined from each perpendicular line of the set S_i, one semi-line represents the sub-region R_1, and the other represents the sub-region R_2.

The distance $D_{(p,q)}$ of the semi-line for both sub-regions (R_1, R_2) is computed for each perpendicular, where p is a pixel of the diagonal d and q is a pixel of the border. The ratio between the shortest and longest distances based on the semi-lines (R_1, R_2) from

each perpendicular line of set S_i are computed. The ratio between the two semi-lines can determine whether the lesion area is more symmetric or more asymmetric to a particular pixel of the longest diagonal. Three features are extracted to represent the lesion asymmetry: average μ_s, variance s_s^2 and the standard deviation s_s from the ratios between the two semi-lines based on all perpendicular lines of set S_i.

2.1.3. Border irregularity

The border is represented by pixels that make up the lesion boundary. A one-dimensional border of the lesion under analysis is defined to extract features based on this property. The number of peaks, valleys and straight lines of the border is computed using the vector product and inflexion point descriptors from the one-dimensional border, according to Oliveira et al. [12]. The inflexion point descriptor aims to analyse border pixels P_i to define which pixels show a change of direction. On the other hand, the vector product descriptor aims to analyse the border pixels to identify peaks and valleys with substantial irregularities. Six features are extracted to represent border irregularities: 1) the number of peaks p_S, valleys v_S and straight lines l_S based on small irregularities of the border using the inflexion point descriptor; and 2) the number of peaks p_L, valleys v_L and straight lines l_L based on large irregularities of the border using the vector product descriptor.

2.2. Colour spaces

Several different colour spaces, described in the literature, are used to obtain more specific information about the colours of a lesion [6]. Here, for the extraction of colour and texture features, four colour spaces were used: *RGB*, *HSV*, *CIE Lab* and *CIE Luv*, which correspond to 12 channels $n = 12$. These colour spaces are commonly used to represent the colours of skin lesions [2,8,13,14]. *HSV*, *CIE Lab* and *CIE Luv* colour spaces represent colours based on human perception. Furthermore, *CIE Lab* and *CIE Luv* are unified colour spaces and can simplify the identification of colour properties [15].

a) *RGB* colour space: This colour space represents the numerical values of the red, green and blue channels, and is widely used, since the images are originally obtained with this colour space. Moreover, the original *RGB* colour image can be used for conversion to other colour spaces. Although this colour space presents some disadvantages such as high correlation between the channels and no perceptual uniformity [16], several studies have achieved good results from it [2,8].

b) *HSV* colour space: This colour space represents the hue, saturation and value channels, which define the perceived colour of an area, the purity of colour and the brightness of colour, respectively. The conversion from the *RGB* colour space to the *HSV* colour spaces is given by:

$$V = \max(R, G, B)\ ,$$

$$S = \begin{cases} [V - \min(R, G, B)]/V, & if\ V \neq 0 \\ 0\ , & if\ V = 0 \end{cases},$$

$$H = \begin{cases} 60(G - B)/[V - \min(R, G, B)], & if\ V = R \\ 120 + 60(B - R)/[V - \min(R, G, B)], & if\ V = G \\ 240 + 60(R - G)/[V - \min(R, G, B)], & if\ V = B \end{cases},$$

$$H = H + 360,\ if\ H < 0\ , \tag{5}$$

where $0 \leq H \leq 360$, $0 \leq S \leq 1$ and $0 \leq V \leq 1$, and the separation of each channel corresponds to $H = H/2$, $S = 255S$ and $V = 255V$.

c) *CIE Lab* and *CIE Luv* colour spaces: These colour spaces were proposed by the International Commission on Illumination (*CIE*, in French), whose main goal was to provide a uniform colour space. This means that the distance between two colours in such a colour space is strongly correlated with the human visual perception. Another advantage of these colour spaces is the separation of the luminance component *L* from the chrominance channels (*a*, *b*) and (*u*, *v*). A difference between these two colour spaces is that the *CIE Lab* colour space normalizes the values by division with the white colour point of the *CIE XYZ* colour space, whereas the *CIE Luv* colour space normalizes the values by the subtraction of such a white colour point [15,16]. The conversion from *RGB* colour space to the *CIE Lab* and *CIE Luv* colour spaces is based on the *CIE XYZ* colour space. Considering the values X_n, Y_n, and Z_n as being the white colour points, the *CIE Lab* colour space is computed by the following equations:

$$L = \begin{cases} 116(Y/Y_n)^{1/3} - 16\ , & for\ Y > 0.008856 \\ 903.3\ Y/Y_n\ , & for\ Y \leq 0.008856 \end{cases},$$

$$a = 500\left[(X/X_n)^{1/3} - (Y/Y_n)^{1/3}\right],$$

$$b = 200\left[(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}\right], \tag{6}$$

where $0 \leq L \leq 100$, $-127 \leq a \leq 127$ and $-127 \leq b \leq 127$, and the separation of each channel corresponds to $L = L * 255/100$, $a = a + 128$ and $b = b + 128$. And finally the *CIE Luv* colour space is computed by the following equations:

$$L = \begin{cases} 116(Y/Y_n)^{1/3} - 16\,, & for\ Y > 0.008856 \\ 903.3\,Y/Y_n\,, & for\ Y \le 0.008856 \end{cases},$$

$$u = 13L(u' - u_n),\ v = 13L(v' - v_n)\,,$$

$$u' = 4X/X + 15\mathrm{Y} + 3\mathrm{Z}\,,\ v' = 9Y/X + 15\mathrm{Y} + 3\mathrm{Z}\,,$$

$$u_n = 4X_n/X_n + 15Y_n + 3Z_n,\ v_n = 9Y_n/X_n + 15Y_n + 3Z_n\,, \quad (7)$$

where $0 \le L \le 100$, $-134 \le u \le 220$ and $-140 \le v \le 122$, and the separation of each channel corresponds to $L = L * 255/100$, $u = 255/354(u + 134)$ and $v = 255/262(v + 140)$.

2.3. Colour variation

Statistical measures based on several colour spaces are commonly applied to the feature extraction from the lesion region [2,8,14]. Furthermore, these measures are also applied to other regions associated with the lesion border. The background skin [8] and surrounding skin (inner or outer peripheral regions) [2] are examples of such regions that are considered for feature extraction. Skin lesion features based on relative colours have been proposed [2,8] in order to assess colour features from the different regions associated with the lesion. Basic colours in the skin lesions have also been considered and computed [17].

In order to analyse the colour variation, six statistical measures are computed for each colour channel c of the lesion region using the four-colour spaces as defined earlier, with $c = 1,2,\dots,n$, where n is the number of channels used for the colour feature extraction.

a) Colour average, variance and standard deviation: These measures evaluate the average and the variation of a set of lesion intensity values I_p, of each colour channel c. The average μ_c, variance s_c^2, and standard deviation s_c are computed by the following equations:

$$\mu_c = \frac{1}{N}\sum_{p=1}^{N}(I_p)\,, \quad (8)$$

$$s_c^2 = \frac{1}{N-1}\sum_{p=1}^{N}\left(I_p - \mu_c\right)^2\,, \quad (9)$$

$$s_c = \sqrt{s_c^2}\,, \quad (10)$$

where N is the number of pixels of the ROI in the image.

b) Minimum and maximum colours: These measures define the minimum value, $min_c = min(I_p)$, and the maximum value, $max_c = max(I_p)$ of the set of lesion intensity values I_p, of each colour channel c.

c) Colour skewness: This measure computes the asymmetry SK_c of the data around the set of lesion intensity values I_p:

$$SK_c = \left[\frac{1}{N}\sum_{p=1}^{N}\left(I_p - \mu_c\right)^3\right] / s_c^3 \,, \tag{11}$$

where μ, s are the average and the standard deviation of the set of lesion intensity values I_p, and N is the number of pixels of the ROI in the image.

2.4. Texture analysis

Texture analysis methods are usually categorized as structural, statistical, model-based and transform. Although the structural approach provides a good symbolic description, some extracted features can be more useful for synthesis rather than an analysis task [18]. Among the various statistical methods, the co-occurrence matrix has shown potential for effective texture discrimination of dermoscopic images [8,14,19]. Fractal dimension is a model-based method, which is also potentially useful for texture analysis [3]. Fourier [20], Gabor [21] and wavelet [3] transforms are also applied to extract texture features in skin lesion images. Texture analysis methods based on the Fourier transform may present poor performance, due to its lack of spatial localization, whereas a Gabor filter allows a superior spatial localization. However, the wavelet transform presents several advantages compared to the Gabor transform, i.e., varying the spatial resolution allows it to represent textures using the most suitable scale, and there are even several possibilities for the wavelet function, in which is possible to choose the best wavelets for a given application [18].

In order to obtain the best features to represent the skin lesion texture in this work three different texture analysis methods were chosen. The texture features are computed for each colour channel using the four-colour spaces as defined earlier. Thus, a total of 420 texture features are extracted: 12 features from the fractal dimension analysis [22], 240 features from the discrete wavelet transform [23] and 168 features from the single-channel co-occurrence matrix [24].

2.4.1. Colour image-based fractal dimensional analysis

In order to extract the texture properties of the skin lesions, fractal dimensions are computed from the image under study using a box-counting method (BCM), since it is simple and effective for skin lesion analysis [3]. A fractal dimension [22] is a procedure for splitting the input image into several quadrants to quantify the irregularity level or self-similarity of the image fractals, according to $D = \log(\mathrm{P})/\log(1/\mathrm{T})$, where P represents the number of elements of the self-similar parts that reconstruct the original image, and T is the number of quadrants corresponding to a fraction of its previous size. The BCM projects a grid over the image; i.e., it divides the image into several squares. The process is iterative, in which the size of each square decreases and the number of squares that covered the fractal is counted at each iteration.

The bi-dimensional fractal dimension D_c^2, which is computed individually for each channel c of the colour spaces, is defined as:

$$D_c^2 = \frac{1}{N}\left(\sum_{i=1}^{rows}\sum_{j=1}^{cols} D_{i,j}\right) + 1, \text{ with } c = 1,2,\dots,n\,, \tag{12}$$

where $D_{i,j}$ is the fractal dimension obtained at each iteration, i.e., it is computed individually for each row i and column j of the image, N is the total number of fractal dimensions, and n is the number of channels used for the texture feature extraction.

2.4.2. Colour image-based wavelet transform

In this work, a bi-dimensional wavelet transform is used to decompose a 2-D image, to which one-dimensional transformations are applied individually along the horizontal and vertical directions of an image [23]. The decomposition of a one-dimensional signal $f(t)$ is based on a family of wavelet functions that usually is defined as complete and with an orthogonal base:

$$W_{a,b} = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt\,. \tag{13}$$

This family is obtained by dilating and translating a single function defined as the mother wavelet ψ:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \tag{14}$$

where a and b are the parameters of dilating and translating, respectively. When a and b are defined for discrete signals, a discrete wavelet transform (DWT) is obtained.

The DWT, based on a multi-resolution, decomposes an input signal in two new signals with different frequencies using quadrature mirror filters. Such signals correspond to low- and high-pass filters that represent the wavelet functions (mother wavelet) $\psi(t)$ and scaling functions (father wavelet) $\phi(t)$, respectively. The low-pass filter corresponds to approximation coefficients, whereas the high-pass filter corresponds to detail coefficients.

The decomposition of a bi-dimensional signal using DWT yields a sub-sample with four sub-bands for one-level of decomposition that are: LL, LH, HL, HH. The sub-band LL corresponds to the clustering of low-pass filtering in the lines and columns. The sub-band LH corresponds to the clustering of low-pass filtering in the lines and high-pass filtering in the columns. The sub-band HL corresponds to the clustering of high-pass filtering in the lines and low-pass filtering in the columns. The sub-band HH corresponds to the clustering of high-pass filtering in the lines and columns. These sub-bands have an equal number of pixels as the original image. A multi-level decomposition can be considered, when the decomposition is applied recursively to the LL sub-band. The result of such decomposition is standard pyramidal wavelet decomposition.

A problem in this wavelet decomposition approach is the large number of features that can be obtained depending on the number of levels used and it can give the classification a high computational cost. Here, a-three-level decomposition was used to decompose the colour images, which is considered sufficient to represent the image texture. Therefore, ten sub-bands were obtained for each channel of the colour spaces. A Haar wavelet filter was used to implement the DWT, with the coefficients defined as $h = \left(1.0/\sqrt{2}, 1.0/\sqrt{2}\right)$. This filter was used since it is simple and has been previously applied to extract texture from skin lesion images [25].

The energy $E(Sb)_c$ and entropy $H(Sb)_c$ measures for the feature extraction from the coefficients obtained by DWT are computed for each sub-band Sb and each colour channel c:

$$E(Sb)_c = \sqrt{\frac{1}{N}\sum_{i=1}^{rows}\sum_{j=1}^{cols}\left(Sb_{i,j}^2\right)}\ , \tag{15}$$

$$H(Sb)_c = \frac{1}{N}\sum_{i=1}^{rows}\sum_{j=1}^{cols}\left[Sb_{i,j}^2 \times \log\left(Sb_{i,j}^2\right)\right], \tag{16}$$

where $Sb_{i,j}$ corresponds to the sub-band coefficient for the pixel i,j and N is the total number of pixels in the sub-band. These measures are commonly used to represent the texture of skin lesion images [3].

2.4.3. Colour image-based co-occurrence matrices

The grey-level co-occurrence matrices (GLCMs) represent the relationship between the intensities of neighbouring pixels to characterize the texture of an image [24]. Such a matrix $m(i,j,d,\theta)$ is obtained by the joint probability of occurrence of grey-levels considering each pair of neighbour pixels i,j of an image, where these pixels are separated by a distance d and in a specific direction θ.

In this study, co-occurrence matrices (CMs) were used for the colour channels. The single-channel co-occurrence matrices (SCMs) were applied separately to each colour channel, with $c = 1,2,\dots,n$, where n is the number of colour channels. The parameters used to set up the matrices are based on Haralick et al. [24]. The intensities of each channel are quantized by an equal probability quantizing algorithm, with $q = 16$. The distance d between one pixel and its neighbours is $d = 1$, and four orientations θ are considered $\theta = (0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ})$. In order to extract rotation invariant features, a normalized SCM is obtained from the SCMs corresponding to the four orientations.

From the normalized SCM, 14 statistical measures based on Haralick's texture features [24] were extracted from the image: angular second moment ASM_c, contrast C_c, correlation CRL_c, variance VAR_c, inverse difference moment IDM_c, sum average SA_c, sum variance SV_c, sum entropy SH_c, entropy H_c, difference variance DV_c, difference entropy DH_c, information measure of correlation 1 $CRL1_c$, information measure of correlation 2 $CRL2_c$, and maximal correlation coefficient MCC_c. These features are expressed in Eq. (17)-(30), where $m_{i,j}$ is the entry value in the position i,j of the normalized matrix, and N is the number of different intensities contained in the quantized image:

$$ASM_c = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(m_{i,j}\right)^2, \tag{17}$$

$$C_c = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[m_{i,j}(i-j)^2\right], \tag{18}$$

$$CRL_c = \left[\sum_{i=1}^{N}\sum_{j=1}^{N}\left(i \times j \times m_{i,j}\right) - \mu_x\mu_y\right]/\sigma_x\sigma_y, \tag{19}$$

where μ_x, μ_y, σ_x and σ_y are the averages and standard deviations of $m_x = \sum_{j=1}^{N}(m_{i,j})$ and $m_y = \sum_{i=1}^{N}(m_{i,j})$; and:

$$VAR_c = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[(i-\mu)^2 m_{i,j}\right], \quad (20)$$

$$IDM_c = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[m_{i,j}/1+(i-j)^2\right], \quad (21)$$

$$SA_c = \sum_{i=2}^{2N}\left(i \times m_{x+y(i)}\right), \quad (22)$$

$$SV_c = \sum_{i=2}^{2N}\left[(i-SE_{ch})^2 m_{x+y(i)}\right], \quad (23)$$

$$SE_c = -\sum_{i=2}^{2N}\left[m_{x+y(i)}\log\left(m_{x+y(i)}\right)\right], \quad (24)$$

$$H_c = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[m_{i,j}\log\left(m_{i,j}\right)\right], \quad (25)$$

$$DV_c = variance\left(m_{x-y}\right), \quad (26)$$

$$DE_c = -\sum_{i=0}^{N-1}\left[m_{x-y(i)}\log\left(m_{x-y(i)}\right)\right], \quad (27)$$

where $m_{x+y(k)} = \sum_{i=1}^{N}\sum_{j=1}^{N}(m_{i,j})$, with $k = 2,3,\ldots,2N$, $i+j=k$, and $m_{x-y(k)} = \sum_{i=1}^{N}\sum_{j=1}^{N}(m_{i,j})$, with $k = 0,1,\ldots,N-1$, $|i-j|=k$; with:

$$CRL1_c = (HXY - HXY1)/max(HX, HY), \quad (28)$$

$$CRL2_c = (1-exp[-2.0(HXY2-HXY)])^{1/2}, \quad (29)$$

where HX and HY are entropies of $m_{x(i)}$ and $m_{y(j)}$, $HXY = -\sum_{i=1}^{N}\sum_{j=1}^{N}[m_{i,j}\log(m_{i,j})]$, $HXY1 = -\sum_{i=1}^{N}\sum_{j=1}^{N}[m_{i,j}\log(m_{x(i)}m_{y(j)})]$, and $HXY2 = -\sum_{i=1}^{N}\sum_{j=1}^{N}[m_{x(i)}m_{y(j)}\log(m_{x(i)}m_{y(j)})]$, and:

$$MCC_c = (second\ largest\ eigen\ value\ of\ Q)^{1/2}, \quad (30)$$

where $Q_{i,j} = \sum_{k}^{N}\left[\left(m_{i,k}m_{j,k}\right)/\left(m_{x(i)}m_{y(k)}\right)\right]$.

3. Skin lesion classification

Here, first the set of features for skin lesion diagnosis are constructed, and then classified. The classification process must be accurate, since it is used to assist dermatologists in their diagnosis; however, the accuracy of the classification depends on several factors, such as a reliable dataset. The pre-processing step in this study included data normalization, dataset balancing and feature selection. The classification was carried out using the Weka library [26].

3.1. Data pre-processing

The data pre-processing step, which precedes the classification process, normalizes the dataset values from the feature extraction process as they contain different ranges, and some classifiers cannot handle such differences. The normalization procedure scales all numeric values in the dataset to within the same interval [0,1] by computing:

$$xn_{im} = [x_{im} - \min(x_{im})]/[\max(x_{im}) - \min(x_{im})]\,, \tag{31}$$

where x_{im} is the actual value of the feature m in the sample i, with the minimum and maximum values of features of all the samples, and xn_{im} is the normalized value of the same feature m in the same sample i.

Another problem is unbalanced datasets. Here the dataset was composed of 916 samples of benign lesions and 188 samples of malignant lesions. These unbalanced datasets with different numbers of samples in each class can decrease the accuracy of the evaluation result, since classifiers tend to prioritize classes with the highest occurrence. Sampling methods are used to overcome such problems [27], and in this work a resampling procedure was applied to the dataset [26]. This procedure produced a random subsample of the dataset using sampling with replacement and a uniform distribution of the samples was made.

Another problem that also affects the performance of classifiers is the choice of meaningful features to represent the images. Therefore, feature selection algorithms that aim to find the best feature subset are used. These algorithms are usually a combination of both search and evaluation methods. Search strategies can be applied to select a candidate subset from extracted features of skin lesions, which is evaluated and compared to the previous best subset until a given stopping criterion is reached. The feature subset selection consists of finding features through search or ranking methods in order to identify a candidate feature subset for evaluation process.

The ranking method produces a ranked list of features based on the evaluation process. The search method influences the search direction and execution time of the selection process depending on the search strategy adopted, which can be complete, sequential, or random [28]. The sequential search strategy is usually used for skin lesion feature selection and it can be by the forward, backward or bi-direction selection. 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 evaluation method, whereas the

backward selection process starts with all features and the worst features are removed at each iteration. The bi-direction selection combines both the forward and backward searches.

The evaluation process using filters allows for assessing the quality of selected features without using any classification algorithms. Each candidate subset is evaluated by applying an independent criterion, which can be based on several measures to compare it with the best current subset previously established. If the new evaluated subset is considered better then it becomes the best current subset. These measures can be defined as [29]:

- Distance measures that try to find the feature that can separate the classes as far as possible from each other;
- Information measures that establish the information gain from a feature and the feature with the most information is preferred; and
- Dependency measures that are also known as correlation measures applied to evaluate the ability to predict the value of one feature from the value of another, or how strongly a feature is in regard to the class.

In this study, six feature selection algorithms, based on the measures discussed above and on a feature transformation algorithm, were used to generate different subsets of features. These six algorithms are commonly used for the selection of skin lesion features [6], since they present several advantages over others, such as computationally efficient, simpler and faster algorithms, independent evaluation criteria, and ability to overcome over-fitting.

a) Relief-F feature selection [30]: This algorithm is an extension of the relief algorithm to deal with noise and multi-class problems. The dataset samples are randomly defined. For each sample that is defined, the closest samples of the same and different classes are selected using a nearest-neighbour algorithm [31]. The quality of each feature is estimated, according to its value in regard to these closest samples.

b) Information gain-based feature selection [32]: This algorithm estimates the quality of a feature, according to its information gain in regard to the class. The information gain between each feature F and the class C is measured by the entropy H, according to the information theory criteria [33]. Therefore, the features that have high information gain $Ig_{(C,F)}$ are considered the most relevant, where $Ig_{(C,F)} = H(C) - H(C|F)$.

c) Gain ratio-based feature selection (GRFS) [26]: This algorithm is also based on the entropy H and it estimates the quality of a feature F, according to its gain ratio in regard to the class C. Therefore, the features that have high gain ratio $Gr_{(C,F)}$ are considered the most relevant, where $Gr_{(C,F)} = [H(C) - H(C|F)]/H(F)$.

d) Correlation coefficient-based feature selection [32]: This algorithm estimates the quality of a feature, according to its Pearson's correlation coefficient in regard to the class. The correlation coefficient is computed by a covariance and variance between the features and the class.

e) Correlation-based feature selection (CFS) [34]: This algorithm tries to find a set of features that are highly correlated with a class and with low inter-correlation between them. The degree of correlation between the features is computed by a symmetrical uncertainty, which is a modified version of the information gain measure.

f) Principal-component analysis (PCA) [35]: Here the features are transformed to a PC based on a correlation matrix, where eigenvectors (vectors of features) are defined, according to some percentage of the variance in the original data. The worst eigenvectors are removed and the new features are ranked, according to the best eigenvalues.

All feature selection algorithms discussed above are single-feature evaluators, with exception of CFS that is a feature subset evaluator. The single-feature evaluators are used with a ranking method, where the features are ranked individually, according to their evaluation, i.e., the most relevant. The number of features to retain is previously defined. The feature subset evaluators measure a subset of features and they return a value that is used in the search [26]. In this study, both the greedy stepwise and best first search methods were applied. The greedy stepwise method searches feature subsets in either forward or backward directions in a greedy way. The selection process must stop when the addition or removal of any feature worsens the outcome of the best-found subset until that moment. The best first method searches the feature subsets by greedy hill-climbing, and the search direction can be forward, backward or bi-direction.

3.2. Classification

In this study, the focus is on models with a single classifier that can choose the best classification using different datasets, e.g., using a stratified k-fold cross-validation procedure [26]. This approach splits the training set in k subsets of equal size and the

procedure is repeated k times. In each procedure, one subset is employed as a test set while the others are used as the training set. The best model is chosen, according to its performance, which is measured by averaging the accuracy obtained from each trial. This procedure can be applied to avoid over-fitting while testing the capacity of the classifier to generalize. In addition, this approach has shown good results compared with other procedures [36].

Six different categories of classifier were applied in this work to evaluate the dataset from the extracted features: the k-nearest neighbours (KNN) [31], Bayes networks (Bayes Net) [37], C4.5 decision tree [38], multilayer perceptron (MLP) [39], support vector machine (SVM) [40] which is the most commonly used classifiers, according to the categories presented by Oliveira et al. [6]. In addition, the optimum-path forest (OPF) classifier [41] was also used in this study. To the best our knowledge, no previous study has used this later classifier to identify skin lesions in images.

a) kNN: Here, a search algorithm and a distance function are used to assess which sample of the training set is closest to an unknown sample and then assigning the unknown sample to the class with the majority of k-nearest neighbours. The main advantages of these classifiers are their simplicity to implement and the possibility to add new samples to the training set at any time.

b) Bayes Net: This is a Bayesian learning-based algorithm [37] that computes the probability of a given set of features to belong to each class, assuming that the features are independent. The Bayes Net learning uses search algorithms and quality measures, which provide a network structure and conditional probability distributions.

c) C4.5: This algorithm is used to create a decision tree [42] that has a structure similar to a flowchart, in which each internal node (non-leaf) represents a test of a feature, each branch represents the result of the test, and each external node (leaf) indicates a prediction of the class. A complete decision tree can contain unnecessary structures, and strategies of pre-pruning and post-pruning can be performed to simplify its structure. Pre-pruning involves decision making during the tree building process, whereas in the post-pruning this is done afterwards. The C4.5 algorithm divides the features at the nodes based on information gain. It prevents overfitting which is also a form of pre-pruning. The post-pruning in C4.5 yields a dense decision tree very quickly. It can also deal with situations in which two features that individually present no contribution, but are powerful predictors when combined [26].

d) MPL: This algorithm is one of the most commonly used architectures of artificial neural network (ANNs) [39] that 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 MPL can include one or more layers of processing, also called hidden layers, placed between the input and output layers. Back-propagation is a supervised learning method widely used in the MLP architecture, which consists of forward and backward processes applied to adjust the weight values of the connections. The MLP algorithm has good capability and flexibility to overcome various non-separable problems.

e) SVM: This classifier is used to build a hyper-plane to separate data, according to the defined classes. This kind of classifier has been commonly applied to classify skin lesions due to its good overall properties. Furthermore, kernel functions simplify the process of separating the non-linear data using a simple hyper-plane in a high dimension feature space. The radial basis function (RBF) and polynomial kernels have been frequently used in several different studies [6]. For the SVM classifier, Platt's [43] sequential minimal optimization algorithm was used.

f) OPF: This is 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 classifier shows some interesting properties, such as speed, simplicity, ability to deal with multi-class classification and overlapping between classes, parameter independence and no assumptions are based on the shape of the classes. For the application of the OPF classifier, it was used the Weka library based on LibOPF [41] as proposed by Amorim et al. [44].

The performance of the classification was evaluated using accuracy (ACC), sensitivity (SE) and specificity (SP) measures, which are based on outcomes of classifiers, according to the predicted class and known class. These outcomes represent the number of correct (true) and incorrect (false) classification for each class, positive and negative. These measures are commonly used according to Oliveira et al. [6] and they are defined as:

$$ACC = \frac{TP+TN}{P+N} \times 100\%, \quad (32)$$

$$SE = \frac{TP}{TP+FN} \times 100\%, \quad (33)$$

$$SP = \frac{TN}{TN+FP} \times 100\%, \quad (34)$$

where P is the number of positive samples and N is the number of negative samples of the dataset. Here, the positive samples represent the benign lesions and the negative samples the malignant lesions. Therefore, TP (true positive) is the number of correctly classified benign lesions, TN (true negative) is the number of correctly classified malignant lesions, FP (false positive) is the number of incorrectly classified malignant lesions, and FN (false negative) is the number of incorrectly classified benign lesions.

A cost function C adopted from Barata et al. [13] is used to deal with the trade-off between SE and SP, which is defined as:

$$C = \frac{c_{10}(1-SE)+c_{01}(1-SP)}{c_{10}+c_{01}}, \quad (35)$$

where c_{10} is the cost of an incorrectly classified benign lesion (FN), and c_{01} is the cost of an incorrectly classified malignant lesion (FP). The costs used to evaluate the classification were $c_{10} = 1$ and $c_{01} = 1.5$, since an incorrect classification of a malignant lesion is more critical. The lower the value of the cost C, the better is the classification performance.

4. Experiment and Discussion

In order to evaluate the proposed feature extraction in the classification of benign and malignant skin lesions, two experiments were performed. First, the experiments for the skin lesion classifications using all features of the dataset are presented. Second, the experiments for the feature selection of skin lesions are presented as well as these for the lesion classification. In this section, classification results are described and discussed. In addition, the image dataset used to evaluate the results is presented, as well as the computational time of the system.

4.1. Dermoscopic image dataset

The dermoscopic images of pigmented skin lesions used to evaluate the extraction of features were collected from the International Skin Imaging Collaboration (ISIC) dataset

[45]. Examples of these images are shown in Figure 2. In addition, the images are paired with the expert manual that contains the skin lesion diagnoses, as well as the ground-truth lesion segmentations in the form of binary masks. In this study, a feature extraction approach, based on shape properties, colour variation and texture analysis, is proposed. Moreover, since the shape properties are obtained from the lesion borders, only the images where the lesion fitted completely within the image frame were selected so that the features could be extracted with greater precision. A total of 1,104 images were selected from the original dataset. Of these, 916 images were benign lesions and 188 images were malignant lesions. The images of the dataset were resized to an average resolution of 400×299 pixels to simplify their processing.

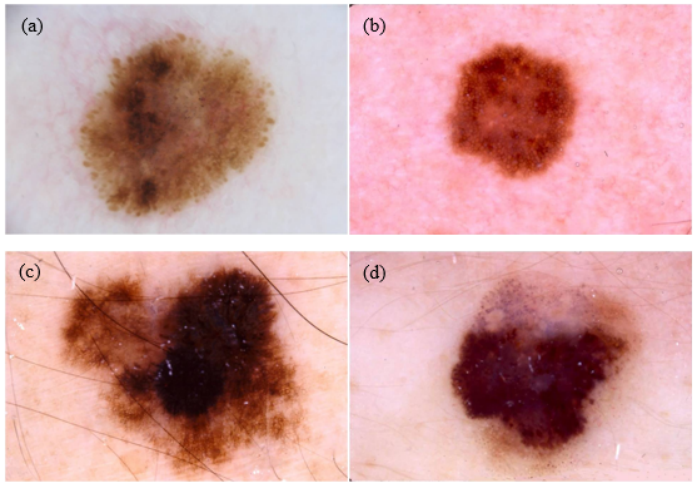

Figure 2: Four examples of dermoscopic images: (a) and (b) are benign lesions, (c) and (b) are malignant lesions.

4.2. Evaluation of the proposed feature extraction

The performance of the classification using all extracted features was evaluated by different classifiers, which were described in the previous section. Each classifier was used with different parameters to find the best results with a ten-fold cross-validation procedure. A set of parameters was defined based on previous studies that had used these classifiers for skin lesion classifications [13,14,46-48]. The best parameters from each classifier, according to the experiments are summarized in Table 2, and consequently, these parameters were used for all other experiments in this study.

The kNN classifier used a linear nearest neighbour search algorithm and three distance functions were compared, i.e., Euclidean, Chebyshev and Manhattan, to find the nearest neighbours. Different values of k were applied for each distance function and the number of neighbours used was $k = \{5,7,\dots,25\}$. The Bayes net classifier used a hill-climbing

search algorithm to find the network structures, and a simple estimator to estimate the conditional probabilities of a network. The parameter alpha for the simple estimator was settled with the following values: $A = \{0.1,0.2, \dots ,0.9\}$. The C4.5 classifier used two sets to define the minimum number of samples per leaf, $M_1 = \{2,4, \dots ,20\}$ and $M_2 = \{82,84, \dots ,100\}$, and the values of the confidence factor used for pruning were $CF = \{0.1,0.2, \dots ,0.9\}$.

Table 2: The best parameters for each classifier.

Classifier	Parameters
k-nearest neighbours	k:5 search algorithm: Linear NN Search (distance function: Manhattan)
Bayes networks	estimator: Simple Estimator (alpha: 0.1) search algorithm: hill-climbing
C4.5 decision tree	confidence factor: 0.3 minimum number: 2
Multilayer perceptron	one hidden layer: $(features + classes)/2$ learning rate: 0.3 momentum: 0.2
Support vector machine	complexity parameter: 10 kernel: RBF (gamma: 0.1)
Optimum-path forest	distance function: Euclidean

The MPL classifier analysed two values: one hidden layer of the neural network, with $H_1 = (features + classes)/2$ and the other $H_2 = classes$. The learning rate $L = 0.3$ is the number of the weights that were updated, and the momentum $M = 0.2$ was applied to the weights when updating. The SVM classifier analysed two kernels: the polynomial and RBF kernels. In the RBF kernel, the parameter gamma was carried out with different values of $G = \{0.001,0.002, \dots ,0.1\}$, and the complexity parameter $C = \{1,2, \dots ,10\}$ was applied to both kernels. And finally the OPF classifier compared three distance functions: Euclidean, Chebyshev and Manhattan, in order to find the distances between the feature vectors. Table 3 shows that good results were achieved using the proposed extracted features, mainly for the specificity of the malignant lesion classification (SP).

Table 3: Performance results for each classifier using all features.

Classifier	ACC	SE	SP	C
kNN	75.8%	69.4%	82.2%	0.229
Bayes Net	68.2%	54.0%	82.4%	0.290
C4.5	86.9%	82.2%	91.5%	0.122
MLP	74.5%	61.2%	87.7%	0.229
SVM	**91.7%**	**87.1%**	**96.2%**	**0.074**
OPF	**92.3%**	**87.5%**	**97.1%**	**0.067**

The best results were obtained by the OPF and SVM classifiers as shown in Table 3 (in bold), where both classifiers achieved a good generalization between the classes. Despite the fast training of the Bayes Net classifier, the classification results were not so expressive, as this classifier is sensitive to redundant features as it assumes that the features should be independent. The kNN classifier did not make a good distinction between the benign and malignant classes. This classifier is sensitive to the existence of irrelevant features, which explains these results. Although the MLP classifier is competent to solve several non-separable problems, it was not able to make a good distinction between the classes. Furthermore, this type of classifier needs a long training time for the size of the feature set. The C4.5 classifier, on the other hand, resulted in a more balanced classification result between the two classes. However, this classifier can have difficulties in dealing with correlated features. All these classifiers can achieve superior results using feature selection algorithms.

4.3. Performance evaluation using feature selection

In order to improve the classification results and to avoid over-fitting caused by a large number of features, different feature selection algorithms were performed to find the best features for the classification process. The single-feature evaluators that use a ranking method, i.e., the correlation coefficient, GRFS, information gain, relief-F and PCA, apply a certain number of features empirically defined by $N = \{25,50,75\}$, with the exception of the PCA that chooses enough eigenvalues to rank the new transformed features. The maximum number of features $F = 5$ was used in order to include the transformed features, and the proportion of variance $V = 0.95$ was used to retain a sufficient number of PC features. Accordingly, 31 eigenvalues are selected by the PCA algorithm to represent the vector with the new features. The number of nearest neighbours for the relief-F was defined as $k = 10$ for the feature estimation.

In the case of the feature subset evaluator, i.e., CFS, the greedy stepwise search method, in either forward or backward directions, was applied until the addition or removal of any feature caused a lower evaluation. This resulted in 37 features selected by the forward direction and 50 by the backward direction. The best first search method was also performed in the directions: forward, backward or bi-direction. However, experimental results, using the classifiers discussed in the previous section, showed that this second method did not improve the classification performance over that obtained

using the stepwise search method alone. Therefore, only the stepwise method was used with CFS for comparison with the other feature selection algorithms.

Figure 3 shows the percentage of selected features for each feature selection algorithm. The features were divided into five categories: shape, colour, fractal texture, wavelet texture and Haralick's texture; the percentage was computed individually for each category. Only the best configurations achieved in the classification results were used for each feature selection algorithm. Therefore, 75 features for the correlation coefficient, GRFS, information gain and relief-F algorithms, 31 features for the PCA algorithm, and 50 features for the CFS algorithm were selected and used.

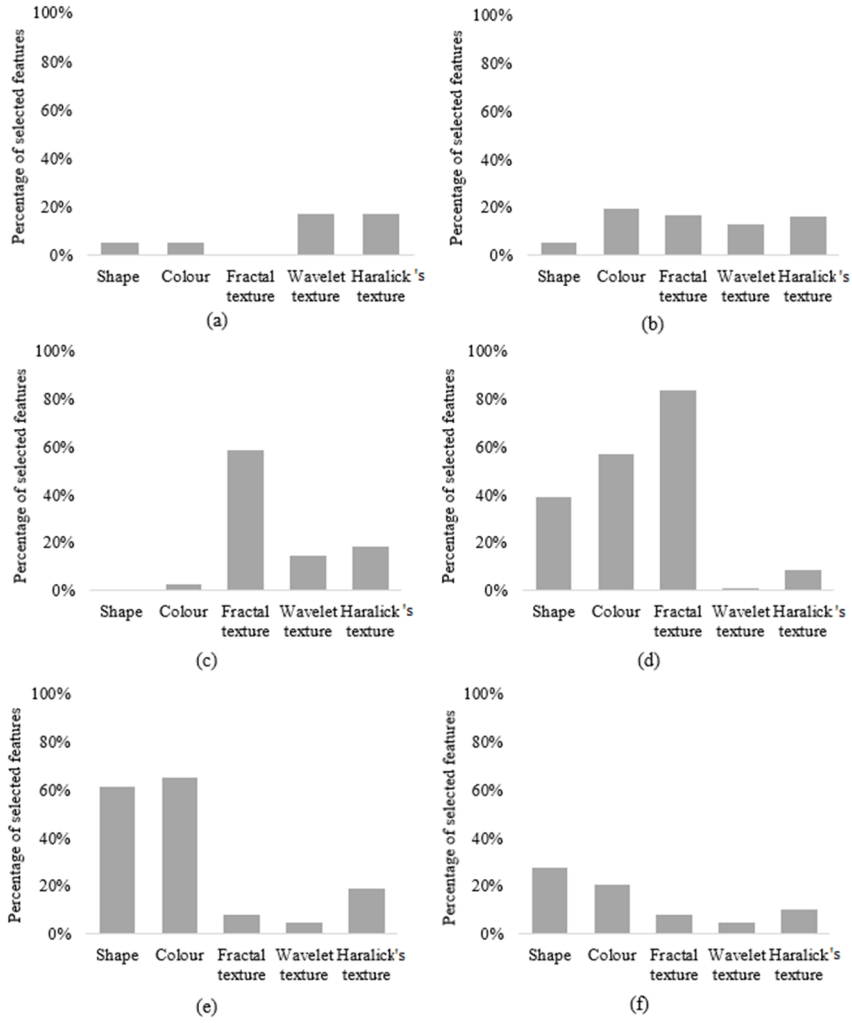

Figure 3: Percentage of selected features after applying feature selection algorithms: (a) correlation coefficient, (b) GRFS, (c) information gain, (d) relief-F, (e) PCA, and (f) CFS.

There were large differences between the feature selection algorithms. The correlation coefficient and information gain were the only algorithms that did not select features from all the categories. The PCA algorithm selected the greatest percentage of features from

the shape and colour categories, whereas the information gain algorithm selected the greatest percentage of texture features. The relief-F algorithm selected over 80% of the fractal texture, but it did not select the wavelet and Haralick's texture features proportionally. On the other hand, the GRFS and CFS algorithms selected features from amongst all the categories in a more uniform way.

Table 4 shows the best classification results using feature selection algorithms. These results, show that the OPF classifier with the features selected by the CFS algorithm and the MPL classifier with the features selected by the GRFS algorithm achieved superior results compared to the others, as presented in Table 4 (in bold). In addition, the features selected by the CFS and GRFS algorithms obtained better results for the classifiers than the other algorithms. As mentioned earlier, these algorithms selected the features more uniformly, which explains these results. The features selected by the PCA algorithm also obtained good results among the classifiers, despite the fact that it selected fewer features and with the C4.5 classifier it had a high SP result. However, this classifier did not stand out as much as the OPF and MPL classifiers, i.e., the C4.5 classifier had a higher classification cost.

Table 4: The best classification results using feature selection algorithms.

Classifier	FS algorithm (Search)	Features	ACC	SE	SP	C
kNN	CFS (Backward stepwise)	50	75.8%	67.8%	83.9%	0.225
Bayes Net	CFS (Forward stepwise)	37	74.4%	64.3%	84.4%	0.236
C4.5	PCA (Ranker)	31	89.7%	83.5%	**95.8%**	0.091
MLP	GRFS (Ranker)	75	**90.6%**	**86.6%**	94.6%	**0.086**
SVM	Relief-F (Ranker)	75	80.1%	76.1%	84.1%	0.191
OPF	CFS (Backward stepwise)	50	**91.6%**	**87.0%**	**96.2%**	**0.075**

The classification results are presented in more details in Figure 4, where it is possible to analyse the variation of the accuracy, sensitivity and specificity, according to the number of ranked features defined by the correlation coefficient, GRFS, information gain and relief-F algorithms. Figure 5 shows the variation of the results for the features selected by the PCA and CFS algorithms. In addition, the classification results for each feature selection are compared with the results using the entire set of features. From the feature selection, the OPF and kNN classifiers maintained their results, but they did not achieve better results. The MPL, C4.5 and Bayes Net classifiers had better results with the feature selection, whereas the SVM classifier achieved much better results with the entire set of features.

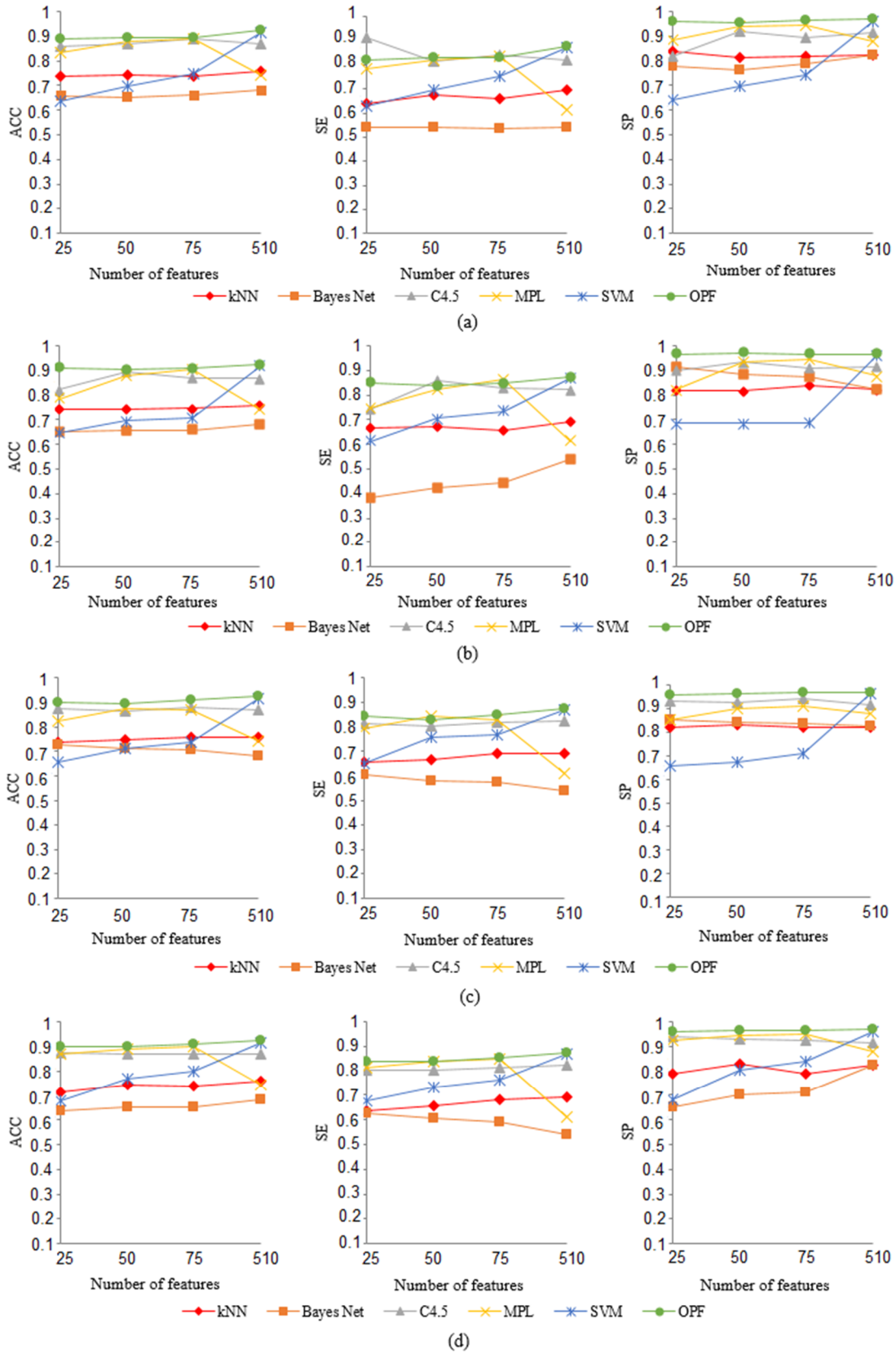

Figure 4: Variation of the classification measures, according to the number of features defined by the ranker of each feature selection algorithm for all features of the dataset: (a) correlation coefficient, (b) GRFS, (c) information gain, and (d) relief-F.

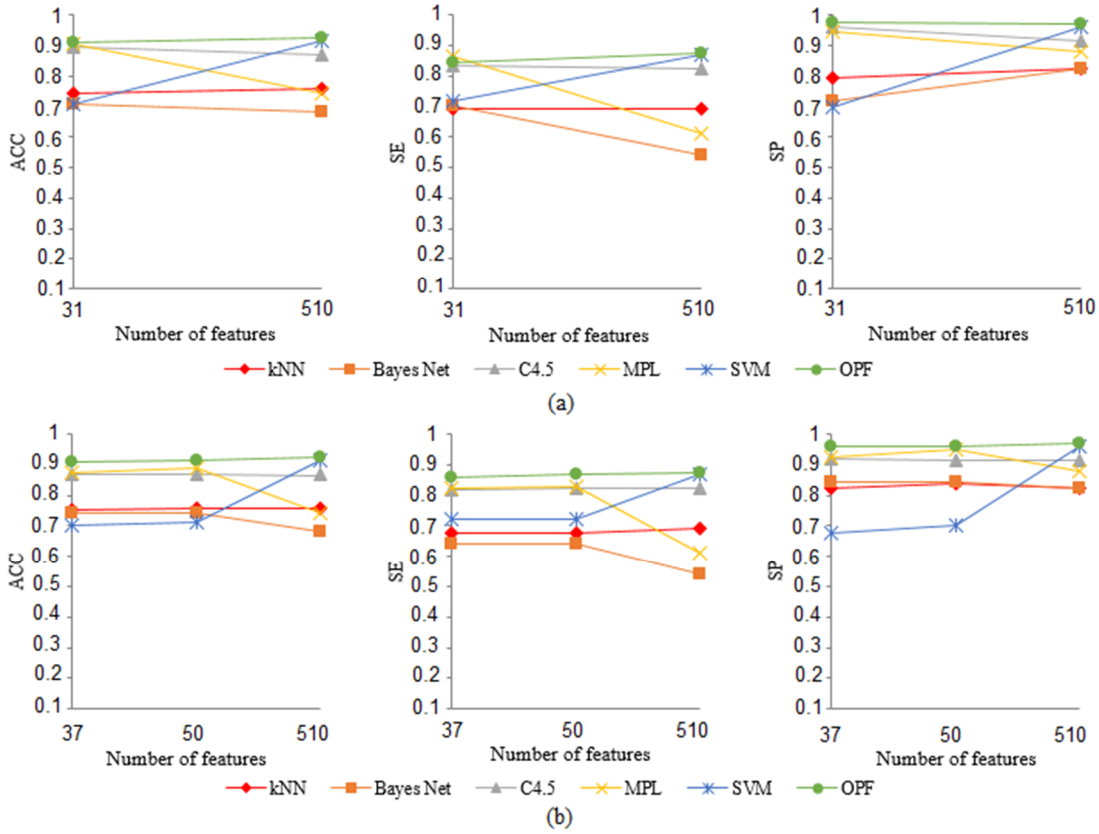

Figure 5: Variation of the classification measures, according to the automatic number of features established by the feature selection algorithms for all features of the dataset: (a) PCA and (b) CFS.

Several automatic diagnosis systems have been proposed using models with a single classifier for the skin lesion classification, as was used in this study. In Celebi et al. [2], the proposed classification model based on the SVM classifier achieved SE = 93.33% and SP = 92.34% in a dataset of 564 dermoscopic images. The authors extracted 11 shape, 354 colour and 72 texture features. In Abbas et al. [7], the proposed system obtained SE = 88.2% and SP = 91.3% in a dataset of 120 dermoscopic images. These authors applied the SVM classifier to distinguish between benign and malignant lesions using asymmetry, border quantification, colour and differential structure features; however, the number of features was not mentioned. Zortea et al. [49] proposed a computational system to differentiate benign lesions and melanoma using a discriminant analysis classifier, which achieved SE = 86% and SP = 52% in a dataset of 206 dermoscopic images. The feature extraction in this later work used 6 asymmetry, 11 colour, 3 border, 3 geometry and 30 texture features of skin lesions. Moreover, the above-mentioned studies also presented the lesion segmentation processes.

The lack of a lesion segmentation process may be considered a limitation of the present study; however ground-truth lesion segmentation masks were used in order to obtain a

more accurate computational system. A segmentation approach as presented by Ma and Tavares [50] can be used to evaluate the effectiveness of the proposed classification model in the segmented images. However, this study did not use all the images of the original dataset as mentioned earlier, the results cannot be compared with the results obtained in the studies using this same dataset and the ground-truth lesion segmentation masks presented in Gutman et al. [45]. These studies considered a set of 1279 images partitioned into training and test sets. The best results were achieved by Lequan et al. [51], with ACC = 0.855, SE = 0.547 and SP = 0.931. The authors proposed a novel method for melanoma recognition by leveraging very deep convolutional neural networks.

4.4. Computational time

The proposed approach was developed using: 1) Visual Studio Express 2012 environment, C/C++ and OpenCV 2.4.9 library for the feature extraction algorithms; and 2) Eclipse IDE 4.6.1 environment, java 1.8.0_111, and Weka 3.8 library for the classification algorithms. Table 5 shows the computational time of the processing of all images for each task, which includes feature extraction, and classification with and without feature selection using the best classification model. These values show that the feature extraction step was the most time-consuming; however, the computation time required by this step can be considerably decreased using optimized C/C++ implementations. All algorithms were performed on an Intel(R) Core(TM) i5 CPU 650 @ 3.20 GHz with 8 GB of RAM, running Microsoft Windows 7 Professional 64-bits.

Table 5: Computational time for the feature extraction and classification tasks considering all images.

Task	Features	Time
Shape feature extraction	18	10.26 min
Colour feature extraction	72	10.12 min
Fractal feature extraction	12	26.79 min
Wavelet feature extraction	240	34.37 min
Haralick's feature extraction	168	29.48 min
Classification (without feature selection)	510	8.01 sec
Classification (with feature selection)	50	5.91 sec

5. Conclusion and future works

In this paper, a combination of features based on shape properties, colour variation and texture analysis using different feature extraction methods was presented. Geometrical

properties, lesion asymmetry and border irregularity were used for the extraction of the shape properties. Statistical measures were used to analyse the colour features. The fractal dimension analysis, discrete wavelet transform and co-occurrence matrix methods were applied to obtain the texture features. Four colour spaces, i.e., *RGB*, *HSV*, *CIE Lab* and *CIE Luv*, were used for the extraction of both colour and texture properties. For the evaluation of the proposed feature extraction method, six different categories of classifiers were adopted; namely: kNN, Bayes networks, C4.5 decision tree, MLP, SVM and OPF. Furthermore, the classification performance was also evaluated using six different feature selection algorithms, which were: correlation coefficient, GRFS, information gain, relief-F, PCA and CFS.

Promising results were obtained with the proposed feature extraction for all the models evaluated. The best classification results were from the OPF classifier when all the features were used. The OPF results were: ACC = 92.3%, SE = 87.5% and SP = 97.1%. The OPF classifier also obtained the best classification results using feature selection algorithms for the skin lesion computational diagnosis system and achieved: ACC = 91.6%, SE = 87% and SP = 96.2%, when 50 features were selected using a CFS algorithm. It should be noted that the OPF classifier did not achieve better results by applying the feature selection algorithms, but it maintained the good results obtain when using all features. Moreover, the feature selection step reduced the computational time for the skin lesion classification. Another interesting result is that in most cases, the performance of the classifiers tends to improve when a percentage of features of all categories are select, i.e., shape, colour, fractal texture, wavelet texture and Haralick's texture by feature selection algorithms.

Future studies regarding the pigmented skin lesion classification of dermoscopic images should involve searching for new methods aiming to develop more efficient and effective systems for better skin lesion diagnoses. However, the classification results can be improved with ensemble methods [26]. Such methods consist of combining the results of several classification models in order to develop a more robust system that provides more accurate results than using a single classifier. Another solution to improve the classification results would be using deep learning architectures [52], since these architectures have shown that they have the capacity to learn from a large dataset in an unsupervised way.

Acknowledgments

The first author would like to thank CNPq ("Conselho Nacional de Desenvolvimento Científico e Tecnológico"), in Brazil, for her PhD grant. Authors gratefully acknowledge the funding of Project NORTE-01-0145-FEDER-000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, co-financed by "Programa Operacional Regional do Norte" (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional" (FEDER).

Conflict of interest statement

The authors report no conflict of interest.

References

[1] Scharcanski, J.; Celebi, M. E. (2013). *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, Berlin, Heidelberg.

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

[3] Garnavi, R.; Aldeen, M.; Bailey, J. (2012). Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 16 (6):1239-1252.

[4] Webb, A. R. (2003). *Statistical pattern recognition*. 2 edn. John Wiley & Sons, England.

[5] Guyon, I.; Gunn, S.; Nikravesh, M.; Zadeh, L. (2006). *Feature extraction: foundations and applications*, vol 207. Studies in Fuzziness and Soft Computing. Springer-Verlag, Berlin, Heidelberg.

[6] Oliveira, R. B.; Papa, J. P.; Pereira, A. S.; Tavares, J. M. R. S. (2016). Computational Methods for Pigmented Skin Lesion Classification in Images: Review and Future Trends. *Neural Computing and Applications*, 27:1-24.

[7] Abbas, Q.; Celebi, M. E.; Garcia, I. F.; Ahmad, W. (2013). Melanoma recognition framework based on expert definition of ABCD for dermoscopic images. *Skin Research and Technology*, 19 (1):e93-e102.

[8] Iyatomi, H.; Norton, K.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Classification of melanocytic skin lesions from non-melanocytic lesions*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5407-5410.

[9] Clawson, K. M.; Morrow, P.; Scotney, B.; McKenna, J.; Dolan, O. (2009). *Analysis of pigmented skin lesion border irregularity using the harmonic wavelet transform*. In: 13th International Machine Vision and Image Processing Conference Dublin, September 2-4 2009. IEEE, pp 18-23.

[10] Zhou, Y.; Smith, M.; Smith, L.; Warr, R. (2010). A new method describing border irregularity of pigmented lesions. *Skin Research and Technology*, 16 (1):66-76.

[11] Lee, T. K.; McLean, D. I.; Atkins, M. S. (2003). Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. *Medical image analysis*, 7 (1):47-64.

[12] Oliveira, R. B.; Marranghello, N.; Pereira, A. S.; Tavares, J. M. R. S. (2016). A computational approach for detecting pigmented skin lesions in macroscopic images. *Expert Systems with Applications*, 61:53-63.

[13] Barata, C.; Ruela, M.; Francisco, M.; Mendonça, T.; Marques, J. S. (2013). Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features. *IEEE Systems Journal*, 8 (3):965-979.

[14] Maglogiannis, I.; Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine*, 13 (5):721-733.

[15] Lissner, I.; Urban, P. (2012). Toward a unified color space for perception-based image processing. *IEEE Transactions on Image Processing*, 21 (3):1153-1168.

[16] Tkalcic, M.; Tasic, J. F. (2003). *Colour spaces: perceptual, historical and applicational background*. In: Proceedings in the IEEE Region 8 EUROCON 2003: Computer as a Tool Ljubljana, Slovenia, September 22-24 2003. IEEE, pp 304-308.

[17] Silva, C. S.; Marcal, A. R. (2013). Colour-based dermoscopy classification of cutaneous lesions: an alternative approach. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1 (4):211-224.

[18] Materka, A.; Strzelecki, M. (1998). *Texture analysis methods: a review*. COST B11 report, Technical University of Lodz. Brussels.

[19] Celebi, M. E.; Iyatomi, H.; Stoecker, W. V.; Moss, R. H.; Rabinovitz, H. S.; Argenziano, G.; Soyer, H. P. (2008). Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, 32 (8):670-677.

[20] Leo, G. D.; Paolillo, A.; Sommella, P.; Fabbrocini, G. (2010). *Automatic Diagnosis of Melanoma: A Software System Based on the 7-Point Check-List*. In: 43rd International Conference on System Sciences, Hawaii January 5-8 2010. IEEE, pp 1-10.

[21] Yuan, X.; Yang, Z.; Zouridakis, G.; Mullani, N. (2006). *SVM-based texture classification and application to early melanoma detection*. In: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, New York, August 30 - September 3 2006. IEEE, pp 4775-4778.

[22] Al-Akaidi, M. (2004). *Fractal speech processing*. Cambridge university press, New York.

[23] Scheunders, P.; Livens, S.; Van de Wouwer, G.; Vautrot, P.; Van Dyck, D. (1998). Wavelet-based texture analysis. *International Journal on Computer Science and Information Management*, 1 (2):22-34.

[24] Haralick, R. M.; Shanmugam, K.; Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3 (6):610-621.

[25] Abedini, M.; Chen, Q.; Codella, N. C. F.; Garnavi, R.; Sun, X. (2015). *Accurate and Scalable System for Automatic Detection of Malignant Melanoma*. In: Celebi, M. E., Mendonca, T., Marques, J. S. (eds) Dermoscopy Image Analysis. CRC Press, Boca Raton, pp 293-343.

[26] Witten, I. H.; Frank, E.; Hall, M. A. (2011). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco.

[27] Chawla, N. V. (2005). *Data mining for imbalanced datasets: An overview*. In: Maimon, O., Rokach, L. (eds) Data mining and knowledge discovery handbook. Springer, New York, pp 853-867.

[28] Dash, M.; Liu, H. (1997). Feature selection for classification. *Intelligent data analysis*, 1 (3):131-156.

[29] Liu, H.; Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, 17 (4):491-502.

[30] Kononenko, I. (1994). *Estimating attributes: Analysis and extensions of RELIEF*. In: Bergadano, F., De Raedt, L. (eds) Machine Learning: ECML-94, vol 784. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 171-182.

[31] Cover, T.; Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13 (1):21-27.

[32] Guyon, I.; Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157-1182.

[33] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27 (1):379-423.

[34] Hall, M. A. (2000). *Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning*. In: Proceedings of the 17th International Conference on Machine Learning, San Francisco, USA, June 29 - July 02 2000. Morgan Kaufmann Publishers Inc., pp 359-366.

[35] Hand, D.; Mannila, H.; Smyth, P. (2001). *Principles of Data Mining*. The MIT Press, London.

[36] Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection*. In: Proceedings of the 14th International Joint Conference on

Artificial Intelligence, Quebec, Canada, August 20-25 1995. Morgan Kaufmann Publishers Inc., pp 1137-1145.

[37] Congdon, P. (2007). *Bayesian statistical modelling*, vol 704. 2 edn. John Wiley & Sons, Chichester.

[38] Quinlan, J. R. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., USA.

[39] Haykin, S. S. (1999). *Neural networks: a comprehensive foundation*. Prentice Hall, Englewood Cliffs, USA.

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

[41] Papa, J. P.; Falcao, A. X.; Suzuki, C. T. (2009). Supervised pattern classification based on optimum-path forest. *International Journal of Imaging Systems and Technology*, 19 (2):120-131.

[42] Han, J.; Kamber, M. (2006). *Data Mining: concepts and techniques*. Elsevier, San Francisco.

[43] Platt, J. C. (1999). *Fast training of support vector machines using sequential minimal optimization*. In: Advances in kernel methods. MIT Press Cambridge, USA, pp 185-208.

[44] Amorim, W. P.; Falcão, A. X.; de Carvalho, M. H. (2014). *Semi-supervised pattern classification using optimum-path forest*. In: 27th SIBGRAPI Conference on Graphics, Patterns and Images, Rio de Janeiro, August 26-30 2014. IEEE, pp 111-118.

[45] Gutman, D.; Codella, N. C. F.; Celebi, E.; Helba, B.; Marchetti, M.; Mishra, N.; Halpern, A. C. (2016). *Skin Lesion Analysis toward Melanoma Detection: A Challenge* at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC), arXiv preprint arXiv:1605.01397.

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

[47] Maglogiannis, I.; Delibasis, K. K. (2015). Enhancing classification accuracy utilizing globules and dots features in digital dermoscopy. *Computer Methods and Programs in Biomedicine*, 118 (2):124-133.

[48] Rahman, M. M.; Bhattacharya, P.; Desai, B. C. (2008). *A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions*. In: 8th IEEE International Conference on International Conference on BioInformatics and BioEngineering, Athens, October 8-10 2008. IEEE, pp 1-6.

[49] Zortea, M.; Schopf, T. R.; Thon, K.; Geilhufe, M.; Hindberg, K.; Kirchesch, H.; Møllersen, K.; Schulz, J.; Skrøvseth, S. O.; Godtliebsen, F. (2014). Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists. *Artificial intelligence in medicine*, 60 (1):13-26.

[50] Ma, Z.; Tavares, J. M. R. S. (2016). A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model. *IEEE Journal of Biomedical and Health Informatics*, 20 (2):615-623.

[51] Lequan, Y.; Chen, H.; Dou, Q.; Qin, J.; Heng, P. A. (2016). Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. *IEEE Transactions on Medical Imaging*:1-11.

[52] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2 (1):1-127.

PART B - ARTICLE 5

SKIN LESION COMPUTATIONAL DIAGNOSIS OF DERMOSCOPIC IMAGES: ENSEMBLE MODELS BASED ON INPUT FEATURE MANIPULATION

Roberta B. Oliveira, Aledir S. Pereira and João Manuel R. S. Tavares

Submitted to an international journal (under review), 2017

ABSTRACT

Background and Objectives: The number of deaths worldwide due to melanoma has risen in recent times, partly because this type of skin cancer is the most aggressive. Computational systems have been developed to assist dermatologists in early diagnosis of skin cancer, or even to monitor skin lesions. However there still remains a challenge to improve classifiers for the diagnosis of such skin lesions. The main objective of this article is to evaluate different ensemble classification models based on input feature manipulation to diagnose skin lesions. *Methods:* Input feature manipulation processes are based on feature subset selections from shape properties, colour variation and texture analysis to generate diversity for the ensemble models. Three subset selection models are presented here: 1) a subset selection model based on specific feature groups, 2) a correlation-based subset selection model, and 3) a subset selection model based on feature selection algorithms. Each ensemble classification model is generated using an optimum-path forest classifier and integrated with a majority voting strategy. The proposed models were applied on a set of 1104 dermoscopic images using a cross-validation procedure. *Results:* The best results were obtained by the first ensemble classification model that generates a feature subset ensemble based on specific feature groups. The skin lesion diagnosis computational system achieved 94.3% accuracy, 91.8% sensitivity and 96.7% specificity. *Conclusions:* The input feature manipulation process based on specific feature subsets generated the greatest diversity for the ensemble classification model with very promising results.

Keywords: Image Classification; Feature extraction; Feature Selection; Ensemble of Classifiers; Computational Diagnosis.

1. Introduction

Skin cancer is one of the most common cancers worldwide, and its incidence has increased in recent years [1]. Computational diagnosis systems have been developed to assist dermatologists in early diagnosis of skin cancer from dermoscopic images. The search for more efficient classifiers for these computational systems is a challenging task. Several studies have proposed an ensemble of classifiers, commonly known as a multiple classifier system or an ensemble classification model to improve skin lesion classifications from dermoscopic images [2-4]. An ensemble of classifiers consists of

integrating several classification models in order to develop a more robust system that provides more accurate results than by using a single classifier [5]. Voting methods [6] are some examples of the integration strategies based on the outputs of the input classifiers for ensemble classification models, e.g., majority voting that counts the votes for each class of all the input classifiers and then designates the class with the majority votes as the classification result. Statistical methods, such as average, sum, product and median can also be used for this same purpose [7]; such as for cases of numeric predictions.

One important requisite for constructing ensembles is to ensure diversity between the classification models, which can be obtained by manipulating the modelling process or the input data. Manipulation the modelling process consists of constructing the classification models by using either different learning algorithms or a single learning algorithm but with different parameters. The more popular approaches for input data manipulation are to manipulate the training samples and the input features. Algorithms used to manipulate the training samples can generate multiple hypotheses, in which a learning algorithm is applied to different subsets of the training samples. Bagging and boosting algorithms are the traditional ways to manipulate the training samples [5], and their hypotheses are integrated by a vote method. The bagging algorithm consists of randomly splitting the original dataset in several training subsets of the same size based on sampling with replacement, which can be applied to any learning algorithm. Likewise, the boosting algorithm combines the classification outputs using the same learning algorithm; however, this type of algorithm is iterative, where each new model is based on the result of the previously built one.

Algorithms for manipulating the input features generate ensembles based on different feature subsets available to the learning algorithm. This process can be, for example, the random splitting of a set of features into subsets [8], or by using a feature selection algorithm combined with manipulation of the training samples [4]. One challenge that affects the performance of classifiers is how to define which features are meaningful to describe the patterns of interest. Consequently, feature selection algorithms [9] can be used for the ensemble construction in order to achieve superior performance for skin lesion classifications.

This article presents ensemble classification models based on input feature manipulation to improve skin lesion computational diagnosis from dermoscopic images.

Two examples of pigmented skin lesions in dermoscopic images are shown in Figure 1. The main contributions of this study are the feature subset selection models based on specific feature groups and the feature selection algorithms for the input feature manipulation. To the best of our knowledge, few studies based on ensemble models and feature manipulation for skin lesion classification have been presented with successful results [10,11].

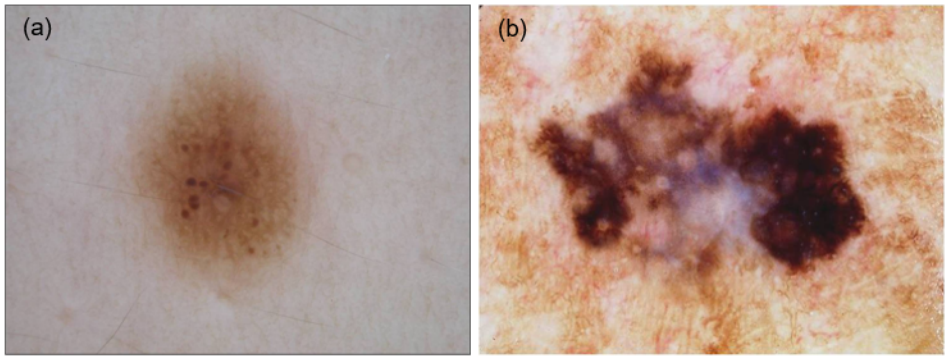

Figure 1: Two examples of pigmented skin lesions: (a) benign lesion and (b) malignant lesion.

This article is organized as follows: Studies relating to the ensemble methods for skin lesion classification are discussed in Section 2. The proposed ensemble classification models based on input feature manipulation are presented in Section 3. The experimental results and their discussion, which include the evaluation process, feature subset and feature selection evaluations, ensemble classification model evaluation and comparison between the classification algorithms used are given in Section 4. Finally, the conclusions drawn for the proposed ensemble classification models and future works about the skin lesion classification are pointed out in Section 5.

2. Related studies

An overview of computational methods for pigmented skin lesion classification in images, which addresses the feature extraction and selection, and classification steps, is presented in Oliveira et al. [12]. The ensemble of classifiers based on input data manipulation has been recently adopted for skin lesion classification to achieve better results than single classifiers. Several algorithms can be used for constructing ensembles; e.g., the AdaBoost [13], which is a popular boosting algorithm that maintains a set of weighting systems for the training samples according to a computed error rate. In Barata et al. [2], the proposed classification system using AdaBoost achieved a sensitivity of 96% and specificity of 80% using a dataset of 176 dermoscopic images. The authors

obtained the best results by using colour features and with combinations of two to five base classifiers for the detection of melanomas and nevi.

Random forest [14] is another ensemble algorithm used for skin lesion computational diagnosis. This algorithm is a variation of the bagging algorithm that is used to create an ensemble of decision trees that ensure the diversity by using a random selection of features to split each tree node. Its error rates are comparable to AdaBoost, but are more robust with respect to noise. Rastgoo et al. [15] proposed an automatic system to differentiate melanoma from dysplastic nevi by using texture features and random forest. This system achieved a sensitivity of 98% and specificity of 70% in a dataset of 180 dermoscopic images. Barata et al. [10] built a system for melanoma detection using the random forest algorithm and obtained a sensitivity of 98% and specificity of 90% in a dataset of 200 dermoscopic images, and a sensitivity of 83% and specificity 76% in a dataset of 482 images. This system is based on the global and local feature fusion of colour and texture properties.

The random forest algorithm also obtained the best results in a system proposed by Garnavi et al. [16]. The authors developed an optimized selection and integration of features derived from texture, border and geometrical properties. This system achieved an accuracy of 91.26% in a dataset of 289 dermoscopic images. Rastgoo et al. [11] proposed an automatic framework based on ensemble methods to differentiate melanoma from dysplastic and benign lesions. This framework used a random forest algorithm and a combination of colour and texture features based on global features, and obtained a sensitivity of 94% and specificity of 92% in a dataset of 193 dermoscopic images.

Other ensemble classification models have also been proposed for skin lesion classification. In Abedini et al. [17], an ensemble model, based on feature random subsets, a linear support vector machine (SVM) classifier and forward model selection for the ensemble fusion, was proposed. The best results were obtained by concatenating the pattern prediction values, which are considered middle-level features. This model achieved an accuracy of 91%, sensitivity of 97% and specificity of 65% for malignant and benign lesions in a dataset of 200 dermoscopic images. Schaefer et al. [4] proposed a multiple classifier system to deal with imbalanced classes. Such a system consists of a random under-sampling method, an SVM using a polynomial kernel, and a neural network for the classifier fusion. In addition, a feature selection algorithm is applied to each classifier, and a diversity measure is used for pruning a pool of classifiers. The

authors used features based on shape, colour and texture properties for the melanoma and benign lesion classification, and they obtained an accuracy of 93.83%, sensitivity of 93.76% and specificity of 93.84% in a dataset of 564 dermoscopic images.

3. Description of the proposed ensemble classification models

In this section, the ensemble classification models based on input feature manipulation for skin lesion computational diagnoses, as well as the dermoscopic image dataset used are presented. Figure 2 gives an overview of three different models for the input feature manipulation in order to generate diversity for the ensembles of classifiers. Given a dataset $T = \{x_{\mathrm{p}}, y_p\}$, with $p = 1,2 \dots, n$, according to the number of images n, where x_{p} is a sample, and y_p is the class to which it belongs. Each sample x_{p} is composed of a set of features F_{pq}, where $q = 1,2 \dots, m$, and m is the number of features. An ensemble $P = \{C_1, C_2, \dots, C_i\}$, with $i = 1,2, \dots, E$, and E is the ensemble size, where C_i is composed of the classification models obtained with the input feature manipulation, a base classifier using optimum-path forest (OPF) [18] and an integration strategy. One classification model is obtained in each iteration i by a subset of feature S_i that is sampled from F_{pq} based on specific feature groups or with a feature selection algorithm (Figures 2a and 2b, respectively). The classification models are also obtained by applying several feature selection algorithms A_i from F_{pq} (Figure 2c).

3.1. Dermoscopic image dataset

The dermoscopic image dataset is composed of pigmented skin lesions, which were collected from International Skin Imaging Collaboration (ISIC) dataset [19]. In addition, the images are paired with an expert manual that contains the skin lesion diagnoses, as well as the ground-truth lesion segmentations in the form of binary masks. In this study, the extracted features from the images are based on shape properties, colour variation and texture analysis. The images in which the lesion did not fully fit within the image frame were removed from the original dataset, since the shape properties are obtained from the lesion borders. In the end, a total of 1104 images were used from the original dataset. Of these, 916 images were benign lesions and 188 images were malignant lesions. The images of the dataset were proportionally resized to an average resolution of 400×299 pixels to simplify their processing.

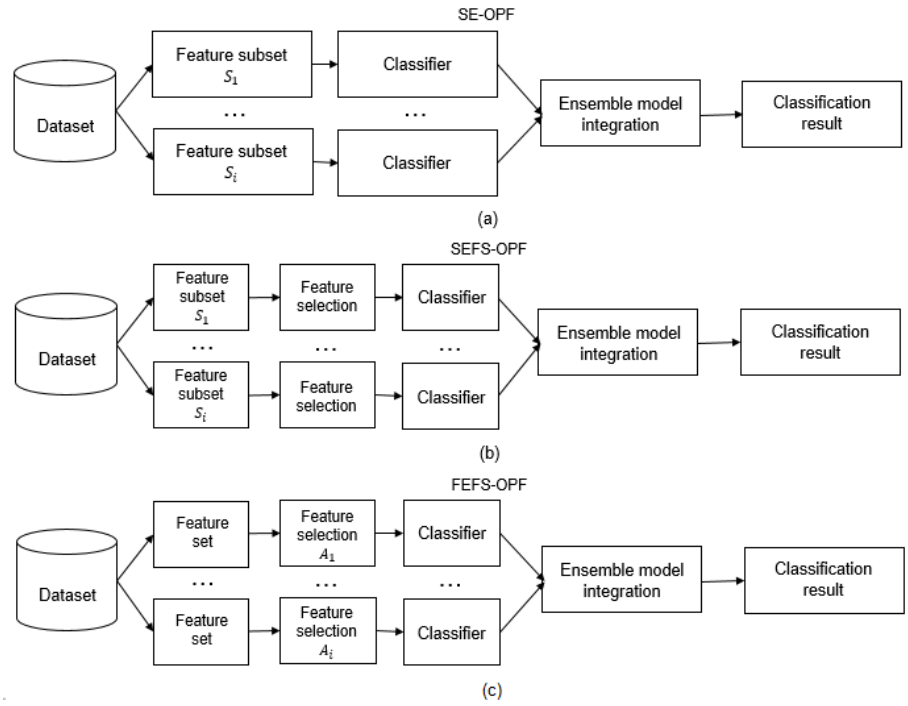

Figure 2: Overview of the proposed ensemble classification models based on input feature manipulation for the skin lesion computational diagnosis: (a) feature subset ensemble (SE-OPF), (b) feature subset ensemble with a feature selection algorithm (SEFS-OPF), and (c) feature set ensemble with feature selection algorithms (FEFS-OPF).

3.2. Feature extraction and data pre-processing

The feature extraction process is based on the intensities of the pixels belonging to the binary masks defined by specialists, in which the non-zero pixels belong to the lesion, and the others to the background skin. A combination of features, based on shape properties, colour variation and texture analysis using different feature extraction methods, were used in this study. A total of 512 features were extracted for each skin lesion image. Of these, 18 features were related to the shape properties, 72 features to colour variation, and 420 features to the texture analysis.

a) Shape properties: shape measures are computed based on the geometrical properties, lesion asymmetry and border irregularity. To assess the geometrical properties of the lesion, the area, perimeter, equivalent diameter, compactness, circularity, solidity, rectangularity, aspect ratio and eccentricity [20-22] were computed. To assess the lesion asymmetry, three features were computed from the lesion, i.e., the average, variance and standard deviation. These features were obtained from the ratios between the shortest and longest distances of each pair of the semi-lines that represent the perpendicular lines by overlapping the two sub-regions of the lesion along an axis [23]. To assess the border irregularity, a number of peaks, valleys and straight lines

of the border were computed using the vector product and inflexion point descriptors based on small and large irregularities of the border from a one-dimensional border [23].

b) Colour variation: the *RGB*, *HSV*, *CIE Lab* and *CIE Luv* colour spaces [24] were used to analyse the colour variation of the skin lesions. The *RGB* colour space is commonly used, as the images are originally obtained using this space. Moreover, the original *RGB* colour image can be used for conversion to other colour spaces, and several studies have achieved good results from this colour space [20,25]. The *HSV*, *CIE Lab* and *CIE Luv* colour spaces represent colours based on human perception. Furthermore, *CIE Lab* and *CIE Luv* are unified colour spaces and can simplify the identification of colour properties, as it is easy to maintain colour-difference ratios [26]. Six statistical measures, i.e., average, variance, standard deviation, minimum and maximum colours, and colour skewness, are computed for each colour channel in the region of the lesion using the aforementioned four-colour spaces that correspond to 12 channels.

c) Texture analysis: three different texture analysis methods were adopted to obtain the best features to represent the skin lesion texture based on colour images; namely, fractal dimension analysis [27], discrete wavelet transform (DWT) [28] and co-occurrence matrix [29]. The *RGB*, *HSV*, *CIE Lab* and *CIE Luv* colour spaces were also used for the texture analysis. The bi-dimensional fractal dimension using a box-counting method [27] is computed individually for each channel of the colour spaces. The energy and entropy measures from the coefficients obtained by DWT are computed for each of the 10 Haar wavelet sub-bands obtained by a-three-level decomposition, as well as for each channel of the colour spaces. Co-occurrence matrices were obtained for each channel of the colour spaces, and the intensities of each channel were quantized with 16 intensity levels. The distance between each reference pixel and its neighbours was one pixel, and four orientations $\theta = (0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ})$ were used. A normalized matrix was obtained from the matrices corresponding to the four orientations. From the normalized co-occurrence matrix, 14 statistical measures based on Haralick's texture features [29] were extracted from the image. These measures are the angular second moment, contrast, correlation, variance, inverse difference moment, sum average, sum variance, sum entropy, entropy, variance difference, entropy difference, information measure of correlation 1, information measure of correlation 2, and the maximal correlation coefficient.

Therefore, 12 features were extracted from the fractal dimension analysis, 240 features were extracted from the discrete wavelet transform, and 168 features were extracted from the co-occurrence matrix.

As the values of the dataset obtained by feature extraction contain different ranges they were normalized into the same interval [0,1] for the skin lesion classification process. The normalization procedure scales all numeric values in the dataset by computing:

$$xn_{pq} = \frac{x_{pq} - \min(x_{pq})}{\max(x_{pq}) - \min(x_{pq})}, \quad (1)$$

where $p = 1,2 \ldots, n$, $q = 1,2 \ldots, m$, n is the number of samples and m is the number of features. Thus, x_{pq} is the actual value of feature q in the sample p, with the minimum and maximum values of features of all the sets of samples, and xn_{pq} is the normalized value of same feature q in the same sample p. In addition, the unbalanced dataset problem is considered in this study, since the dataset is composed of 916 samples of benign lesions and 188 samples of malignant lesions. These unbalanced datasets concerning the number of samples in each class can decrease the accuracy of the evaluation results, since the classification tends to be based on the classes with the largest number of occurrences. Different sampling methods [30] have been used to solve such classification problems [4,31]. Here, the resampling procedure was applied to the dataset [5]. This procedure produces a random subsample of the dataset using sampling with replacement and the class distribution is made into a uniform distribution.

3.3. Feature selection

The feature selection process aims to find the best feature subsets to generate the ensembles of classifiers. Feature selection algorithms are usually a combination of both search and evaluation methods [9]. Search methods can be applied to select a candidate subset from extracted features of skin lesions, which is evaluated and compared to the previous best subset until a given stopping criterion is reached. In this study, six feature selection algorithms were applied to generate different feature subsets for the ensemble of classifiers; namely, Pearson's correlation coefficient [32], gain ratio-based feature selection (GRFS) [5], information gain-based feature selection [32], relief-F [33], principal-component analysis (PCA) [34] and correlation-based feature selection (CFS) [35]. These algorithms have been commonly used for skin lesion feature selections [12], since they have several advantages, such as computationally efficiency, are simple and

fast algorithms, independent evaluation criteria, and have the ability to overcome over-fitting.

All feature selection algorithms mentioned earlier are single-feature evaluators, with the exception of CFS that is a feature subset evaluator. The single-feature evaluators are used with a ranking method, where the features are ranked individually, according to their evaluation, i.e., the most relevant. The number of features to be maintained is previously defined. The feature subset evaluators measure a subset of features and they return a value that is used in the search [5]. In this study, both the greedy stepwise and best first search methods were adopted.

The greedy stepwise method searches feature subsets in either the forward or backward directions in a greedy way. The selection process must stop when the addition or removal of any feature occurs that worsens the outcome of the best-found subset until that moment. The best first method searches the feature subsets by greedy hill-climbing, and the search direction can be forward, backward or bi-direction. 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 evaluation method, whereas the backward selection process starts with all features and the worst features are removed at each iteration. The bi-direction selection combines both the forward and backward searches.

3.4. Base classifier and integration strategy

In this study, the focus is on homogeneous ensemble methods that are built with only one base classifier through input feature manipulation, and the classification model results are combined by an integration strategy. The number of base classifiers used defines the ensemble size. An OPF classifier [18] based on input feature manipulation for a set of training data was used to generate the ensemble classification models in this work.

The OPF classifier is 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 links of all pairs of nodes. The arcs are weighted by the distances between the feature vectors of their corresponding nodes. The Euclidean $D_E(i,j)$, Chebyshev $D_C(i,j)$ and Manhattan $D_M(i,j)$ distance functions [5] were used to measure the distances between the feature vectors:

$$D_E(i,j) = \sqrt{\sum_{q=1}^{m} \left| x_{iq} - x_{jq} \right|^2}, \quad (2)$$

$$D_C(i,j) = \sum_{q=1}^{m} \left| x_{iq} - x_{jq} \right|, \quad (3)$$

$$D_M(i,j) = \max_{q=\{1,2,..,m\}} \left| x_{iq} - x_{jq} \right|, \quad (4)$$

where x_{iq} is the feature value of a sample i, x_{jq} is the feature value of a sample j, $q = 1,2 \ldots, m$, and m is the number of features.

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 classifier shows some interesting properties, such as speed, simplicity, ability to deal with multiclass classifications and overlapping between classes, parameter independence and no assumption is based on the shape of the classes. Ensembles of OPF classifiers for reducing the size of the training set using under-sampling were proposed by Ponti Jr and Rossi [36]. For the application of the OPF classifier, it was used the Weka library based on LibOPF [18] as proposed by Amorim et al. [37].

Applying a good integration method is also important for the performance of the ensemble model. The challenge is how to integrate the results produced by the base classifiers. Here, the majority voting method [6] combines the classification results to generate an ensemble model. This method analyses which class receives the majority votes based on the results of all base classifiers and therefore the ensemble model must have an odd number of classifiers.

3.5. Input feature manipulation for the ensemble classification models

The input feature manipulation process aims to generate diversity for an ensemble classification model with the combination of the best feature subsets for the base classifier. In this section, three different models for the feature manipulation of skin lesions are presented. These models are based on specific feature groups and feature selection algorithms in order to create different feature subsets.

3.5.1. Feature subset selection model based on specific feature groups

Different groups based on feature categories and feature extraction algorithms are used to select the feature subsets. The extracted features are divided in three categories: shape,

colour and texture. Also the texture extraction algorithms, i.e., fractal texture, wavelet texture and Haralick's texture, are analysed, as well as the colour and texture extraction algorithms for each colour space separately. In this study, the feature groups constructed for feature manipulation were:

- Group 1: shape, colour, and texture;
- Group 2: fractal texture, wavelet texture, and Haralick's texture;
- Group 3: shape + colour, shape + texture, and colour + texture;
- Group 4: *RGB* colour, *HSV* colour, *LAB* colour, and *LUV* colour;
- Group 5: *RGB* texture, *HSV* texture, *LAB* texture, and *LUV* texture;
- Group 6: shape + *RGB* algorithms, shape + *HSV* algorithms, shape + *LAB* algorithms, and shape + *LUV* algorithms; and
- Group 7: *RGB* algorithms, *HSV* algorithms, *LAB* algorithms, and *LUV* algorithms.

The effectiveness of the feature groups are also evaluated individually in the experimental results section. The feature subset selection model generates a feature subset ensemble (SE-OPF). Algorithm 1 shows the procedure to set up this ensemble classification model, which was used for the input feature manipulation based on the feature groups and was also used by the OPF classifier [18] and majority voting [6].

Algorithm 1 SE-OPF

Require:

Ensemble size E, training sample set T, feature set F, group-based feature subsets S_i' from the feature set F

Procedure:

1. **for** $i = 1$ to E **do**
2. Select one feature subset S_i from S_i'
3. Train the OPF classifier C_i using T with the selected feature subset S_i
4. **end for**
5. **for** each new sample **do**
6. Compute the majority voting V of all classification models of the ensemble C_i
7. **end for**

3.5.2. Correlation-based feature subset selection model

The correlation-based feature subsets were set up using the feature groups discussed in the previous section and a CFS algorithm for the feature selection. The CFS algorithm [35] tries to find a set of features that are highly correlated with the class and have low inter-correlation between them. The degree of correlation between the features is computed by a symmetrical uncertainty, which is a modified version of the information

gain measure. Such an algorithm is adopted for this subset selection model, since experimental results using the OPF classifier [18] showed that this algorithm improved the classification performance more than the other feature selection algorithms.

The correlation-based subset selection model generates a feature subset ensemble with a feature selection algorithm (SEFS-OPF). Algorithm 2 shows the procedure to set up this ensemble model, which was used for feature input manipulation based on feature groups and the CFS algorithm, as well as the OPF classifier [18] and majority voting [6].

Algorithm 2 SEFS-OPF

Require:

Ensemble size E, training sample set T, feature set F, group-based feature subsets S_i' from the feature set F

Procedure:

1. **for** $i = 1$ to E **do**
2. Select one feature subset S_i from S_i'
3. $FS \leftarrow$ Selected features from S_i using a CFS algorithm A
4. Train the OPF classifier C_i using T with the selected features FS
5. **end for**
6. **for** each new sample **do**
7. Compute the majority voting V of all classification models of the ensemble C_i
8. **end for**

3.5.3. Subset selection model based on feature selection algorithms

All features discussed in the previous sections were used to generate the feature subsets. The diversity for an ensemble classification model is obtained by using different feature selection algorithms; namely, correlation coefficient [32], GRFS [5], information gain [32], relief-F [33], PCA [34] and CFS [35]. This subset selection model generates a feature set ensemble with feature selection algorithms (FEFS-OPF). Algorithm 3 shows the procedure to set up this ensemble model, which was used for the input feature manipulation based on the feature selection algorithms, and with the OPF classifier [18] and majority voting [6].

4. Experimental results and Discussion

In this section, the classification results are described and discussed. In order to evaluate the effectiveness of the ensemble models for the classification of benign and malignant skin lesions, three experiments were performed. First, the experiments for the feature subset and feature selection evaluations; second, the experiments for the ensemble

classification model evaluation; and finally, the experiments to compare the results with the classification methods reported in the literature. In addition, the evaluation process used to evaluate the results is introduced.

Algorithm 3 FEFS-OPF

Require:

Ensemble size E, training sample set T, feature set F

Procedure:

1. **for** $i = 1$ to E **do**
2. $FS \leftarrow$ Selected features from F using one feature selection algorithm A_i
3. Train the OPF classifier C_i by using T with the selected features FS
4. **end for**
5. **for** each new sample **do**
6. Compute the majority voting V of all classification models of the ensemble C_i
7. **end for**

4.1. Evaluation process

The performance of the ensemble classification models based on the input feature manipulation as described in the previous section was evaluated by using a stratified k-fold cross-validation procedure [5]. This kind of procedure consists of splitting the training set in k subsets of equal size; the procedure being repeated k times. In each procedure, one subset is used as a test set while the others are used as the training set. The best model based on its performance is chosen. Performance is the average accuracy obtained from each trial. The k-fold cross-validation procedure can be applied to avoid over-fitting while testing the capacity of the classifier to generalize. In addition, it has shown good results compared with other procedures [38].

The measures used to evaluate the performance of the classification are accuracy (ACC), sensitivity (SE) and specificity (SP), which are based on outcomes of the ensemble of classifiers, according to the majority voting. These outcomes represent the number of correct and incorrect classifications for each class, positive (benign) and negative (malignant). These measures are commonly used [12] and they are defined as: SE is the percentage of correctly classified positive samples with respect to all positive samples, SP is the percentage of correctly classified negative samples with respect to all negative samples, and ACC is the percentage of correctly classified positive and negative samples based on all samples.

A cost function C adopted from Barata et al. [2] is used to deal with the trade-off between SE and SP, which is defined as:

$$C = \frac{c_{10}(1-SE)+c_{01}(1-SP)}{c_{10}+c_{01}}, \quad (5)$$

where c_{10} is the cost of an incorrectly classified benign lesion (FN), and c_{01} is the cost of an incorrectly classified malignant lesion (FP). The costs used to evaluate the classification were $c_{10} = 1$ and $c_{01} = 1.5$, since an incorrect classification of a malignant lesion is more critical. The lower the value of cost C, the better the classification is.

4.2. Evaluation of the feature subset and feature selection

In order to define the best feature subsets for the ensemble classification models, several subsets based on specific feature groups discussed in the previous section were evaluated. Table 1 shows the results for each feature subset using the OPF classifier. Three distance functions, i.e., Euclidean, Chebyshev and Manhattan were compared using this classifier, in order to find the distances between the feature vectors. The Euclidean distance was the best distance function for this classifier, according to the experiments using all features, which achieved an ACC = 92.3%. Consequently, this distance function was used for all other experiments in this study.

Table 1: Performance results for the feature subsets compared to different feature groups (best result for each group is in bold).

Group	Feature subset	ACC
1	Shape	89.1%
	Colour	91.0%
	Texture	**91.6%**
2	Fractal texture	89.9%
	Wavelet texture	**90.7%**
	Haralick's texture	88.3%
3	Shape and colour	90.5%
	Shape and texture	91.3%
	Colour and texture	**91.7%**
4	*RGB* colour	90.6%
	HSV colour	**92.0%**
	LAB colour	90.3%
	LUV colour	90.3%
5	*RGB* texture	**91.8%**
	HSV texture	91.1%
	LAB texture	91.2%
	LUV texture	90.8%
6	Shape and *RGB* algorithms	91.6%
	Shape and *HSV* algorithms	**93.0%**
	Shape and *LAB* algorithms	92.7%
	Shape and *LUV* algorithms	91.7%
7	*RGB* algorithms	90.8%
	HSV algorithms	91.2%
	LAB algorithms	**92.5%**
	LUV algorithms	91.4%

The results in Table 1 show that there is diversity between the feature subsets. The three best feature subsets were the shape combined with the *HSV* algorithms, the *LAB* algorithms, and the *HSV* colour. The shape, colour and texture features provided an improvement to the classification when they were combined. The texture features, i.e., the fractal, wavelet and Haralick's features, achieved better results when the features were combined than when they were used individually. The feature extraction algorithms for each colour space provided better results when combined with the shape features.

The diversity for an ensemble classification model is also obtained by using different feature selection algorithms. Such algorithms were used to find the best features for the classification process. The single-feature evaluators use a ranking method, i.e., the correlation coefficient, GRFS, information gain, relief-F and PCA, and a set of retained number of features is empirically defined by $N = \{25,50,75\}$, with the exception of PCA that chooses a sufficient number of eigenvalues to rank the new transformed features. The maximum number of features $F = 5$ was used to include the transformed features, and the proportion of variance $V = 0.95$ was used to retain a sufficient amount of PC features. Accordingly, 31 eigenvalues were selected by the PCA algorithm to represent the vector with the new features. The feature estimation defined the number of nearest neighbours $k = 10$ for the relief-F.

In the case of the feature subset evaluator, i.e., CFS, the greedy stepwise search method, in either forward or backward directions, is applied until the addition or removal of any feature produces a decrease in evaluation. Consequently, 37 features were selected in the forward direction and 50 in the backward direction. The best first search method was also carried out until five consecutive non-improving features, in the directions: forward (37 features), backward (50 features) or bi-direction (37 features) were found. However, experimental results, using the OPF classifier as discussed in the previous section, showed that this method did not improve the classification when applied with the stepwise search method. Therefore only the stepwise search method was used with CFS and compared with the other feature selection algorithms.

Table 2 shows the best classification results using the feature selection algorithms. Although all the feature selection algorithms obtained good results, the OPF classifier using the features selected by the CFS algorithm achieved the best results, as shown in Table 2. These algorithms were applied to generate the feature subsets for the ensemble classification models.

Table 2: Comparing several feature selection algorithms (best result is in bold).

Feature selection	Features	ACC
Correlation coefficient	75	89.6%
GRFS	25	91.1%
Information gain	75	90.8%
Relief-F	75	91.0%
PCA	31	91.0%
CFS	50	**91.6%**

4.3. Evaluation of the ensemble classification models

The performance of the three ensemble classification models based on the input feature manipulation, OPF classifier and majority voting; namely, the SE-OPF, SEFS-OPF, FEFS-OPF algorithms, were evaluated using ten-fold cross-validation. Four ensembles of classifiers were generated for each algorithm, where $E = \{3,5,7,9\}$ describes the ensemble size, i.e., the number of base classifiers. Several subsets based on combination of the specific feature groups were performed for the feature manipulation using the SE-OPF algorithm. The best subsets of the specific feature groups were performed for the feature manipulation based on the SEFS-OPF algorithm using the CFS algorithm for each ensemble. In addition, all extracted features were performed using different feature selection algorithms for the feature manipulation based on the FEFS-OPF algorithm. Table 3 shows the combination of the subsets and feature selection algorithms for each ensemble that achieved the best classification results.

Table 4 shows the best classification results for each ensemble model. The SE-OPF algorithm achieved its best classification results using $E = 9$. Likewise, 9 classifiers for the ensemble yielded the best results for the SEFS-OPF algorithm, whereas the FEFS-OPF algorithm obtained its best results using $E = 3$. Although the SE-OPF algorithm did not have all the best classification measures, it resulted in a more balanced classification between the benign and malignant classes, i.e., with a lower classification cost. The classification results are presented in more details in Figure 3, which shows the variation of the accuracy, sensitivity and specificity measures, according to the ensemble size defined for each ensemble classification model.

Table 3: Combination of the subsets and feature selection algorithms for each ensemble.

Ensemble classification model	Number of classifier	Feature subsets
SE-OPF	3	Shape, colour and texture
	5	Shape, *RGB* algorithms, *HSV* algorithms, *LAB* algorithms and *LUV* algorithms
	7	Shape, colour, texture, shape + *RGB* algorithms, shape + *HSV* algorithms, shape + *LAB* algorithms and shape + *LUV* algorithms
	9	Shape, *RGB* colour, *HSV* colour, *LAB* colour, *LUV* colour, *RGB* texture, *HSV* Texture, *LAB* texture and *LUV* texture
SEFS-OPF	3	Shape + CFS, colour + CFS and texture + CFS
	5	Shape + CFS, *RGB* algorithms + CFS, *HSV* algorithms + CFS, *LAB* algorithms + CFS and *LUV* Algorithms + CFS
	7	Shape + CFS, colour + CFS, texture + CFS, shape + *RGB* algorithms + CFS, shape + *HSV* algorithms + CFS, shape + *LAB* algorithms + CFS and shape + *LUV* algorithms + CFS
	9	Shape + CFS, *RGB* colour + CFS, *HSV* colour + CFS, *LAB* colour + CFS, *LUV* colour + CFS, *RGB* texture + CFS, *HSV* texture + CFS, *LAB* texture and *LUV* texture + CFS
FEFS-OPF	3	All features + PCA, all features + CFS and all features + GRFS
	5	All features + PCA, all features + correlation coefficient, all features + GRFS, all features + information gain and all features + relief-F
	7	All features + PCA, all features + correlation coefficient, all features + GRFS, all features + information gain, all features + relief-F, all features + CFS (best first) and all features + CFS (stepwise)
	9	All features + PCA + OPF (ED), all features + CFS + OPF (ED), all features + GRFS + OPF (ED), all features + PCA + OPF (CD), all features + CFS + OPF (CD), all features + GRFS + OPF (CD), all features + PCA + OPF (MD), all features + CFS + OPF (MD) and all features + GRFS + OPF (MD)

ED: Euclidean distance, CD: Chebyshev distance and MD: Manhattan distance

Table 4: Classification results for the ensemble classification models (best results are in bold).

Ensemble classification model	ACC	SE	SP	C
SE-OPF (feature subsets + OPF)	**94.3%**	**91.8%**	96.7%	**0.053**
SEFS-OPF (feature subsets + CFS + OPF)	93.9%	**91.8%**	96.0%	0.057
FEFS-OPF (all features + FS + OPF)	93.7%	90.4%	**96.9%**	0.057

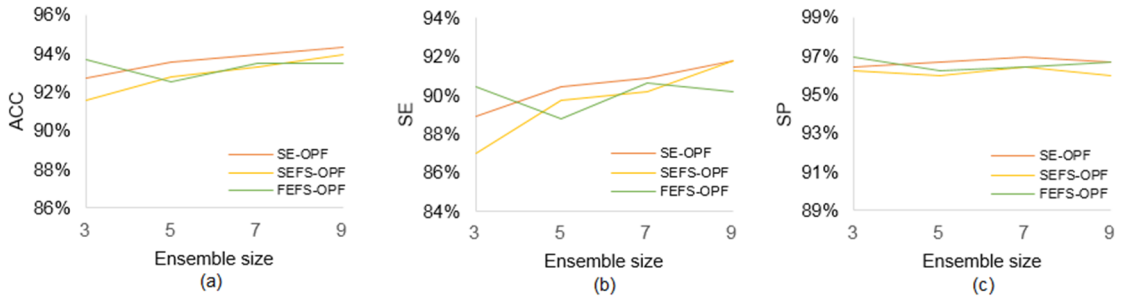

Figure 3: Variation of the classification measures, according to the ensemble size established for each ensemble classification model: (a) accuracy, (b) sensitivity and (c) specificity.

4.4. Comparison between classification algorithms

The classification results achieved by the best ensemble model proposed here, based on the input feature manipulation as previously discussed, were compared against the ones obtained using three different ensemble algorithms. These algorithms are commonly used in the literature; namely, bagging [6], AdaBoost [13] and random forest [14]. The proposed ensemble model was also compared to the individual OPF classifier [18] to analyse the effectiveness of the ensemble algorithms. In addition, this classifier was adopted as a base classifier for the bagging and AdaBoost algorithms, since these algorithms can be used with any learning algorithm. The classification algorithms were applied with and without feature selection based on all the extracted features. The CFS algorithm was used in these experiments, since it improved the classification more than the other feature selection algorithms, as mentioned previously. Table 5 shows the results using different classification methods, as well as the results of the proposed model; the best results for each measure are shown in bold.

Table 5: Comparative results between classification algorithms (best results are in bold).

Classification algorithms	ACC	SE	SP	C
OPF	92.3%	87.5%	**97.1%**	0.067
OPF (CFS)	91.6%	87.0%	96.2%	0.075
Bagging (OPF)	89.7%	85.9%	93.5%	0.095
Bagging (CFS + OPF)	91.8%	88.4%	95.3%	0.075
AdaBoostM1 (OPF)	92.3%	**92.3%**	92.3%	0.077
AdaBoostM1 (CFS + OPF)	91.6%	87.0%	96.2%	0.075
Random forest	93.9%	91.3%	96.6%	0.055
Random forest (CFS)	93.7%	90.4%	96.9%	0.057
Proposed model (feature subsets + OPF)	**94.3%**	91.8%	96.7%	**0.053**

The results in Table 5 show that only the bagging and AdaBoostM1 algorithms achieved better results by using the features selected by the CFS algorithm rather than without the feature selection. Although the AdaBoostM1 algorithm without the feature selection yielded a better accuracy and achieved an average distinction between the benign and malignant classes, the cost was higher because the specificity was not very expressive. On the other hand, the random forest algorithm was more effective without the feature selection, since it obtained a better accuracy and a lower cost between the sensitivity and specificity. In addition, this algorithm obtained better classification results than the bagging and AdaBoostM1 algorithms. Likewise, the individual OPF classifier without the feature selection achieved better results than the bagging and AdaBoostM1

algorithms. Nevertheless, the accuracy obtained by the OPF classifier was not better than the random forest algorithm and the proposed model. Moreover, the classification cost was higher between the sensitivity and specificity.

The proposed model showed good generalization between the benign and malignant classes. Furthermore, this model achieved a better accuracy and lower cost compared to other classification algorithms used in the literature. Since this study did not use all the images from the original dataset as mentioned previously, the results cannot be compared with the results obtained in the studies using this same dataset and the ground-truth lesion segmentation masks presented in Gutman et al. [19]. These studies used the full set of images from the data set which consisted of 1279 images and they divided them into test and training sets. The best results were achieved by Lequan et al. [39] using the whole dataset, obtaining ACC = 0.855, SE = 0.547 and SP = 0.931. These latter authors proposed a novel method for melanoma recognition by leveraging very deep convolutional neural networks.

The proposed ensemble classification model based on input feature manipulation was developed using: 1) Visual Studio Express 2012 environment, C/C++ and OpenCV 2.4.9 library for the feature extraction algorithms; and 2) Eclipse IDE 4.6.1 environment, java 1.8.0_111, and Weka 3.8 library for the classification algorithms. The feature extraction times for all the images from the binary masks were: shape - 10.26 min; colour - 10.12 min; fractal texture - 26.79 min; wavelet texture - 34.37 min; and Haralick's texture - 29.48 min. Finally, the best ensemble classification model required a total of 60.09 s to process all the samples. These values show that the feature extraction step was the most time-consuming; however, the computation time required by this step can be considerably decreased using optimized C/C++ implementations. All algorithms were performed on an Intel(R) Core(TM) i5 CPU 650 @ 3.20 GHz with 8 GB of RAM, running Microsoft Windows 7 Professional 64-bits.

5. Conclusion and future works

In this article, three ensemble classification models based on input feature manipulation from the shape properties, colour variation and texture analysis, were presented; namely, the SE-OPF, SEFS-OPF and FEFS-OPF algorithms. The first model manipulates the features by using different subsets based on specific feature groups. The second model manipulates the features by using the CFS algorithm for the feature

selection from the subsets defined in the first model. Finally, the third model manipulates the features by using different feature selection algorithms, i.e., correlation coefficient, GRFS, information gain, relief-F, PCA and CFS, from all extracted features. Each ensemble model was generated by using the OPF base classifier and integrated with the majority voting strategy. The effectiveness of the feature groups and feature selection algorithms used were individually evaluated to find the best features for the classification process, as well as to generate diversity for the ensemble classification models.

Promising results were achieved with the proposed ensemble classification models. The best classification results were obtained by the feature subset selection model based on feature groups (SE-OPF algorithm). Nine base classifiers were used for this model based on shape, *RGB* colour, *HSV* colour, *LAB* colour, *LUV* colour, *RGB* texture, *HSV* texture, *LAB* texture and *LUV* texture, which yielded the following results: ACC = 94.2%, SE = 91.7% and SP = 96.7%. The feature manipulation process based on these specific feature subsets also provided an excellent generation of diversity for the ensemble classification model.

The lack of a lesion segmentation process can be seen as a limitation of the present study, although ground-truth lesion segmentation masks were used in order to obtain a more accurate computational system. A segmentation approach as presented in Ma and Tavares [40] can be used to evaluate the effectiveness of the proposed ensemble classification model in the segmented images using a level-set approach. Although the ensemble algorithms improve accuracy by combining the different classification models, these algorithms can present a high computational complexity and are rather hard to analyse [5]. Comprehensible models [41], which can be used to solve such problems, aim to produce a single classification model from an ensemble model without losing too much accuracy compared to using the integrated hypothesis model.

Future studies for pigmented skin lesion classification from dermoscopic images should search for new methods to develop more efficient and effective systems. In order to approach other challenges of dermoscopy image diagnoses, the proposed ensemble classification models should be taken into account in future works to identify the presence of global and local patterns. Discriminating between benign and malignant skin lesions is a challenging task for pattern analysis [42]. Essentially, the classification results can be improved by using deep learning architectures [43], since these architectures have revealed their capacity to learn from large amounts of data. Therefore, deep learning

architectures should be taken into account in future works concerning skin lesion classification in dermoscopic images.

Acknowledgments

The first author would like to thank CNPq ("Conselho Nacional de Desenvolvimento Científico e Tecnológico"), in Brazil, for her PhD grant. Authors gratefully acknowledge the funding of Project NORTE-01-0145-FEDER-000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, co-financed by "Programa Operacional Regional do Norte" (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional" (FEDER).

References

[1] American Cancer Society (2015). *Global Cancer Facts & Figures*. 3 edn. American Cancer Society, Atlanta.

[2] Barata, C.; Ruela, M.; Francisco, M.; Mendonça, T.; Marques, J. S. (2013). Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features. *IEEE Systems Journal*, 8 (3):965-979.

[3] Abedini, M.; Codella, N. C. F.; Connell, J. H.; Garnavi, R.; Merler, M.; Pankanti, S.; Smith, J. R.; Syeda-Mahmood, T. (2015). A generalized framework for medical image classification and recognition. *IBM Journal of Research and Development*, 59 (2-3):1-18.

[4] Schaefer, G.; Krawczyk, B.; Celebi, M. E.; Iyatomi, H. (2014). An ensemble classification approach for melanoma diagnosis. *Memetic Computing*, 6 (4):233-240.

[5] Witten, I. H.; Frank, E.; Hall, M. A. (2011). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco.

[6] Dietterich, T. G. (2000). *Ensemble methods in machine learning*. In: Multiple Classifier Systems, vol 1857. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 1-15.

[7] Rahman, M. M.; Bhattacharya, P.; Desai, B. C. (2008). *A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions*. In: 8th IEEE International Conference on International Conference on BioInformatics and BioEngineering, Athens, October 8-10 2008. IEEE, pp 1-6.

[8] Blachnik, M. (2014). Ensembles of instance selection methods based on feature subset. *Procedia Computer Science*, 35:388-396.

[9] Dash, M.; Liu, H. (1997). Feature selection for classification. *Intelligent data analysis*, 1 (3):131-156.

[10] Barata, C.; Emre Celebi, M.; Marques, J. S. (2015). *Melanoma detection algorithm based on feature fusion*. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society Milan, August 25-29 2015. IEEE, pp 2653-2656.

[11] Rastgoo, M.; Morel, O.; Marzani, F.; Garcia, R. (2015). *Ensemble approach for differentiation of malignant melanoma*. In: The International Conference on Quality Control by Artificial Vision 2015, Le Creusot, France 2015. International Society for Optics and Photonics, pp 953415-953419.

[12] Oliveira, R. B.; Papa, J. P.; Pereira, A. S.; Tavares, J. M. R. S. (2016). Computational Methods for Pigmented Skin Lesion Classification in Images: Review and Future Trends. *Neural Computing and Applications*, 27:1-24.

[13] Freund, Y.; Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55 (1):119-139.

[14] Breiman, L. (2001). Random forests. *Machine learning*, 45 (1):5-32.

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

[16] Garnavi, R.; Aldeen, M.; Bailey, J. (2012). Computer-Aided Diagnosis of Melanoma Using Border- and Wavelet-Based Texture Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 16 (6):1239-1252.

[17] Abedini, M.; Chen, Q.; Codella, N. C. F.; Garnavi, R.; Sun, X. (2015). *Accurate and Scalable System for Automatic Detection of Malignant Melanoma*. In: Celebi, M. E., Mendonca, T., Marques, J. S. (eds) Dermoscopy Image Analysis. CRC Press, Boca Raton, pp 293-343.

[18] Papa, J. P.; Falcao, A. X.; Suzuki, C. T. (2009). Supervised pattern classification based on optimum-path forest. *International Journal of Imaging Systems and Technology*, 19 (2):120-131.

[19] Gutman, D.; Codella, N. C. F.; Celebi, E.; Helba, B.; Marchetti, M.; Mishra, N.; Halpern, A. C. (2016). *Skin Lesion Analysis toward Melanoma Detection: A Challenge* at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC), arXiv preprint arXiv:1605.01397.

[20] Iyatomi, H.; Norton, K.; Celebi, M. E.; Schaefer, G.; Tanaka, M.; Ogawa, K. (2010). *Classification of melanocytic skin lesions from non-melanocytic lesions*. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society Buenos Aires, August 31 - September 4 2010. IEEE, pp 5407-5410.

[21] Maglogiannis, I.; Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine*, 13 (5):721-733.

[22] Celebi, M. E.; Iyatomi, H.; Stoecker, W. V.; Moss, R. H.; Rabinovitz, H. S.; Argenziano, G.; Soyer, H. P. (2008). Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, 32 (8):670-677.

[23] Oliveira, R. B.; Marranghello, N.; Pereira, A. S.; Tavares, J. M. R. S. (2016). A computational approach for detecting pigmented skin lesions in macroscopic images. *Expert Systems with Applications*, 61:53-63.

[24] Tkalcic, M.; Tasic, J. F. (2003). *Colour spaces: perceptual, historical and applicational background*. In: Proceedings in the IEEE Region 8 EUROCON 2003: Computer as a Tool Ljubljana, Slovenia, September 22-24 2003. IEEE, pp 304-308.

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

[26] Lissner, I.; Urban, P. (2012). Toward a unified color space for perception-based image processing. *IEEE Transactions on Image Processing*, 21 (3):1153-1168.

[27] Al-Akaidi, M. (2004). *Fractal speech processing*. Cambridge university press, New York.

[28] Scheunders, P.; Livens, S.; Van de Wouwer, G.; Vautrot, P.; Van Dyck, D. (1998). Wavelet-based texture analysis. *International Journal on Computer Science and Information Management*, 1 (2):22-34.

[29] Haralick, R. M.; Shanmugam, K.; Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3 (6):610-621.

[30] Chawla, N. V. (2005). *Data mining for imbalanced datasets: An overview*. In: Maimon, O., Rokach, L. (eds) Data mining and knowledge discovery handbook. Springer, New York, pp 853-867.

[31] Rastgoo, M.; Lemaitre, G.; Massich, J.; Morel, O.; Marzani, F.; Garcia, R.; Meriaudeau, F. (2016). *Tackling the problem of data imbalancing for melanoma classification*. In: 3rd International Conference on Bioimaging, Rome, Italy, February 21-23 2016.

[32] Guyon, I.; Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157-1182.

[33] Kononenko, I. (1994). *Estimating attributes: Analysis and extensions of RELIEF*. In: Bergadano, F., De Raedt, L. (eds) Machine Learning: ECML-94, vol 784. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 171-182.

[34] Hand, D.; Mannila, H.; Smyth, P. (2001). *Principles of Data Mining*. The MIT Press, London.

[35] Hall, M. A. (2000). *Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning*. In: Proceedings of the 17th International Conference on

Machine Learning, San Francisco, USA, June 29 - July 02 2000. Morgan Kaufmann Publishers Inc., pp 359-366.

[36] Ponti Jr, M. P.; Rossi, I. (2013). *Ensembles of Optimum-Path Forest classifiers using input data manipulation and undersampling*. In: Zhou, Z.-H., Roli, F., Kittler, J. (eds) Multiple Classifier Systems, vol 7872. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 236-246.

[37] Amorim, W. P.; Falcão, A. X.; de Carvalho, M. H. (2014). *Semi-supervised pattern classification using optimum-path forest*. In: 27th SIBGRAPI Conference on Graphics, Patterns and Images, Rio de Janeiro, August 26-30 2014. IEEE, pp 111-118.

[38] Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection*. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence, Quebec, Canada, August 20-25 1995. Morgan Kaufmann Publishers Inc., pp 1137-1145.

[39] Lequan, Y.; Chen, H.; Dou, Q.; Qin, J.; Heng, P. A. (2016). Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. *IEEE Transactions on Medical Imaging*:1-11.

[40] Ma, Z.; Tavares, J. M. R. S. (2016). A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model. *IEEE Journal of Biomedical and Health Informatics*, 20 (2):615-623.

[41] Ferri, C.; Hernández-Orallo, J.; Ramírez-Quintana, M. J. (2002). *From ensemble methods to comprehensible models*. In: Lange, S., Satoh, K., Smith, C. H. (eds) Discovery Science, vol 2534. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 165-177.

[42] Argenziano, G.; Soyer, H. P.; Chimenti, S.; Talamini, R.; Corona, R.; Sera, F.; Binder, M.; Cerroni, L.; De Rosa, G.; Ferrara, G.; Hofmann-Wellenhof, R.; Landthaler, M.; Menzies, S. W.; Pehamberger, H.; Piccolo, D.; Rabinovitz, H. S.; Schiffner, R.; Staibano, S.; Stolz, W.; Bartenjev, I.; Blum, A.; Braun, R.; Cabo, H.; Carli, P.; De Giorgi, V.; Fleming, M. G.; Grichnik, J. M.; Grin, C. M.; Halpern, A. C.; Johr, R.; Katz, B.; Kenet, R. O.; Kittler, H.; Kreusch, J.; Malvehy, J.; Mazzocchetti, G.; Oliviero, M.; Özdemir, F.; Peris, K.; Perotti, R.; Perusquia, A.; Pizzichetta, M. A.; Puig, S.; Rao, B.; Rubegni, P.; Saida, T.; Scalvenzi, M.; Seidenari, S.; Stanganelli, I.; Tanaka, M.; Westerhoff, K.; Wolf, I. H.; Braun-Falco, O.; Kerl, H.; Nishikawa, T.; Wolff, K.; Kopf, A. W. (2003). Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet. *Journal of the American Academy of Dermatology*, 48 (5):679-693.

[43] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2 (1):1-127.