Series in BioEngineering

Jacob Scharcanski
M. Emre Celebi *Editors*

Computer Vision Techniques for the Diagnosis of Skin Cancer

Springer

Series in BioEngineering

For further volumes:
http://www.springer.com/series/10358

Jacob Scharcanski · M. Emre Celebi
Editors

Computer Vision Techniques for the Diagnosis of Skin Cancer

Springer

Editors
Jacob Scharcanski
Federal University of Rio Grande do Sul
Porto Alegre
Brazil

M. Emre Celebi
Department of Computer Science
Louisiana State University
Shreveport, LA
USA

ISSN 2196-8861 ISSN 2196-887X (electronic)
ISBN 978-3-642-39607-6 ISBN 978-3-642-39608-3 (eBook)
DOI 10.1007/978-3-642-39608-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013949422

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Malignant melanoma is one of the most rapidly increasing cancers in the world. Invasive melanoma alone has an estimated incidence of 76,690 and an estimated total of 9,480 deaths in the United States in 2013 [1]. Early diagnosis is particularly important since melanoma can be cured with a simple excision if detected early.

In the past, the primary form of diagnosis for melanoma has been unaided clinical examination. In recent years, dermoscopy has proved valuable in visualizing the morphological structures in pigmented lesions. However, it has also been shown that dermoscopy is difficult to learn and subjective. Newer technologies such as infrared imaging, multispectral imaging, and confocal microscopy, have recently come to the forefront in providing greater diagnostic accuracy. These imaging technologies can serve as an adjunct to physicians and provide automated skin cancer screening. Although computerized techniques cannot as yet provide a definitive diagnosis, they can be used to improve biopsy decision making as well as early melanoma detection, especially for patients with multiple atypical nevi.

The goal of this volume is to summarize the state of the art in the utilization of computer vision techniques in the diagnosis of skin cancer and provide future directions for this exciting subfield of medical image analysis. The intended audience includes researchers and practicing clinicians, who are increasingly using digital analytic tools.

The volume opens with six chapters on dermoscopy images. In "Pigment Network Detection and Analysis," Sadeghi et al. describe a novel graph-based method to extract pigment networks from dermoscopy images. Laplacian of Gaussian edge detection output is converted to a graph from which cyclic subgraphs that correspond to skin texture structures are extracted. Subgraphs for round structures such as globules, dots, and bubbles are eliminated based on their size and color. Another higher-level graph is then created from the remaining subgraphs, where each node represents a hole in the pigment network. Finally, the image is classified according to the density ratio of this graph using the LogitBoost classifier. The authors obtain promising results on a set of 436 dermoscopy images.

In "Pattern Analysis in Dermoscopic Images," Sáez et al. propose a model-based approach for the classification of dermoscopy images based on five global patterns (reticular, globular, cobblestone, homogeneous, and parallel) defined in the commonly used Pattern Analysis scheme. The color information in the images

is represented in a three-channel color space such as RGB, YIQ, HSV, and L*a*b*. Each of these channels is then modeled as a Markov Random Field following a Finite Symmetric Conditional Model. Coupling among color channels is taken into account by assuming that the model features follow a multivariate normal distribution. The authors obtain promising results on a set of 100 dermoscopy images.

In "A Bag-of-Features Approach for the Classification of Melanomas in Dermoscopy Images: The Role of Color and Texture Descriptors," Barata et al. investigate the applicability of the bag-of-features (BOF) paradigm based on color and texture features to the classification of dermoscopy images. Based on experiments on a set of 176 dermoscopy images, the authors conclude that the BOF paradigm provides an accurate representation of dermoscopy images and that the color descriptors outperform texture descriptors with respect to classification accuracy.

In "Automatic Diagnosis of Melanoma Based on the 7-Point Checklist," Fabbrocini et al. present a fully automated computer-aided diagnosis system for melanoma based on the 7-point checklist. The system involves border detection using thresholding, extraction of various low- and high-level dermoscopic features, and classification using logistic model trees.

In "Dermoscopy Image Processing for Chinese," Xie et al. elaborate on a computer-aided diagnosis system for melanoma specifically designed for Chinese patients. The system involves hair removal using partial differential equation-based inpainting, border detection using self-generating neural networks combined with genetic algorithms, extraction of various shape, color, and texture features, and classification using a neural network ensemble. The authors obtain promising results on a set of 70 dermoscopy images.

The final dermoscopy chapter, "Automated Detection of Melanoma in Dermoscopic Images" by Arroyo and Zapirain, describes a computer-aided diagnosis system for melanoma based on the ABCD rule of dermoscopy. The system involves removal of artifacts such as black frames, hairs, and bubbles, border detection using thresholding, extraction of various low- and high-level dermoscopic features, and classification using decision trees.

The volume continues with four chapters on clinical (macroscopic) images. In "Melanoma Decision Support Using Lighting-Corrected Intuitive Feature Models," Amelard et al. propose a framework that performs illumination correction and feature extraction on photographs of lesions acquired using standard consumer-grade cameras. They also discuss how these lighting-corrected intuitive feature models can be used to classify skin lesions with high accuracy.

In "Texture Information in Melanocytic Skin Lesion Analysis Based on Standard Camera Images," Cavalcanti and Scharcanski describe five representative sets of visual features commonly used for the representation of texture information in melanocytic lesions, and analyze how these features distinguish between malignant and benign lesions using popular classifiers.

In "Recovering Skin Reflectance and Geometry for Diagnosis of Melanoma," Sun presents a method for estimating and using the skin reflectance recovered as a

replacement for conventional photographs in the evaluation of the ABCD criteria. They also discuss how to use the geometry of a lesion's surface to characterize the topography disruption within the lesion. In order to demonstrate the effectiveness of their approach, the authors compare features derived from reflectance and geometrical information with two-dimensional skin line patterns.

A chapter on classification entitled "Melanoma Diagnosis with Multiple Decision Trees" by Zhou and Song completes the volume. The authors discuss the decision tree model as a mechanism that mimics the clinical diagnostic rules. They also compare the performance of various decision trees via experiments, demonstrating that decision trees can be effective in melanoma diagnosis.

As editors, we hope that this volume focused on analysis of skin lesion images will demonstrate the significant progress that has occurred in this field in recent years. We also hope that the developments reported in this volume will motivate further research in this exciting field.

Jacob Scharcanski
M. Emre Celebi

Reference

1. Siegel, R., Naishadham, D., Jemal, A.: Cancer statistics. CA Cancer J. Clin. **63**(1), 11–30 (2013)

Contents

Pigment Network Detection and Analysis

Maryam Sadeghi, Paul Wighton, Tim K. Lee, David McLean, Harvey Lui and M. Stella Atkins

Abstract We describe the importance of identifying pigment networks in lesions which may be melanomas, and survey methods for identifying pigment networks (PN) in dermoscopic images. We then give details of how machine learning can be used to classify images into three classes: PN Absent, Regular PN and Irregular PN.

Keywords Dermoscopic structures · Pigment network · Melanoma · Computer-aided diagnosis · Machine learning · Graph-based analysis

Introduction

Malignant melanoma, the most deadly form of skin cancer, is one of the most rapidly increasing cancers in the world. Melanoma is now the fifth most common malignancy in the United States [1], with an estimate of 9,180 deaths out of 76,250 incidences in the United States during 2012 [2]. Metastatic melanoma is very difficult to treat, so the best treatment is still early diagnosis and prompt surgical excision of the primary cancer so that it can be completely excised while it is still localized. Unlike many cancers, melanoma is visible on the skin; up to 70 % of all melanomas are first identified by the patients themselves (53 %) or close family members (17 %) [3].

M. Sadeghi · T. K. Lee · D. McLean · H. Lui · M. S. Atkins (✉)
Department of Dermatology and Skin Science, University of British Columbia, Vancouver, Canada
e-mail: Stella@cs.sfu.ca

M. Sadeghi · T. K. Lee · M. S. Atkins
School of Computing Science, Simon Fraser University, Vancouver, Canada

M. Sadeghi · T. K. Lee
Cancer Control Research Program, BC Cancer Research Center, Vancouver, Canada

P. Wighton
Martinos Center for Biomedical Imaging, Harvard Medical School, Boston, USA

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_1,

Therefore, advances in computer-aided diagnostic methods based on digital images, as aids to self-examining approaches, may significantly reduce the mortality.

In almost all the clinical dermoscopy methods, dermatologists look for the presence of specific visual features for making a diagnosis of melanoma. Then, these features are analyzed for irregularities and malignancy [4–7]. The most important diagnostic feature of melanocytic lesions is the pigment network, which consists of pigmented network lines and hypo-pigmented holes [7]. These structures show prominent lines, homogeneous or inhomogeneous meshes. The anatomic basis of the pigment network is either melanin pigment in keratinocytes, or in melanocytes along the dermoepidermal junction. The reticulation (network) represents the rete-ridge pattern of the epidermis. The holes in the network correspond to tips of the dermal papillae and the overlying suprapapillary plates of the epidermis [8, 9].

A pigment network can be classified as either *Typical* or *Atypical*, where the definition of a *Typical* pigment network is: "a light-to-dark-brown network with small, uniformly spaced network holes and thin network lines distributed more or less regularly throughout the lesion and usually thinning out at the periphery" [7]. For an *Atypical* pigment network, the definition is: "a black, brown or gray network with irregular holes and thick lines" [7]. The goal is to automatically classify a given image to one of three classes: *Absent*, *Typical*, or *Atypical*.

Figure 1 illustrates these three classes.

In melanocytic nevi, the pigment network is slightly pigmented. Light brown network lines are thin and fade gradually at the periphery. Holes are regular and narrow. In melanoma, the pigment network usually ends abruptly at the periphery and has irregular holes, thickened and darkened network lines, and treelike branching at the periphery where pigment network features change between bordering regions [10]. Some areas of malignant lesions manifest as a broad and prominent pigment network, while others have a discrete irregular pigment network. The pigment network also may be absent in some areas or the entire lesion.

To simulate an expert's diagnostic approach, an automated analysis of dermoscopy images requires several steps. Delineation of the region of interest, which has been widely addressed in the literature, is always the first essential step in a computerized analysis of skin lesion images [11, 12]. The border characteristics provide essential information for an accurate diagnosis [13]. For instance, asymmetry, border irregularity, and abrupt border cutoff are some of the critical features calculated based on the lesion border [14]. Furthermore, the extraction of other critical clinical indicators and dermoscopy structures such as atypical pigment networks, globules, and blue-white areas depend on the border detection. The next essential step is the detection and analysis of the key diagnostic features of specific dermoscopic structures, such as pigment networks, discussed in this chapter.

We first review published methods for computer-aided detection and analysis of pigment networks, and then provide details of a successful approach for quantifying the irregularity of these pigment networks.

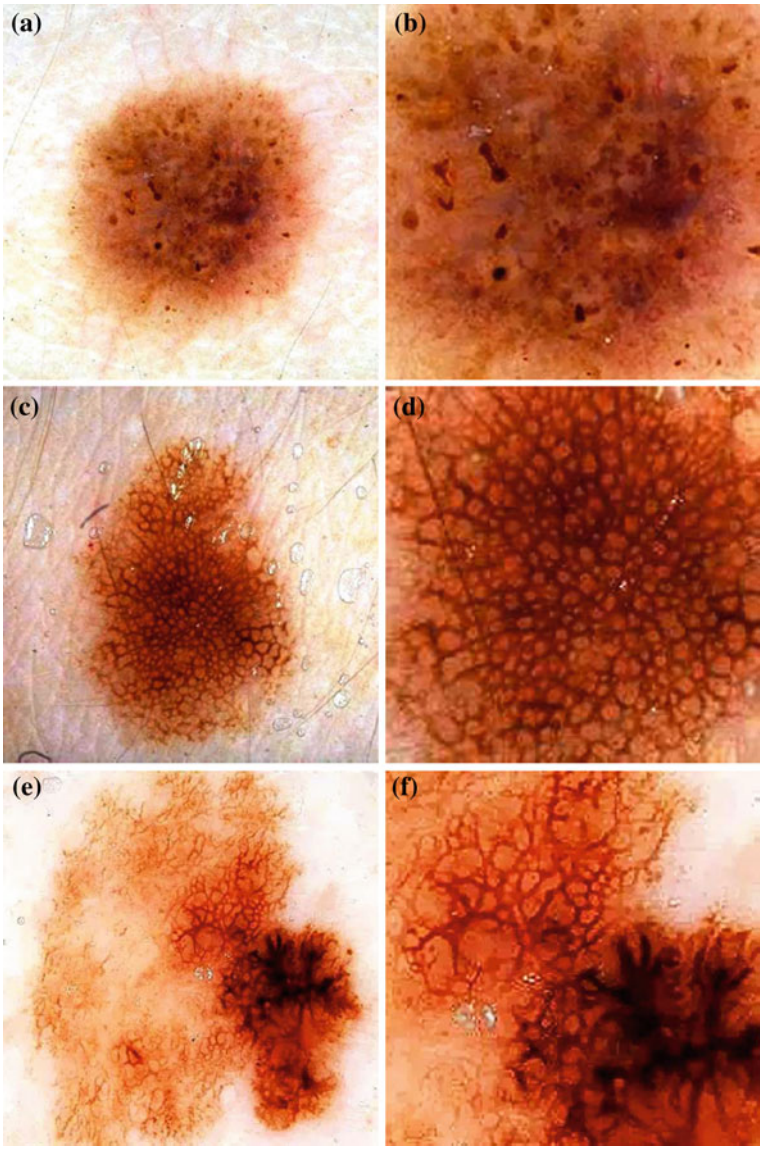


Fig. 1 The three classes of the dermoscopic structure pigment network: **a**–**b** *Absent*; **c**–**d** *Typical*; and **e**–**f** *Atypical*. **b**, **d**, **f** are magnifications of (**a**, **c**, **e**) respectively

Pigment Network Detection

The automated detection of pigment network has received recent attention [15–24], and there is a very recent comprehensive review of computerized analysis of pigmented skin lesions, by Korotkov and Garcia [25]. Fleming et al. [15] report

techniques for extracting and visualizing pigment networks via morphological operators. They investigated the thickness and the variability of thickness of network lines; the size and variability of network holes; and the presence or absence of radial streaming and pseudopods near the network periphery. They use morphological techniques in their method and their results are purely qualitative. Fischer et al. [17] use local histogram equalization and gray level morphological operations to enhance the pigment network. Anantha et al. [18] propose two algorithms for detecting pigment networks in skin lesions: one involving statistics over neighboring gray-level dependence matrices, and one involving filtering with Laws energy masks. Various Laws masks are applied and the responses are squared. Improved results are obtained by a weighted average of two Laws masks whose weights are determined empirically. Classification of these tiles is done with approximately 80 % accuracy.

Betta et al. [19] begin by taking the difference of an image and its response to a median filter. This difference image is thresholded to create a binary mask which undergoes a morphological closing operation to remove any local discontinuities. This mask is then combined with a mask created from a high-pass filter applied in the Fourier domain to exclude any slowly modulating frequencies. Results are reported visually, but appear to achieve a sensitivity of 50 % with a specificity of 100 %.

Di Leo et. al. [20] extend this method and compute features over the 'holes' of the pigment network. A decision tree is learned in order to classify future images and an accuracy of 71.9 % is achieved. Shrestha et. al. [21] begin with a set of 106 images where the location of the atypical pigment network (APN) has been manually segmented. If no APN is present, then the location of the most 'irregular texture' is manually selected. They then compute several texture metrics over these areas (energy, entropy, etc.) and employ various classifiers to label unseen images. They report accuracies of approximately 95 %.

There are three works where supervised learning has been used to detect the dermoscopic structure pigment network [22, 24, 26]. Serrano and Acha [22] use Markov random fields in a supervised setting to classify 100 tiles (sized 40×40) that have been labeled with one of five global patterns: reticular, globular, cobblestone, homogeneous, and parallel. In the context of the study, a reticular pattern can be considered equivalent to pigment network. Using tenfold cross validation, they achieve an impressive overall accuracy of 86 %, considering the difficulty of a five-class problem. It is unclear, however, how the tiles were selected. It could be the tiles were difficult, real-world examples, or that they were text-book-like definitive exemplars. Nowak et al. [24] have developed a novel method for detecting and visualizing pigment networks, based on an adaptive filter inspired by Swarm Intelligence with the advantage that there is no need to preprocess the images.

Wighton's method, published in [26], used machine learning to analyze a dataset of 734 images from [8] and classify the images into *Absent*/*Present*. Labels of either *Absent* or *Present* for the structure pigment network are derived from the atlas. A custom training set is created consisting of 20 images where the pigment network is present across the entire lesion and 20 images absent of pigment network. Pixels are assigned a label from the set $L = \{background, absent, present\}$ as follows: for each

image, pixels outside the segmentation are assigned the label *background*, while pixels inside the segmentation are assigned either the label *absent* or *present*. By considering these three labels, they simultaneously segment the lesion and detect the structure pigment network. A feature-set consisting of Gaussian and Laplacian of Gaussian filter-banks was employed. They present visual results in Fig. 2 by plotting $L = background$, $L = absent$, and $L = present$ in the red, green, and blue channels, respectively.

Although these studies have made significant contributions, there has yet to be a comprehensive analysis of pigment network detection on a large number of dermoscopic images. All other work to date has either: (1) not reported quantitative validation [15, 17, 24, 26]; (2) validated against a small ($N < 100$) number of images [19]; (3) only considered or reported results for the 2-class problem (e.g.

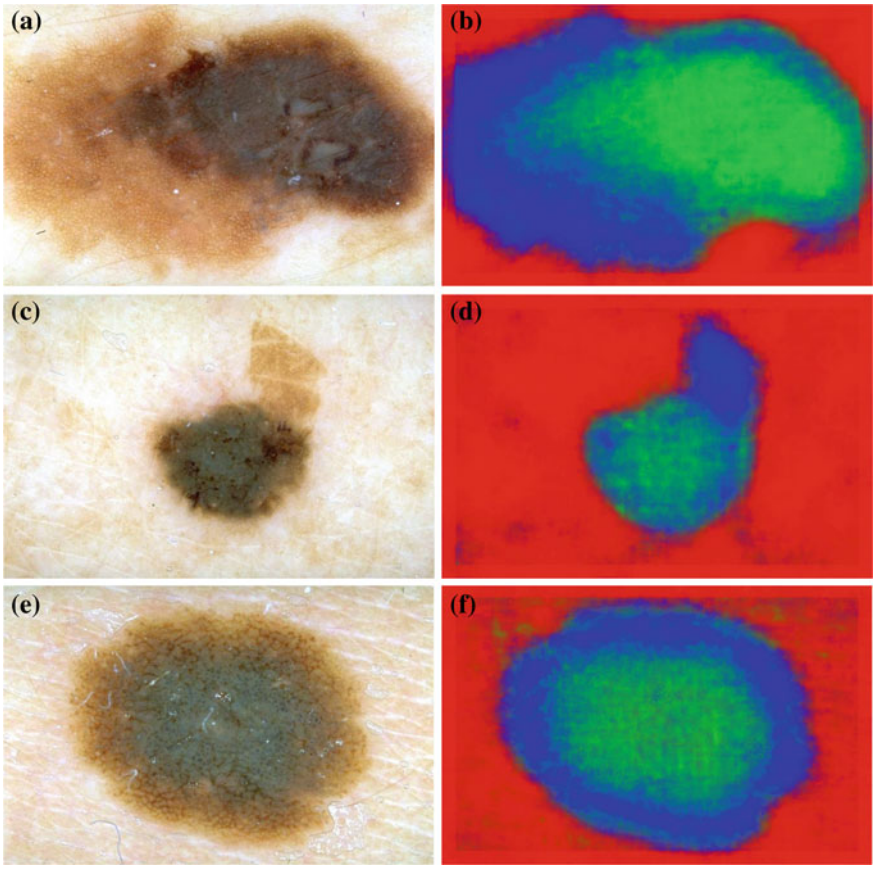


Fig. 2 Qualitative results of pigment network detection from [26]. *First column* original dermoscopic images. *Second column red*, *green*, and *blue* channels encode the likelihood that a pixel is labeled as *background*, *absent*, and *present*, respectively

Absent/*Present* rather than *Absent*/*Typical*/*Atypical*) [18–21, 26]; (4) not explicitly identified the location of the network [18]; or (5) has made use of unrealistic exclusion criteria and other manual interventions [21].

Now we describe our successful approach to analyze the texture in dermoscopy images, to detect regular and irregular pigment networks in the presence of other structures such as dots and globules. The method is based on earlier work on the 2-class problem (*Absent* and *Present*) published in [27, 28] and our work on the 3-class problem (*Absent*, *Typical*, and *Atypical*) published in [29] (see Fig. 1).

Pigment Network Analysis: Overview

We subdivide the structure into the darker mesh of the pigment network (which we refer to as the 'net') and the lighter colored areas the net surrounds (which we refer to as the 'holes'). After identifying these substructures we use the clinical definitions previously mentioned to derive several structural, geometric, chromatic and textural features suitable for classification. The result is a robust, reliable, automated method for identifying and classifying the structure pigment network. Figure 3 illustrates an overview of our approach to irregular pigment network detection. After pre-processing, we find the 'hole mask' indicating the pixels belonging to the holes of the pigment network. Next, a 'net mask' is created, indicating the pixels belonging to the net of the pigment network. We then use these masks to compute a variety of features including structural (which characterizes shape), geometric (which characterizes distribution and uniformity), chromatic and textural features. These features are fed into a classifier to classify unseen images into three classes of *Absent*, *Typical* and *Atypical*. The major modules in Fig. 3 are explained in the following sub-sections.

Pre-processing

In order to prevent unnecessary analysis of the pixels belonging to the skin, the lesion is first segmented. Either manual segmentation or our automatic segmentation method [12] was used. Next the image is sharpened using the MATLAB Image Processing Tool Box function *Unsharp* mask, one of the most popular tools for image sharpening [30]. A two-dimensional high-pass filter is created using Eq. 1. This high-pass filter sharpens the image by removing the low frequency noise. We use the default parameters of MATLAB in our experiments ($\alpha = 3$). Figure 4b shows the result of the sharpening step.

$$SharpeningFilter(\alpha) = \left(\frac{1}{\alpha+1}\right)\begin{vmatrix} -\alpha & \alpha-1 & -\alpha \\ \alpha-1 & \alpha+5 & \alpha-1 \\ -\alpha & \alpha-1 & -\alpha \end{vmatrix}. \tag{1}$$

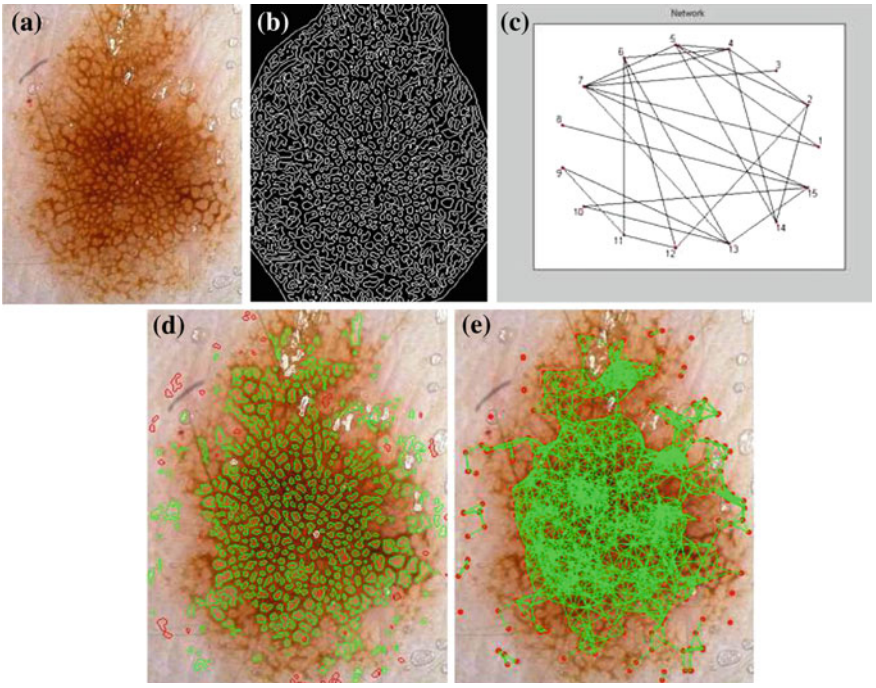


Fig. 3 Steps of the proposed algorithm for hole detection. **a** Original image. **b** LoG response. **c** Image to graph conversion. **d** Cyclic subgraphs. **e** Graph of holes

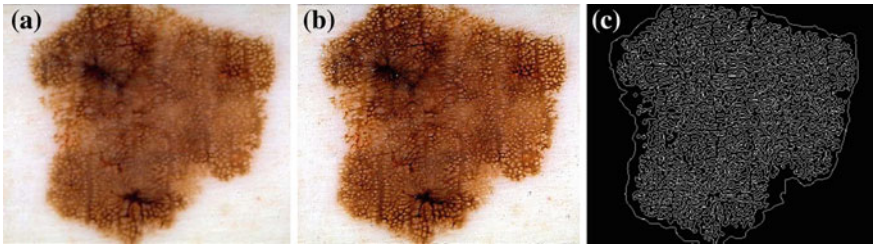


Fig. 4 **a** A given skin lesion image. **b** Sharpened image. **c** Result of the edge detection after segmenting the lesion

To investigate structures of the skin texture, it was necessary to reduce the color images to a single plane before applying our algorithm. Various color transforms (NTSC, L*a*b, Red, Green, and Blue channels separately, Gray(intensity image), etc) were investigated for this purpose. After the training and validation step, we selected the green channel as the luminance image. Results of the different color transformations are reported in the result section of this chapter.

Hole Detection

As discussed previously, a pigment network is composed of holes and nets. We first describe the detection of the holes. Figure 3 shows steps of our novel graph-based approach to hole detection. After the pre-processing step described above, sharp changes of intensity are detected using the Laplacian of Gaussian (LoG) filter. The result of this edge detection step is a binary image which is subsequently converted into a graph to find holes or cyclic structures of the lesion. After finding loops or cyclic subgraphs of the graph, noise or undesired cycles are removed and a graph of the pigment network is created using the extracted cyclic structures. According to the density of the pigment network graph, the given image can be classified into *Present* or *Absent* classes, but for irregularity analysis we also need to extract more features and characteristics of the net of the network.

We used the LoG filter to detect the sharp changes of intensity along the edge of the holes inside the segmented lesion. Because of the inherent properties of the filter, it can detect the "light-dark-light" changes of the intensity well. Therefore it is good choice for blob detection and results in closed contours. The detection criterion of the edge of a hole is set to the zero crossing in the second derivative with the corresponding large peak in the first derivative. We follow the MATLAB implementation of the LoG edge detection which looks for zero crossings and their transposes. All zeros are kept and edges lie on the zero points. If there is no zero, an edge point is arbitrarily chosen as a negative second derivative point. Therefore when all "zero" responses of the filtered image are selected, the output image includes all closed contours of the zero crossing locations inside a segmented lesion. An example of the edge detection step is shown in Figs. 4c and 3b. This black and white image captures the potential holes of the pigment network.

Now, we consider the steps necessary to extract the holes accurately. In previous works [15, 16, 19], these structures usually are found by morphologic techniques and a sequence of closing and opening functions applied to the black and white image. We did not use this approach because using morphologic techniques is error-prone in detecting the round shaped structures. Instead, the binary image is converted to a graph (G) using 8-connected neighbors. Each pixel in the connected component is a node of G and each node has a unique label according to its coordinate.

To find round texture features (i.e. holes), all cyclic subgraphs of G are detected using the Iterative Loop Counting Algorithm (ILCA) [31]. This algorithm transforms the network into a tree and does a depth first search on the tree for loops.

After finding cyclic subgraphs which may represent the holes of a pigment network, these subgraphs were filtered and noise or wrongly detected structures (globules and dots) were removed according to parameters learned in a training and validation step.

Pigment network holes should have higher mean intensity than the border; on the other hand the reverse is true for globules and brown dots. Therefore we thresholded the difference between the average intensity of inner pixels and the average intensity of the border to discriminate globules from holes of the pigment network. First, we

remove all detected cycles which are shorter than 7 pixels and longer than 150 pixels. These parameters can be set for a given data set according to the scale, magnification and resolution of images. The atlas image set [8] used in the experiment does not provide precise information about the resolution and magnification of the image set which is used in our experiment. Furthermore, we are uncertain if the resolution and magnification is the same for all images of the atlas. Therefore, to play safe, we set a wide range (7–150) to find as many holes of pigment networks as possible. In order to belong to a pigment network, a detected cyclic subgraph should have a higher intensity on the area contained by the network structures (holes) than on the network itself (lines), but in globules and brown dots, the mean intensity of the area inside the structure is lower than the intensity of border pixels so we can discriminate them. We also have to deal with oil bubbles and white cysts and dots. These structures are similar to holes of the pigment network in terms of the mean intensity of the inside being lighter than the border area, but they are much brighter inside. So, if there is at least one pixel with high intensity (set to 0.8, on a scale of 0–1) in the inside area of a hole, it will be colored as white representing oil bubbles, white cysts or dots. Therefore, these wrongly detected round structures of brown dots and globules, white dots, white cysts and oil bubbles are removed from the rest of the analysis. We colored these noise structures with red and white in Fig. 5. Thus, the multi-level thresholds, determined by the training step, are set up as:

$$Color = \begin{cases} Green & 0.01 < (I - B) < 0.1 \\ White & 0.1 < (I - B) \ \wedge \ \exists P_I > 0.8 \\ Red & Otherwise \end{cases} \quad (2)$$

where I is mean intensity of the inside area, B is the mean intensity of the border or outside area, and P_I is a pixel inside the hole. Figure 5 shows three examples of skin lesions with filtered cyclic subgraphs overlaid. These structures can be used later for the analysis of globules and dots which are other important structures of the skin lesion texture.

In order to visualize the location of a pigment network based on the detected holes, we created a new higher-level graph G whose nodes are centers of the holes belonging to the pigment network (green colours). Nodes within a maximum distance threshold (MDT) are connected together. However, there is not a minimum node distance threshold. The value of the MDT is computed based on the average diameter of all holes in the image. Based on the pigment network definition, holes of a regular network are uniformly spaced. To consider this spatial arrangement, the MDT should be proportional to the size of holes and is defined as alpha (set to 3) times the average diameter of holes.

Figure 6 illustrates two examples of skin lesions with their graphs of holes overlaid in green. The first column shows a *Present* image and the second one shows an *Absent* image, both of which are classified correctly using the only the graph of the pigment network holes.

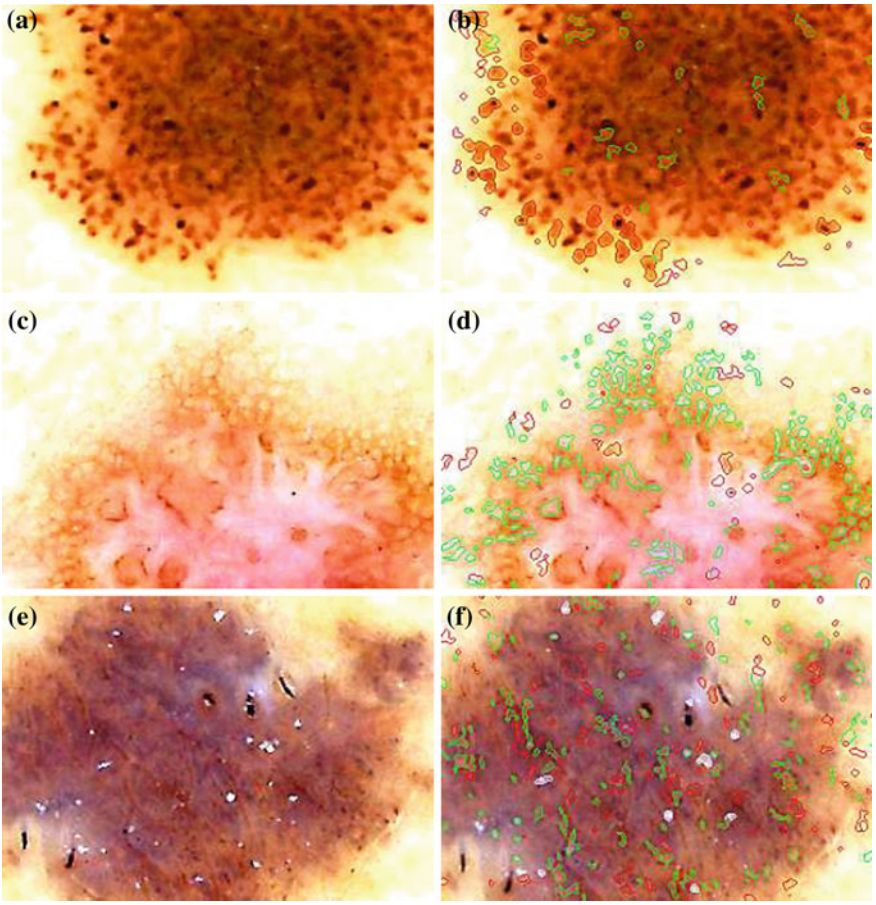


Fig. 5 Detected cyclic subgraphs are filtered based on their inside-outside intensity differences. **a**, **c** and **e** show original skin lesions. **b**, **d**, **f** show *green*, *red* and *white* colours overlaid; the *red* colours mostly belong to globules and *brown dots*. *White dots* and oil bubbles are colored as *white* and holes of the pigment network are visualized as green

Net Detection

In order to identify the net of a pigment network, we apply the Laplacian of Gaussian (LoG) filter to the green channel of the image. The LoG filter identifies high frequency components of an image and therefore makes an ideal net detector. The major issue with applying this operator is that its response is strongly dependent on the relationship between the frequency of the structures and the size of the Gaussian kernel used. We used $\sigma = 0.15$, which is an appropriate value for images of the two atlases used in our experiment [8, 9], however it can be tuned for a given image-set according to scale and magnification. In our experiment, we observed that the

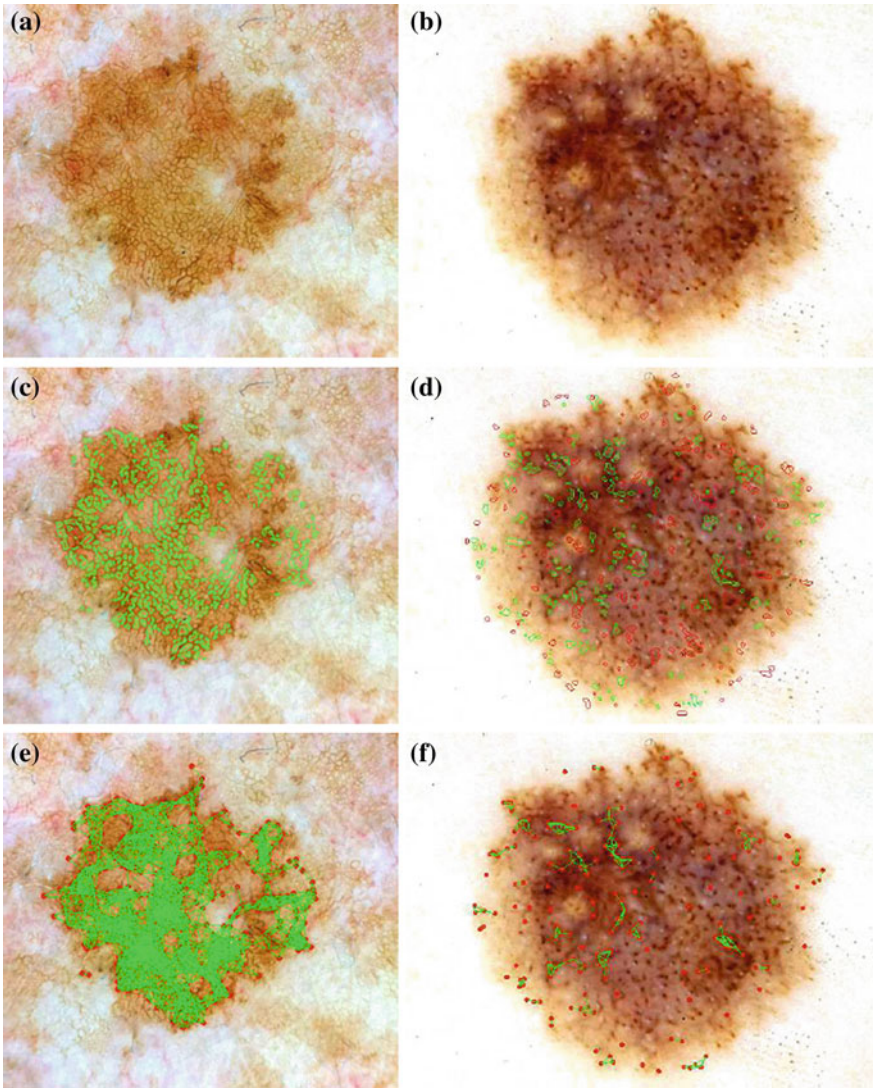


Fig. 6 Results of applying our approach to two *Present* and *Absent* dermoscopic images; (**a**) and (**b**) are skin lesions, (**c**) and (**d**) show cyclic subgraphs, the *green lines* represent potential holes of the pigment network and *red lines* show holes that did not pass the test of belonging to the pigment network, and (**e**) and (**f**) visualize the pigment network over the image. **a** Original image. **b** Original image. **c** Cyclic subgraphs. **d** Cyclic subgraphs. **e** Present. **f** Absent

average thickness of the pigment network is almost proportional to the average size of holes of the network in *Typical* lesions. We therefore set the size of the LoG window to half of the average hole size in the image. The average window size over all images of our data set is 11 pixels. We then threshold the filter response automati-

cally, resulting in a 'net mask' which indicates which pixels belong to the net of the pigment network. Furthermore, we skeletonize this mask, resulting in a 'skeleton mask'. Figure 7 illustrates the net extraction process.

Qualitative results of detecting pigment network 'net' and 'holes' ares illustrated in Figs. 8 and 9 shows the result of pigment network detection on a dermoscopic image taken by an iPhone dermoscope (HandyScope).

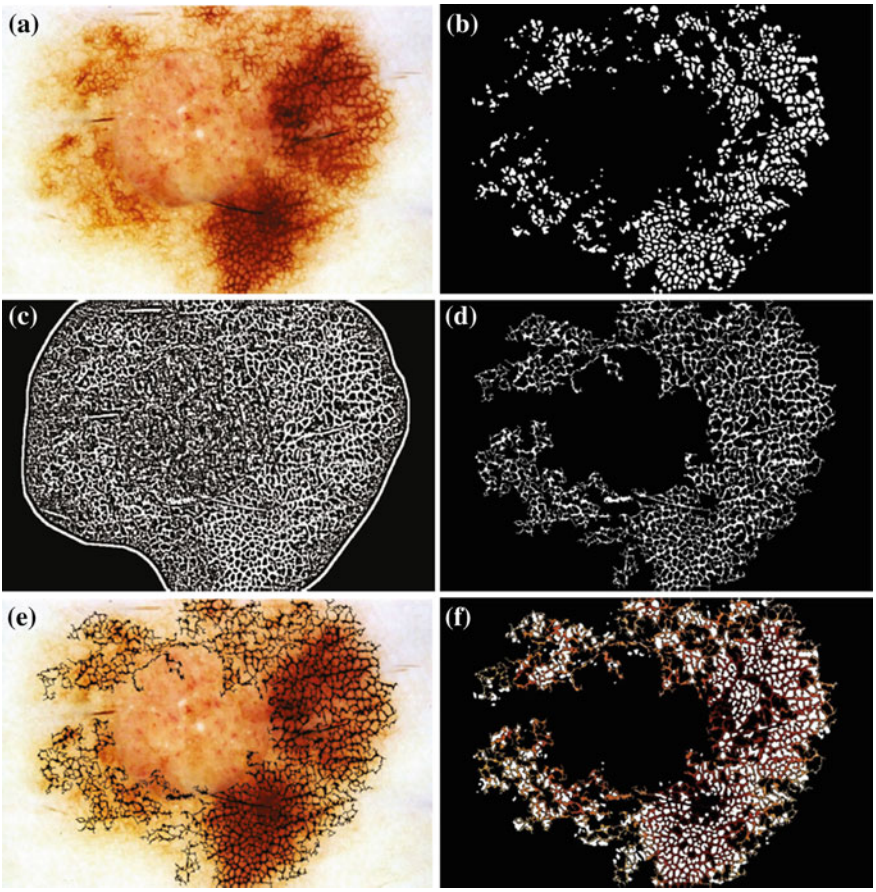


Fig. 7 Net detection. **a** A dermoscopic image, **b** detected holes in the previous step, **c** response of the LoG filter, **d** the resulting 'net mask', **e** the extracted net of the pigment network overlaid on the original image, and **f** the segmented pigment network

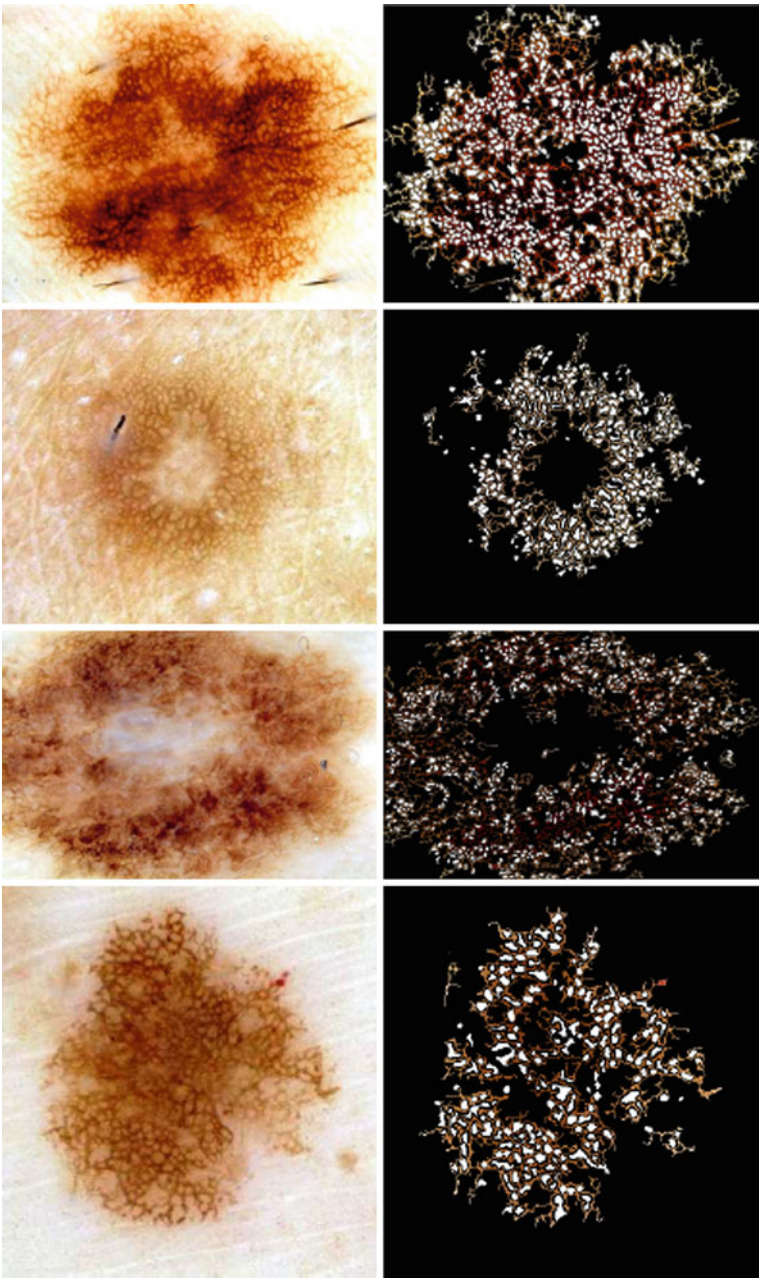


Fig. 8 Four images of the image set: the *left column* shows the original images and the *right column* shows their corresponding pigment networks (*brown*) and holes (*white*)

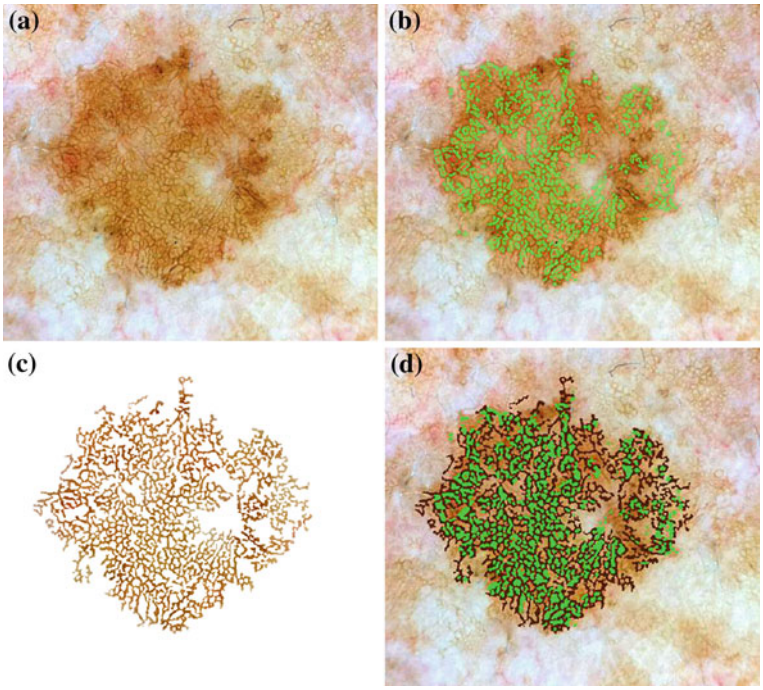


Fig. 9 Pigment network detection on a dermoscopic image taken by an iPhone dermoscope. **a** A dermoscopy image, **b** detected holes in the previous step, **c** the resulting 'net mask', and **d** the segmented pigment network

Feature Extraction

Based on the definitions of *Typical* and *Atypical* pigment networks, we use the results of the hole and net detection to propose a set of features capable of discriminating among the three classes (*Absent*, *Typical* and *Atypical*). We propose a set of structural (shape), geometric (spatial) chromatic and textural features.

Structural Features (20 Features)

Diagnostically important characteristics of a network include the thickness of the nets as well as the size of the holes.

For each spatially disjoint section of the net mask, we compute its size (number of pixels in the net mask) and length (number of pixels in the skeleton mask). Our features are then the mean, standard deviation and coefficient of variation ($mean/std$) of the sizes and lengths of the nets. Thickness is also computed by measuring the distance from each pixel in the net mask to the closest pixel in the skeleton mask. The

mean, standard deviation and ratio of thickness as well as a 6-bin thickness histogram are also included as features. For each spatially disjoint section of the hole mask, we compute the size (number of pixels) and include as features the mean, standard deviation and coefficient of variation ($mean/std$) of hole size as well as the total number of holes.

We also include the ratio of the network size (number of pixels in the net and hole masks) to the lesion size (number of pixels in the segmentation mask).

Geometric Features (2 Features)

We have defined a new feature called 'Density Ratio' of holes which is useful in discriminating between the absence and presence of a pigment network. This feature is defined as

$$Density = \frac{|E|}{|V| * \log(LesionSize)} \tag{3}$$

where $|E|$ is the number of edges in the graph G, $|V|$ is the number of nodes of the graph and *LesionSize* is the size of the area of the image within the lesion boundary, being investigated for finding the pigment network. The rationale of Eq. 3 is that a bigger $|E|$ means that more holes are closer than the MDT. Also, having a smaller $|V|$ for a fixed $|E|$ means that nodes or holes are uniformly spaced close to each other and the graph of the pigment network is dense. Therefore, based on the pigment network definition, having a high 'Density Ratio' is a requirement for being *Present*. *LesionSize* is used to normalize the ratio $|E|/|V|$. For example, a fixed number of vertices and edges in a small lesion is more likely representing *Present* than in a relatively big lesion. However, since there is not a linear relationship between the size of a lesion and the probability of being *Present* or *Absent*, we found experimentally that the logarithm of *LesionSize* is more appropriate.

Clinically, there is an emphasis on the 'uniformity' of the network in order to differentiate between *Typical* and *Atypical*. We expect that lesions with a *Typical* pigment network have a higher number of edges due to uniformly spaced holes. Therefore, in addition to the 'Density Ratio' of holes as a feature, we included another feature, which is the number of edges in the graph G.

Chromatic Features (37 Features)

Color also plays a crucial role in clinical diagnosis. We therefore convert the image to HSV colour space [32] and compute features over each channel as well as the original green channel of the image. In each channel, for the hole, net and lesion masks respectively we compute the mean, standard deviation and coefficient of variation ($mean/std$) of the intensity values. Additionally, we also propose a new chromatic

feature called the 'atypicality measure' which is the sum of the intensity values over the green channel of the pixels in the net mask normalized by the number of the pixels in the net mask.

Textural Features (10 Features)

We use five of the classical statistical texture measures of Haralick et al. [33]: entropy, energy, contrast, correlation and homogeneity which are derived from a grey level co-occurrence matrix (GLCM). The GLCM is a tabulation of how often different combinations of pixel luminance values (gray levels) occur in a specific pixel pairing of an image. We construct 2 GLCMs (in the four directions of 0, 45, 90, 135 and within the distance of 4 pixels using 8 gray levels averaged to obtain a single GLCM) and extract the 5 texture metrics from each. The first GLCM is constructed over the entire lesion (using the pixels in the lesion mask) and the second is constructed over the pigment network (using the pixels in the net and hole masks).

Evaluation for *Absent*/*Present* Classification Using 'Density Ratio'

To measure the strength of our proposed feature, 'Density Ratio', we applied our method to a set of dermoscopic images taken from Argenziano et al. Interactive Atlas of Dermoscopy [8]. We tuned the parameters and thresholds of our proposed method according to a set of 100 images (50 *Absent* and 50 *Present*) of size 768×512. Then we tested the method for another set of 500 images (250 *Absent*, 250 *Present*) randomly selected from the atlas. We classified the unseen images by feeding the 'Density Ratio' into the SimpleLogistic [34] classifier implemented in Weka [35] (a general data mining tool developed by University of Waikato in New Zealand) which uses a powerful boosting algorithm, LogitBoost [36]. Boosting is a method for combining the performance of many weak features to produce a powerful classifier [36]. SimpleLogistic fits logistic models by applying LogitBoost with simple regression functions as base learners. Some of these images were challenging due to acquisition parameters such as lighting and magnification, being partial (entire lesion was not visible), or due to the presence of an unreasonable amount of occlusion by either oil or hair. These challenging images are usually discarded from test sets in the previous work. However, these images were kept in our test set.

Table 1 Correct classification rates (Accuracy) of different colour transformations for N = 500 images

	R	G	B	YIQ	Gray	L*a*b
Correct classification (accuracy)	90.7	**94.3**	90.1	92.6	91.1	89.7

Evaluation for Different Color Spaces

Table 1 shows the percentage of correct classifications (*Present* or *Absent*) for the 500 test images, using different color transformations. It is seen that the green channel gives the best classification. Comparing our results to Anantha et al. method [18] (achieving 80 % accuracy), we achieved a better result, however the same gold standard is not used and the image sets are different. Therefore, a direct comparison is impossible due to different images and ground truths. Note that we deliberately created a difficult dataset by not excluding oily, hairy and low-contrast images. Our method also locates the pigment network and provides a qualitative analysis which can be used for extraction of pigment network characteristics to discriminate typical pigment networks from atypical ones.

Interestingly the Y channel of YIQ (the transformation used for NTSC systems) has the second best result. The Y channel transformation is defined as:

$$Y = 0.299R + 0.587G + 0.114B \tag{4}$$

where R, G, B are the red, green, and blue color components, respectively. To compute the luminance Y, the green channel has larger weight than the other channels so is the likely reason the Y channel works well. In the gray-scale experiment, the intensity image is calculated by $(R + G + B)/3$ and in the L* experiment, the L* component of the L*a*b space is used as the intensity image.

Evaluation for *Absent/Typical/Atypical* Classification

In another experiment, we evaluated the whole feature set (69 features) on the three class problem using the SimpleLogistic classifier.

Since we have not performed any artifact (hair and oil bubble) detection and removal algorithm, in this evaluation we excluded oily and hairy images and we applied the method described above to a set of 436 dermoscopic images taken from two atlases of dermoscopy [8, 9]. Among these images, a clean subset of 400 images, from the set of 600 images used in our *Absent/*Present evaluation from [8], is used. Each image is labeled as *Absent*, *Typical* or *Atypical*, representing the presence and the regularity of the pigment network. The other 36 images are from [9] and have been labeled by 40 experts, each one assigning a label of either *Absent*, *Typical* or *Atypical*

to each image. Overall labels for these images are generated by majority voting. In total, our dataset consists of 436 images (161 *Absent*, 154 *Typical*, 121 *Atypical*). We compute results for both the 3-class (*Absent*, *Typical* or *Atypical*) and 2-class problems (*Absent*, *Present*). Ten-fold cross validation was used to generate all results. Table 2 summarizes these results in terms of Precision (Positive Predictive value), Recall (True Positive Rate or Sensitivity), F-measure, and Accuracy. F-measure is a measure of a test's accuracy that considers both the Precision and the Recall of the test to compute the score where

$$\text{Precision} = \frac{TruePositive}{TruePositive + FalsePositive} \tag{5}$$

$$\text{Recall} = \frac{TruePositive}{TruePositive + FalseNegative} \tag{6}$$

$$\text{F-measure} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

and accuracy is computed as:

$$\text{Accuracy} = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative} \tag{8}$$

Comparing our results with the results generated by the others using different datasets is not possible, and the only work that we could reproduce is the method by Di Leo et al. [20]. For comparison, the feature set described in [20] was also implemented and results over our image sets are computed. As can be seen, this work outperforms

Table 2 Comparing accuracy, precision, recall and f-measure of our proposed features with Di Leo et al. features [20] using the same set of 436 images

Absent-Typical-Atypical classification					
	Precision	Recall	F-measure	Accuracy	N
Absent	0.905	0.950	0.927	–	161
Typical	0.787	0.792	0.790	–	154
Atypical	0.750	0.694	0.721	–	121
Our weighted avg	**0.820**	**0.823**	**0.821**	**0.823**	**436**
Di Leo et al. [20]	0.709	0.711	0.709	0.719	436
Absent-Present classification					
Absent	0.893	0.932	0.912	–	161
Present	0.959	0.935	0.947	–	275
Our weighted avg	**0.935**	**0.933**	**0.934**	**0.933**	**436**
Di Leo et al. [20]	0.875	0.876	0.875	0.876	436

the previous work [20] on the 2-class problem and is the only one to date that reports quantitative results for the 3-class problem.

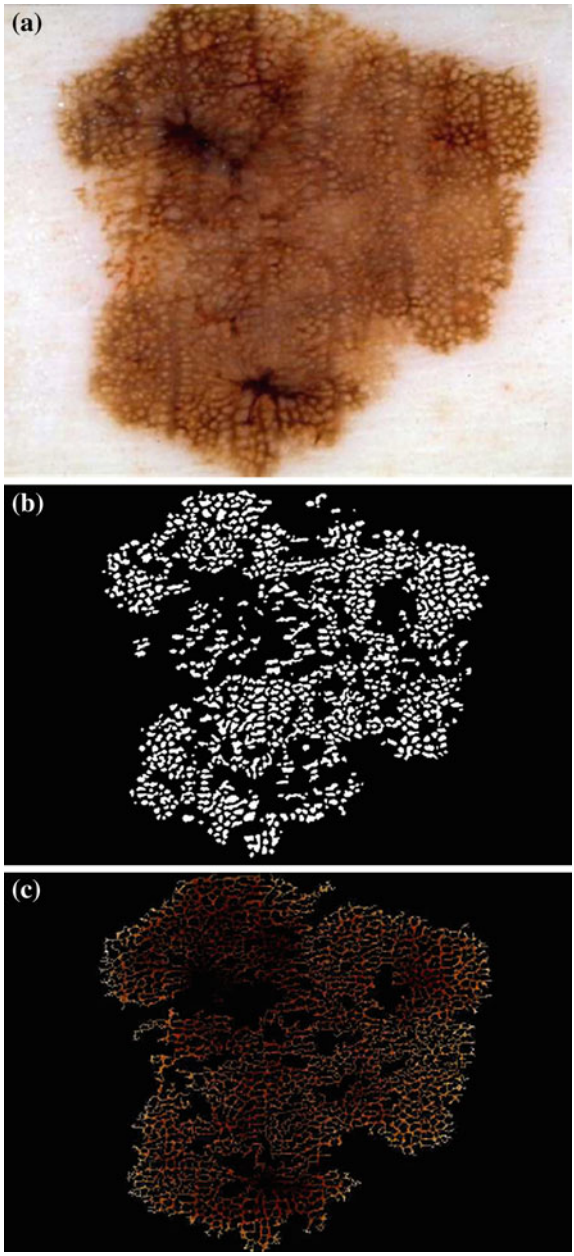


Fig. 10 Pigment network detection on a challenging image with an inter-expert agreement of 82.5 % favoring the *Atypical* diagnosis where 17.5 % of experts classify the pigment network as a *Typical* structure. This image was misclassified as *Typical* by the proposed algorithm

The proposed method is able to correctly classify lesions where the pigment network structure is present with the accuracy of 94.3 %. However, it has more difficulties (with the accuracy of 82.3 %) on dealing with *Typical/Atypical* classification. In some images, it is not easy to detect the atypicality of the pigment network even by experts. Figure 10 shows one of these challenging images where an *Atypical* pigment network is misclassified as a *Typical* pigment network by the proposed algorithm. Please note how regular is the network size and distribution over the lesion, however it is classified as *Atypical* by 82.5 % of the experts.

Summary

We reviewed the motivation and methods for identifying the structure Pigment Network in dermoscopic images, and described in detail a successful, graph-based method for classifying and visualizing pigment networks in real dermoscopic images. In this method, a set of clinically motivated features was identified over these substructures suitable for classification, and verified by evaluating its ability to classify and visualize the structure. The feature set has proven to be robust, outperforming previous work on a large dataset consisting of 436 images, which is the only validation to date on the 3-class problem.

The accuracy of the system is 94.3 % in classifying images to one of two classes of *Absent* and *Present* over a large and inclusive dataset consisting of 500 images. The method was also validated on a different set and achieved an accuracy of 82.3 % discriminating between three classes (*Absent*, *Typical* or *Atypical*).

This method can be used as a part of an automatic diagnosis system for classifying moles and detecting skin cancer. Furthermore, the same idea with different features may also be applied for extracting other diagnostic skin patterns such as globules and streaks.

References

1. Siegel, R., Naishadham, D., Jemal, A.: Cancer statistics. CA Cancer J. Clin. **62**(1), 10–29 (2012)
2. Siegel, R., Naishadham, D., Jemal, A.: Cancer facts and figures, pp. 1–68. American Cancer Society, Atlanta. http://www.cancer.org/research/cancerfactsfigures/acspc-031941 (2012). Accessed 3 Dec 2012
3. Koh, H.K., Miller, D.R., Geller, A.C., Clapp, R.W., Mercer, M.B., Lew, R.A.: Who discovers melanoma?: patterns from a population-based survey. J. Am. Acad. Dermatol. **26**(6), 914–919 (1992)
4. Stolz, W., Riemann, A., Cognetta, A.B., Pillet, L., Abmayr, W., Holzel, D., Bilek, P., Nachbar, F., Landthaler, M., Braun-Falco, O.: ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur. J. Dermatol. **4**(7), 521–527 (1994)
5. Menzies, S.W., Ingvar, C., Crotty, K.A., et al.: Frequency and morphologic characteristics of invasive melanoma lacking specific surface microscopy features. Arch. Dermatol. **132**, 1178–1182 (1996)

6. Kenet, R., Fitzpatrick, T.: Reducing mortality and morbidity of cutaneous melanoma: a six year plan. b) identifying high and low risk pigmented lesions using epiluminescence microscopy. J. Dermatol. **21**(11), 881–884 (1994)
7. Argenziano, G., Soyer, H.P., et al.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. J. Am. Acad. Dermatol. **48**(5), 679–693 (2003)
8. Argenziano, G., Soyer, H.P., De Giorgio, V., Piccolo, D., Carli, P., Delfino, M., Ferrari, A., Hofmann-Wellenhof, V., Massi, D., Mazzocchetti, G., Scalvenzi, M., Wolf, I.H.: Interactive Atlas of Dermoscopy (Book and CD-ROM). Edra medical publishing and new media, Milan (2000)
9. Soyer, H.P., Argenziano, G., Chimenti, S., Menzies, S.W., Pehamberger, H., Rabinovitz, H.S., Stolz, W., Kopf, A.W.: Dermoscopy of Pigmented Skin Lesions: An Atlas Based on the Consensus Net Meeting on Dermoscopy 2000. Edra, Milan (2001)
10. Stanganelli, I.: Dermoscopy. http://emedicine.medscape.com/article/1130783-overview (2010). Accessed 12 May 2011
11. Celebi, M.E., Aslandogan, Y.A., Stoecker, W.V., Iyatomi, H., Oka, H., Chen, X.: Unsupervised border detection in dermoscopy images. Skin Res. Technol. **13**(4), 454–462 (2007)
12. Wighton, P., Sadeghi, M., Lee, T.K., Atkins, M.S.: A fully automatic random walker segmentation for skin lesions in a supervised setting. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 1108–1115 (2009)
13. Celebi, M.E., Iyatomi, H., Schaefer, G., Stoecker, W.V.: Lesion border detection in dermoscopy images. Comput. Med. Imaging Graph. **33**(2), 148–153 (2009)
14. Lee, T.K., McLean, D.I., Atkins, M.S.: Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. Med. Image Anal. **7**(1), 47–64 (2003)
15. Fleming, M.G., Steger, C., et al.: Techniques for a structural analysis of dermatoscopic imagery. Comput. med. imaging graph. **22**(5), 375–389 (1998)
16. Grana, C., Cucchiara, R., Pellacani, G., Seidenari, S.: Line detection and texture characterization of network patterns. In: Proceedings of 18th International Conference on Pattern Recognition, ICPR 2006, vol. 2, pp. 275–278. IEEE, Washington (2006)
17. Fischer, S., Guillod, J., et al.: Analysis of skin lesions with pigmented networks. In: Proceeding of International Conference Image Processing, pp. 323–326 (1996)
18. Anantha, M., Moss, R.H., Stoecker, W.V.: Detection of pigment network in dermatoscopy images using texture analysis. Comput. Med. Imaging Graph. **28**(5), 225–234 (2004)
19. Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., Sommella, P.: Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. In: IEEE International Workshop on Medical Measurement and Applications, pp. 63–67. IEEE Computer Society, Washington (2006)
20. Di Leo, G., Liguori, C., Paolillo, A., Sommella, P.: An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin lesions. In: IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems, pp. 190–194 (2008)
21. 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., et al.: Detection of atypical texture features in early malignant melanoma. Skin Res. Technol. **16**(1), 60–65 (2010)
22. Serrano, C., Acha, B.: Pattern analysis of dermoscopic images based on markov random fields. Pattern Recogn. **42**(6), 1052–1057 (2009)
23. Barata, C., Marques, J.S., Rozeira, J.: A system for the detection of pigment network in dermoscopy images using directional filters. IEEE Trans. Biomed. Eng. **59**(10), 2744–2754 (2012)
24. Nowak, L.A., Ogorzałek, M.J., Pawłowski, M.P.: Pigmented network structure detection using semi-smart adaptive filters. In: IEEE 6th International Conference on Systems Biology (ISB), pp. 310–314 (2012)
25. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. Artif. Intell. Med. **56**(2), 69–90 (2012)
26. Wighton, P., Lee, T.K., Lui, H., McLean, D.I., Atkins, M.S.: Generalizing common tasks in automated skin lesion diagnosis. IEEE Trans. Inf. Technol. Biomed. **15**(4), 622–629 (2011)

27. Sadeghi, M., Razmara, M., Ester, M., Lee, T.K., Atkins, M.S.: Graph-based pigment network detection in skin images. In: Proceeding of the SPIE Medical Imaging Conference, vol. 7623 (2010)
28. Sadeghi, M., Razmara, M., Lee, T.K., Atkins, M.S.: A novel method for detection of pigment network in dermoscopic images using graphs. Comput. Med. Imaging Graph. **35**(2), 137–143 (2011)
29. Sadeghi, M., Razmara, M., Wighton, P., Lee, T.K., Atkins, M.S.: Modeling the dermoscopic structure pigment network using a clinically inspired feature set. In: Medical Imaging and Augmented Reality, vol. 6326, pp. 467–474 (2010)
30. Pratt, W.K.: Digital Image Processing, 2nd edn. Wiley, New York (1991)
31. Kirk, J.: Count loops in a network. http://www.mathworks.com/matlabcentral/fx_files/10722/1/content/html/run_loops_html.html (2007). Accessed 12 May 2009
32. Shih, T.Y.: The reversibility of six geometric color spaces. Photogram. Eng. Remote Sens. **61**(10), 1223–1232 (1995)
33. Haralick, R.M., Dinstein, I., Shanmugam, K.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
34. Landwehr, N., Hall, M., Frank, E.: Logistic model trees. Mach. Learn. **59**(1), 161–205 (2005)
35. Goebel, M.: A survey of data mining and knowledge discovery software tools. ACM Special Interest Group Knowl. Discov. Data Min. Explor. Newsl. **1**(1), 20–33 (1999)
36. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Ann. Stat. **28**(2), 337–407 (2000)

Pattern Analysis in Dermoscopic Images

Aurora Sáez, Begoña Acha and Carmen Serrano

Abstract In this chapter an extensive review of algorithmic methods that automatically detect patterns in dermoscopic images of pigmented lesions is presented. Pattern Analysis seeks to identify specific patterns, which may be global and local. It is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma. In this chapter, a description of global and local patterns identified by pattern analysis is presented as well as a brief explanation of algorithmic methods that carry out the detection and classification of these patterns. Although the 7-Point Checklist method corresponds to a different diagnostic technique than pattern analysis, it can be considered as a simplification of it as it classifies seven features related with local patterns. For this reason, the main techniques to automatically assess the 7-Point Checklist are briefly explained in this chapter.

Keywords Pattern analysis · Dermoscopy · Texture descriptors · Local patterns detection · Global patterns detection · Classification · Machine learning

Introduction

The medical term melanoma refers to a malignant tumor developed from melanocytic cells. Melanoma generally appears de novo, and less frequently as the evolution of acquired benign melanocytic nevi. In the last decades, mainly due to sun exposure,

A. Sáez (✉) · B. Acha · C. Serrano
Dpto. Teoría de la Señal y Comunicaciones, Universidad de Sevilla. Camino de los Descubrimientos s/n, Seville, Spain
e-mail: aurorasaez@us.es

B. Acha
e-mail: bacha@us.es

C. Serrano
e-mail: cserrano@us.es

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_2,

the incidence of melanoma has dramatically increased, particularly in young white population. If diagnosed and treated early, the mean life expectancy of individuals suffering from melanoma can be increased by at least 25 years [11].

A non-invasive technique to assist dermatologists in the diagnosis of melanoma is dermoscopy, which is an epiluminescence light microscopy, that magnifies lesions and enables examination down to the dermo-epidermal junction. There are four main diagnosis methods from dermoscopic images: ABCD rule, pattern analysis, Menzies method and 7-Point Checklist. These methods were evaluated during the 2,000 Consensus Net Meeting on Dermoscopy (CNMD) [5] by experts from all over the world. A 2-step procedure was used to facilitate the diagnosis:

1. To decide whether the lesion is melanocytic or non-melanocytic.
2. To decide whether the melanocytic lesion is benign, suspect, or malignant.

Pattern Analysis, considered as the classic approach for diagnosis in dermoscopic images, was deemed superior to the other algorithms. The favorable results of pattern analysis were not unexpected, because this method probably reflects best the way the human brain is working when categorizing morphological images [5]. Pattern Analysis, set forth by Pehamberger and colleagues in 1987 [36], was updated by the Consensus Net Meeting of 2000 [5]. This methodology defines the significant dermatoscopic patterns of pigmented skin lesions. Currently, it is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma [38].

Pattern Analysis seeks to identify specific patterns, which may be global or local. The global features allow a quick preliminary categorization of a given pigmented skin lesion prior to more detailed assessment, and they are presented as arrangements of textured patterns covering most of the lesion. The local features represent individual or grouped characteristics that appear in the lesion. Some conclusions from this methodology were extracted in the consensus mentioned [5]; the global feature most predictive for the diagnosis of melanoma was the multicomponent pattern, whereas the globular, cobblestone, homogeneous, and starburst patterns were most predictive for the diagnosis of benign melanocytic lesions. Regarding to local features, atypical pigmented network, irregular streaks, and regression structures were the features that showed the highest association with melanoma, followed by irregular dots/globules, irregular blotches, and blue-whitish veil. Vascular structures were not found to be significantly associated with melanoma. On the contrary, typical pigmented network, regular dots/globules, regular streaks, and regular blotches were mostly associated with benign melanocytic lesions.

Although the 7-Point Checklist method corresponds to a different diagnostic technique than pattern analysis, it can be considered as a simplification of it, as it classifies seven features related with local patterns [3]. Such simplified algorithm was designed to prevent non-experts from missing the detection of melanomas, even at the cost of decreased specificity.

Due to the proven benefits of applying digital imaging to dermatology [5, 49], image processing research has directed a strong effort to develop Computer Aided Diagnosis (CAD) tools to assist physicians in their task of analyzing pigmented lesions, especially because a dermatologist is not always the physician that analyzes

them in a first trial. In 2009 Maglogianis and Doukas [33] presented an overview of CAD systems, describing how to extract features through digital image processing methods and techniques for skin lesion classification. The special issue -Advances in skin cancer image analysis [14]—edited by Celebi, Stoecker and Moss in 2011 summarized the progress that has taken place in this field, including works related to multispectral imaging system, enhancement of dermoscopy images, detection of lesion border and feature extraction. And the recent work from Korotkov and Garcia [31] presents an extensive review of computerized analysis of pigmented skin lesions applied to microscopic (dermoscopic) and macroscopic (clinical) images.

In this chapter an exhaustive review of methods devoted to quantify features in pattern analysis is presented. A brief summary of the main techniques focused on the extraction of local patterns [2, 13, 30, 34, 42, 43, 45], including those that quantify the 7-Point Checklist [8, 9, 17, 19, 20] , will be presented in section "Local Pattern Analysis". Section "7-Point Checklist Method" will be devoted to explain the main algorithms that implement the 7-Point Checklist method. When dealing with the detection and/or classification of global patterns, a few methods have been published in the literature [1, 25, 40, 44]. In section "Global Pattern Analysis" a description of these methods is introduced, describing in detail Serrano's work [44], where reticular, globular, cobblestone, homogeneous and parallel patterns are classified.

Local Pattern Analysis

The presence of specific dermoscopic features in different regions of the same lesion contributes to make a diagnosis of melanocytic lesions and are called local patterns. They are dermoscopic structures such as pigment network, dots and globules, streaks, blue-whitish veil, regression structures, hypopigmentation and vascular structures, whose appearance description is presented below [4]. The predominant presence of some of these local patterns can determine some global patterns [4, 5].

- Pigment network. Delicate, regular grid of brownish lines over a diffuse light-brown background.
- Dots/globules. Sharply circumscribed, usually round or oval, variously sized black, brown or grey structures.
- Streaks. Brownish-black linear structures of variable thickness.
- Blue-whitish veil. Grey-blue to whitish-blue, diffuse pigmentation associated with pigment network alterations, dots/globules and/or streaks.
- Pigmentation. Dark-brown to grey-black, diffuse area that precludes recognition of more subtle dermoscopic features such as pigment network or vascular structures.
- Hypopigmentation. Diffuse area of decreased pigmentation within an otherwise ordinary pigmented lesion.
- Regression structures. White areas, blue areas and a combination of both. Virtually indistinguishable from the blue-whitish veil.
- Vascular structures.

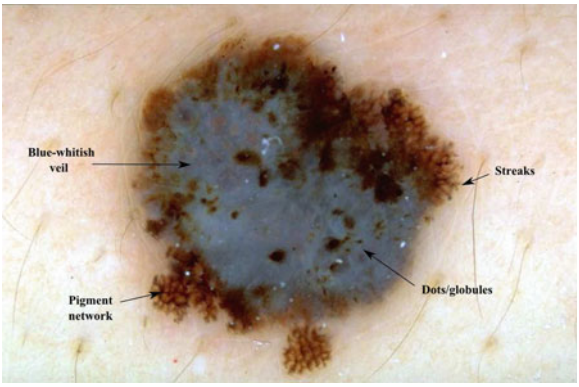


Fig. 1 Example of local patterns

- Other criteria, such as milia-like cysts, comedo-like openings, blotches, Lacunas network, etc.

Local patterns can be presented in the lesion with an irregular/regular or atypical/typical nature, implying malignancy or not. Figure 1 shows some examples of local patterns.

In the literature we can find numerous works that are focused on the automatic identification of local features. They are briefly explained in the subsections below.

Pigment Network

The pigment network is the most studied local pattern. The reason is that it is the most common local pattern in melanocytic lesions, and the identification of melanocytic lesions is the first step in the procedure of pigmented skin lesion diagnosis as explained in the Introduction Section. A pigment network can be typical, when the pattern is regularly meshed, narrowly spaced and its distribution is more or less regular, or atypical, characterized by a black, brown, or grey, irregular network, distributed irregularly throughout the lesion. An atypical network signals malignancy [4]. Figure 2 shows the variability of its appearance.

In the last years several authors have focused on the automatic detection of this pattern.

Anantha et al. [2] in 2004 compared two statistical texture identification methods for detecting the pigment network. The first method was the neighboring grey-level dependence matrix (NGLDM), and the second method used the lattice aperture waveform set (LAWS). They analyzed images of 64×64 pixels. The authors concluded that both methods detect grossly any pigment network with reasonable accuracy, with slightly better results obtained by the latter. 155 dermatoscopic images were analyzed,

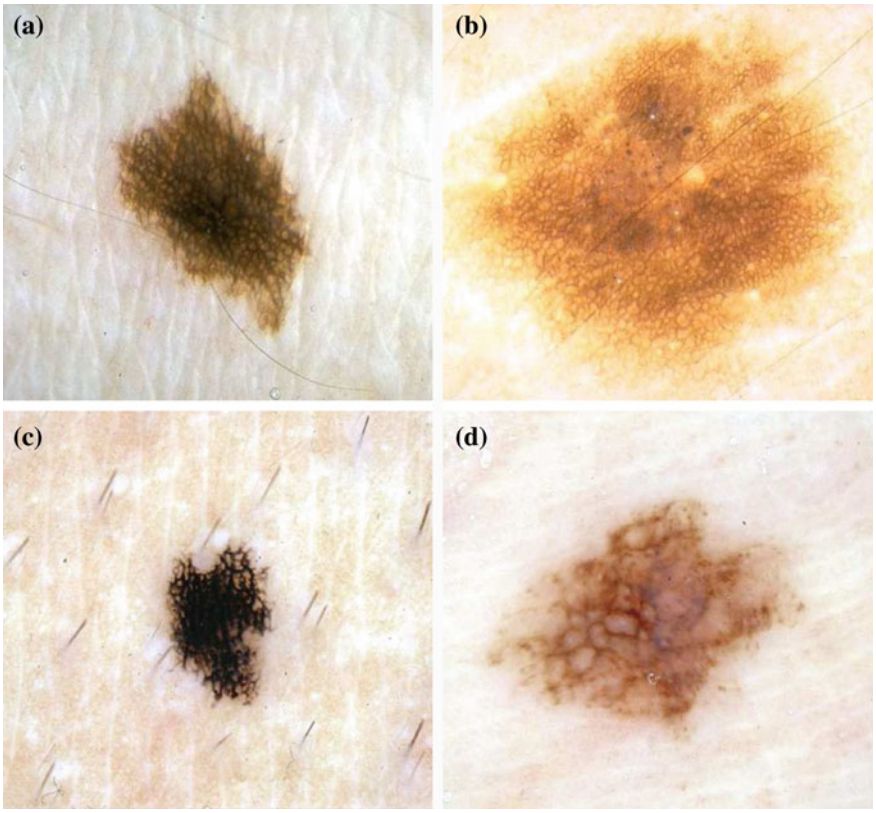


Fig. 2 Example of lesions with pigment network. **a** and **b** present typical pigment network, whereas **c** and **d** atypical

including 62 malignant melanomas and 93 benign lesions. The success classification percentage was around 78 % and 65 % for LAWS and NGLDM, respectively.

Grana et al. [23] in 2006 presented an approach that addressed the problem of detecting the pigment network based on the work of linear structure identification presented in [48], applied to dermatological images by Fleming et al. [21] in 1998. Line points detection was satisfied considering the lines of the pigment network as ridges. As a consequence this set of points must satisfy at the same time two conditions: the first order derivative should be zero, while the second order derivative should have a high module value. After the detection of lines from the zeros of the first derivative, Fleming et al. [21], following the procedure presented in [48], closed lines through an analysis of the second derivative. However, Grana et al. [23] made use of a set of morphological masks that rotate in different directions in order to identify the terminations of the lines and, subsequently, line linking. The thinning approach worked by selectively eroding line points which matched with exactly only one of the morphological masks checking for the presence of an 'L' shaped point

which is not an 8-connection between two different network segments. 'T' shaped connections were eroded with another set of masks. After the network extraction, the image was divided into eight sectors oriented along the principal axes, in order to provide some statistics on the network characteristics of the whole lesion and of every eighth thereof counting the number of meshes, along with the number of unclosed terminations and the average line width. A set of 60 selected lesions was examined. Interestingly, the authors classified each lesion with regard to the distribution of the pigment network as no network pattern, partial network pattern if the lesion is partially covered with pigment network and complete network pattern. An overall 88.3 % network detection performance, without failed detections, was achieved.

Shrestha et al. [45] presented a study in 2010, whose purpose was to identify a method that could discriminate malignant melanoma with an irregular texture, most commonly an atypical pigment network (APN), from benign dysplastic nevi, which generally do not have an APN, using texture measurements alone. In this study, a gray-level co-occurrence matrix (GLCM) is constructed from the luminance plane. Five different GLCMs were constructed for each image using pixel distances (d-values) of 6, 12, 20, 30, and 40. Five classical statistical texture measures were calculated from each GLCM: energy, inertia, correlation, inverse difference, and entropy. Both the average and the range of each of these measures were computed, yielding 10 parameters related to texture. These parameters fed six different classifiers (BayesNet, ADTree, DecisionStump, J48, NBTree, and Ran- dom Forest) in order to determine whether an image presented pigment network or not. The method was tested with 106 dermoscopy images including 28 melanomas and 78 benign dysplastic nevi. The dataset is divided into APN areas and non-APN area. 10-fold validation is employed to validate the method. The correlation average provided the highest discrimination accuracy(95.4 %). The best discrimination of melanomas is attained for a d-distance of 20.

Sadeghi et al. [42] proposed a method to detect and classify the dermoscopic structure pigment network. The method was based on the detection of the 'holes' of the network and follows the next steps: image enhancement, pigment network detection, feature extraction, and classification in three classes. First, a two-dimensional high-pass filter was applied to highlight texture features. Then, the lesion was segmented using Wighton et al. method [54] which employed supervised learning and the random walker algorithm. In the pigment network detection step, a Laplacian of Gaussian (LOG) filter was used to detect sharp changes of intensity. Then, the resulting binary image was converted into a graph using 8-connected neighbouring. Cyclic structures were found in this graph, and noise or undersired cycles were removed. Lines and holes of the pigment network were identified and 69 clinically inspired features were extracted: 20 structural features, including network thickness and its variation within the lesion, as well as size of the holes and its variation along the network; 2 geometric features to study the 'uniformity' of the network; 37 chromatic features; and 10 textural features, using the five classical statistical texture measurements, also proposed in [45]. This allowed to classify the network into typical or atypical type. These 69 features were fed into a classifier based on a powerful boosting algorithm LogitBoost. A dataset consisting of 436 images (161 Absent, 154

Typical network, 121 Atypical network) was used. The authors computed results for both the 3-class (Absent, Typical or Atypical) and 2-class problems (Absent, Present). Ten-fold cross validation was used to generate all results. An accuracy of 82 % discriminating between three classes and an accuracy of 93 % discriminating between two classes were achieved. In [43], the same authors, according to the density of the pigment network graph, classified a given image into Present or Absent. The method was evaluated with 500 images obtaining an accuracy of 94.3 %.

In 2011 Wighton et al. [53] proposed the use of supervised learning and MAP estimation for automated skin lesion diagnosis. The authors applied this method to three task: segmentation, hair detection and identification of pigment network. The method was divided into three main steps. First, in a feature extraction stage, images were converted to CIE $L^*a^*b^*$ [37], and each color channel was filtered with a series of Gaussian and Laplacian of Gaussian filters at various scale ($\sigma = 1.25, 2.5, 5, 10, 20$), so that a total of 30 features were obtained for each pixel. Secondly, after feature extraction, Linear Discriminant Analysis (LDA) was used to reduce the dimensionality. Finally, the posterior probabilities $P(p|l_i)$ ($p = pixel$, $l_i = class$) in this subspace were modelled as multivariate Gaussian distributions. In the training phase, parameters for multivariate Gaussian distributions of each class were estimated. And in the labelling stage, individual pixels from previously unseen images were assigned a label using MAP estimation. A training dataset consisting in 20 images where pigment network was present across the entire lesion and 20 images absent of pigment network was employed. All the images belonged to the dermoscopy atlas presented in [4], where labels of 'present' or 'absent' of pigment network are supplied for each image. Pixels from the training images were assigned a label as 'background', 'absent' or 'present'. To label a new unseen image, features were computed as in the training phase and the dimensionality of the feature space is reduced. To estimate the probability that a pixel p was labelled l_i ($P(l_i|p)$) the authors assigned the most probable label according to MAP estimation.

Skrovseth et al. [47] also proposed a pattern recognition technique with supervised learning to identify pigment network. They selected a training set consisting of a large number of small images containing either a sample of network or a sample of other textures, consisting of both skin and lesion regions. 20 different texture measures were analyzed and the three that contributed maximally to separate the two classes with a linear classifier were selected. A new image is divided into overlapping subimages of the same size as the training images. A pixel is classified as network if at least one of the subimages it belongs to is classified as it.

Barata et al. [7] presented a work focused on the detection of pigment network in 2012. The method was based on the use of directional filters. The first step was to convert the color image into a grey scale one to remove two types of artifacts: hair and reflections caused by the dermatological gel. An inpainting technique was applied. In the second step, regions with pigment network were detected using two of its distinctive properties: intensity and geometry or spatial organization. A bank of directional filters was applied to perform an enhancement of the network. The spatial organization was implemented by connectivity among pixels. The result was a binary net-mask. The final step aimed to assign a binary label to each image: with or

without pigment network. To accomplish this objective, features which characterize the topology of the detected regions in a given image were extracted and used to train a classifier using a boosting algorithm. The algorithm was tested on a dataset of 200 dermoscopic images (88 with pigment network and 112 without) achieving a sensitivity of 91.1 %, a specificity of 82.1 % and an accuracy of 86.2 % in the classification with or without pigment network.

Betta et al. [9] proposed a method for the detection of atypical pigment network. The method was based on their previous work [8], where the pigment network was detected but not classified as atypical/typical. The authors combined two different techniques: structural and spectral analysis. The structural analysis searched for primitive structures such as lines and/or points. To identify these local discontinuities the monocromatic image was compared with a median filtered version of it. In the spectral analysis the Fourier transform of the monocromatic image was performed in order to determine the spatial period of the texture. In this way, local discontinuities, not clearly associated with the network, were disregarded. The result of this phase was a 'regions with network' mask. This mask in conjunction with 'local discontinuities' image provided a 'network image', where the areas belonging to the lesion and constituting the pigment network were highlighted. Two indices related to the spatial and chromatic variability of these areas were presented to quantify the possible atypical nature of the network. 30 images were processed to assess the performance of this detection.

Di Leo et al. [19] extended the work proposed by Betta et al. [9] to detect atypical pigmented network. First, the pigment network was detected following [9]. Then, 13 color and geometric features were extracted. C4.5 algorithm was used as classifier. 173 digital dermoscopy images (77 atypical pigment network, 53 typical pigment network and 43 absent pigment network) obtained from the Interactive Atlas of Dermoscopy [4] were used. 90 images were used for training and 83 images for testing. Sensitivity and specificity values greater than 85 % were reached.

In Table 1 the classification results of the works reported in this section are summarized.

Table 1 Results of **pigment network** detection

Algorithm	Year	Classification	Accuracy (%)	No. images
Anantha et al. [2]	2004	Absent/present	78	155
Grana et al. [23]	2006	No/Partial/complete	88.3	60
Shrestha et al. [45]	2010	Melanoma/no	95.4	106
Sadeghi et al. [42]	2010	Absent/present	93	436
Sadeghi et al. [42]	2010	Absent/atypical/typical	82	436
Skrovseth et al. [47]	2010	Absent/present (per-pixel)	–	–
Wighton et al. [53]	2011	Absent/present (per-pixel)	–	734
Barata et al. [7]	2012	Absent/present	86.2	200
Betta et al. [9]	2006	Atypical/typical	–	30
Di Leo et al. [19]	2008	Atypical/typical	85	173

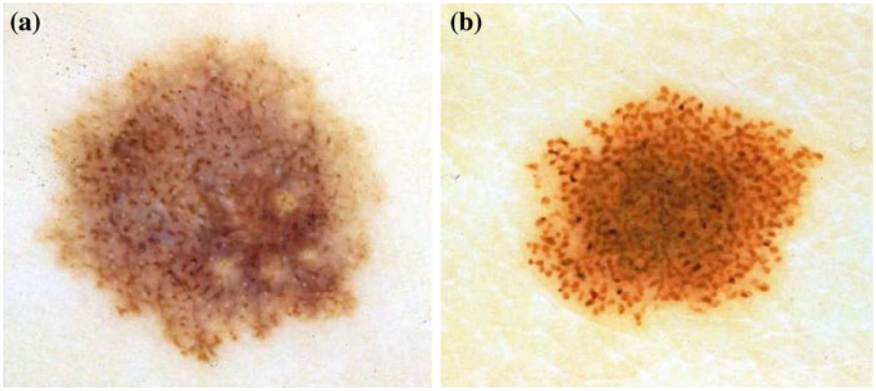


Fig. 3 Example of lesions with *dots/globules*

Table 2 Results of **dots/globules** detection

Algorithm	Year	Classification	Accuracy	No. images
Yoshino et al. [56]	2004	Absent/Present	–	–
Skrovseth et al. [47]	2010	Absent/Present	–	–

Dots and Globules

Dots and globules are round or oval, variously sized black, brown or grey structures, as it has already mentioned in Local Pattern Analysis. It is another dermoscopic structure which is difficult to discriminate from pigment network [42]. This could be the reason why there are so few works in the literature focused on its identification. Some examples of lesions with this structures are shown in Fig. 3.

Based on the classification described in [35], Skrovseth et al. [47] computed a score for each pixel in a gray scale image. Given P surrounding pixels with values g_k, $k = 1, ..., P$ at a radius R of a central pixel with gray value g_c, the score of the central pixel is calculated as $S_c = \Sigma_{k=1}^{P}(g_c - g_k)$. The authors argue that this score will be large for a dark spot, and therefore, a simple thresholding would give the position of the dot.

Yoshino et al. [56] presented an algorithm that use morphological closing operation to detect dots. The closing operation used a linear structural element. Afterwards, a thresholding is applied to detect dots.

Table 2 shows a summary of the works focused on the globules detection.

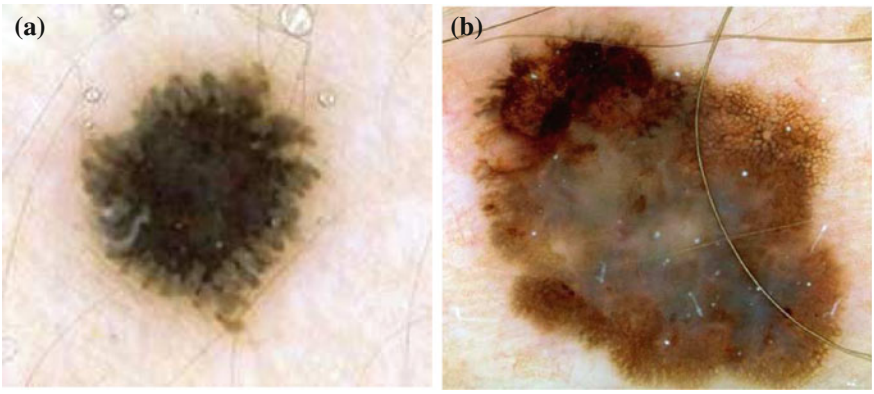


Fig. 4 Example of lesions with (**a**) regular streaks, (**b**) irregular streaks

Streaks

Streaks are brownish-black linear structures of variable thickness that are found in benign and malignant lesions. They are typically placed at the periphery of a lesion and are not necessarily connected to the lines of the pigment network. Streaks can be irregular, when they are unevenly distributed (malignant melanoma), or regular (symmetrical radial arrangement over the entire lesion) [10]. An example of regular and irregular streaks can be found in Fig. 4 .

Mirzaalian et al. [34] proposed a a machine-learning approach to detect streaks which captures the quaternion tubularness in the color dermoscopic images. First, tubularness filters [22] to enhance streak structures in dermoscopic images were used. Given the estimated tubularness and direction of the streaks, a vector field in order to quantify radial streaming pattern of the streaks was defined. Specifically, they computed the amount of flux of the field passing through iso-distance contours of the lesion, where each contour was the loci of the pixels which have equal distance from the outer lesion contour. So, an appearance descriptor based on the mean and variance of the flux through the different concentric bands of the lesion is constructed. The final step is to learn how the extracted descriptors can best distinguish the three different classes: the absence, presence of regular, or presence of irregular streaks in the dermoscopic images. This task is performed with a SVM classifier with a database 99 dermoscopic images.

In [41], the authors followed four steps to locate streaks: preprocessing, blob detection, feature selection and two-class classification (absent-present). In the preprocessing step, lesions were segmented, reoriented so that the major axis was parallel to the x-axis and resized so that its major axis occupied 500 pixels. Lightness component (L*) from L*a*b* color representation was used for the rest of the analysis. Streaks can be modeled as linear structures with a Gaussian cross-section profile near the border. Therefore, in the blob detection, four Laplacian of Gaussian (LOG) filters

with different sizes, hsize = 3, 5, 7, 9, were employed to detect these linear structures. Candidate to streaks were extracted in this step. Once they were detected, their orientations were estimated using the Averaged Squared Gradient Flow (ASGF) algorithm [29]. Then, 25 features were extracted from the candidate linear streak structures and from the lesion: one set of 12 features was based on properties of the detected candidate streak lines and another feature set contained the 13 common color and texture features of the entire lesion. These 25 features were fed to a SimpleLogistic classifier, that classifies a lesion into absence and presence of streaks. The method was tested with a database of 300 dermoscopic images (105 Absent and 195 Present) achieving an accuracy detection of 0.815 using 10-fold cross validation.

Sadeghi et al. [39] recently presented an extension of their previous work [41]. In this version, they proposed an algorithm that classifies a lesion into absence of streaks, regular streaks, and irregular streaks. The work aimed to identify valid streak lines from the set of candidate streak lines obtained in [41] in order to reduce false positive streaks such as hairs and skin lines. The method also extends the analysis to identify the orientation and spatial arrangement of streak lines. These novel geometric features are used to identify not only the presence of streak lines, but whether or not they are Irregular or Regular; important for melanoma diagnosis. Therefore, a total of 31 features are fed into a classifier, achieving an accuracy of 76.1 % when classifying 945 images into the three classes.

Betta et al. [8] identified streaks as finger-like irregularities with uniform brown color at the lesion contour. Therefore, they detected the simultaneous occurrence of two different structures: finger-like track of the lesion contour, and brown pigmentation in the same restricted region. For the first purpose, the color image was converted to an 8-bit gray-level image, and then three different binary images were obtained by applying three different thresholds. The contours of these binary images were extracted by a blob-finding algorithm. The best of the three extracted contours was selected manually and divided into 16 parts. For each part, an irregularity ratio was evaluated. This parameter represented the ratio between number of pixels of the detected contour in this part and the number of pixels in the line connecting the extreme contour points of this part. The contour in that region was assumed as irregular if the ratio was greater than a threshold. On the other hand, brown pigmentation of those 16 subimages is analyzed by thresholding the hue component. Finally, the occurrence of streaks was assumed only if both an irregular contour and a brown pigmentation were found in the same sub-image. The authors presented experimental results for 10 images achieving a 90 % of success rate. A further evaluation was presented by Fabbrocini et al. [20]. They used 23 and 30 images for training and test set, respectively. The two thresholds mentioned above were determined by a Receiver Operating Characteristic curve (ROC curve) on the training image set. A sensitivity and a specificity of 86 and 88 %, respectively, were achieved.

Table 3 summarizes the classification results of the works reported in this section.

Table 3 Results of **streaks** detection following pattern analysis and 7-Point Checklist

Algorithm	Year	Classification	Accuracy (%)	No. images
Mirzaalian et al. [34]	2012	Absent/regular/irregular	91	99
Sadeghi et al. [39]	2013	Absent/present	78.3	945
Sadeghi et al. [39]	2013	Absent/regular/irregular	76.1	945
Betta et al. [8]	2005	Absent/present	90	10
Fabbrocini et al. [20]	2010	Absent/present	86	30

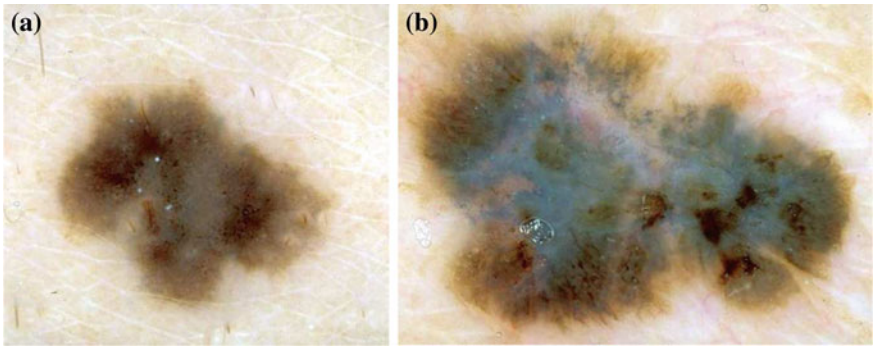


Fig. 5 Example of lesions with *blue–whitish* veil

Blue–Whitish Veil

Blue-whitish veil is characterized by a grey-blue to whitish-blue diffuse pigmentation. Some examples of lesions that present this pattern are shown in Fig. 5.

Celebi et al. [13] proposed a machine learning approach to detect blue-white veil in dermoscopy images based on their early work [12]. Fifteen color features and three texture features were extracted. The color features involved absolute color features and relative color features when compared to the average color of the background skin. The texture features were based on the gray level co-ocurrence matrix (GLCM). The classifier used was C4.5 algorithm. Only 2 out of the 18 features were finally selected for the classification model, both belonging to color features. The classification results for manually selected test pixels yield a sensitivity of 84.33 % and a specificity of 96.19 %. In a second experiment, the authors aimed to discriminate between melanoma and benign lesions based on the area of the blue–white veil detected. They extracted three numeric values from the detected blue-white region: area, circularity and ellipticity. A new classification model based on these features was generated using C4.5 algorithm and 10-fold-cross validation. A sensitivity of 69.35 % and a specificity of 89.97 % for the entire image set (545 images) were obtained.

In a recent work, Arroyo et al. [6] also proposed supervised machine learning techniques to detect blue-white veil. To this aim, color features were extracted from

Table 4 Results of **blue-whitish veil** detection

Algorithm	Year	Classification	Sensitivity (%)	Specificity (%)	No. images
Celebi et al. [12]	2006	Absent/present (per pixel)	84.33	96.19	100
Celebi et al. [13]	2008	Melanoma/Benign	69.35	89.97	545
Arroyo et al. [6]	2011	Melanoma/Benign	80.50	90.93	887
Di Leo et al. [17]	2009	Absent/present	87	85	135

each individual pixel and the classifier used was C4.5 algorithm, that generated the decision tree. Candidate areas selected in the previous step were subsequently classified as melanoma with blue-white veil. For this purpose 12 features were extracted from the candidate area such as area, solidity or ellipticity. The authors used a database consisting of 887 images. 120 images were selected to obtain the training data, 60 corresponding to melanoma with blue-white veil pattern and 60 corresponding to other cases. The method achieved a sensitivity of 80.50 % and a specificity of 90.93 %

Di Leo et al. [17] focused on the detection of two different patterns, blue-whitish veil and regression structures. Firstly, the lesion was subdivided into regions. The color image is converted via principal component analysis (PCA) and a two dimensional (2-D) histogram is computed with the two first principal components. The most significant peaks in the 2-D histogram were found as representative of color regions in the input image. All the pixels in the lesion were assigned to one the main peaks via clustering, so that a lesion map was created. Regions in the lesion map were subsequently classified as present or absent of blue whitish veil and regression. To this aim geometric and color features were extracted and a logistic model tree (LMT) was proposed as classifier. 210 digital dermoscopic images obtained from the Interactive Atlas of Dermoscopy [4] were used. 70 and 50 cases corresponding to the presence of Blue Veil and Regression area respectively were used as training set. 65 cases of Blue Veil and 40 cases of Regression structures were utilized for the test set. A sensitivity of 0.87 and a specificity of 0.85 were obtained for the detection of blue veil and a sensitivity and a specificity both equal to 0.85 for regression structures.

Table 4 shows the classification results of the mentioned works.

Blotches

Blotches are dark structureless areas within pigmented lesions [5]. Blotches that are located asymmetrically within a lesion are indicative of malignant melanoma [30].

The chapter of Stoecker et al. [50] studied the effectiveness of the absolute and relative color blotch features for melanoma/benign lesion discrimination over a dermoscopy image set containing 165 melanomas and 347 benign lesions using a neural network approach. The authors proposed two approaches to detect the blotchy areas. The first method used thresholds placed upon the values of the red, green, and blue (RGB) components of the pixels within the lesions. The second method used

relative color thresholds, subtracting the observed pixel value within the lesion from the background skin color before applying relative thresholds. Then, several blotch indices were computed, including the scaled distance between the largest blotch centroid and the lesion centroid, ratio of total blotch areas to lesion area, ratio of largest blotch area to lesion area, total number of blotches, size of largest blotch, and irregularity of largest blotch. It was determined that relative color were more effective than absolute color giving a diagnostic accuracy of about 77 %.

Khan et al. [30] investigated new and existing blotch features for melanoma discrimination. Four experiments were performed to achieve this aim. Blotches are first extracted using absolute and relative color thresholds to construct blotch masks proposed in [50]. Then, fuzzy logic techniques for extracting blotches based on blotch size were studied, where a fixed minimum blotch size was fuzzified to detect an area as a blotch only if its size exhibits a certain degree of association with a fuzzy set representative of blotch size. To compute the second fuzzy set, the relative color values at each pixel position inside of the blotchy areas were extracted from melanoma lesions belonging to a training set of images. This fuzzy set provided the basis for differentiating between melanoma and benign skin lesions. A third fuzzy set was constructed similarly, but using separate relative color histograms for the red, green and blue color planes. These sets were also used for melanoma discrimination. Finally, a new set of four asymmetry features were computed. The lesion border mask was divided into four quadrants and a set of asymmetry features was computed. The work concluded that features computed from blotches using the fuzzy logic techniques based on three plane relative color and blotch size yielded the highest diagnostic accuracy of 81.2 %. 424 dermoscopy images (134 melanoma and 290 benign images) from three different sources were used.

Madasu and Lovell [32] proposed an extension of fuzzy co-clustering algorithm for images (FCCI) technique [24] for detecting blotches. Madasu and Lovell extended FCCI technique to include texture features as additional clustering parameters. Texture features were computed using the normalized entropy function. A set of 50 images were used for testing the proposed algorithm. The authors claimed that the blotches are accurately located independently of their shape, size or location within the image.

A summary of classification results of the three works presented is shown in Table 5.

Table 5 Results of **blotches** detection

Algorithm	Year	Classification	Accuracy (%)	No. images
Stoecker et al. [50]	2005	Melanoma/Benign	77	512
Khan et al. [30]	2008	Melanoma/Benign	81.2	424
Madasu and Lovell [32]	2009	Absent/Present	–	50

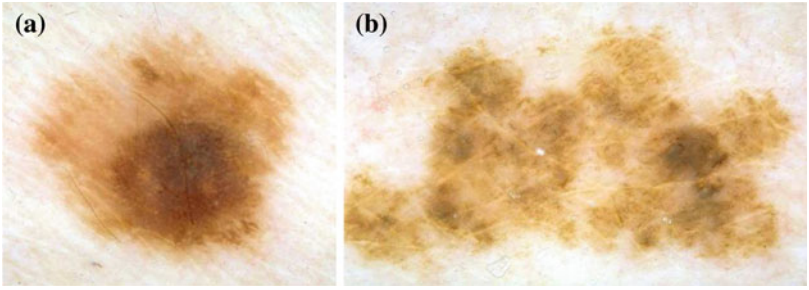


Fig. 6 **a** Nevu with hypopigmentation. **b** Melanoma with hypopigmentation

Table 6 Results of **hypopigmentation** detection

Algorithm	Year	Classification	AUC (%)	No. images
Dalal et al. [15]	2011	Melanoma/Benign	95.2	244

Hypopigmentation

Hypopigmentation represents a diffuse area of decreased pigmentation within an otherwise ordinary pigmented lesion. White areas in a melanoma have an eccentric location and an irregular shape. White areas in a nevus are located in the lesion periphery.

Dalal et al. [15] proposed a method to discriminate melanomas from benign nevi by automatically detecting white areas and measuring features of these white areas. In order to identify white and hypopigmented areas, thresholds were determined for each color plane based on color histogram analysis using a training set of images. The lesion was segmented in concentric deciles. Overlays of the white areas on the lesion deciles were determined. Nine indices were calculated to characterize the automatically detected white areas in a lesion. These indices included lesion decile ratios, normalized number of white areas, absolute and relative size of largest white area, relative size of all white areas, and white area eccentricity, dispersion, and irregularity. A neural network was selected as classifier. 244 benign and malignant dermoscopy images with white areas were selected. The methods used a randomly selected training set of 75 lesions and a test set of 169 lesions.

Regression Structures

A regression structure is a white scarlike depigmentation irregularly distributed within the lesion.

Table 7 Results of **regression structures** detection

Algorithm	Year	Classification	Sensitivity (%)	Specificity (%)	No. images
Di Leo et al. [17]	2009	Absent/present	85	85	90

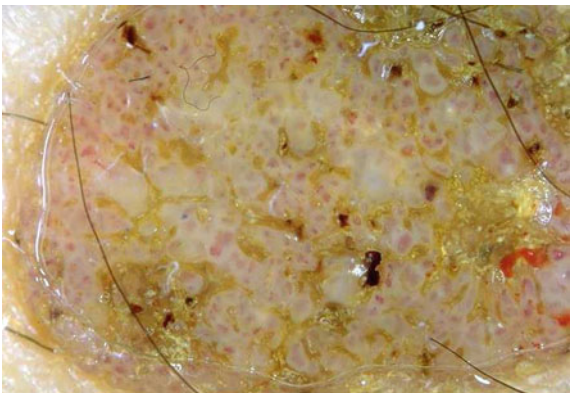


Fig. 7 Example of lesion with *dotted* vessels

As mentioned aboved, Di Leo et al. in their work [17] detected regression structures and blue-whitish veil. The method is reported in section "Local Pattern Analysis". Table 7 shows the classification result of this work.

Vascular Pattern

A vascular pattern, and more specifically, with atypical nature presents linear-irregular or dotted vessels not clearly combined with regression structures and associated with pigment network alterations, dots/globules and/or streaks.

In Betta et al. work [9], reported in section "Pigment Network", a method for the detection of atypical vascular pattern is also proposed. Due to its difficult to obtain a relevant number of ELM images with the occurrence of this criterion, the training set was constituted by N pixels selected as vascular patterns in a set of images containing occurrences of this criterion. The Hue, Saturation and Luminance components was evaluated and the frequency histograms corresponding to the three color planes were determined. The pixel classification depended on the value of its HSL components. However, the authors warned that in some cases the algorithm gave rise to wrong detection, evidencing a low specificity.

General Feature Extraction

Situ [46] formulate the problem of local dermoscopic feature detection as a multi-instance learning (MIL) problem and not as the identification of each local pattern. The first step was to segment each lesion image into five homogeneous regions by the graph cut method. Each of these regions was considered an instance. In a standard multi-instance learning (MIL) problem a lesion is positive if and only if at least one of its instances is positive. The authors employed the method of diverse density (DD) and evidence confidence (EC) function to convert MIL to a single-instance learning (SIL) problem. Both texture and color descriptors were extracted and a SVM classifier was used. The performance of this MIL approach was compared with its boosting version and a baseline method without using MIL, where the descriptors were extracted from the whole lesion. The authors concluded that MIL methods can be helpful in recognizing certain local features that are important for melanoma detection.

7-Point Checklist Method

Although the 7-Point Checklist method is a different diagnosis algorithm, it is considered a simplification of the classic pattern analysis due to the low number of features to identify. This algorithm is applied once the lesion is diagnosed as melanocytic. It was developed by Argenziano et al. [3]. The 7-Point Checklist is a score system. This method uses seven specific criteria for melanoma. It includes three major criteria:

- Atypical pigment network
- Blue-whitish veil
- Atypical vascular pattern

to which 2 points are attributed to each of them, and four minor criteria:

- Irregular streaks
- Irregular pigmentation
- Irregular dots/globules
- Regression structures

to which 1 point is attributed to each of them. A score of 3 or greater is associated with a high likelihood of melanoma diagnosis at pathology evaluation.

Some works focused on the detection of one or several specific criteria can be found in the literature [8, 9, 17, 19, 20]. These works have been explained in section "Local Pattern Analysis".

Di Leo et al. in [18] and [16], joined some methods also mentioned in section "Local Pattern Analysis" in order to present an automatic implementation of the 7-Point Checklist method. In [18] the authors focused on the detection of five criteria. It is in [16] where the seven criteria of the method were addressed: Atypical pigment

network and Irregular dots/globule were detected by the methods proposed in [19] and [9]; Blue-whitish veil, Regression structures and Irregular pigmentation detection followed the steps proposed in [17]; atypical vascular pattern was detected with [9]; and Irregular streaks with the method presented in [8]. 300 images were used for the evaluation. For each dermoscopic criterion a training and testing set were selected from the database in order to train a classifier and/or carry out a statistical analysis. The system distinguished between Melanoma and Benign lesions. The performance of the automatic system was estimated through a comparison with the application of the 7-Point Checklist diagnostic method by expert dermatologists to 287 images of the database. The global sensitivity and specificity values of the software diagnostic tool were 0.83 and 0.76, respectively.

Other authors [11, 26] focused their works on machine learning approaches, where the feature extraction step was inspired on the 7 point checklist criteria.

Global Pattern Analysis

Global features allow a quick preliminary categorization of a given pigmented skin lesion prior to more detailed assessment, and they are presented as arrangements of textured patterns covering most of the lesion.

The main global patterns are [4]:

- Reticular pattern. The most common global feature in melanocytic lesions is characterized by a pigment network covering most parts of a given lesion. Basically, the pigment network appears as a grid of thin brown lines over a diffuse light brown background.
- Globular pattern. It is characterized by the presence of numerous, variously sized, round to oval structures with various shades of brown and grey-black coloration.
- Cobblestone pattern. It is quite similar to the globular one but is composed of closely aggregated, larger somewhat angulated globules resembling a cobblestone.
- Homogeneous pattern. It appears as a diffuse, brown, grey-blue to grey-black or reddish-black pigmentation in the absence of pigmented network or other distinctive local features.
- Starbust pattern. It is characterized by the presence of pigmented streaks in a radial arrangement at the edge of a given pigmented skin lesion.
- Parallel pattern. It is found exclusively in melanocitic lesions on skin of palms and soles due to particular anatomic structures inherent to this location.
- Multicomponent pattern. It is a combination of three or more distinctive dermoscopic structures within a given lesion.
- Lacunar pattern. It is characterized by various to numerous, smooth-bordered, round to oval, variously sized structures called red lacunas, whose morphologic hallmark is their reddish, blue-purplish or black coloration.
- Unspecific pattern. In some instances, a pigmented lesion cannot be categorized into one of the global patterns listed above, because its overall morphologic aspect

does not fit at all into these artificial, albeit rather distinctive categories. For this type of lesion the term unspecific pattern is used.

In [51] and [52] Tanaka et al. presented an analysis of texture to classify a pattern into three categories: homogeneous, globular and reticular. The lesion area was divided into small regions. 110 texture features of each sub-image were calculated. These features were based on intensity histogram information, differential statistical features, Fourier power spectrum, run-length matrix, coocurrence matrix and connected components. 35 features were selected by discriminant analysis. As result, the patterns could be classified correctly into three categories at the ratio of 94% of classification success rate.

In [44] Serrano and Acha proposed a model-based technique to automatically classify five types of global patterns (reticular, globular, cobblestone, homogeneous and parallel). In the model-based methods, image classification was treated as an incomplete data problem, where the value of each pixel was known and the label, which designated the texture pattern the pixel belongs to, is missing. In such techniques, images were modelled as random fields and the segmentation/classification problem was posed as an statistical optimization problem. Most of the existing techniques use the spatial interaction models like Markov Random Field (MRF) or Gibbs Random Field (GRF) to model digital images. Following Xia et al. model [55], Serrano and Acha considered an image as a random field G, defined on a $W \times H$ rectangular lattice, where W and H represented the image dimensions. The lattice was denoted by $S = \{(i, j) : 1 \leq i \leq W, 1 \leq j \leq H\}$, which was indexed by the coordinate (i, j). The color values were represented by $G = \{\mathbf{G_s} = \mathbf{g_s} : \mathbf{s} \in \mathbf{S}\}$, where $s = (i, j)$ denoted a specific site and the random variable $\mathbf{G_s}$ represented a color pixel in the $L^*a^*b^*$ color space. An observed image was an instance of G. It could be described by a finite symmetric conditional model (FSCM) [28] as follows:

$$\mathbf{g}_s = \boldsymbol{\mu}_s + \sum_{\mathbf{t} \in \nu_\mathbf{g}} \boldsymbol{\beta}_{\mathbf{s},\mathbf{t}} [(\mathbf{g_{s+t}} - \boldsymbol{\mu}_{\mathbf{s+t}}) + (\mathbf{g_{s-t}} - \boldsymbol{\mu}_{\mathbf{s-t}})] + \mathbf{e_s} \tag{1}$$

where $\nu_g = \{(0, 1), (1, 0), (1, 1), (-1, 1)\}$ is the set of shift vectors corresponding to the second order neighbourhood system, $\mu_s = [\mu_L, \mu_a, \mu_b]^T$ is the mean of the color pixels in a window centred in site s, $\{\boldsymbol{\beta}_{s,t} : t \in \nu_g\}$ is a diagonal matrix whose elements are the set of correlation coefficients associated with the set of translations from the central site s and $\{\mathbf{e}_s\}$ is a stationary Gaussian noise color sequence with a diagonal covariance matrix Σ as:

$$\Sigma = \begin{bmatrix} \sigma_L^2 & 0 & 0 \\ 0 & \sigma_a^2 & 0 \\ 0 & 0 & \sigma_a^2 \end{bmatrix} \tag{2}$$

Thus, an 18-component feature vector characterized the random field G:

$$\mathbf{f} = (\mu_L, \mu_a, \mu_b, \sigma_L^2, \sigma_a^2, \sigma_b^2, \beta_{L,t}, \beta_{a,t}, \beta_{b,t} : t \in \nu_g) \tag{3}$$

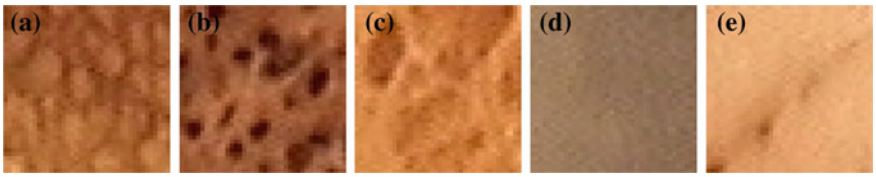


Fig. 8 Example of global patterns. 40×40 dermoscopic image sample representing global patterns: **a** reticular pattern; **b** globular pattern; **c** cobblestone pattern; **d** homogeneous pattern and (**e**) parallel pattern

In order to estimate the parameter vector $\mathbf{f}$, the least-squares estimation method was applied.

Features were supposed to follow a Normal mixture distribution, with a different mean vector and covariance matrix depending on the global pattern each particular image was representing. In other words, if λ represented the pattern class, features will follow the distribution:

$$N(\mathbf{M}_\lambda, \Sigma_\lambda) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_\lambda|}} \exp\left(-\frac{1}{2}(\mathbf{f} - \mathbf{M}_\lambda)^T \Sigma_\lambda^{-1} (\mathbf{f} - \mathbf{M}_\lambda)\right) \quad (4)$$

In order to find the optimum label the pattern belongs to, the *maximum a-posteriori* (MAP) criterion was applied together with the assumption that the five possible global patterns (reticular, cobblestone, homogeneous, parallel and globular) were equally probable, what resulted in the *maximum likelihood* (ML) criterion. Then:

$$\hat{\lambda} = \arg\max_{\lambda \in \Lambda} P(\mathbf{F} = \mathbf{f} | \lambda) \quad (5)$$

This ML problem can then be solved by minimizing the following energy:

$$\hat{\lambda} = \min_{\lambda \in \Lambda} E_G(\mathbf{f}, \lambda) = \min_{\lambda \in \Lambda} \{((\mathbf{f} - \mathbf{M}_\lambda)^T \Sigma_\lambda^{-1} (\mathbf{f} - \mathbf{M}_\lambda) + \ln\left((2\pi)^{18} |\Sigma_\lambda|\right)\} \quad (6)$$

The proposed algorithm was tested on a database containing 100 40×40 image samples of the five types of patterns (see Fig. 8). For each type of them 20 images were used. 10-fold cross-validation was performed: 90 % of the total set of images was used to train (90 images) and 10 % to validate (10 images). The total set was divided into 10 groups for the testing. Each time one different testing group was employed to validate the algorithm. In this way, 18 images of each pattern were used to train and two images of each type were used to validate. The authors compared different color spaces obtaining the success classification percentages described in Table 8.

Since Serrano and Acha developed their methodology to analyze global pattern in pigmented lesions in 2009, different researchers have followed their steps.

Gola et al. [25] presented a method which, in conjuction with the ABCD rule, tried to detect three global patterns in order to increase diagnostic accuracy of pigmented

Table 8 Percentage of success clasification in five global patterns for different color spaces

	Reticular	Globular	Cobblestone	Homogeneous	Parallel	Average
RGB	85	85	50	95	85	80
YIQ	75	90	70	95	70	80
HSV	70	80	65	95	95	81
L*a*b*	90	80	80	90	90	86

lesions. To this aim they developed three different algorithms. In the two first ones, based on edge detection and mathematical morphology, they detected globular and reticular patterns. In the third one, they performed color analysis in the RGB color space with the aim of detecting the blue veil pattern. The algorithms were tested with a database consisting of 20 images per global pattern. The proposed algorithms produced an average accuracy above 85 %.

In [40] the authors detected and classified global dermoscopic patterns as well. They investigated texture analysis and classified five classes of global lesion patterns (reticular, globular, cobblestone, homogeneous, and parallel ones). To this purpose, a statistical approach based on texton classification was followed, where texture features were modelled by the joint probability distribution of filter responses. This distribution was represented by texton (cluster center) frequencies, and textons and texture models were learnt from training images. The classification of an unseen image proceeded by mapping the image to a texton distribution and comparing this distribution to the learnt models. The texton-based classification was performed in the L*a*b* color space and in the gray-scale image using different filter banks. The procedure was divided into two steps:

- Learning stage

1. A set of 81×81 pixel images representing the five patterns were assembled.
2. Training images were convolved with a filter bank to generate filter responses.
3. Exemplar filter responses were chosen as textons and were used to label each filter response in the training images.
4. The histogram of texton frequencies was used to form models corresponding to the training images.

- Classification stage

1. The same procedure as in the training stage was followed to build the histogram corresponding to the unseen image.
2. This histogram was then compared with the models of the texton dictionary.
3. A nearest neighbor classifier was used and the Chi-square statistic was employed to measure distances.

The proposed set of filters was a filter bank composed by $18 + 18 + 2$ filters to detect average intensity, edges, spots, wave, meshes and ripples of dermatoscopic structures. The correct classification rate attained was 86.8 %.

Table 9 Classification results in [1]

Pattern	SE (%)	SP (%)	AUC
Reticular	87.11	97.96	0.981
Globular	86.25	97.21	0.997
Cobblestone	87.76	93.23	0.990
Homogeneous	90.47	95.10	0.996
Parallel	85.25	89.50	0.989
Starburst	89.62	90.14	0.966
Multicomponent	98.50	93.11	0.989

SE Sensitivity, *SP* specificity, *AUC* Area under the Receiver operating characteristic (ROC) curve

Table 10 Classification results in [27]

Classifier	SE (%)	SP (%)	AUC
Melanoma	100	95.9	0.993
Parallel ridge pattern	93.1	97.7	0.985
Parallel furrow pattern	90.4	85.9	0.931
Fibrillar pattern	88	77.9	0.89

SE Sensitivity, *SP* specificity, *AUC* Area under the Receiver operating characteristic (ROC) curve

In their work, Abbas et al. [1] extracted color and texture features from the dermoscopic image in order to classify it into its global pattern. Color related images were extracted from the CIECAM02 representation of the color image. Texture was analysed by means of the steerable pyramids transform (SPT). Both groups of features fed an AdaBoost MC classifier, which classified pigmented lesions into seven different groups of global patterns: (a) Reticular pattern or pigmented network, (b) Globular pattern, (c) Cobblestone pattern, (d) Homogeneous pattern, (e) Parallel pattern, (f) Starburst pattern, (g) Multicomponent pattern. In Table 9 results are summarized.

Iyatomi et al. [27] focused only on the detection of parallel pattern. 428 image features were extracted, which included color-related features, symmetry features, border-related features and texture features. Then, using principal component analysis (PCA), these features were transformed into 198 orthogonal principal components (PCs) without information loss. The first 10 PCs more discriminative were selected. Four linear classifiers were used for parallel ridge, parallel furrow and fibrillar pattern detection. In addition, acral volar melanoma was also classified. The achieved results are shown in Table 10.

Discussion

Pattern analysis is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma [38]. In fact, it was deemed superior to the other algorithms (i.e., ABCD Dermoscopy Rule, Menzies score, 7-Point Checklist) for diagnostic

efficiency by experts from all over the world in the 2000 Consensus Net Meeting on Dermoscopy (CNMD) [5].

Pattern analysis aims to detect local or global patterns in a pigmented lesion to determine if it is melanocytic and, in such a case, its malignancy [5]:

1. Detection of pigment network, aggregate globules, streaks, homogeneous blue or parallel pattern are signs of melanocytic lesions.
2. Atypical pigment network, dots or streaks irregularly distributed, blue-white veil or regression may be signs of melanoma.

The presence of specific dermoscopic features in different regions of the same lesion contributes to make a diagnosis of melanocytic lesions and these features are called local patterns. The predominant presence of some of these local patterns determines some global patterns.

In the literature, much more works devoted to detect local patterns than global patterns are published. Most of the papers related to local patterns address the problem of the detection of the pigmented network, which is the most common local pattern appearing in melanocytic lesions.

Regarding global patterns, Serrano's article [44] was the first work that automatically detected global patterns in dermoscopic images. Nevertheless, multicomponent pattern was not considered in this paper. Later, other works have provided new methods to classify global patterns, included multicomponent, with high success classification rate.

References

1. Abbas, Q., Celebi, M.E., Serrano, C., Garcia, I.F., Ma, G.: Pattern classification of dermoscopy images: a perceptually uniform model. Pattern Recognit. **46**, 86–97 (2013)
2. Anantha, M., Moss, R.H., Stoecker, W.V.: Detection of pigment network in dermatoscopy images using texture analysis. Comput. Med. Imaging Graph. **28**(5), 225–234 (2004)
3. Argenziano, G., Fabbrocini, G., Carli, P., De Giorgi, V., Sammarco, E., Delfino, M.: 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. Arch. Dermatol. **134**, 1563–1570 (1998)
4. Argenziano, G., Soyer, H., De Giorgio, V., Piccolo, D., Carli, P., Delfino, M., Ferrari, A., Hofmann-Wellenhof, R., Massi, D., Mazzocchetti, G., Scalvenzi, M., Wolf, I.H.: Interactive atlas of dermoscopy. EDRA-Medical Publishing and New Media, Milan (2000)
5. Argenziano, G., Soyer, H.P., Chimenti, S., Talamini, R., Corona, R., Sera, F., Binder, M., Kopf, A.W.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. J. Am. Acad. Dermatol. **48**(5), 680–693 (2003)
6. Arroyo, J. L. G., Zapirain, B. G., Zorrilla, A. M.: 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, ISSPIT 2011, art. no. 6151559, 196–201 (2011)
7. Barata, C., Marques, J. S., Rozeira, J.: A system for the detection of pigment network in dermoscopy images using directional filters. IEEE Trans. Biomed. Eng. **59**(10), art. no. 100, 2744–2754 (2012)

8. Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., Scalvenzi, M.: Automated application of the “7-point checklist" diagnosis method for skin lesions: estimation of chromatic and shape parameters. Conference Record —IEEE Instrumentation and Measurement Technology Conference 3, art. no. 1604486, 1818–1822 (2005)
9. Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., Sommella, P.: Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. IEEE International Workshop on Medical Measurement and Applications, MeMeA, art. 1644462, 63–67 (2006)
10. Braun, R.P., Rabinovitz, H.S., Oliviero, M., Kopf, A.W., Saurat, J.-H.: Dermoscopy of pigmented skin lesions. J. Am. Acad. Dermatol. **52**(1), 109–121 (2005)
11. Capdehourat, G., Corez, A., Bazzano, A., Alonso, R., Musé, P.: Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions. Pattern Recognit.Lett. **32**, 2187–2196 (2011)
12. Celebi, M. E., Kingravi, H. A., Aslandogan, Y. A., Stoecker, W. V.: Detection of blue–white veil areas in dermoscopy images using machine learning techniques. Progress in Biomedical Optics and Imaging—roceedings of SPIE 6144 III, art. no. 61445T (2006)
13. Celebi, M.E., Iyatomi, H., Stoecker, W.V., Moss, R.H., Rabinovitz, H.S., Argenziano, G., Soyer, H.P.: Automatic detection of blue–white veil and related structures in dermoscopy images. Comput. Med. Imaging Graph. **32**(8), 670–677 (2008)
14. Celebi, M.E., Stoecker, W.V., Moss, R.H.: Advances in skin cancer image analysis. Comput. Med. Imaging Graph. **35**(2), 83–166 (2011)
15. 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.: Concentric decile segmentation of white and hypopigmented areas in dermoscopy images of skin lesions allows discrimination of malignant melanoma. Comput. Med. Imaging Graph. **35**(2), 148–154 (2011)
16. Di Leo G., Paolillo A., Sommella, P., Fabbrocini, G., Rescigno, O.: A software tool for the diagnosis of melanomas. In: 2010 IEEE Instrumentation and Measurement Technology Conference, pp. 886–891 (2010)
17. Di Leo, G., Fabbrocini, G., Paolillo, A., Rescigno, O., Sommella, P.: Toward an automatic diagnosis system for skin lesions: estimation of blue–whitish veil and regresin structures. In: 6th International Multi-Conference on Systems, Signals and Devices (2009)
18. Di Leo, G., Fabbrocini, G., Paolillo, A., Sommella, P.: Automatic Diagnosis of Melanoma: a Software System based on the 7-Point Check- List. In: Proceedings of the 43rd Annual Hawaii International Conference on System Sciences, 5–8 Jan, Computer Society Press (2010)
19. Di Leo, G., Paolillo, A., Sommella, P., Liguori, C.: An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin lesions. In: Proceedings of IEEE Conference Virtual Environments Human-Computer Interfaces and Measurement Systems, pp. 190–194 (2008)
20. Fabbrocini, G., Betta, G., Di Leo, G., Liguori, C., Paolillo, A., Pietrosanto, A., et al.: Epiluminescence image processing for melanocytic skin lesion diagnosis based on 7-point check-list: a preliminary discussion on three parameters. Open Dermatol. J. **2010**(4), 110–115 (2010)
21. Fleming, M.G., Steger, C., Zhang, J., Gao, J., Cognetta, A.B., Pollak, L., Dyer, C.R.: Techniques for a structural analysis of dermatoscopic imagery. Comput. Med. Imaging Graph. **22**(5), 375–389 (1998)
22. Frangi, A., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI, pp. 130–137. Springer, Heidelberg (1998)
23. Grana, C., Cucchiara, R., Pellacani, G., Seidenari, S.: Line detection and texture characterization of network patterns. In: Proceedings—International Conference on Pattern Recognition, vol. 2(1699200), pp. 275–278 (2006)
24. Hanmandlu, M., Seba, S., Madasu, V.K., Lovell, B. C.: Fuzzy co-clustering of medical images using Bacterial Foraging. In: 23rd International Conference Image and Vision Computing New Zealand, pp. 1–6. IEEE press, New York (2008)

25. Isasi Gola, A., Garcia Zapirain, B., Mendez Zorrilla, A.: Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. Comput. Biol. Med. **41**, 742–755 (2011)
26. Iyatomi, H., Oka, H., Celebi, M.E., Tanaka, M. Ogawa K.: Parameterization of dermoscopic findings for the internet-based melanoma screening system. In: Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007), pp. 189–193 (2007)
27. Iyatomi, H., Oka, H., Celebi, M.E., Ogawa, K., Argenziano, G., Soyer, H.P., Koga, H., Saida, T., Ohara, K., Tanaka, M.: Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. J. Invest. Dermatol. **128**(8), 2049–2054 (2008)
28. Kashyap, R.: Chellappa, R.:Estimation and choice of neighbors in spatial interaction models of images. IEEE Trans. Inf. Theory **1**, 60–72 (1983)
29. Kass, M., Witkin, A.: Analyzing oriented patterns. Comput. Vis. Graph. Image Proces. **37**(3), 362–385 (1987)
30. Khan, A., Gupta, K., Stanley, R.J., Stoecker, W.V., Moss, R.H., Argenziano, G., Soyer, H.P., Cognetta, A.B.: Fuzzy logic techniques for blotch feature evaluation in dermoscopy images. Comput. Med. Imaging Graph. **33**(1), 50–57 (2009)
31. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. Artif. Intel. Med. **56**(2), 69–90 (2012)
32. Madasu, V. K., Lovell, B.C.: Blotch detection in pigmented skin lesions using fuzzy co-clustering and texture segmentation. DICTA 2009—Digital Image Computing: Techniques and Applications, art. no. 5384959, 25–31 (2009)
33. Maglogiannis, I., Doukas, C.N.: Overview of advanced computer vision systems for skin lesions characterization. IEEE Trans. Inf. Technol. Biomed. **13**(5), 721–733 (2009)
34. Mirzaalian, H., Lee, T.K., Hamarneh, G.: Learning features for streak detection in dermoscopic color images using localized radial flux of principal intensity curvature. In: Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis, 6164758, pp. 97–101 (2012)
35. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intel. **24**(7), 971–987 (2002)
36. Pehamberger, H., Steiner, A., Wolff, K.: In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions. J. Am. Acad. Dermatol. **17**, 571–583 (1987)
37. Rangayyan, R.M., Acha, B., Serrano, C.: Color image processing with biomedical applications. SPIE Press, Bellingham (2011)
38. Rezze, G.G., De Sa, B.C.S., Neves, R.I.: Dermoscopy: the pattern analysis. Anais Brasileiros de Dermatologia **81**(3), 261–268 (2006)
39. Sadeghi, M., Lee, T., Lui, H., McLean, D., Atkins, S.: Detection and analysis of irregular streaks in dermoscopic images of skin lesions. IEEE Trans. Med. Imaging (2013, in press)
40. Sadeghi, M., Lee, T.K., McLean, D., Lui, H., Atkins, M.S.: Global pattern analysis and classification of dermoscopic images using textons. Progress in Biomedical Optics and Imaging—Proceedings of SPIE 8314, art. no. 83144X (2012)
41. Sadeghi, M., Lee, T.K., McLean, D., Lui, H., Atkins, M.S.: Oriented pattern analysis for streak detection in dermoscopy images. Medical image computing and computer-assisted intervention : MICCAI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, 15(Pt 1), pp. 298–306 (2012)
42. Sadeghi, M., Razmara, M., Wighton, P., Lee, T.K., Atkins, M.S.: Modeling the dermoscopic structure pigment network using a clinically inspired feature set. Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics. LNCS, vol. 6326, pp. 467–474. Springer, New York (2010)
43. Sadeghi, M., Razmara, M., Lee, T.K., Atkins, M.S.: A novel method for detection of pigment network in dermoscopic images using graphs. Comput. Med. Imaging Graph. **35**(2), 137–143 (2011)

44. Serrano, C., Acha, B.: Pattern analysis of dermoscopic images based on Markov random fields. Pattern Recognit. **42**, 1052–1057 (2009)
45. 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., Argenziano, G., Soyer, H.P.: Detection of atypical texture features in early malignant melanoma. Skin Res. Technol. **16**(1), 60–65 (2010)
46. Situ, N., Yuan, X., Zouridakis, G.: Boosting instance prototypes to detect local dermoscopic features. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10, art. no. 5626776, pp. 5561–5564 (2010)
47. Skrovseth, S.O., Schopf, T.R., Thon, K., Zortea, M., Geilhufe, M., Mllersen, K., Kirchesch, H.M., Godtliebsen, F.: A computer aided diagnostic system for malignant melanomas. In: 2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies. ISABEL, 2010, 5702825 (2010)
48. Steger, C.: An unbiased detector of curvilinear structures. IEEE Trans. Pattern Anal. Mach. Intel. **20**(2), 113–125 (1998)
49. Stoecker, W.V., Moss, R.H.: Editorial: digital imaging in dermatology. Comput. Med. Imaging Graph. **16**(3), 145–150 (1992)
50. Stoecker, W.V., Gupta, K., Stanley, R.J., Moss, R.H., Shrestha, B.: Detection of asymmetric blotches (asymmetric structureless areas) in dermoscopy images of malignant melanoma using relative color. Skin Res. Technol. **11**(3), 179–184 (2005)
51. Tanaka, T., Torii, S., Kabuta, I., Shimizu, K., Tanaka, M., Oka, H.: Pattern classification of nevus with texture analysis. In: Annual International Conference of the IEEE Engineering in Medicine and Biology—Proceedings 26 II, pp. 1459–1462 (2004)
52. Tanaka, T., Torii, S., Kabuta, I., Shimizu, K., Tanaka, M.: Pattern classification of nevus with texture analysis. IEEJ Trans. Electr. Electron. Eng. **3**(1), 143–150 (2008)
53. Wighton, P., Lee, T.K., Lui, H., McLean, D.I., Atkins, M.S.: Generalizing common tasks in automated skin lesion diagnosis. IEEE Trans. Information Technology in Biomed. **15**(4), 5763779, 622–629 (2011)
54. Wighton, P., Sadeghi, M., Lee, T.K., Atkins, M.S.: A fully automatic random walker segmentation for skin lesions in a supervised setting. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), LNCS (PART 2), vol. 5762 , pp. 1108–1115. Springer, Heidelberg (2009)
55. Xia, Y., Feng, D., Zhao, R.: Adaptive segmentation of textured images by using the coupled Markov random field model. IEEE Trans. Image Proces. **11**, 3559–3566 (2006)
56. Yoshino, S., Tanaka, T., Tanaka, M., Oka, H.: Application of morphology for detection of dots in tumor. In: Proceedings of the SICE Annual Conference, WPI-3-3, pp. 407–410 (2004)

A Bag-of-Features Approach for the Classification of Melanomas in Dermoscopy Images: The Role of Color and Texture Descriptors

Catarina Barata, Margarida Ruela, Teresa Mendonça and Jorge S. Marques

Abstract The identification of melanomas in dermoscopy images is still an up to date challenge. Several Computer Aided-Diagnosis Systems for the early diagnosis of melanomas have been proposed in the last two decades. This chapter presents an approach to diagnose melanomas using Bag-of-features, a classification method based on a local description of the image in small patches. Moreover, a comparison between color and texture descriptors is performed in order to assess their discriminative power. The presented results show that local descriptors allow an accurate representation of dermoscopy images and achieve good classification scores: Sensitivity = 93 % and Specificity = 88 %. Furthermore it shows that color descriptors perform better than texture ones in the detection of melanomas.

Keywords Melanoma diagnosis · Dermoscopy · Bag-of-features · Feature extraction · Feature analysis · Color features · Texture features

Introduction

Dermoscopy is a widely used microscopy technique for the in-vivo observation of skin lesions. A magnification instrument is used to increase the size of the lesion and a liquid (oil, alcohol or water) is placed on top of the lesion prior to the

C. Barata (✉) · M. Ruela · J. S. Marques
Institute for Systems and Robotics, Instituto Superior Técnico, Lisboa, Portugal
e-mail: ana.c.fidalgo.barata@ist.utl.pt

M. Ruela
e-mail: margarida.ruela@ist.utl.pt

J. S. Marques
e-mail: jsm@isr.ist.utl.pt

T. Mendonça
Faculdade de Ciências da Universidade do Porto, Porto, Portugal
e-mail: tmendo@fc.up.pt

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_3,

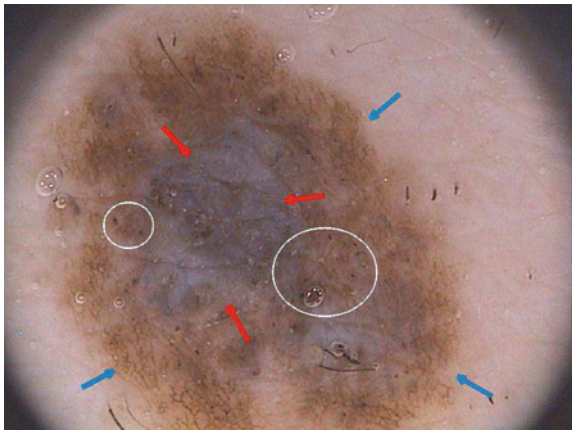


Fig. 1 Melanoma with specific dermoscopic features: blue-whitish veil (*red arrows*); pigment network (*blue arrows*), dots and globules (*white circles*)

observation to eliminate surface reflection. This step makes the cornified layer of the skin translucent, allowing a better visualization of several pigmented structures located within the epidermis, dermis and dermoepidermal junction [3]. Several magnification instruments are currently used by dermatologists: dermatoscope, stereomicroscope or a digital acquisition system. The later allows the attainment of dermoscopy images, that can be processed and analyzed by a Computer Aided-Diagnosis (CAD) system.

The diagnosis of pigmented skin lesions using dermoscopy is based on medical algorithms: ABDC rule [47], 7-point checklist [2] and Menzies' method [29]. All of these methods have in common a set of dermoscopic criteria that can be divided in two groups. The first group is called global features and allows a preliminary and quick categorization of a skin lesion. Global features are a set of patterns (reticular, cobblestone, globular, parallel, etc) that can be found in different pigmented skin lesions. The other group of dermoscopic criteria are the local features (pigment network, dots and globules, streaks, pigmentation related structures, vascular pattern, etc). These features are sometimes called the letters of the dermoscopic alphabet since they are the cues that allow a final diagnosis of the lesion (melanoma or not) [3]. Figure 1 illustrates some of the local dermoscopic features.

Both global and local features play an important role in the diagnosis of melanomas. Some dermatologists perform an analysis of skin lesions using as reference only the global dermoscopic features. This global evaluation method is called pattern analysis and has received some attention in the skin research area, such as [1, 36, 40], which try to reproduce the medical analysis. The published works focus on the identification of the different patterns but do not perform a diagnosis of a skin lesion. However, it is undeniable that local features are the backbone of the common medical algorithms since ABCD rule, 7-point checklist and Menzies' method use these features and their properties (shape or number of colors) to score a skin lesion, thus diagnosing it as melanoma or not. There are several studies which focus on detecting one or more of these dermoscopic criteria, such as pigment network

[6, 37], irregular coloration [10, 31, 45], irregular streaks [35] and granularity [46]. However, as far as the authors know only one study combines a set of detectors and the 7-point checklist algorithm in a CAD system to perform a diagnosis using dermoscopy images [15]. Two reviews on state-of-the art methods can be found in [11, 25].

Most of the CAD systems found in literature use a different procedure, following a pattern recognition approach to classify dermoscopy images [9, 18, 21, 34]. These works have successfully exploited a global representation of the lesion using features inspired by the ABCD rule (color, shape, texture and symmetry). Most of the extracted features are able to perform a good description of the lesion regarding its shape and global color distribution. However, localized texture and color patterns associated to differential structures (e.g., pigment network, dots, streaks or blue-whitish veil) might be missed since a global analysis is being performed. To overcome this situation, this chapter describes a different approach for the analysis of dermoscopy images. Since experts usually try to characterize local structures in the image, the described strategy will try to mimic this behavior and represent the image by a set of local features, each of them associated to a small region in the image. The local features used describe the texture and color of each region and a comparative study between the two types of descriptive features is performed, in order to assess their degree of discrimination.

Bag-of-Features

The description of an image with local features have been successfully used in several complex image analysis problems, such as scene recognition and object-class classification [22, 23, 27, 42, 44, 50]. The used approach is called Bag-of-Features (BoF) [42, 44] and it is inspired by the bag-of-words (BoW) [5], which is a well known text retrieval algorithm. The procedure used by BoW to model documents evolves in three different steps. The first step consists of parsing the documents of the dataset into words, i.e., dividing the documents to smaller components. Images can also be sampled into smaller regions (patches). Two sampling strategies are commonly used in BoF: sparse and dense sampling. Sparse sampling is performed by detecting a set of informative keypoints (e.g., corners) and their respective support regions (square patches). This detection can be done using one or more of the several detectors proposed in literature (e.g. Difference of Gaussian [28] or Harris-Laplace [30]). A comparative study between the six most popular keypoint detectors can be found in [22]. For dense sampling it is assumed that each keypoint is the node of a regular grid defined in the image domain. The patches associated with the keypoints are extracted by sampling uniformly over the grid. Both sampling methods have been used in different works and a comparison between the two strategies was performed by van de Sande et al. [38]. Their experimental results showed that dense sampling outperformed sparse sampling. The BoF approach proposed in this work uses dense

sampling to extract the patches from a given image and only patches whose area is more than 50 % inside the lesion are considered.

The second step in BoW document analysis is to represent each word by its stem. The equivalent for the image analysis case is to represent each patch by a feature vector $x_i \in \mathbb{R}^n$. Different features can be used to locally describe the patches. This chapter focus on two specific kinds of features, color and texture, which will be addressed in section "Local Features". N square patches are extracted from each image I on the dataset. Therefore, a family of local features will be associated with I as follows

$$F = \{x_1, \ldots, x_N\}, \quad x_i \in \mathbb{R}^n. \tag{1}$$

The last step of both BoW and BoF corresponds to the training process. In the first, each discriminative word receives an unique label. Very common words, which occur in most documents, are rejected and do not receive a label. This process can be seen as the creation of a dictionary of representative words. Then, each document is analyzed separately and its discriminative words are compared with the ones from the dictionary. From this comparison will result an histogram of the frequency of occurrence of the dictionary words within the document. This histogram will represent the document and will be used to compare different documents and assess their degree of similarity. Reproducing this histogram representation in the BoF case requires some extra effort. First, assuming that there is a dataset of L images, this dataset has associated with it set of all the extracted local features

$$\mathscr{F} = \bigcup_{k=1}^{L} F^{(k)}. \tag{2}$$

In practice, the set $\mathscr{F}$ has many thousands (or even millions) of feature vectors. Therefore, in order to obtain a visual dictionary (analogous to the dictionary of BoW), this set has to be approximated by a collection of prototypes $c_1, \ldots, c_K$, called *visual words*. The *visual words* are obtained using a clustering algorithm (in this work K-means is used). After obtaining a visual dictionary, all feature vectors in the training set are classified in the nearest *visual word* and a label

$$l_i^{(k)} = \arg\min_j \|x_i^{(k)} - c_j\|, \tag{3}$$

which identifies a specific *visual word*, assigned to each feature vector $x_i^{(k)}$. The final step is to characterize each image $I^{(k)}$ by a histogram of visual words frequency

$$h^{(k)}(l) = \frac{1}{N^{(k)}} \sum_{i=1}^{N^{(k)}} \delta(l_i^{(k)} - l), \tag{4}$$

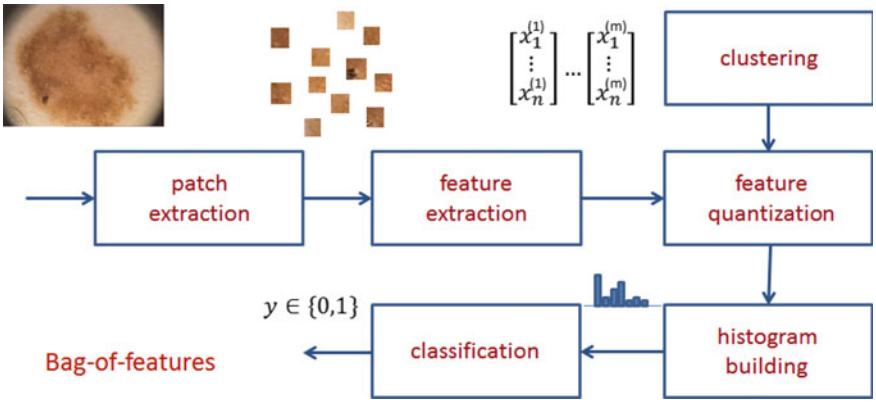


Fig. 2 Block diagram of the BoF classification system

where $\delta(.)$ denotes the Kronecker delta ($\delta(x) = 1$, if $x = 0$; $\delta(x) = 0$, otherwise). As in BoW, this histogram will act as the feature vector that describes the image and the set of L feature vectors obtained this way will be used to train a classifier.

For each new image I to be classified the process is similar to the one described previously. The image is sampled and local features are extracted. Then, the local features are compared with the dictionary of visual words obtained in the training phase and, finally, the histogram of visual words frequency is computed. The image is classified using the computed histogram and the classifier learned using the training set.

All the steps of the BoF strategy described previously are summarized in Fig. 2.

There are several factors that can impact the performance of BoF. Following the blocks sequence on Fig. 2 these factors are: the size of the regular grid (δ) used in the patch extraction step, the type and quantity of extracted features, the size of the dictionary (K), and the classification algorithm used. All these factors are thoroughly analyzed in this work. Several values for δ and K are tried and three classification algorithms with different properties are tested: k-nearest neighbor (kNN) [16], AdaBoost [17] and support vector machines (SVM) [8, 12].

Local Features

The local dermoscopic criteria used by dermatologists to diagnose skin lesions can be represented by two different kinds of image features: texture and color features. Local dermoscopic structures such as pigment network, dots and streaks can be characterized by texture features since these features represent the spatial organization of intensity in an image, allowing the identification of different shapes. Color features describe the color distribution, thus they are able to characterize particular pigmented regions such as blue-whitish veil or regression areas.

In theory both of these features provide a good description of the extracted patches (see Fig. 2) and both play an important role in the final classification. One of the main objectives of this chapter is to determine if the previous hypothesis are correct. This objective is accomplished by assessing the performance of color and texture features separately and by combining both of them. Moreover, since both color and texture features can be extracted using different types of descriptors, a comparison between some of them is also performed. The several texture and color descriptors tested are described in the next sections.

Texture Descriptors

Texture features characterize the intensity of an image. Therefore, it is necessary to convert the original RGB image into a gray level one before extracting texture descriptors. This is done by selecting the color channel with the highest entropy [41].

Texture descriptors can be divided into several categories depending on the methodology used. This chapter focus on three different methods: statistical, signal processing and gradient [33]. In statistical methods, the features are extracted by computing neighbor pixel statistics. A very well known method for computing these statistics is the gray level co-occurrence matrix (GLCM) proposed by Haralick et al. [20]. This matrix stores the relative frequencies of gray level pairs of pixels at a certain relative displacement and can then be used to compute several statistics which will be the elements of the feature vector. The results presented in this chapter are obtained using five of the most common statistics: contrast, correlation, homogeneity, energy and entropy. The performance of these features is directly related with GLCM, since it has been already proved that the performance of a classification system is influenced by the number of gray levels (G) used as well as the way of combining the orientations of the nearest neighbors [13]. Therefore, several values of G and two ways of combining the orientations (average GLCM versus four GLCM) are tested, according to what is proposed in [13].

Signal processing approaches have in common three sequential steps. First, the image $I(x, y)$ is convolved with a bank of N filters, with a certain impulse response $h_i(x, y), \quad i = 1, \ldots, N$. Then, an energy measure of the output $J_i(x, y) \quad i = 1, \ldots, N$, is performed

$$E_i = \sum_x \sum_y |J_i(x, y)|. \tag{5}$$

Finally, the energy content is used to computed statistics that are the components of the feature vector [33]. The two statistics computed in this chapter are the mean μ_i and standard deviation σ_i

$$\mu_i = \frac{E_i}{M}, \quad \sigma_i = \sqrt{\frac{E_i - \mu_i}{M}}, \tag{6}$$

where M is the number of pixels (x, y) in J_i.

Several filter banks can be found on literature [33]. This chapter compares two of the most well known: Laws [26] and Gabor [4] filter masks. The filter masks proposed by Laws [26] have been widely used for texture characterization. These masks can have a dimension 3×3 or 5×5 and result from convolving two of of the five possible 1-D kernels. Each 1-D kernel focus on specific textural characteristics like edges, waves or ripples. In this chapter only three kernels will be used: L, which computes the average grey Level, E that extracts Edges (describe linear structures, such as pigment network) and S that extracts Spots (describe circular structures, such as dots). The 1-D kernel values are the following: $L_3 = [1\ 2\ 1]$, $E_3 = [1\ 0\ -1]$, $S_3 = [1\ -2\ 1]$, $L_5 = [1\ 4\ 6\ 4\ 1]$, $E_5 = [-1\ -2\ 0\ 2\ 1]$ and $S_5 = [-1\ 0\ 2\ 0\ -1]$. All the possible combinations of 1-D kernels are considered, thus the filter bank has a dimension $N = 9$. Since it is not known the dimension of the masks that leads to the best results, both 3×3 and 5×5 filter banks are tested.

Gabor filters have been used for texture classification [4] and edge detection [19]. Therefore, they can be used to characterize dermoscopic structures that have a linear shape (e.g. pigment network or streaks). The impulse response of a Gabor filter is the following

$$h_i(x, y) = e^{\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma_G^2}} \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right), \tag{7}$$

where γ is an aspect ratio constant. σ_G is the standard deviation, λ is the wavelength, φ is the phase of the filter and $(\tilde{x}, \tilde{y})$ are obtained from rotating (x, y) as follows [19]

$$\tilde{x} = x \cos\theta_i + y \sin\theta_i, \qquad \tilde{y} = -x \sin\theta_i + y \cos\theta_i. \tag{8}$$

The angle amplitude $\theta_i \in [0, \pi]$ determines the orientation of the filter h_i and the step between two consecutive filters is $\frac{\pi}{N_\theta}$, where N_θ is the number of filters in the filter bank [19]. This descriptor depends on several parameters. In this chapter two of them are varied: $\sigma_G \in \{1, 2, \ldots, 5\}$ and $N_\theta \in \{2, 3, \ldots, 10\}$. All the others are kept constant and equal to: $\gamma = 0.5$, $\varphi = 0\ rad$ and $\frac{\tilde{x}}{\lambda} = 0.56$ [19].

Gradient features such as gradient histograms have also been successfully used in several classification problems [14, 28]. In this work, two gradient histograms are used: amplitude and orientation. The image gradient $g(x, y) = [g_1(x, y)\ \ g_2(x, y)]$ is computed using Sobel masks. Then, gradient magnitude and orientation are respectively computed as follows

$$\| g(x, y) \| = \sqrt{g_1(x, y)^2 + g_2(x, y)^2}, \phi(x, y) = \tan^{-1}\left(\frac{g_2(x, y)}{g_1(x, y)}\right). \tag{9}$$

Finally, the histograms of gradient amplitude and orientation are obtained

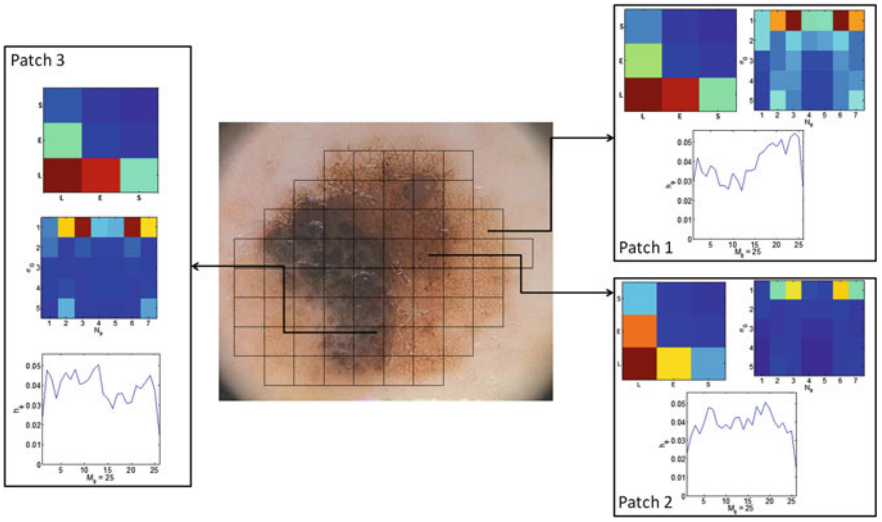


Fig. 3 Texture features for three different 60×60 patches: energy content for Laws (5×5 masks) and Gabor ($N_\theta = 1, 2, \ldots, 7$ and $\sigma_G = 1, 2, \ldots, 5$) filters and histogram of the gradient phase ($M_\phi = 25$)

$$h_a(i) = \frac{1}{N} \sum_x \sum_y b_i(\|g(x, y)\|), \qquad i = 1, \ldots, B_a,$$
$$h_\phi(i) = \frac{1}{N} \sum_x \sum_y \tilde{b}_i(\phi(x, y)), \qquad i = 1, \ldots, B_\phi, \tag{10}$$

where N is the number of pixels inside the patch and B_a, B_ϕ are the number of bins of the magnitude and orientation histograms, respectively. Finally, $b_i(.)$, $\tilde{b}_i(.)$ are the characteristic functions of the ith histogram bin

$$b_i(a) = \begin{cases} 1 \text{ if} a \text{belongs to the} i\text{th amplitude bin} \\ 0 \text{ otherwise} \end{cases}$$
$$\tilde{b}_i(\phi) = \begin{cases} 1 \text{ if} \phi \text{belongs to the} i\text{th orientation bin} \\ 0 \text{ otherwise} \end{cases}. \tag{11}$$

The parameter varied for both gradient features is the number of bins of the histograms ($B_a \in \{15, 25, 35, 45\}$ and $B_\phi \in \{15, 25, 35, 45\}$).

Figure 3 shows some of the extracted texture features for three different patches within the same lesion. The exemplified patches were selected in order to include a specific dermoscopic structure: pigment network (patch 1), dots (patch 2) and globules (patch 3). It is clear that the extracted descriptors (Laws, Gabor and h_ϕ) are different for each patches, which demonstrates that different dermoscopic structures have different textural properties and, therefore, must be described separately.

Color Descriptors

Several color descriptors, such as histograms and mean color, have been used in object and scene recognition problems [39]. The descriptors are usually computed over one or more color spaces like RGB, HSV/I [49], CIE La*b* and L*uv [49] and the biologically inspired opponent color space (Opp) [7]. These six color spaces have different properties, thus they might provide different information for the melanoma classification problem addressed in this chapter. For this reason the six previous color spaces are tested.

For each color space a set of three histograms is computed (one for each of the three color components). For each patch, the histogram associated with the color channel $I_c,\ c \in \{1, 2, 3\}$ is given by

$$h_c(i) = \frac{1}{N} \sum_{x,y} b_c(I_c(x, y))\ i = 1, \ldots, B_c, \tag{12}$$

where N is the number of pixels inside the patch, i is the histogram bin, B_c is the number of bins and $b_c(.)$ is the characteristic function of the ith bin

$$b_c(I_c(x, y)) = \begin{cases} 1 \text{ if} I_c(x, y) \text{belongs to the} i\text{th color bin} \\ 0 \text{ otherwise} \end{cases}. \tag{13}$$

The bins are defined by dividing the color component range into intervals with the same width. For all histograms, the number of intervals $B_c \in \{15, 25, 35, 45\}$ is a tested parameter.

Another common color descriptors are color moments [48]. Color moments result from assuming that the distribution of color in an image can be seen as a probability distribution. Since probability distributions are usually characterized by a set of unique moments, they can be used as color features. The general definition of the $1st$ order color moment for the color channel $I_c,\ c \in \{1, 2, 3\}$ is the following

$$M_c^1 = \frac{\sum_x^N I_c(x)}{N}, \tag{14}$$

where N is the number of pixels inside the patch. Higher order (p) color moments are defined by [48]

$$M_c^p = \left(\frac{\sum_x^N (I_c(x) - M_c^1)^p}{N} \right)^{\frac{1}{p}}. \tag{15}$$

The first three order moments are used in this chapter. These moments correspond to: mean (M^1), standard deviation (M^2) and skewness (M^3). Therefore, each patch will be represented by a total of nine color moments (three for each color component).

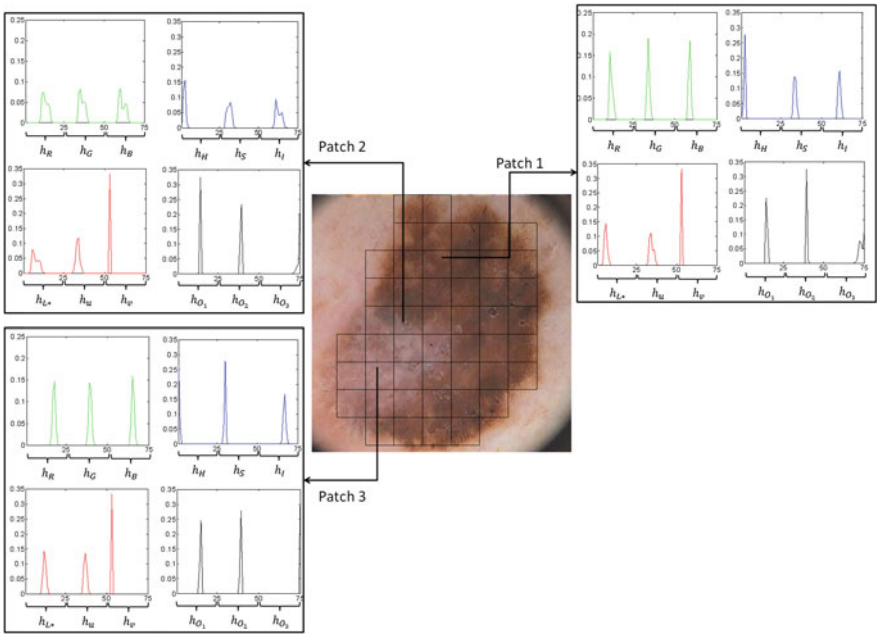


Fig. 4 Color histograms ($M_c = 25$) for three different 60×60 patches: h_{RGB} (*green*), h_{HSI} (*blue*), h_{L^*uv} (*red*) and h_{Opp} (*black*)

Figure 4 exemplifies some of the extracted color histograms (RGB, HSI, L*uv and Opp) for three different patches. Each patch was selected in order to represent a different color section of the lesion: patch 1 was extracted from the brown region, patch 2 was extracted from the brown-white transition region and patch 3 was extracted from the white region. As in the texture features case, the three feature vectors are different, thus each color region is characterized differently.

Experimental Results

The proposed method was evaluated with a dataset of 176 dermoscopy images (25 melanomas and 151 nevi). These images were taken during clinical exams performed at Hospital Pedro Hispano, Matosinhos, with a digital acquisition system that allows a magnification of $20\times$. Images are stored in BMP and $JPEG$ formats and their average resolution is 573×765. Each image was classified by an experienced dermatologist as *melanoma* or *non-melanoma* (ground truth label).

The evaluation metrics used are the Sensitivity (SE) and Specificity (SP). These two measures are combined in a cost function (C)

$$C = \frac{c_{10}(1 - SE) + c_{01}(1 - SP)}{c_{10} + c_{01}}, \tag{16}$$

where c_{10} is the cost of an incorrectly classified melanoma and c_{01} is the cost of an incorrectly classified non-melanoma. This cost function represents the trade-off between SE and SP. In this chapter it is assumed that an incorrect classification of a melanoma is worse and, therefore, the classification error costs are defined as: $c_{10} = 1.5c_{01}$ and $c_{01} = 1$. The selected classifiers are those which achieve the lowest values of C.

Since the dataset is small, the different possibilities are tested using a stratified 10-fold cross validation method. Both classes were evenly distributed by the ten folds. To decrease the impact of class unbalance, local features associated with each melanoma in the different training sets were repeated. Gaussian noise ($w \sim N(0, \sigma_n^2)$, with $\sigma_n = 0.0001$) was added to each repeated local features to prevent exact matching between feature vectors on the training set.

BoF depends on several parameters. The local feature extraction process relies on the size of the patches (δ) while the classification process depends both on the size of the codebook (K) and the classification algorithm used. The best δ is searched in the set $\{20, 40, \ldots, 100\}$ and K in the set $\{100, 200, 300\}$. Each one of the tested classification algorithms depends on several parameters as well. In the kNN case, the parameters tested are the number of neighbors ($k \in \{5, 7, \ldots, 25\}$) and the distance d used to compare the feature vectors. For two vectors $\mathbf{x}$ and $\mathbf{y}$ these distances are computed as follows

- **Euclidean**

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \tag{17}$$

- **Histogram Intersection**

$$d(\mathbf{x}, \mathbf{y}) = \sum_i min(x_i, y_i) \tag{18}$$

- **Kullback-Leibler**

$$d(\mathbf{x}, \mathbf{y}) = \sum_i \log\left(\frac{y_i}{x_i}\right) y_i \tag{19}$$

For AdaBoost, the parameter tested is the number of weak classifiers ($W \in [2, 300]$). SVM is tested using the BoF default kernel χ^2, defined for two vectors $\mathbf{x}$ and $\mathbf{y}$ as

$$Kernel_{\chi^2} = e^{-\rho d_{\chi^2}(\mathbf{x},\mathbf{y})} \tag{20}$$

where width parameter ρ is searched in the set $\{2^{-9}, 2^{-8}, \ldots, 2^8, 2^9\}$ and

$$d_{\chi^2}(\mathbf{x}, \mathbf{y}) = \sum_i \frac{(x_i - y_i)^2}{x_i + y_i} \tag{21}$$

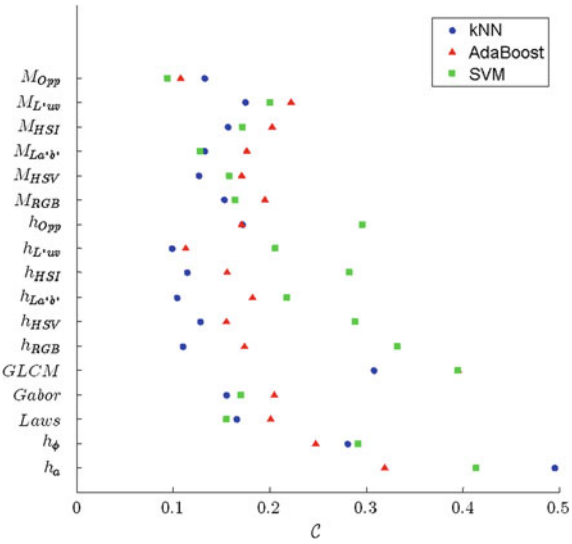


Fig. 5 Best cost results (x axis) for single features using kNN (*blue circle*), AdaBoost (*red triangle*) and SVM (*green square*)

The optimal parameters and results are computed using a total of more than 4,90,000 possible combinations.

Single Features Comparison for the Best Configuration of Classifiers

Figure 5 shows the cost results obtained for the different local features and classifiers. These results show that good classification scores can be achieved using single features (e.g., $C = 0.094$ for M_{Opp}, $C = 0.099$ for h_{L^*uv} and $C = 0.104$ for $h_{La^*b^*}$).

For texture features, the two best descriptors are the signal processing ones. Both Laws and Gabor texture descriptors achieve promising results: $C = 0.155$ for Gabor and $C = 0.166$ for Laws. GLCM and gradient features achieve worse results, which suggests that filter descriptors provide more discriminative information regarding local dermoscopic features.

Color space histograms outperform the corresponding color moments for kNN and AdaBoost classifiers. In the case of SVM, color moments perform much better than their corresponding histograms and the best single descriptor classification result is achieved in this case: M_{Opp}, $C = 0.094$, that corresponds to $SE = 94\,\%$ and $SP = 88\,\%$. The best results achieved with color features outperform those obtained with texture features, which suggest that the former are more discriminative.

Table 1 shows the performance measures and the best configurations for some of the best texture and color features. Good results are achieved both with kNN

Table 1 Classification results and configurations for best texture and color features

Features	SE (%)	SP (%)	C	Feature parameters	Classifier parameters
Laws	100	61	0.155	$\delta = 40\ 5 \times 5$	$K = 100$, SVM: $\rho = 2^{-5}$
Gabor	98	64	0.155	$\delta = 80$, $\sigma_G = 4$ $N_\theta = 5$	$K = 100$, kNN: $k = 19$, Histogram intersection
h_{L^*uv}	100	75	0.099	$\delta = 80$, $B_c = 15$	$K = 300$, kNN: $k = 13$, Kullback-Leibler
$h_{La^*b^*}$	93	85	0.104	$\delta = 80$, $B_c = 25$	$K = 300$, kNN: $k = 19$, Kullback-Leibler
M_{Opp}	93	88	0.094	$\delta = 40$	$K = 100$, SVM: $\rho = 2^{-4}$

and SVM and, for some features, both classifiers lead to very similar results (see Fig. 5). However, kNN appears to achieve the best overall results. It is interesting to notice that for Gabor, h_{L^*uv} and $h_{La^*b^*}$ the best comparative distance are the statistical ones: Kullback-Leibler or Histogram Intersection. This occurrence is also noted for the other tested features and can be explained by the fact that the actual features provided for training and classification are the histograms of *visual words* frequencies, i.e., distributions.

Fusion of Color and Texture Features

Combining different descriptors of the same class may improve the results. To test this hypothesis the two best texture descriptors (Gabor and Laws) and the two types of color descriptors (moments and histograms) were combined. The fusion strategy used is early fusion where the feature vectors are concatenated into a single one [43]. Figure 6 shows the results achieved for each pair of descriptors using kNN and SVM (similar performance is achieved with AdaBoost). The results for the pairs color moments/histograms are identified with the respective color space, whilst the pair Gabor/Laws is labeled Tex. Finally, the performance achieved with each descriptor of the pair is also represented in the graphics (red asterisks). These results show that the fusion of descriptors can either improve the overall results (more evident on the SVM case) or improve the results when compared with the worst descriptor of the pair. Moreover, the fusion of descriptors approximate the results achieved with color and texture descriptors by significantly improving the last ones.

Table 2 shows the best results achieved for each pair as well as the configurations that led to those results. It is interesting to notice that the best global results are achieved with RGB despite the drawbacks of this color space. Nonetheless, this result is worse than the one achieved with the best single descriptor (see Table 1).

Combining different classes of descriptors is an usual approach in BoF and other pattern recognition methods. In this chapter, the best color and texture descrip-

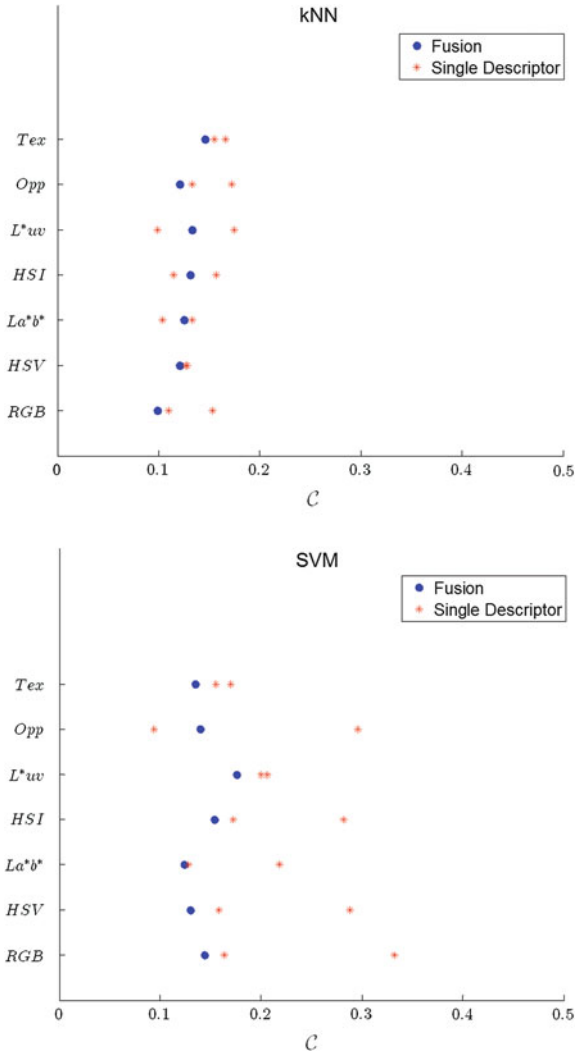


Fig. 6 Early fusion cost results (x axis) for kNN (*top*) and SVM (*bottom*). The performance of each descriptor of the pair is also shown (*red asterisk*)

tors were combined. Due to the different properties of the feature sets and to the large dimension of the feature vector that would result from an early fusion, a late fusion strategy was applied in this case [43]. In this method, the final decision is made by combining the outputs of different classifiers (in this case one classifier for color and other for texture descriptors). The SVM classifier trained using the M_{Opp} (see Table 1) descriptor was combined with two classifiers that achieved similar classification results: kNN ($C = 0.145$) and SVM ($C = 0.135$), both trained using the early fusion of Laws and Gabor descriptors (see Fig. 6). Late fusion strategies use the posteriori probability of each class and one of several possible rules [24] to make

Table 2 Classification results and configurations for descriptors fusions

Fusion	SE (%)	SP (%)	C	Feature parameters	Classifier parameters
Tex	91	79	0.135	$\delta = 40, 5 \times 5$ $\sigma_G = 1$ $N_\theta = 5$	K = 100, SVM: $\rho = 2^{-3}$
Opp	96	77	0.121	$\delta = 20$, $B_c = 25$	$K = 100$, kNN: $k = 15$, Histogram intersection
L^*uv	93	75	0.141	$\delta = 20$, $B_c = 15$	$K = 200$, kNN: $k = 9$, Kullback-Leibler
HSI	100	67	0.131	$\delta = 60$, $B_c = 25$	$K = 300$, kNN: $k = 23$, Histogram intersection
La^*b^*	89	86	0.124	$\delta = 80$, $B_c = 25$	$K = 100$, SVM: $\rho = 2^{-3}$
HSV	98	73	0.121	$\delta = 60$, $B_c = 45$	$K = 300$, kNN: $k = 11$, Histogram intersection
RGB	100	75	0.099	$\delta = 80$, $B_c = 45$	$K = 200$, kNN: $k = 15$, Histogram intersection

the final decision. For SVM, the probabilities are computed using the Platt's method [32] while for kNN, the posterior probabilities are computed as follows:

$$P(w|x) = \frac{k_w}{k}, \tag{22}$$

where w represents the class that can be either 0 or 1, x is a pattern to be classified and k_w is the number of patterns amongst the total number of neighbors k that belong to class w. The combination is computed using the Sum-Rule [24].

The best fusion result was achieved using the kNN classifier trained with texture descriptors combined with the M_{Opp}: $C = 0.097$, SE $= 96\,\%$, SP $= 82\,\%$. Since the color descriptor performs much better than the texture descriptors (Gabor and Laws), it is understandable that the performance of the fusion is slightly inferior to the color descriptor alone.

Visual Words Analysis

The analysis of the identified *visual words* can provide important information regarding the most distinctive dermoscopic structures. Therefore, a simple study of the *visual words* that lead to the best results is performed in this chapter.

This study is done using the information of local color features M_{Opp}, using the best configuration for these features (see Table 1). The average *visual words* histograms for melanomas and non-melanomas were computed using all the histograms of the dataset. Figure 7 shows the obtained average histograms. These two

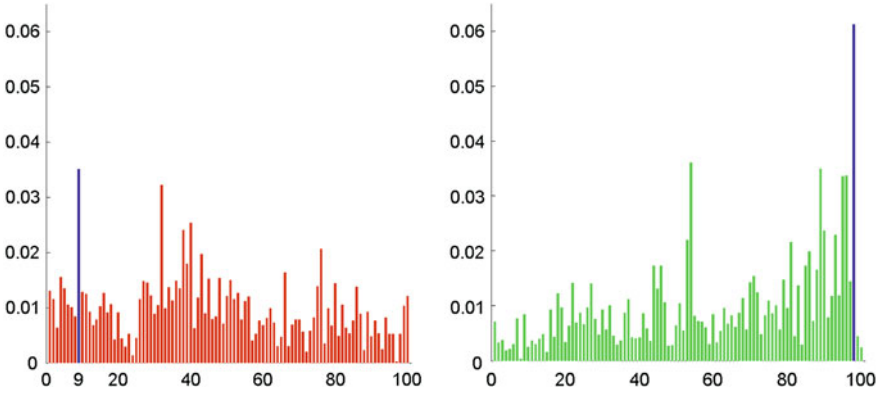


Fig. 7 Average visual words histograms for melanoma (*red*) and non-melanoma (*green*), obtained using the local color features M_{Opp}. Most frequent *visual word* is highlighted (*blue*)

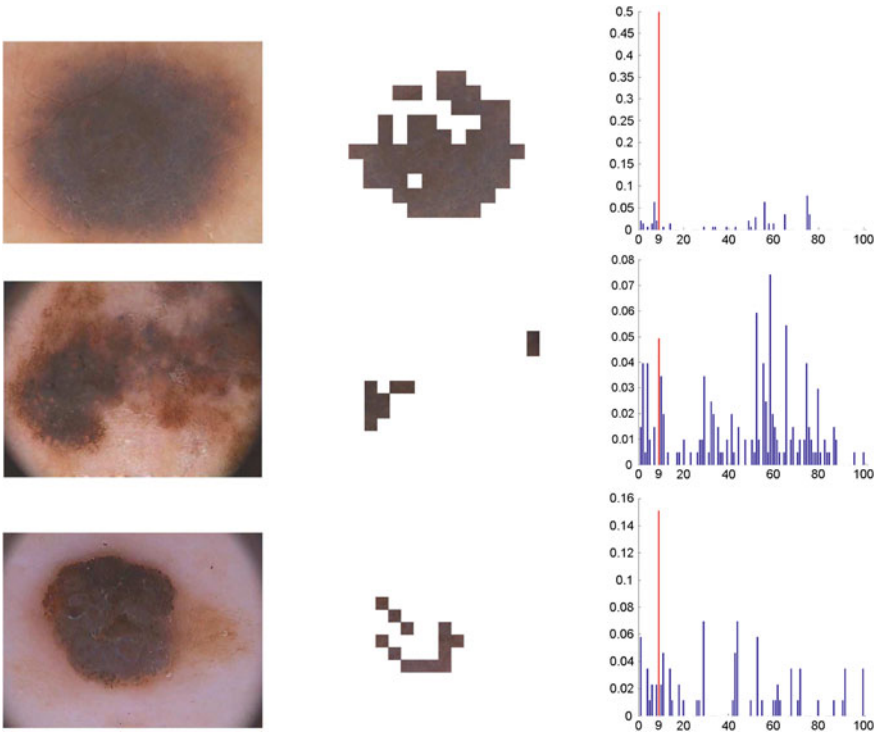


Fig. 8 Example of the most selected color visual word in melanomas: melanomas (*left*); 9th bin visual word (*mid*); visual words histograms, with the 9th bin visual word highlighted (*right*)

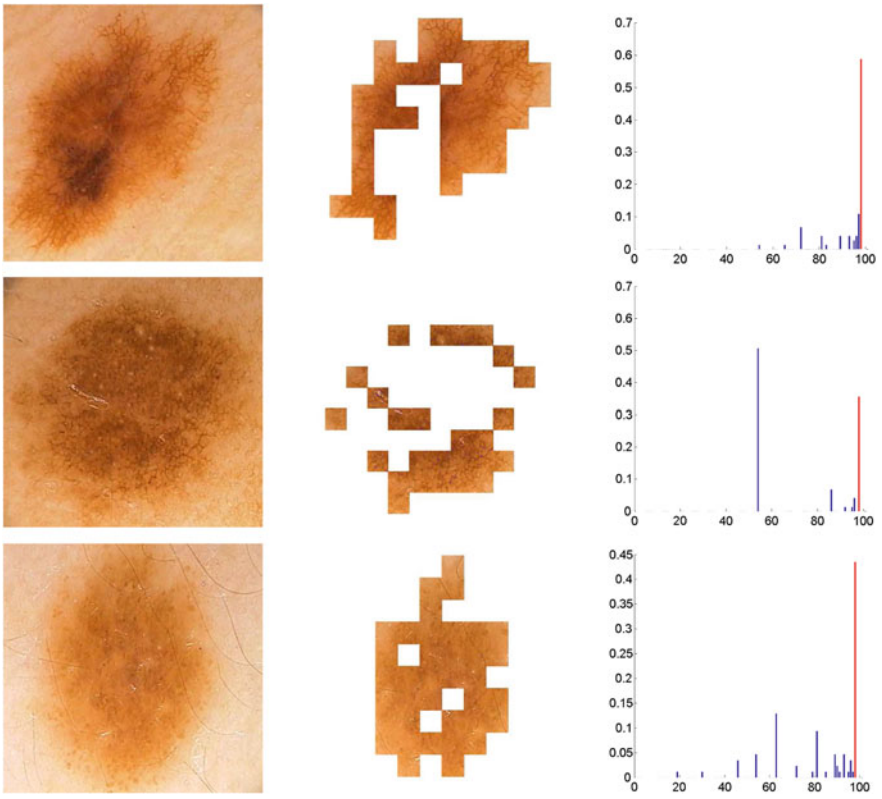


Fig. 9 Example of the most selected color visual word in non-melanomas: non-melanomas (*left*); 98th bin visual word (*mid*); visual words histograms, with the 98th bin visual word highlighted (*right*)

histograms are significantly different and it is interesting to notice that there are some *visual words* which are more common than others.

The next step is to select the most frequent *visual word* for melanomas (see Fig. 6, highlighted 9th bin) and assess the patches associated with this word. Figure 8 shows three melanomas from the dataset, the patches associated with the *visual word* and the corresponding histograms. Although each lesion is described by a different histogram, the same *visual word* (9th bin) is present in all of them and is one of the most frequent. The detected patches are extracted from a blue-gray region (blue-whitish veil), which is one of the atypical pigmentations associated with melanoma [3].

A similar analysis can be performed using the most frequent *visual word* of non-melanomas (see Fig. 7). As before, the patches identified as being this specific visual word were extracted from examples of benign lesions (see Fig. 9). This *visual word* corresponds to a healthy light brown region in all the exemplified lesions. Moreover,

this *visual word* is associated with a considerable number of patches, which suggests that these lesions have a more or less uniform pigmentation. Experts usually associate these two evidences (light brown and uniform pigmentation) with benign lesions [3]. It is interesting to notice that, as in the previous analysis, the observations are consistent with the medical knowledge.

The *visual words* analysis performed is simple. However, the results are interesting and in accordance to what is expected to observe in both melanoma and non-melanoma lesions. This suggests that BoF can be used to classify melanomas and to identify specific dermoscopic features and patterns [3] by associating them to *visual words*. Future work should focus on this task, which can be seen as a multiple object recognition problem.

Conclusions

This chapter investigates the applicability of local color and texture features to the melanoma classification problem. Several factors associated with the performance of BoF were tested, namely the type of descriptors used and the classification algorithm.

The results show that individually color descriptors perform better than texture descriptors and that good classification results can be achieved using kNN (SE = 93 %, SP = 85 % with $h_{La^*b^*}$ and SE = 100 %, SP = 75 % with h_{L*uv}) and SVM (SE = 93 %, SP = 88 % with M_{Opp}). The fusion of color and texture descriptors also achieved good results, with a score of SE = 96 %, SP = 82 % for the combination of Opp moments with Gabor and Laws texture descriptors.

A simple analysis of the *visual words* showed that the dictionary obtained using BoF has potential to be used as a detector/identifier for specific dermoscopic features and patterns. Future work will rely on testing this hypothesis in order to develop a more medical oriented system. Moreover, sparse sampling methods should be tested in order to compare their performances with that of the dense sampling used in this chapter. Finally, high-level descriptors should be tested as well.

Acknowledgments The authors thank to Dr. Jorge Rozeira for providing the dermoscopy images. This work was supported by Fundação Ciência e Tecnologia in the scope of the grant SFRH/BD/84658/2012 and projects PTDC/SAU-BEB/103471/2008 and PEst-OE/EEI/LA0009/2011.

References

1. Abbas, Q., Celebi, M.E., Serrano, C., García, I.F., Ma, G.: Pattern classification of dermoscopy images: a perceptually uniform model. Pattern Recogn. **46**, 86–97 (2013)
2. Argenziano, G., Fabbrocini, G., Carli, P., De Giorgi, V., Sammarco, E., Delfino, M.: 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. Arch. Dermatol. **134**, 1563–1570 (1998)

3. Argenziano, G., Soyer, H., De Giorgi, V., Carli, P., Delfino, M., Ferrari, A., Hofmann-Wellenhof, R., Massi, D., Mazocchetti, G., Scalvenzi, M., Wolf, I.: Interactive atlas of dermoscopy. Edra Medical Publishing and New Media, Milan (2000). http://www.dermoscopy.org/atlas/
4. Arivazhagana, S., Ganesanb, L., Priyala, S.P.: Texture classification using gabor wavelets based rotation invariant features. Pattern Recogn. Lett. **27**, 1976–1982 (2006)
5. Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. ACM Press, New York (1999)
6. Barata, C., Marques, J.S., Rozeira, J.: A system for the detection of pigment network in dermoscopy images using directional filters. IEEE Trans. Biomed. Eng. **59**(10), 2744–2754 (2012)
7. Bratkova, M., Boulos, S., Shirley, P.: oRGB: A pratical opponent color space for computer graphics. IEEE Comput. Graphics Appl. **29**, 42–55 (2009)
8. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Disc. **2**, 121–167 (1998)
9. Celebi, M.E., Kingravi, H.E., Uddin, B., Iyatomi, H., Aslandogan, Y., Stoecker, W.V., Moss, R.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imag. Graphics **31**(6), 362–373 (2007)
10. Celebi, M.E., Iyatomi, H., Stoecker, W., Moss, R.H., Rabinovitz, H., Soyer, H.P.: Automatic detection of blue-white veil and related structures in dermoscopy images. Comput. Med. Imag. Graph. **32**(8), 670–677 (2008)
11. Celebi, M.E., Stoecker, W.V., Moss, R.H.: Advances in skin cancer image analysis. Comput. Med. Imag. Graph. **35**, 83–84 (2011)
12. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**, 27:1–27:27 (2011). Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm
13. Clausi, D.: An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. **28**(1), 45–62 (2002)
14. Dala, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and. Pattern Recogn. **1**, 886–893 (2005)
15. Di Leo, G., Paolillo, A., Sommella, P., Fabbrocini, G.: Automatic diagnosis of melanoma: a software system based on the 7-point check-list. In: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences, pp. 1818–1823 (2010)
16. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley, New York (1999)
17. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. **55**, 119–139 (1997)
18. Ganster, H., Pinz, A., Wildling, E., Binder, M., Kittler, H.: Automated melanoma recognition. IEEE Trans. Med. Imag. **20**(3), 233–239 (2001)
19. Grigorescu, C., Petkov, N., Westenberg, M.A.: Contour detection based on nonclassical receptive field inhibition. IEEE Trans. Image Process. **12**, 729–739 (2003)
20. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**, 610–621 (1973)
21. Iyatomi, H., Oka, H., Celebi, M.E., Hashimoto, M., Hagiwara, M., Tanaka, M., Ogawa, K.: An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. Comput. Med. Imag. Graphics **32**(7), 566–579 (2008)
22. Jiang, Y.G., Ngo, C.W., Yang, J.: Towards optimal bag-of-features for object categorization and semantic video retrieval. In: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, pp. 494–501 (2007)
23. Khan, F.S., van de Weijer, J., Vanrell, M.: Top-down color attention for object recognition. In: Proceedings of the IEEE 12th International Conference on Computer Vision, pp. 979–986 (2009)
24. Kittler, J., Hatef, M., Duin, R.P.W., Matas, J.: On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. **20**, 226–239 (1998)
25. Korotkov, K., Garcia, R.: Computerized analysis of pigment skin lesions: a review. Artif. Intell. Med. **56**, 69–90 (2012)

26. Laws, K.I.: Rapid texture identification. In: Proceedings of SPIE Conference on Image Processing for Missile Guidance (1980)
27. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2169–2178 (2006)
28. Lowe, D.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
29. Menzies, S., Ingvar, C., Crotty, K., McCarthy, W.H.: Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. Arch. Dermatol. **132**, 1178–1182 (1996)
30. Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. Int. J. Comput. Vis. **60**(1), 63–86 (2004)
31. Pellacani, G., Grana, C., Cucchiara, R., Seidenari, S.: Automated extraction of dark areas in surface microscopy melanocytic lesion images. Dermatology **208**(1), 21–26
32. Platt, J.: Probabilities for sv machines. In: Advances in Large Margin Classifiers, pp. 61–74. MIT Press, Cambridge (2000)
33. Randen, T., Husoy, J.H.: Filtering for texture classification: A comparative study. IEEE Trans. Pattern Anal. Mach. Intell. **21**, 291–310 (1999)
34. Rubegni, P., Cevenini, G., Burroni, M., Perotti, R., Dell'Eva, G., Sbano, P., Miracco, C.: Automated diagnosis of pigmented skin lesions. Int. J. Cancer **101**(6), 576–580 (2002)
35. Sadeghi, M., Lee, T.K., McLean, D., Lui, H., Atkins, M.S.: Detection and analysis of irregular streaks in dermoscopic images of skin lesions. IEEE Trans. Med. Imag. **32**(5), 849–861. doi:10.1109/TMI.2013.2239307
36. Sadeghi, M., Lee, T.K., McLean, D., Lui, H., Atkins, M.S.: Global pattern analysis and classification of dermoscopic images using textons. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 168–173 (2012)
37. Sadeghi, M., Razmara, M., Wighton, P., Lee, T.K., Atkins, M.S.: A novel method for detection of pigment network in dermoscopic images using graphs. Comput. Med. Imag. Graph. **35**(2), 137–143 (2011)
38. Van de Sande, K.E.A., Gevers, T., Snoek, C.G.M.: A comparison of color features for visual concept classification. In: Proceedings of the 2008 international conference on Content-based image and video retrieval (2008)
39. van de Sande, K.E.A., Gevers, T., Snoek, C.G.M.: Evaluating color descriptors for object and scene recognition. IEEE Trans. Pattern Anal. Mach. Intell. **32**, 1582–1593 (2010)
40. Serrano, C., Acha, B.: Pattern analysis of dermoscopic images based on markov random fields. Pattern Recogn. **42**, 1052–1057 (2009)
41. Silveira, M., Nascimento, J.C., Marques, J.S., Marçal, A.R.S., Mendonça, T., Yamauchi, S., Maeda, J.: Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. IEEE J. Sel. Top. Sign. Process. **3**, 35–45 (2009)
42. Sivic, J., Zisserman, A.: Video google: A text retrieval approach to object matching in videos. In: Proceedings of the 9th IEEE International Conference on Computer Vision, pp. 1470–1477 (2003)
43. Snoek, C.G.M.: Early versus late fusion in semantic video analysis. In: In ACM Multimedia, pp. 399–402 (2005)
44. Squire, D.M., Mller, W., Mller, H., Raki, J.: Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback. Pattern Recogn. Lett. **21**(13–14), 143–149 (1999)
45. Stoecker, W.V., Gupta, K., Stanley, R.J., et al.: Detection of asymmetric blotches in dermoscopy images of malignant melanomas using relative color. Skin Res. Technol. **11**(3), 179–184 (2005)
46. Stoecker, W.V., Wronkiewicz, M., Chowdhury, R., Stanley, R., Xu, J., Bangert, A., Shrestha, B., Calcara, D., Rabinovitz, H., Oliviero, M., Ahmed, F., Perry, L., Drugge, R.: Detection of granularity in dermoscopy images of malignant melanoma using color and texture features. Comput. Med. Imag. Graph. **35**(2), 144–147 (2011)

47. Stolz, W., Riemann, A., Cognetta, A.B.: ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur. J. Dermatol. **4**, 521–527 (1994)
48. Stricker, M., Orengo, M.: Similarity of color images. In: Proceedings SPIE, vol. 2420, pp. 381–392 (1995)
49. Tkalcic, M., Tasicl, J.F.: Colour spaces: perceptual, historical and applicational background. In: Proceedings of the IEEE Region 8 EUROCON 2003. Computer as a Tool, vol. 1, pp. 304–308 (2003)
50. Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: An in-depth study. Tech. Rep. 5737, Institut National De Recherche en Informatique et en Automatique (2005)

Automatic Diagnosis of Melanoma Based on the 7-Point Checklist

Gabriella Fabbrocini, Valerio De Vita, Sara Cacciapuoti, Giuseppe Di Leo, Consolatina Liguori, Alfredo Paolillo, Antonio Pietrosanto and Paolo Sommella

Abstract An image based system implementing a well-known diagnostic method is disclosed for the automatic detection of melanomas as support to clinicians. The software procedure is able to recognize automatically the skin lesion within the digital image, measure morphological and chromatic parameters, carry out a suitable classification for detecting the dermoscopic structures provided by the 7-Point Checklist. Advanced techniques are introduced at different stages of the image processing pipeline, including the border detection, the extraction of low-level features and scoring of high order features.

Keywords Melanoma · Pigmented lesions · Dermoscopy · Image analysis · Semi-automatic diagnosis

Introduction

Malignant melanoma is currently one of the leading cancers among many fair skinned populations around the world. Change of recreational behavior together with the increase in ultraviolet radiation have caused dramatic increase in the number of melanomas diagnosed [1–3].

Currently, the possibility to increase the accuracy of diagnosis of melanoma is one of the most important tools to reduce the mortality rate of this tumor. The aim

G. Fabbrocini · V. De Vita · S. Cacciapuoti · G. Di Leo
Dermatology Research, University Federico II of Naples, Via S. Pansini 5, 80131 Naples, Italy

G. Di Leo · C. Liguori(✉) · A. Paolillo · A. Pietrosanto · P. Sommella(✉)
Department of Industrial Engineering (DIIn), University of Salerno, Via Giovanni Paolo II 132, 84084 Fisciano, Sa, Italy
e-mail: tliguori@unisa.it

P. Somella
e-mail: psommella@unisa.it

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_4,

of prevention campaigns is to increase public awareness of early warning signs. The possibility of new diagnostic methodologies and algorithms can contribute to perform an earlier diagnosis and to reduce the metastatic risk. Investigations have shown that the curability rate of thin melanomas (<1 mm) is between 91.8 and 98.1 % [4]. Epiluminescence microscopy (ELM) is an *in vivo*, non invasive technique that has disclosed a new dimension of the clinical morphologic features of pigmented skin lesions, using various incident light magnification systems with an oil immersion technique [2, 3]. Results of previous studies demonstrated that ELM improves accuracy in diagnosing pigmented skin lesions from 10 to 27 % when compared to clinical diagnosis by the naked eye [5].

Three diagnostic models with similar reliability have become more widely accepted by clinicians: 1. pattern analysis, which is based on the "expert" qualitative assessment of numerous individual ELM criteria; 2. the ABCD-rule of dermoscopy which is based on a semi-quantitative analysis of the following criteria: asymmetry (A), border (B), color (C) and different dermoscopic (D) structures; 3. the ELM 7-Point Checklist scoring diagnosis analysis, proposed by Argenziano et al. defining seven standard ELM criteria: Atypical pigment network, Blue-whitish veil, Atypical vascular pattern, Irregular streaks, Irregular pigmentation, Irregular dots/globules, Regression structures. The ELM 7-Point Checklist provides a simplification of standard pattern analysis and, if compared to ABCD, allows less experienced observers to achieve higher diagnostic accuracy values [6].

The 7-Point Checklist

The 7-point checklist is a diagnostic method that requires the identification of only 7 dermoscopic criteria to help clinicians to use dermoscopy. This simplified algorithm has been shown to be reproducible with non-expert dermatologists, who were able to classify a high percentage of melanomas [6].

In the original paper on the 7-point checklist, dermoscopic images of melanocytic skin lesions were studied to evaluate the incidence of 7 standard criteria. These features were selected for their frequent association with melanoma.

The 7 standard criteria are briefly defined in Table 1 along with the corresponding histological correlates and the scoring system (the 3 major criteria have a score of 2 points and the 4 minor criteria a score of 1 point). The differences between melanomas and nevi were evaluated by a univariate statistical test and the significant variables were used for stepwise logistic regression analysis to determine their diagnostic weights in the diagnosis of melanoma, as expressed by odds ratios. Using the odds ratios calculated with multivariate analysis, a score of 2 was given to the 3 criteria with odds ratios >5, termed "major" criteria, and a score of 1 to the 4 criteria with odds ratios <5, termed "minor" criteria. The total score for the lesion is obtained by simple addition of the individual scores for each detected criterion. In order to diagnose a melanoma, the identification of at least 2 melanoma-specific dermoscopic criteria is required, i.e. a minimum total score of 3 (1 major plus 1 minor or 3 minor

Table 1 Dermoscopic criteria and scores according to the 7 point checklist method

ELM criterion	Definition	Score
Atypical pigment network	Prominent (hyper-pigmented and broad) network with irregular shape or distribution	2
Blue-whitish veil	Irregular, confluent, gray-blue to whitish blue pigmentation not associated with red-blue lacunes or maple leaf pigmentation	2
Atypical vascular pattern	Linear, dotted or globular red structures irregularly distributed outside areas of regression	2
Irregular streaks	Radially and asymmetrically arranged linear or extensions at the edge of the lesion	1
Irregular pigmentation	Brown, gray and black areas of diffuse pigmentation with irregular shape or disruption and abrupt end	1
Irregular dots/globules	Black, brown or blue round structures irregularly distributed within the lesion	1
Regression structures	White areas, peppering, multiple blue-gray dots associated	1

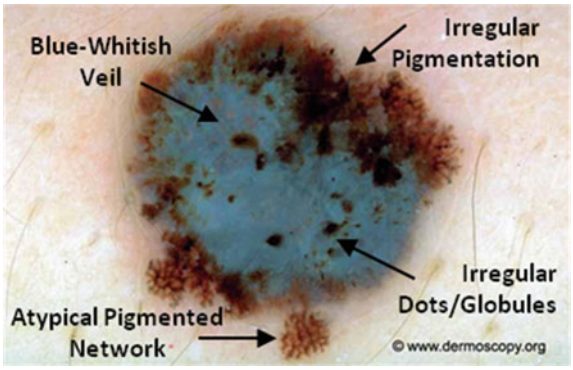


Fig. 1 Application of the 7-point checklist to the digital ELM image of a pigmented skin lesion: blue-whitish veil (score 2) + atypical pigment network (score 2) + irregular pigmentation (score 1) + irregular dots/globules (score 1) = total score 6. In order to diagnose a melanoma, the identification of at least 2 melanoma-specific dermoscopic criteria is required: in fact, a minimum total score of 3 is required (1 major plus 1 minor or 3 minor criteria). Histology confirmed the diagnosis of melanoma. Source of this figure: www.dermoscopy.org

criteria). Figure 1 shows a diagnosis example where different structures are present and scores corresponding to melanomas are computed.

Like most other diagnostic algorithms, the 7-point checklist was developed and validated retrospectively on a set of 342 dermoscopic images of histologically proven

melanocytic skin lesions (the sensitivity and specificity of melanoma detection were 95 and 75 %, respectively). Argenziano et al. showed that the ELM 7-point checklist, in the hands of experienced observers, gave the greatest sensitivity value (95 %), especially in the subgroup of early cutaneous melanoma. Compared with overall ELM diagnosis, the specificity was lower (75 vs. 90 %) because of the tendency to overclassify melanocytic nevi (especially the atypical types) as melanomas with the scoring diagnostic systems. A decrease in specificity may result in some increase in biopsy examinations of benign lesions, but the increase in sensitivity would decrease the chances of missing melanomas. The authors designed a model that requires the identification of only 7 standard ELM criteria, thus enabling even the less experienced clinician to use the method. In fact, this simplified scored pattern analysis was shown to be reproducible not only with a test set performed by experts but also by less experienced dermatologists, who were able to classify a high percentage of melanomas (85–93 %). The lower specificity values (45–48 %) obtained by the less experienced observers could be explained by the fact that most of the non-melanomas used to determine specificity were clinically atypical (leading to the decision to perform a biopsy); thus, they require more experience to perform correct assessments. However, use of the model would have avoided the excision of almost half of those lesions. For a cutaneous melanoma to be diagnosed, identification of at least 1 major and 1 minor ELM criterion (or 3 minor criteria) was required. This confirmed the previously reported rule that a single criterion usually does not suffice to make a diagnosis [6].

More recently, Haenssle et al. have assessed the sensitivity, specificity, and diagnostic accuracy of the 7-Point Checklist in the setting of a prospective long-term study. They have screened patients at increased melanoma risk at regular intervals by naked-eye examination, the dermoscopic 7-point checklist, and digital dermoscopy follow-up (10-year study interval). They have detected 127 melanomas including 50 melanomas in situ. The mean Breslow thickness of invasive melanomas has been 0.57 mm. A total of 79 melanomas were detected using the 7-point checklist melanoma threshold of 3 or more points (62 % sensitivity, compared with 78–95 % in retrospective settings). In all, 48 melanomas have scored fewer than 3 points and have been excised because of complementary information (e.g., lesional history, dynamic changes detected by digital dermoscopy). The specificity of the 7-point checklist has been 97 % (compared with 65–87 % in retrospective settings). Regression patterns, atypical vascular patterns, and radial streaming have been associated with the highest relative risk for melanoma (odds ratio 3.26, 95 % confidence interval 2.05–5.16; odds ratio 3.04, 95 % confidence interval 1.70–5.46; odds ratio 2.91, 95 % confidence interval 1.64–5.15; $P < 0.0003$, respectively). Melanomas thicker than 0.5 mm have exhibited significantly more regression patterns and atypical vascular patterns ($P < 0.02$). The malignant versus benign ratio for all excised lesions has been 1:8.6 (127 melanomas, 1092 non-melanomas). As a consequence, the 7-point checklist has appeared less sensitive but highly specific in this prospective clinical setting. Complementary information has clearly increased sensitivity. Thus, the authors have suggested that regression patterns or radial streaming in nevi of patients

at high risk should raise a higher melanoma suspicion than might be concluded from retrospective studies [7].

Further studies have proposed the use of simplified diagnostic algorithms, such as 3-point. Firstly Soyer et al. have evaluated the diagnostic performance of non-experts using a new 3-point checklist based on a simplified dermoscopic pattern analysis. Clinical and dermoscopic images of 231 clinically equivocal and histopathologically proven pigmented skin lesions were examined by 6 non-experts and 1 expert in dermoscopy. For each lesion the non-experts assessed 3 dermoscopic criteria (asymmetry, atypical network and blue-whitish veil) constituting the 3-point method. In addition, all examiners made an overall diagnosis by using standard pattern analysis of dermoscopy. Asymmetry, atypical network and blue-white structures were shown to be reproducible dermoscopic criteria, with a kappa value ranging from 0.52 to 0.55. When making the overall diagnosis, the expert had 89.6 % sensitivity for malignant lesions (tested on 68 melanomas and 9 pigmented basal cell carcinomas), compared to 69.7 % sensitivity achieved by the non-experts. Remarkably, the sensitivity of the non-experts using the 3-point checklist reached 96.3 %. The specificity of the expert using overall diagnosis was 94.2 % compared to 82.8 and 32.8 % achieved by the non-experts using overall diagnosis and 3-point checklist, respectively. These data showed that the 3-point checklist can be considered a valid and reproducible dermoscopic algorithm with high sensitivity for the diagnosis of melanoma in the hands of non-experts. Thus, the authors suggested that the 3-point checklist might be applied as a screening procedure for the early detection of melanoma [8]. Afterwards, Zalaudek et al. revaluated these preliminary results in a large number of observers independently from their expertise in dermoscopy. The three-point checklist showed good interobserver reproducibility (kappa value: 0.53). Sensitivity for skin cancer (melanoma and basal cell carcinoma) was 91.0 % and this value remained basically uninfluenced by the observers' professional profile. These results confirmed that the three-point checklist was a feasible, simple, accurate and reproducible skin cancer screening tool [9].

In 2010, Gerely et al. compared the sensitivity, specificity, and diagnostic accuracy of the seven-point and three-point checklist methods in the diagnosis of clinically atypical pigmented skin lesions and melanoma. The sensitivity, specificity, and positive and negative predictive values of the seven-point checklist method were 87.50, 16.17, 51.22, and 57.14 %, respectively. The sensitivity, specificity, and positive and negative predictive values of the three-point checklist method were 89.58, 31.25, 56.58, and 75 %, respectively. Thus, this study highlighted that the three-point checklist was observed to be a superior screening test. The seven-point checklist provided a more detailed analysis, especially for thin melanomas. In comparison with the seven-point method, the three-point method may be useful for less experienced observers when they need to obtain greater diagnostic accuracy [10].

However, it has been demonstrated that the 7-point checklist and the other abovecited diagnostic algorithms have actually a lower diagnostic accuracy if they are performed by inexperienced dermatologists [11]. Dermoscopy requires high training to optimize diagnosis of pigmented skin lesions. Indeed, Binder et al. demonstrated

that ELM pattern analysis increases the quality of diagnostic performance of ELM experts but decreases the performance of clinicians not specially trained in ELM [12].

To avoid these problems, to enhance the reproducibility of clinical diagnosis and to help clinicians with poor dermoscopic experience, computer-assisted analysis of dermoscopic images has been investigated. Computer-aided diagnosis can help clinicians in the diagnosis of skin lesions. The aim of these systems is to increase the specificity and the sensitivity in melanoma recognition and reduce unnecessary biopsies. Most of these automated systems are based on the afore-mentioned melanoma diagnosis methods. In general, image processing techniques are used to locate the lesions, extract image parameters describing the dermatological features of the lesions, and, based on these parameters, perform the diagnosis. Their potential benefits are very promising, but there are considerable difficulties involved in their development and their use in clinical practice. It has been widely stated that their accuracy can achieve the same range as dermoscopic diagnosis performed by experts or even that they can obtain better accuracy [13]. Computer-aided diagnosis, based on mathematical analysis of pigmented skin lesions, can be a tool to transform a qualitative evaluation into a quantitative one and to increase sensitivity of dermatologists with low dermoscopic experience. Suggest clinicians step by step the dermoscopic criteria arising from observation can mean to implement the dermoscopy use without reduce its diagnostic accuracy.

Automatic Diagnostic Systems

As previously introduced, there has been much research aimed at obtaining an improved and consistent differentiation between benign and malignant melanocytic skin lesions by means of digital dermoscopy analysis. Computerized dermoscopy image analysis, in fact, adds a quantitative evaluation to the "clinical eye observation" and can be used to improve biopsy decision-making [14].

Therefore, different groups have been developing diagnostic systems of recorded images (slides or digital cameras) techniques to assist clinicians in differentiating early melanoma from benign skin lesions [15].

For example, in [16] an automated melanoma recognition system is proposed taking into account 21 parameters extracted from images.

Schmid [17] proposed a color based segmentation scheme without extracting features, whereas a new procedure based on the Catmull-Rom spline method and the computation of the gray-level gradient of points extracted by interpolation of normal direction on spline points was employed in [18].

A computer algorithm for the diagnosis of melanocytic lesions based on the evaluation of 64 different analytical parameters is described in [19], whereas a software module which automatically evaluates the outline of a lesion providing 50 objective parameters subdivided into three categories (geometries, texture and coloured islands) is developed in [14].

Very interesting summaries of the main researches about the digital dermoscopy (in terms of acquisition, calibration, image datasets and processing methods) are reported in [20, 21]. As a results of these surveys, high accuracy may be achieved by computer aided diagnostic systems employing statistics obtained from low-level features and parameters. Nevertheless, it is not likely that the digital system will completely substitute the expert in dermoscopy.

The Proposed Framework

According to the author's opinion, the automated system should be in fact integrated by higher level features based on a particular diagnostic scheme in order to gain greater clinical acceptance.

More precisely, the software diagnostic system should be able to reproduce the expertise of a well-trained dermatologist and support the clinician in his/her visual inspection and diagnosis according to well-known dermoscopic methods. In detail, three different diagnostic models have become more widely accepted by clinicians for the interpretation of the features inspected by dermoscopy.

Starting from the previous considerations, the authors have tackled the problem of defining suitable image processing algorithms for the automatic implementation of the 7-Point Checklist.

A preliminary study about the image processing techniques for the extraction of the pigmented lesion (from healthy skin) and the detection of chromatic features was reported in [22].

Further studies [23, 24] have led to the introduction of a software framework [25] for the automatic detection of dermoscopic criteria. Following the example of the Computer Aided System architecture proposed for digital ELM images by Schmid [26] and the methodological approach to the classification suggested in [27], the software framework includes all the processing algorithms derived from the clinical knowledge gained by expert dermatologists (well-trained in the 7-Point Checklist application).

In [28] a statistical approach is introduced for the automatic detection of a minor criterion (Irregular dots/globules).

The present chapter reviews the main image processing techniques adopted to provide and improve the diagnostic capability of the automatic tool which implements the 7-Point Check-List. Basic software tasks such as feature identification and high-level classification are deeply investigated with respect to multiple dermoscopic structures. Finally, the experimental results are extended to a large set of pigmented lesion and a comparison among the different techniques is also carried out.

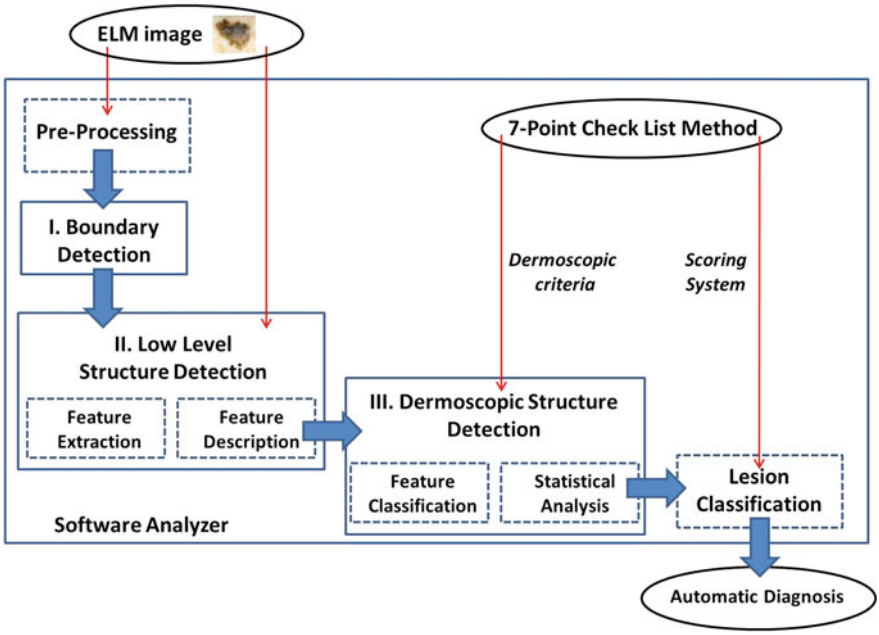


Fig. 2 Scheme of the software framework for the automatic diagnosis of ELM images

Methods

According to the scheme reported in Fig. 2, the software procedure developed for the automatic analysis and diagnosis of dermoscopic images is organized into three main detection stages.

After a preliminary processing stage designed to remove hair and/or artifacts, the *Boundary Detection* (I) allows the pigmented lesion to be extracted from the surrounding healthy skin. Then, the *Low-Level Structure Detection* (II) aims to identify and measure the main morphological and chromatic *features* throughout the lesion. Finally, at *Dermoscopic Structure Detection* (III), the feature classification and analysis are performed in order to detect each ELM criterion (high-level structure) provided by the 7-Point Checklist.

In the following subsections, for each stage of the automated procedure the remarkable literature is reviewed as well as the authors' novel approach is detailed in terms of advanced statistical techniques.

Pre-processing

Lesion segmentation in the presence of hair is usually doomed to failure. Because shaving the lesion area before the acquisition often interferes with the clinical practice, a computer-aided system for the analysis of dermoscopic images should always include an automated hair removal algorithm.

A well known hair removal algorithm was proposed in [29]: it identifies the image segments that approximate the structure of the hair, and then the regions that contain these segments are interpolated using the information of the surrounding pixels.

A similar approach is proposed by Schmid in [30], concerned with uniform color spaces, such as $L^*u^*v^*$ (the main advantage is that color difference can be measured and used for comparisons between pixels or distance measures in the spectral domain). According to the latter approach the morphological closing operator with a spherical structuring element is applied to the luminance component L^*, then the threshold operation is carried on the difference with the original image. More sophisticated techniques [31, 32] have also been introduced using image inpainting. Nevertheless, similar results are achievable. Thus, the algorithm disclosed in [30] has been preferred as preliminary stage in the proposed framework.

Boundary Detection

Boundary detection is a critical problem in ELM images because the transition between the lesion and the surrounding skin is smooth and hard to detect accurately, even for a trained dermatologist. Consequently, different approaches [33] have been developed for automatic detection of lesion border in both clinical and dermoscopy images.

Many studies have introduced techniques based on color clustering [17, 34–36]. For example, in [17] the first two principal components of the CIE $L^*u^*v^*$ color space are adopted to determine a 2D histogram. Then, initial cluster centers are calculated from the peaks using a perceptron classifier and, finally, the lesion image is segmented using a modified version of the fuzzy c-means (FCM) clustering algorithm. Other color clustering algorithms provide for median cut, k-means, FCM and mean-shift [37]. Further approaches investigating on digital lesion images include snakes methods based on gradient vector flow [38, 39], improved region-based active contour algorithms [40], morphological flooding [30] and optimized JESG segmentation [41]. Finally, the Histogram Thresholding represents a widely-adopted strategy, upon which the latest investigations have been focused by introducing color channel optimization, hybrid (i.e. combined global and local) thresholding [42], and/or fusion within Markov Random Field framework [43].

A very interesting comparison of the main proposed approaches is reported in [44], where a new algorithm based on Statistical Region Merging (SRM) is also

introduced. As results of this survey, two approached emerged as the most effective methods: SRM and DTEA (Dermatologist-like Tumor Extraction Algorithm, [45]).

The Statistical Region Merging is a recent technique [46] belonging to the region growing and merging group. The method models segmentation as an inference problem, in which the image is treated as an observed instance I of an unknown theoretical image I^*, whose statistical (true) regions are to be determined. This method is typically adopted for its simplicity, computational efficiency, and excellent performance without the use of quantization or color space transformations. Specifically, each pixel of the true image I^* can be modeled as a set of Q independent random variables whereas the statistical regions represent theoretical objects sharing a common homogeneity property:

– inside any statistical region the pixels have the same expectation for each color channel (for example *Red*, *Green* and *Blue*);
– the expectation of adjacent regions are different for at least one color channel.

Given the homogeneity property the ideal segmentation of the observed image I relies on the frontiers between the statistical regions which are connecting pixels with differences in their color expectation. Figure 3 depicts an example of color segmentation for the ELM image performed through the SRM: each region is displayed according to its mean RGB values (averaged on pixels constituting the region). The parameter Q allows to quantify the statistical complexity of I^*, the generality of the model and finally control the coarseness of the segmentation.

Thus, the lesion map resulting from SRM segmentation can be further investigated in order to detect the inner regions constituting the pigmented lesion to be contoured. According to the method suggested in [44], the background skin color is estimated as mean R, G and B colors of the pixels belonging to four patches (20×20 sized) from the corners of the image. Post-processing provides for the deletion of light-colored and bounding regions (including the regions whose mean color has an Euclidean distance less than 60 to the background skin color, the regions that touch the image frame and those with rectangular borders). The initial border detection result is obtained by removing the isolated regions and then merging the remaining regions. Finally a morphological dilation with a circular structuring element is applied to obtain the automatic border.

The DTEA algorithm is based on thresholding followed by iterative region growing. Following the same approach, the authors suggested in [25] a novel lesion border detection. The developed algorithm, referred to as Adaptive Thresholding, consists of three steps:

i. color to monochrome image conversion;
ii. image binarization using an adaptive threshold;
iii. border identification, based on a blob-finding algorithm.

In the first step, 3 different monochrome images are obtained from the source image (*RGB* standard color) corresponding to the red, green and blue planes. For each component (see Fig. 4a), two modes (classes) are typically evident in the pixel intensity histogram (as depicted in Fig. 4b) corresponding respectively to the pigmented

lesion (the image foreground) and the surrounding skin (the image background). Then, the algorithm introduced by Otsu [47] is adopted to select the optimum threshold S^* for each histogram, thus allowing the image background and foreground to be detected. The adaptive algorithm aims to minimize the intra-class variance σ_W:

$$\sigma^2{}_W(S) = P_0(S)\sigma^2{}_0(S) + P_1(S)\sigma^2{}_1(S) \tag{1}$$

defined as a weighted sum of variances σ_i of the two intensity classes C_i resulting from the S threshold:

$$P_0(S) = \sum_{k=1}^{S} \frac{f_k}{N} \quad P_1(S) = \sum_{k=S+1}^{L} \frac{f_k}{N} \tag{2}$$

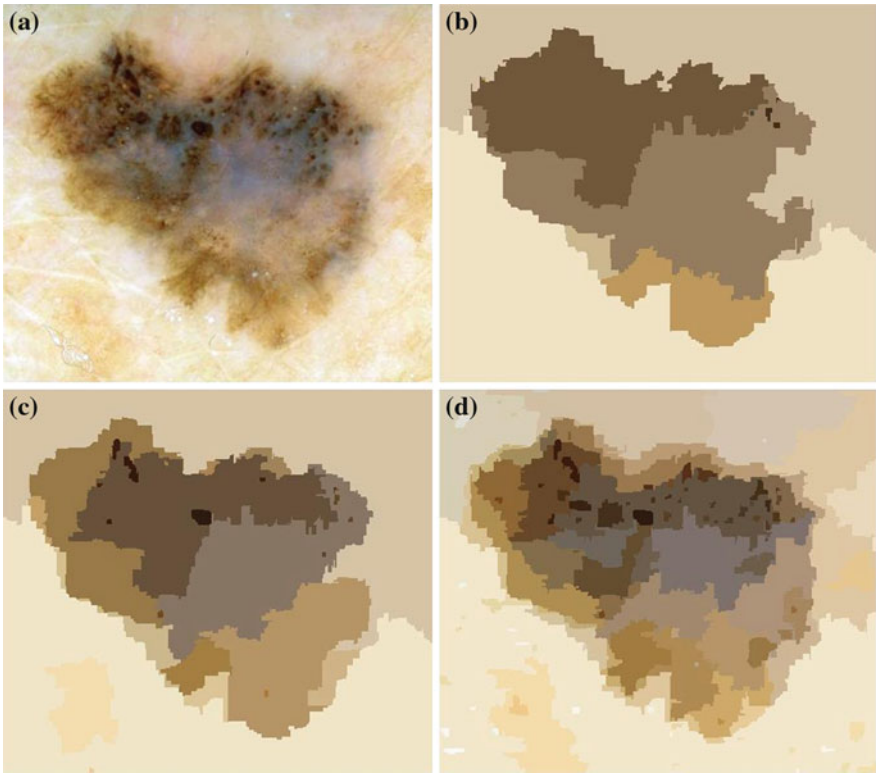


Fig. 3 Segmentation using statistical region merging: **a** ELM image; **b** results for Q = 32; **c** results for Q = 64; **d** results for Q = 256

where P_i is the probability distribution, N is the number of the image pixels, L is the number of histogram bins and f_k the number of pixels associated with k intensity value.

Otsu shows that minimizing the intra-class variance is the same as maximizing the between-class variance σ_B:

$$\sigma^2{}_B(S) = \sigma^2 - \sigma^2{}_W(S) = P_0(S)P_1(S)[\mu_0(S) - \mu_1(S)]^2 \qquad (3)$$

which is expressed in terms of class probabilities P_i and class means μ_i (with $i = 0, 1$).

The adoption of the Otsu's method to RGB color image leads to three histograms and potentially different thresholds values. Since the proposed approach has been

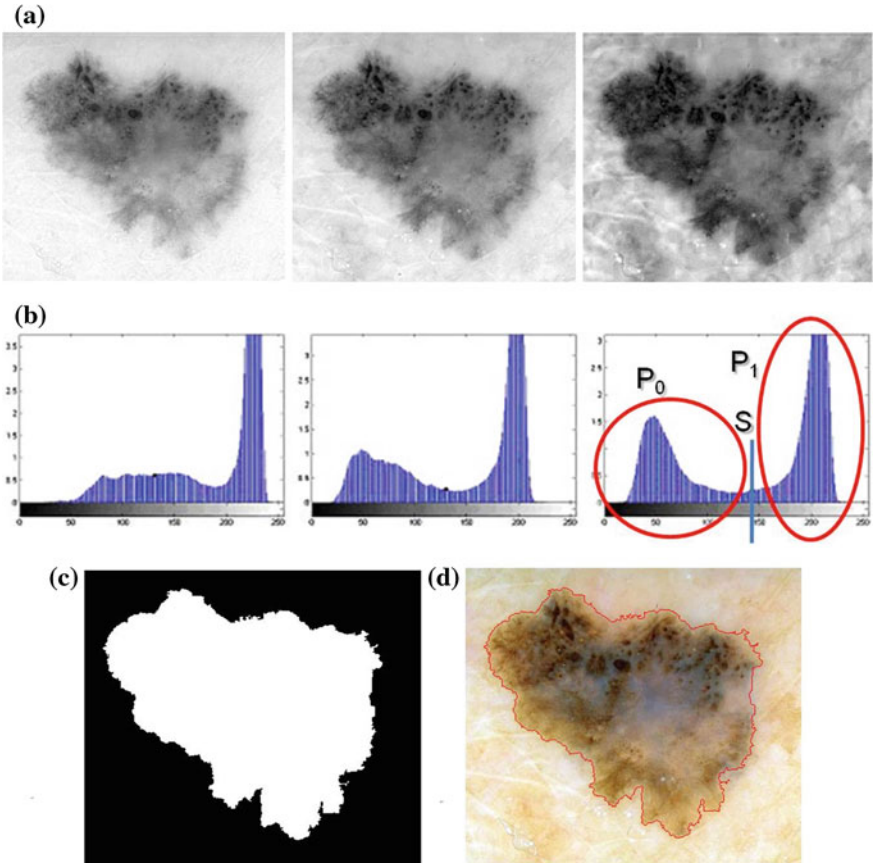


Fig. 4 Example of boundary detection: **a** image conversion (*Red*, *Green*, *Blue planes*); **b** intensity histogram; **c** binary mask; **d** lesion contour

experimentally revealed to be more sensitive to surrounding skin (the image background), the largest binary mask (the image foreground) is considered for next processing. An example of result is shown in Fig. 4c.

Finally, a simple blob-finding algorithm is adopted to extract the contour of the lesion from the binary mask. According to the modified version of *Moore's Neighbor Contour Tracing* proposed in [48], the tracking algorithm collects and sorts the contour lines (single pixel width) of the binary mask into an ordered list (the adopted algorithm also reveals to be computationally efficient by deleting the stopping criterion concerned with the start pixel). At this point, the border is superimposed on the color ELM image and displayed for visual inspection to the diagnostician (Fig. 4d).

Low Level Structure Detection

The dermoscopic criteria as defined by 7-Point Checklist Method are characterized both by chromatic and morphological low level structures (features). Thus, once the lesion is localized, *feature extraction* is performed by adopting suitable statistical techniques, which may be grouped into the following macro-categories:

- color segmentation
- texture analysis

Color Segmentation

Starting from the source image and the binary mask, the color segmentation stage is carried out with the aim of splitting the internal area into multiple chromatically homogenous regions (the lesion map).

To this aim, the SRM algorithm previously introduced may be adopted, by regulating the coarseness through a suitable choice of the Q parameter.

An alternative approach, proposed in [49] and investigated by the authors for dermoscopic images in [25], is represented by the Multi-Thresholding of the color image. In particular the following steps are proposed: (i) Principal Component Analysis (PCA); (ii) 2D histogram construction; (iii) peaks picking algorithm; (iv) histogram partitioning; (v) lesion partitioning.

(i) The Principal Component Analysis (also known as the discrete Karhunen-Loeve Transform or Hotelling Transform [50]) is a technique for reducing the dataset dimensionality while retaining those characteristics that contribute most to dataset variance. As for the application of the Principal Component Analysis to the ELM image, the RGB components of the pixels corresponding to lesion area (selected trough the binary mask obtained as final result of image segmentation) constitute the starting dataset (belonging to a state space with dimension $N = 3$). A new 3D representation of the lesion pixels can be obtained from the Hotelling Transform equation. An example is reported in Fig. 5a–c: the decreasing variability in each

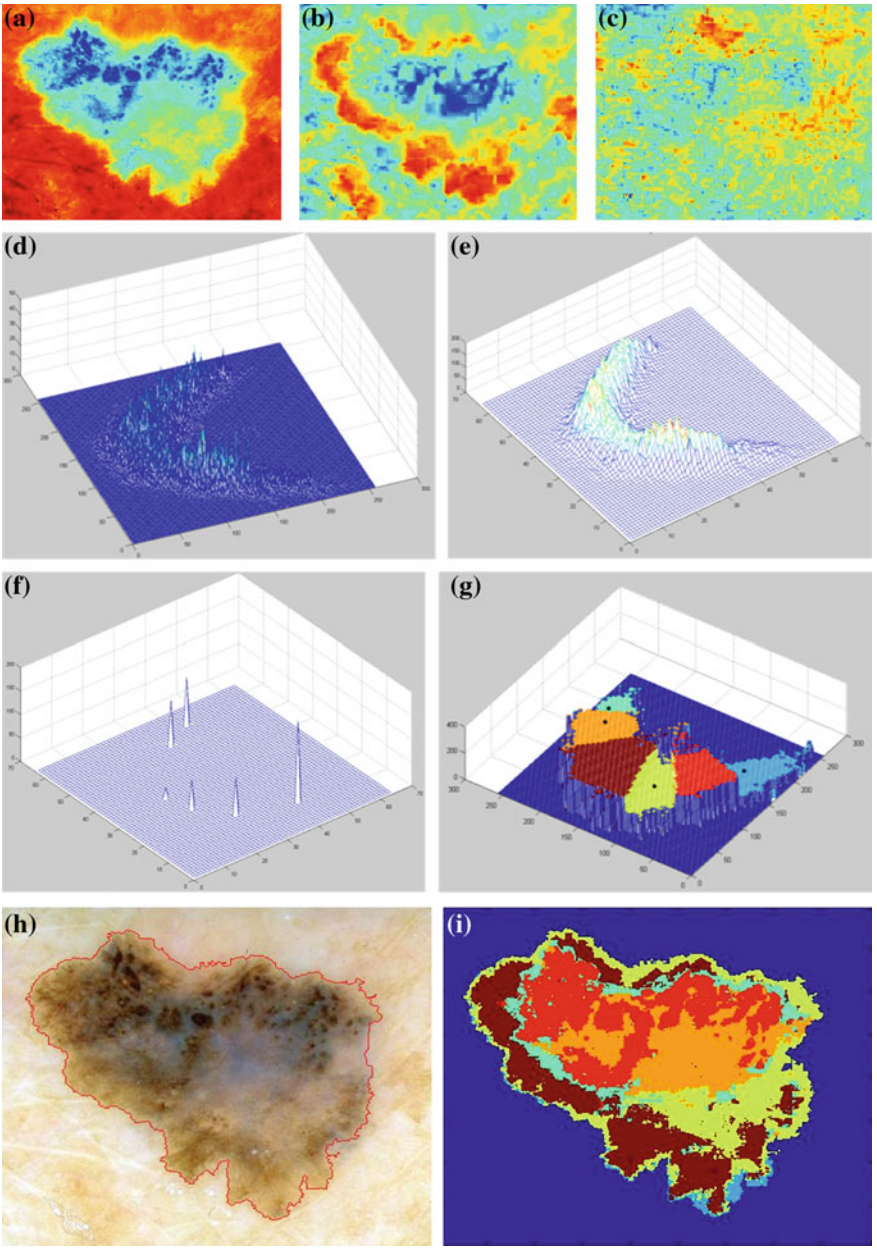


Fig. 5 Color segmentation based on Multi-thresholding: **a** 1st principal component; **b** 2nd principal component; **c** 3rd principal component; **d** joint histogram of the first 2 principle components; **e** down-sampled 2-D histogram; **f** result of peak-picking method; **g** partitioned 2-D histogram; **h** ELM image and lesion contour; **i** lesion map

individual band as long as the order of the principal component increases can be noted easily (for more evidence, PCA is related to all image pixels taking into account also the surrounding skin).

(ii) Since the low order components preserve sufficient information in order to obtain reliable information (Fig. 5a, b) whereas the third component contains most of the image noise (Fig. 5c), a joint histogram is created from the first 2 principle components (referring to which the multithresholding has to be carried out). An example of 2-D histogram is depicted in Fig. 5d. Because the estimated histograms are, in general, noisy due to the scarcity of data, it is advantageous to smooth and down-sample the histograms to eliminate noise effects. In particular the original histogram is reduced from size 256×256 to size 64×64 (see Fig. 5e).

(iii) The multithresholding is carried out by finding peaks in the 2-D histogram with significant mass around them. It is expected that these peaks will correspond to the cluster centroids in 2-D space and consequently will be well-representative of corresponding color regions (or segments) in the starting image. The knowledge of the number of segments is implicit in the peak search, and so is the maximum number K of peaks which have to be determined in the 2-D histogram. In our application the algorithm of Koonty [51] has been considered as peaks-picking method. As an example, in Fig. 5f the result of the peaks-picking algorithm is depicted with reference to the 2-D histogram shown in Fig. 5e when K equal to 10 is selected as maximum number of different color regions.

(iv) Once the peaks are identified, each corresponding hopefully to a segment, the other (non-peaks) histogram bins are attributed to the nearest dominant peak, constituting effectively their domains. Thus, a 2-D histogram is partitioned using its peak bins and an assignment rule (gravity force) which takes into account the strength (height) of the peak and the distance from the pick to the histogram bin under consideration. Figure 5g shows a partitioned 2-D histogram (after the partitioned 64×64 2-D histogram has been sampled back to its original size by simple replication of the bin labels by 4×4 fold): each color represents a histogram region.

Once the partitioned 2-D is computed, each pixel in the starting image (see Fig. 5h) can be directly labeled by taking into account the corresponding values for the two principal components. In particular assigning to each histogram region (or segment) an arbitrary intensity value, a gray-level image (or alternatively a false-color image) can be obtained where different regions are easily identified (see Fig. 5i).

Texture Extraction

As to the search for morphological (low-level) structures within the lesion, several approaches have been proposed in literature, including both structural and spectral methods [52]).

The structural techniques, which are intended to search for primitive structures such as points, lines and circles, have been extensively adopted for automatically detecting texture and/or local networks in dermoscopic images. For example, one of the most recent studies about the pigment network [53] introduces a feature extraction

based on the Laplacian of Gaussian (LOG) filtering. More in detail, the result of the edge detection step is a binary image which is subsequently converted into a graph to find the lesions meshes. Similarly, the detection of cyclic structures representing the pigment network is performed in [54] on the basis of the matching filtering principle and the adoption of suitable directional filters (namely 2-D Gabor filters).

The spectral technique is based on the Fourier analysis of the grey-level image. About computerized dermoscopic analysis, the approach is useful to determine the spatial period of the texture, thus allowing the identification of the regions where typically a network exists.

Thus, in order to disclose the pigment network (the dermoscopic criterion mainly correlated to low level morphological structures), a feature extraction combining structural and spectral methods has been introduced by the authors in [23]. With reference to the diagram in Fig. 6, the proposed algorithm is arranged into two processing paths, which share the input 8-bit grey-level image extracted from the ELM color image at first stage:

(i) The structural technique is proposed in order to identify the main local discontinuities within the image: the monochromatic image is first compared with its version obtained by a suitable median filter, then a close-opening operation is performed, which deletes eventual isolated points.

(ii) A sequence of Fast Fourier Transform (FFT), high-pass filtering, Inverse Fast Fourier Transform (IFFT) and suitable thresholding has been adopted. As goal of the spectral path, the local discontinuities which are not clearly associated to a network are disregarded. The result of this phase is a "regions with network" mask to be applied on the image yielded by the structural technique, in order to remove discontinuities which do not actually belong to the pigment network.

Finally the intermediate results from boundary detection ("lesion" mask), structural path ("local discontinuities" mask), and spectral path ("regions with network" mask) are combined according to the AND logic.

As a final result, a "network image" is achieved, where the areas constituting the pigment network are highlighted.

Feature Description

Following the automatic extraction of low level structures, the feature analysis is proposed in order to determine measurement information in terms of both chromatic and morphological descriptors.

Thus, the lesion map resulting from segmentation (performed through the SRM technique or the Multi-Thresholding algorithm) is investigated in order to identify the most significant descriptors of the local regions (*features*), such as components in the main color spaces in terms of mean value and standard deviation as well as relative difference among neighbors. An example is reported in Fig. 7, where the corresponding feature extraction resulted from a color fine segmentation ($Q = 256$).

More in detail, for each local region the pixel components in *RGB*, *HSI* and *Luv* color space are considered and the corresponding mean value and standard deviation

are computed. A percentage (30 %) dilation is also considered in order to compute further chromatic descriptors as the relative difference (mean value) with respect to neighbor regions. Moreover, the following morphological descriptors are computed for further analysis:

– relative dimension *A* %, defined as the number of the region pixels with respect to the lesion area;
– eccentricity *e* of the ellipse that has the same second-moments as the region; it is computed as the ratio of the distance between the foci of the ellipse and its major axis length with value between 0 and 1 (the degenerate cases corresponding respectively to a circle and a line segment).

About the feature extraction results concerning with the pigment network, some feature descriptors are computed from the statistical distribution of observed intensities in the *network image* at specified positions relative to each other. According to the number of intensity points (pixels) in each combination, statistics may be classified into first-order, second-order and higher-order statistics.

The Gray Level Co-occurrence Matrix (GLCM) method [55] is an extensively adopted way of extracting second order statistical texture features.

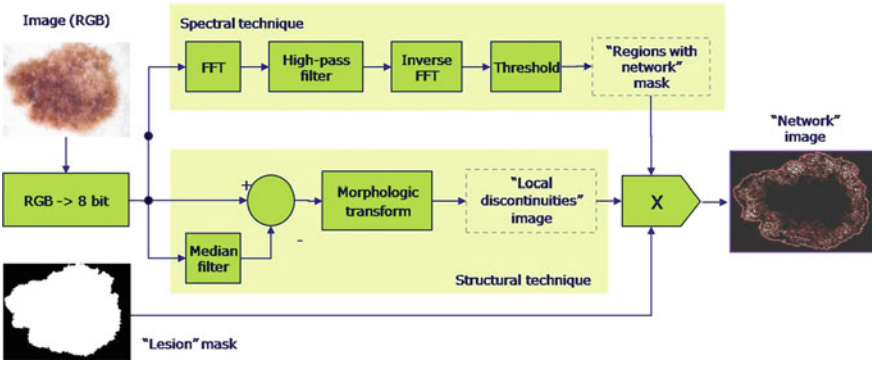
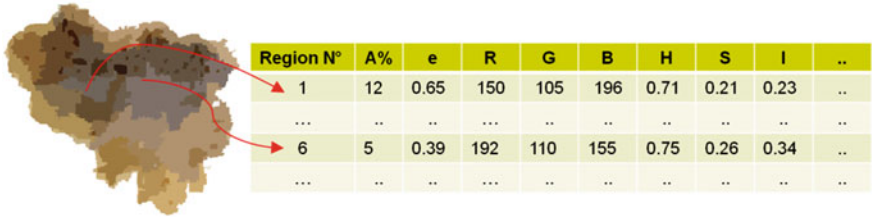


Fig. 6 Proposed scheme for texture analysis of pigmented lesions



Region N°	A%	e	R	G	B	H	S	I	..
1	12	0.65	150	105	196	0.71	0.21	0.23	..
...	..	..	...	..	..	..	..	..	..
6	5	0.39	192	110	155	0.75	0.26	0.34	..
...	..	..	...	..	..	..	..	..	..

Fig. 7 Example of fine color segmentation (SRM, Q = 256) feature description: segment area A % (percentage with reference to the lesion area); segment Eccentricity e, mean value of *Red* (R), *Green* (G), *Blue* (B), *Hue* (H), *Saturation* (S), *Intensity* (I) components (averaged on pixels constituting the segment)

Generally speaking, a GLCM is a matrix where the number of rows and columns is equal to the number of gray levels, G, in the image; each matrix element $P(i, j|d, \theta)$ contains the second order statistical probability values for changes between gray levels i and j at a particular displacement distance d and at a particular angle θ. A very interesting example of GLCM application to the computerized analysis of digital dermoscopic images is reported in [56], where 176 texture descriptors (on a total of 428 objective descriptors including color and asymmetry properties) were derived from 11 different-sized co-occurrence matrices with distance value d ranging from 1/2 to 1/64 of the length L of the major axis of the lesion. Texture descriptors mainly contributed to PCA-based classifiers able to effectively discriminate between melanomas and nevi as well as ridges and furrows.

Following this example, in order to avoid dependency of direction, one may calculate an average (isotropic) matrix out of four matrices ($\theta = 0°, 45°, 90°, 135°$), whereas the parameter d is suitable chosen according to the image resolution (as $d = L/32$). Finally, form the isotropic GLCM, a set of texture descriptors is computed, which includes entropy, inverse difference moment and correlation.

Dermoscopic Structure Detection

At this stage, each (high-level) dermoscopic structure provided by 7-Point Checklist is automatically disclosed within the lesion through suitable classification algorithm and/or statistical analysis, which take into account the features descriptors previously introduced.

Feature Classification

Most literature concerning with computerized dermoscopy has been focused on supervised learning as typical approach to classify discriminative features inspired from both ABCD rule and identification of specific patterns within the lesion. Supervised learning is, in fact, a general technique of estimating model parameters given a set of training examples. Thus, dermoscopic features are fed into a classifier and supervised learning is typically used to diagnose unseen images.

Although a general model using supervised learning and Maximum A Posteriori Probability (MAP) estimation has been recently proposed in [57] to perform common tasks in automated skin lesion diagnosis (also including border detection, artifact detection) with interesting and promising results, it is usually the case that a supervised learning is only performed in the final stage of feature classification.

Following this trend, the classification of the chromatic features (the lesion map) is straightforwardly viewed as a problem of data mining from feature descriptors. In this context a well-known class of solutions is represented by Decision Tree Classifiers, which belong to the Machine Learning techniques [58]. This type of classifiers has been firstly introduced in the computer-assisted analysis of ELM images by

Debeir et al. in [59], where Decision Tree were suitably learned and adopted both for skin-lesion segmentation and pigmented lesion classification (between five lesion patterns). Moreover, decision trees have been successfully adopted as pixel classification technique in [60] in order to automatically detect the blue-white veil areas in dermoscopic images.

A Decision Tree Classifier is a predictive model, trained (or induced) by adopting a suitable dataset with respect to which classification results are already available. More in depth, given a collection of objects (each one described by a set of attributes) a Decision Tree is a graph, wherein each internal node stands for an attribute, each arc toward a child node defines a property related to the parent node and finally a terminal node (or leaf) constitutes a classification result (a single value for the attribute adopted as class discriminator). The paths constituted by internal nodes with a parent-child relationship and the corresponding arcs define the rules of the predictive model that can be adopted for classifying new collections of objects. The Decision Tree Technique can be generally preferred to other solutions (also including Artificial Neural Networks and Support Vector Machines) because Decision Tree Classifiers are often fast to train and apply and generate easy to understand rules. Many induction algorithms have been proposed in literature, which are different for the type (discrete and/or continuous) of attributes they can apply to and the parameter adopted as performance index for the evaluation of the goodness of induction.

Probably the C4.5 algorithm [61] is the most widely adopted for decision tree induction. It can be related to attributes varying into both discrete and continuous range, whereas the information gain (relative entropy or Kullback-Leibler divergence) is considered as leading parameter in the splitting procedure (i.e. identification of a significant attribute and its corresponding optimal value to segment the collection into suitable groups). Moreover the C4.5 algorithm tries to prevent the over-fitting condition by implementing a pruning strategy. Given a large training set, in fact, decision tree classifiers could produce rules that perform well on the training data but do not generalize well to unseen data. In particular the C4.5 is able to identify sub-trees that do not contribute significantly to predictive accuracy and replacing each by a leaf.

Another popular method for classification is instead linear logistic regression. For example, in [54, 62], the SimpleLogistic classifier is proposed to perform the automatic detection of pigment network and irregular streaks respectively. Generally speaking, logistic regression tries to fit a simple (linear) model to the data through a process which typically reveals quite stable, resulting in low variance but potentially high bias. The tree induction exhibits low bias but often high variance because searches a less restricted space of models, allowing it to capture nonlinear patterns in the data, but making it less stable and prone to over-fitting. Consequently a promising way explored by the authors for performing the classification tasks is a combination of a tree structure and logistic regression models resulting in a single tree according to the model proposed in [63].

Thus, the Logistic Model Tree (LMT) has been proposed for classifying the chromatic features (the lesion map resulting from the color segmentation) on the basis of the corresponding descriptors, as detailed further in the text.

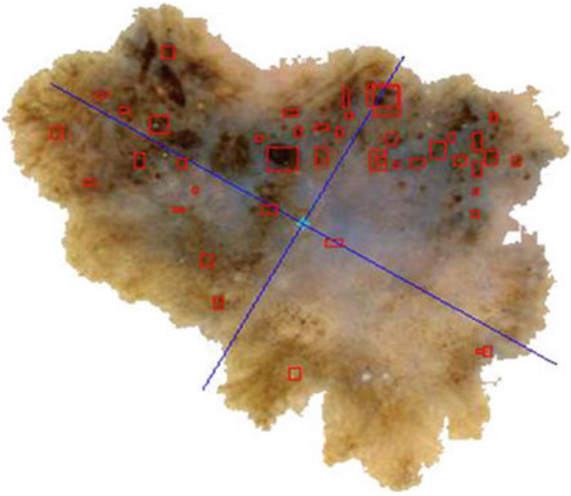


Fig. 8 Result of feature extraction: detection of round items within the lesion area

Statistical Analysis

About the classification of morphological features, a statistical approach based on the Test Hypothesis is proposed in order to verify the irregular distribution of the dermoscopic structures of interest. For better explaining the underlying idea, the method description refers to the example of feature extraction reported in Fig. 8. The symmetry axes (blue lines) of the lesion are computed as the major and minor axis of the ellipse characterized by the same normalized second central moments as the region of interest. Moreover, the main round items highlighted as red boxes are the chromatic and/or morphological features resulting from color segmentation and texture extraction (more details are given further in the text). They could correspond to texture elements of the lesion network and/or isolated dots and globules. Thus, the candidate features could be associated to irregular dermoscopic structures within the lesion (and classified as irregular) if their spatial distribution is not uniform. In the opposite case, i.e if the observed (spatial) round items were randomly scattered within the lesion, the number of elements in each of 4 quadrants (as resulted from the drawing of the main lesion axes) could be modeled according to the Binomial Distribution.

Therefore, a Binomial Test can be performed to estimate the casual distribution of N round objects, once the accepted risk α of Type I Error is fixed. According to the proposed approach, if the paucity or plenty of objects is observed in any quadrant and/or couple of quadrants, the Null Hypothesis (i.e. the spatial symmetry of round items) is refused and the morphological structured are classified as irregular.

The approaches previously introduced have been adopted to perform the automatic detection of dermoscopic high-level structures (criteria) provided by the 7-Point Checklist method.

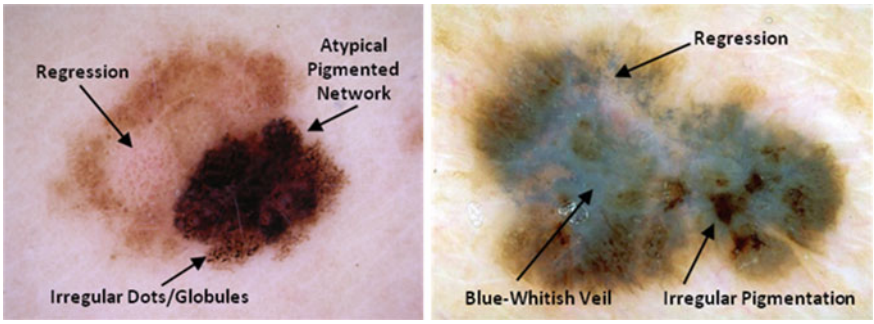


Fig. 9 Detection of dermoscopic criteria according to 7-point checklist

Examples of the diagnosis (by expert dermatologist) about two pigmented lesions are shown in Fig. 9, where the dermoscopic structures are highlighted.

More in detail, Fig. 9a shows a melanoma (total diagnostic score equal to 4) where a major criterion (Atypical Pigment Network) and two minor criteria (Regression and Irregular Dots/Globules) are detected. Similarly, within the melanoma displayed in Fig. 9b a major dermoscopic criterion (Blue-whitish Veil) and two minor structures (Regression and Irregular Pigmentation) are highlighted.

Hereinafter, two performance indexes are considered to estimate the accuracy of each classification algorithm:

- sensitivity, defined as the ratio of correct detection of the high-level structure analyzed and total number of cases where the dermoscopic criterion is present;
- specificity, defined as the ratio of correct decision about the high-level structure and total number of cases where the dermoscopic criterion of interest is absent.

The performance indexes range from 0.0 to 1.0 with the ideal classifier characterized by sensitivity and specificity both equal to the maximum value.

Blue-Whitish Veil, Irregular Pigmentation and Regression

The approach based on the Logistic Model Tree is adopted for the automatic detection of the dermoscopic structures which are more closely dependent on chromatic features. The model can be suitably computed for classify the regions constituting the lesion map which results from the color segmentation.

As already mentioned, for each region the components of the corresponding pixels in the *RGB*, *HSI* (Hue, Saturation and Intensity) and CIE *Luv* color spaces have been considered to compute mean value and standard deviation as feature descriptors (vector x). In addition the area percentage of each region with respect to total area of the lesion is taken into account.

An example of Logistic Model Tree as obtained by training is reported in Fig. 8 with reference to the detection of the Blue-whitish Veil.

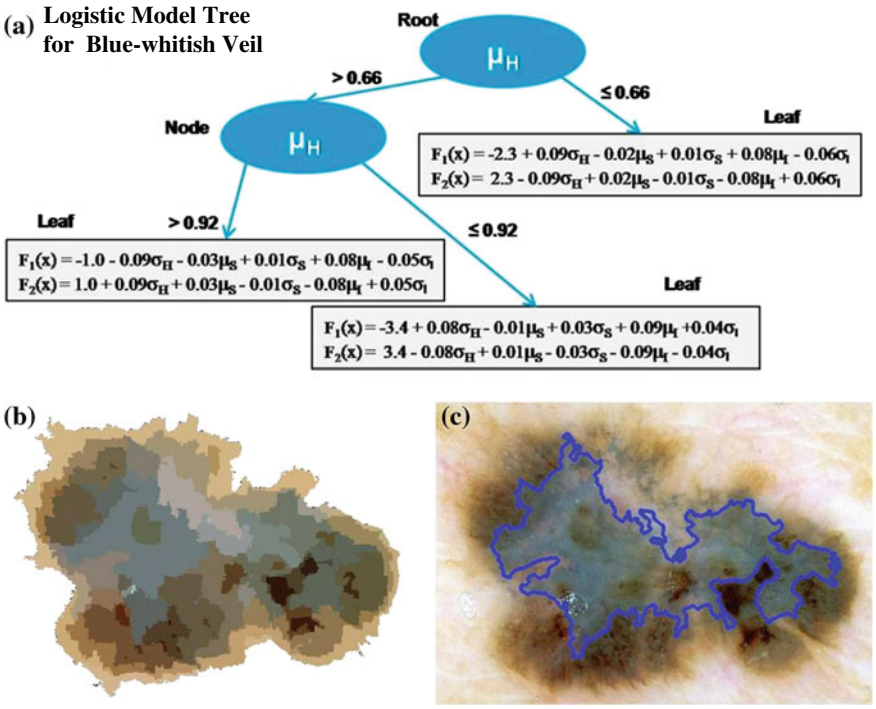


Fig. 10 Detection of the blue-whitish veil: **a** logistic model tree; **b** lesion map (feature extraction); **c** regions classified as blue-whitish veil

As you can see in the scheme (Fig. 10a), three different logistic regression models are computed on the basis of three ranges for the Hue mean value of the region to be analyzed (which can be interpreted as corresponding to blue, red or polychromatic "path"). The regression functions $F_i(x)$ (with $i = 1$, Blue-Veil region and $i = 2$, no Blue-Veil region) take into account the standard deviation for the Hue component, the mean and standard deviation for Saturation and Intensity components, in order to determine the probability that the chromatic features (color regions displayed in Fig. 10b) belong to an area characterized by the Blue Veil (the resulting detection of the criterion is depicted in Fig. 10c).

Analogous LMT models are computed (with reference to the same feature vector x) and adopted to classify the color regions (not detected as Blue-whitish Veil) as area of either Regression or Irregular Pigmentation.

About the automatic detection of regression structures, two different logistic models (see Fig. 11a) have been computed according to the range wherein the mean value for the Saturation component of the region segment falls. Just five feature descriptors are truly significant to determine the probability that the color region belongs to an area characterized by Regression (an example of the resulting detection is also reported in Fig. 11b–d).

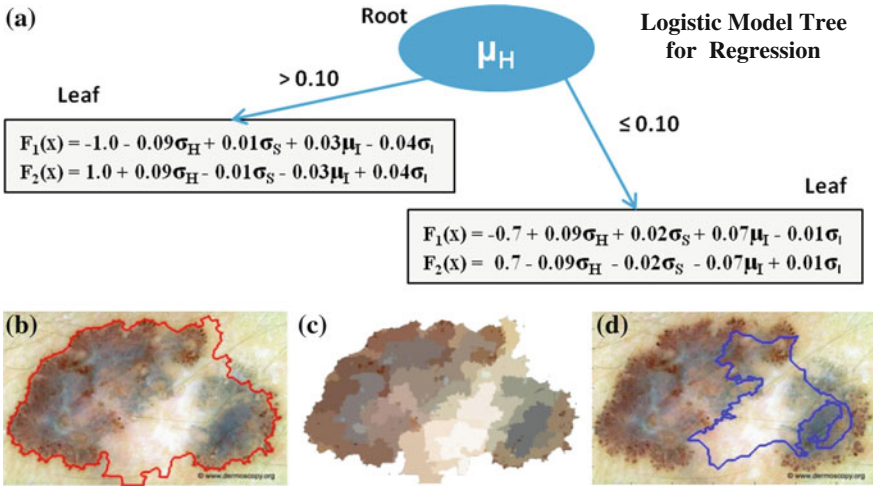


Fig. 11 **a** LMT for classification with respect to regression; **b** ELM image and contour; **c** lesion map; **d** detection results

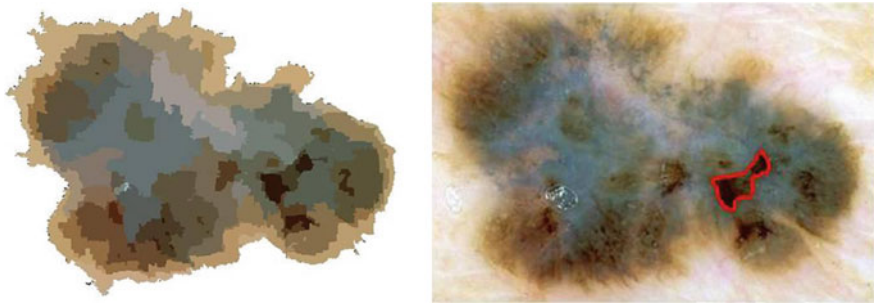


Fig. 12 LMT for automatic detection of irregular pigmentation

About the Irregular Pigmentation, a very simple LMT has been obtained: it computes the class probabilities taking into account the Intensity and L components (mean and standard deviation), and the area percentage measured for each color segment of the image (see Fig. 12).

$$F_1(x) = -0.5 - 0.09\mu_L + 0.07S_{\%} + 0.04\mu_I - 0.02\sigma_i \qquad F_2(x) = -F_1(x)$$

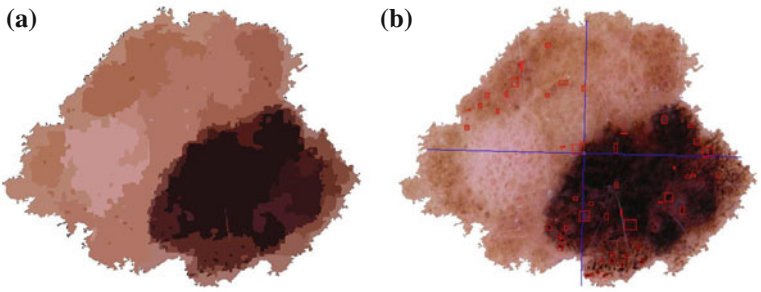


Fig. 13 Detection of irregular dots and globules: **a** color segmentation (SRM); **b** feature extraction: rounded items for hypothesis test

Irregular Dots and Globules

In order to detect the small dark areas of interest, a fine level of color segmentation is required which can be achieved by considering the Statistical Region Merging for high value of Q. As you can easily note in the example reported in Fig. 13a, the darkest segments may be deeply investigated to seek for the structures which represents Irregular Dots and Globules.

A statistical analysis based on the histogram of the SRM image is adopted by considering and ordering the statistical regions with respect to the increasing value of Intensity value (within a suitable range for Hue component). Moreover, the morphological feature descriptors previously introduced (percentage area $A\%$ and eccentricity e) are also compared with corresponding thresholds derived by experimental testing and tuned in order to extract rounded items inside the lesion.

Once the feature identification and analysis is completed with the detection of N round objects (see the items lightened with respect to the main symmetry axes of the lesion in Fig. 13b), according to the statistical approach previously introduced, the casual distribution is considered as Null Hypothesis of a Binomial Test. The following thresholds $k_{1,min}$, $k_{1,MAX}$ and k_2 can be jointly adopted for estimating the irregularity of Dots and Globules:

$$\sum_{k=0}^{k_{1,\min}} \binom{N}{K} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{N-k} + \sum_{k=k_{1,MAX}}^{N} \binom{N}{K} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{N-k} \leq \alpha \quad (4)$$

$$2\sum_{k=0}^{k_2} \binom{N}{K} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{N-k} \leq \alpha \quad (5)$$

where α is the accepted risk of Type I Error.

When the round items observed in each quadrant exceed the calculated thresholds, the Null Hypothesis is refused and the corresponding Dots and Globules are classified as Irregular.

Atypical Pigment Network

The approach based on the Hypothesis Test is also adopted for the classification (in terms of spatial irregularity) of the results from the feature extraction.

As introduced in the previous section, the pigment network within the lesion of interest is detected suitably combining the texture extraction and the color segmentation. Specifically, the areas constituting the network (white objects in Fig. 14a) are matched with the darkest regions of the lesion map computed through the Statistical Region Merging (see Fig. 14b). Then, the *N* objects resulting from the coupled analy-

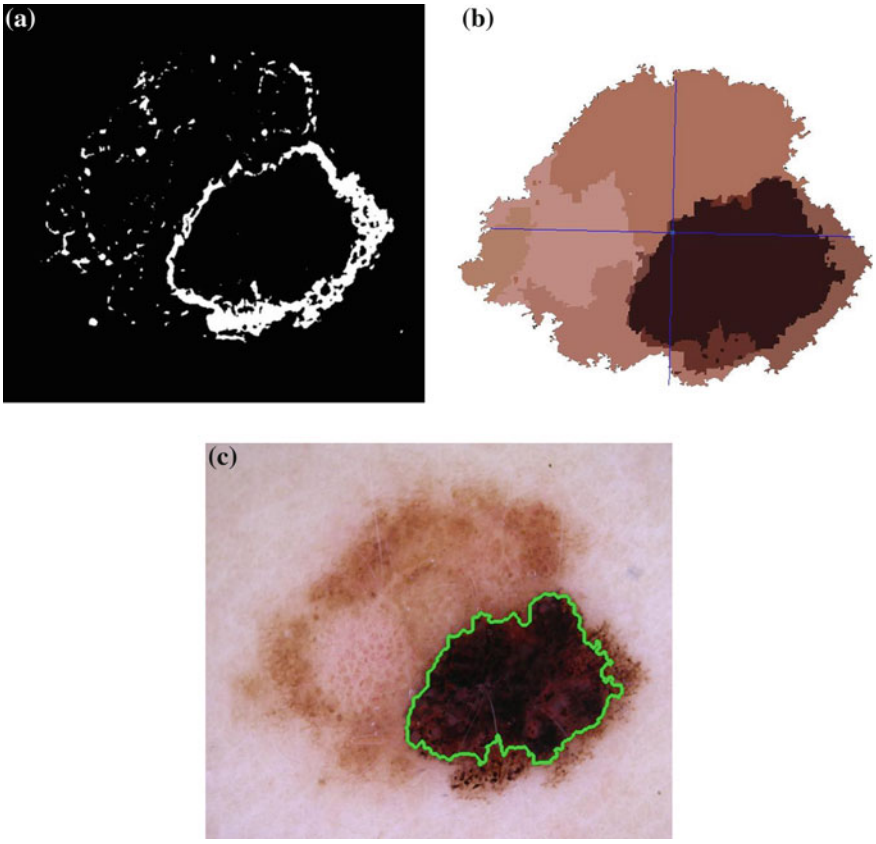


Fig. 14 Detection of atypical pigment network: **a** texture extraction; **b** color segmentation; **c** classification results

sis are classified as irregularly distributed ("atypical") by performing the Hypothesis Test according to the Eqs. (4) and (5). In the example reported in Fig. 14c, the major criterion is detected in the darkest right-bottom area where the pigment network is mainly distributed.

Irregular Streaks

Although the presence of Irregular streaks is highly suggestive for malignancy of a lesion, the modeling, detection and analysis of streak lines and starburst pattern have rarely been used. A summary of the previous studies is reported in [62], where an original graph-based approach and very interesting feature set are also proposed. Nevertheless, the main hypothesis (the lesion modeling as an ellipse) leads to reduced detection accuracy when the algorithm is applied to generalized image set. A shape independent approach is proposed by the authors combining structural technique and color segmentation. The presence of asymmetrically arranged (linear or bulbous) extensions at the edge of the lesion can be detected by searching for the simultaneous occurrence of two different structures:

(i) brown pigmentation localized in the same restricted region, and
(ii) finger-like track of the contour of the lesion.

Both structures are detected by mean a local analysis of the lesion contour, this last is split into 10 equally length segments. A color segmentation of the region of interest is performed through the Statistical Region Merging in order to seek for the black/brown dermoscopic structures. Then a morphological irregularity index is computed and compared with a suitable threshold.

The irregularity index is defined as ratio of number of pixels constituting the lesion contour and the shortest path", where the "shortest path" pixels are the points belonging to the line that connects the farther contour points in the region (see Fig. 15).

Atypical Vascular Pattern

About the automatic detection of Atypical Vascular Pattern, an approach combining color segmentation and structural analysis is proposed similarly to the methodology concerned with the Irregular Dots/Globules. Again the inner area of the lesion (see Fig. 16a) is considered and segmented through the Statistical Region Merging at fine level ($Q = 256$).

The resulting SRM image (see Fig. 16b) is firstly matched with the texture descriptors (entropy, inverse difference moment and correlation) based on the gray level co-occurrence matrix in order to exclude texture areas.

Then, a statistical analysis of the candidate SRM segments is performed by comparing the corresponding Hue range and eccentricity with suitable thresholds (experimentally tuned through a supervised learning approach) in order to detect linear or globular red structures irregularly distributed within the lesion. An example of linear vascular pattern is reported in Fig. 16c.

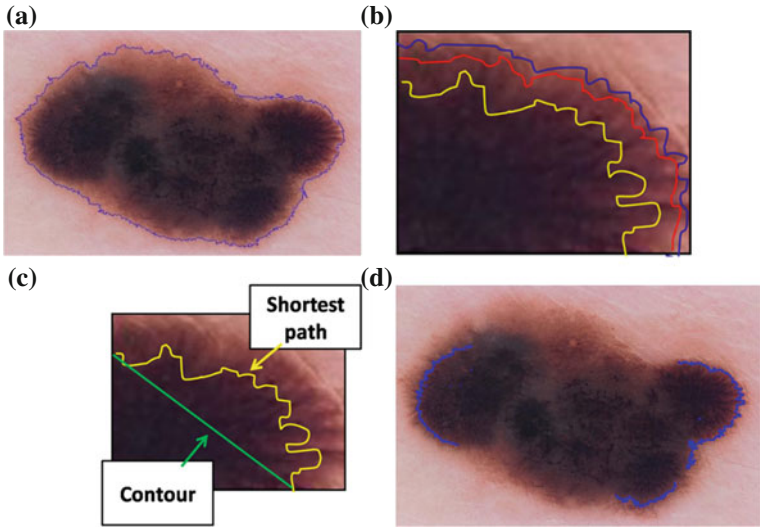


Fig. 15 Detection of irregular streaks: **a** ELM image and lesion contour; **b** the lesion area next to the border segment (*blue line*) is investigated (colour segmentation based on SRM) searching for darkest regions; **c** finger-like structures are detected through a quantitative comparison between the edge (*yellow line*) of the brown pigmentation and the corresponding straight line (*green line*); **d** results of automatic detection

Experimental Results

In order to develop and test the automatic procedure for the diagnosis of pigmented skin lesions, images of benign and malignant lesions were collected and stored in a database.

200 cases were extracted from a dermoscopy atlas [64] and observed by epiluminescence microscopy by two different dermatologists (M.S., G.F.) to evaluate the grade of accuracy in the management of 7 Point Checklist algorithm. Moreover, three dermatologists specifically trained in dermoscopy were asked to assess digital images of 100 melanocitic skin lesions selected among a digital collection of lesions screened between 2010 and 2012 at the *Department of Dermatology of the University of Naples Federico II*. In this department, the imaging is performed by a digital camera (Canon Power-Shot G9 with Heine Dermaphot Optics) that is combined with an epiluminescence microscope in order to produce digitized ELM images of skin lesions. The observers were first asked to use pattern analysis to score each lesion as naevus, melanoma or lesion to be excised, using the individual criteria listed in the seven-point checklist. For each image, the corresponding clinical and/or histological analyses were available. The images were extracted in order to obtain a quite homogeneous diagnosis distribution of the cases with respect to the criteria of interest.

As consequence, the overall database refers both to cutaneous melanomas and melanocytic nevi (also including Clark, Spitz, Reed nevi). About the image qual-

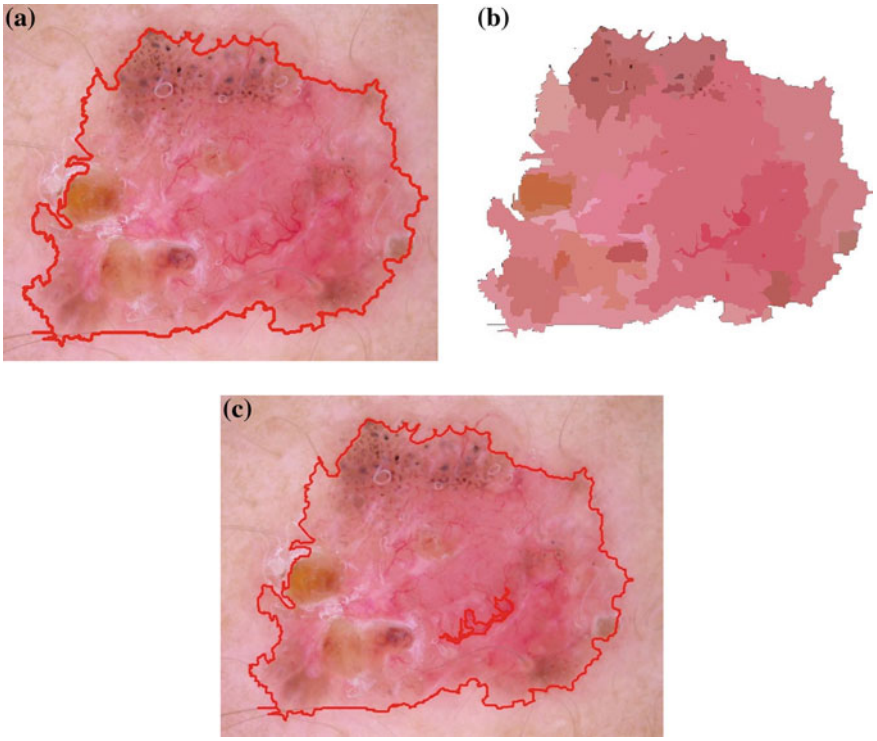


Fig. 16 **a** ELM image and lesion contour; **b** SRM segmentation of the lesion area; **c** automatic detection of atypical vascular pattern

ity, all the pictures are 24-bit RGB color images in JPEG format with dimensions ranging from 700 × 447 to 2272 × 1520 pixels. The lesions are imaged completely with healthy skin visible at margins. As previously mentioned, the image pre-processing strategy based on mathematical morphology [30] has been adopted for artifact removal.

Border Detection

The proposed technique based on the Adaptive Thresholding has been compared with the unsupervised approach based on the SRM algorithm, which was revealed the most effective method [44] for contour detection in dermoscopy images of pigmented skin lesions.

Comparison has taken into account 120 dermoscopy images (60 invasive malignant melanoma and 60 benign) randomly selected from the starting dataset.

As a ground truth for the evaluation of the border detection error, a manual border was obtained by selecting a number of points on the lesion border, connecting these points by a second-order B-spline and filling the resulting closed curve. More in detail, three dermatologists were asked to select the points on the lesion border, then the corresponding binary images were suitable combined. A majority policy is taken into account: only the image pixels resulted as inner points of the lesion by at least two dermatologist are considered as white-value pixels of the ground truth binary image (*Ref_Binary*). Finally the tracing contour algorithm [48] is applied to determine the ground truth manual border.

Using the dermatologist-determined borders, the automatic borders resulting from the Adaptive Thresholding and SRM have been compared using the metric suggested in [65]. Here, the percentage border error is given by:

$$Border\ Error = \frac{(Automatic\_Binary)XOR(Ref\_Binary)}{Area(Ref\_Binary)} 100\,\% \qquad (6)$$

where *Automatic_Binary* is the binary image obtained by filling the computer detected border, the exclusive-*OR* operation gives the pixels for which the *Automatic_Binary* and *Ref_Binary* disagree, and *Area*(*I*) denotes the number of pixels in the binary image *I*.

Table 2 shows the mean and standard deviation border error for the automated methods considered. Although the error rates increase in the melanoma group (due to the presence of higher border irregularity and color variegation in these lesions), the proposed approach has achieved the best results (lowest error values) in terms of both accuracy (mean) and consistency (standard deviation).

An example of automatic contour extraction for a melanoma is reported in Fig. 17a, where the resulting borders are compared with the manual border.

As you can see in the reported details (Fig. 17b, c), the automatic border resulting from the Adaptive Thresholding is able to better match the manual border than the result from SRM. The threshold which takes into account the image as whole is, in

Table 2 Dermoscopic criteria and scores according to the 7-Point Checklist method

Segmentation technique	Benignant lesions		Melanoma		All lesions	
	Mean (%)	Standard deviation (%)	Mean (%)	Standard deviation (%)	Mean (%)	Standard deviation (%)
Adaptive statistical thresholding	7.4	2.3	10.0	6.2	8.7	4.8
Unsupervised approach (SRM)	8.5	3.3	13.1	8.7	10.8	6.9

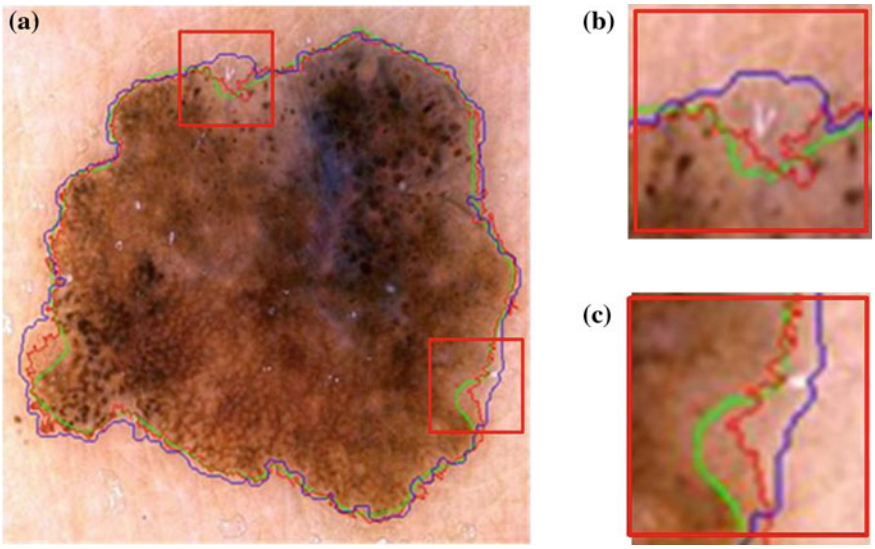


Fig. 17 Comparison between automated procedures for border detection: ground truth (*green line*), adaptive thresholding (*red line*), unsupervised approach (*blue line*)

fact, able to separate lesion and surrounding healthy skin also in critical local regions where the pixels components in RGB space are statistically close.

Automatic Detection of Dermoscopic Structures

As preliminary step a Training and Test Set have been suitably selected from the reference database for each dermoscopic criterion.

As guideline, the Training and Test Set have to share the same case distribution with respect to the criterion of interest. For example, 150 digital images have been adopted to develop the automatic detection of Irregular Dots/Globules, whereas the remaining 137 images have been adopted as Test Set to verify the software procedure.

As result, the Training and Test Set include respectively 45 and 39 skin lesions characterized by dermoscopic structures of interest.

About the color segmentation, a comparison has been carried out between the Statistical Region Merging (controlling the coarseness by varying Q from 32 to 256) and the Multi-Threshold approach based on Principal Component Analysis and 2D-histogram.

The classification results from physicians have been taken into account: 3 expert dermatologists were asked to inspect the results (lesion map) from the color segmentation in order to set the classification attribute for the features (local regions) of each image belonging to the Training Sets.

About the detection of Blue-whitish Veil, Irregular Pigmentation and Regression, multiple Logistic Model Trees (correspondingly to the different color segmentation) have been induced from the Training Set and verified (in terms of classification performance) with respect to the Test Set. Moreover, on the basis of experts' observations concerned with the Training Set, suitable thresholds (about the minimum detection area) have been derived to aggregate the *per-feature labeling* into *per-image classification accuracies*.

About the detection of Atypical Pigment Network Irregular Dots/Globules, suitable thresholds have been determined from the image properties of the Training Set through ROC curves [66] for the quantities introduced in the feature extraction stage (maximum region dimension $A\,\%$, and eccentricity e, range for I component) as well as the classification (minimum number N_0 of round items to perform the statistical test and the risk α).

As an example, the verification of the proposed approach about Irregular Dots/Globules with respect to the previously introduced Test Set has resulted in 35 skin lesions correctly scored (with respect to 39 cases where the minor criterion was present).

Moreover, the classifier lead to 15 false detections (automatic score = 1). Table 3 summarizes the corresponding per-image results (both for Training and Test Set) in terms of sensitivity and specificity. The overtraining has been avoided: similar performance of the classifier have been achieved for the two Image Sets.

Analogous results have been achieved for the detection of all dermoscopic criteria of interest.

Table 4 summarizes the performance indexes of diagnostic algorithms with reference to the images including into the corresponding Test Sets.

Goal of the comparison has been the evaluation of the color segmentation approach (between the SRM technique and Multi-threshold) that better allows to highlight the chromatic and morphological features, on which the classification of the dermoscopic criteria are based. Thus, the per-image performance corresponding to the best segmentation technique are reported in bold.

As you can easily note, the Statistical Region Merging has been revealed as the preferred solution for color segmentation.

A quite coarse segmentation ($Q = 64$) is able to disclose the areas characterized by Atypical Pigment Network, Irregular Pigmentation and Regression. Namely a satisfactory sensitivity (not inferior than 0.80) is achieved without downgrading the specificity (which has to be in special account for the minor criteria).

Table 3 Irregular dots/globules: classifier performance (SRM, Q = 256)

Image set	Sensitivity (%)	Specificity (%)
Training set	80	82
Test set	90	85

Table 4 Automatic diagnosis of pigmented lesions: comparison among per-image classification results (test sets)

ELM criterion	Multi-threshold		SRM ($Q = 64$)		SRM ($Q = 128$)		SRM ($Q = 256$)	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Atypical pigment network	80	82	**82**	**85**	85	78	87	74
Blue-whitish veil	**90**	**93**	66	92	73	90	84	85
Irregular pigmentation	84	77	**85**	**81**	93	66	97	63
Irregular dots/globules	–	–	–	–	87	88	**90**	**85**
Regression structures	72	75	**80**	**83**	83	70	85	67

Table 5 Classifier performance for atypical vascular pattern and irregular streaks

ELM criterion	Sensitivity (%)	Specificity (%)
Atypical vascular pattern	80	78
Irregular streaks	86	88

The finest segmentation ($Q = 256$) has to be preferred (in terms of sensitivity) for the detection of the reduced-size objects (otherwise not revealed) which can be classified as Irregular Dots/Globules.

Finally, the Multi-Threshold approach is the segmentation technique able to better identify the large areas of the lesions characterized by Blue-whitish Veil.

Finally, Table 5 summarizes the classification performance in corresponding Test Sets of the proposed approach for the automatic detection of Atypical Vascular Pattern and Irregular Streaks.

Discussion and Future Work

On the basis of advanced techniques of image processing a Computer-Aided System has been achieved for the analysis of digital dermoscopic images according to the 7-Point Checklist method. The software routines are suitable to carry out: (i) the detection of the lesion contour; (ii) the extraction and measurement of the main chromatic and morphological features within the pigmented lesion; (iii) the classification and scoring of the dermoscopic structures.

The automatic procedures have been be tested with respect to a quite extensive metrological characterization (performance of each classifier estimated in terms of the sensitivity and specificity) and revealed to be a very promising software tool supporting the physician. Using pooled data obtained from expert observation and from computer detection, the diagnostic outcomes of pattern analysis were compared. The *sensitivity* of the system was calculated as the percentage of dermoscopic images scored as melanomas from the computer system and diagnosed as melanoma from expert dermatologists and confirmed by histology (the gold standard): it was resulted 97 %. The *specificity* was calculated as the percentage of dermoscopic images scored both by the observers and by computer analysis as benignant melanocytic and naevi: it was resulted 87 %.

Starting from the present framework, further research efforts will be firstly addressed to compare and integrate the very promising approaches and corresponding feature descriptors reported in the most recent literature [53, 54, 57, 62], in order to improve the classification accuracy of the dermoscopic structures.

Then, the correlations existing among the seven dermoscopic criteria will be deeply investigate and a confidence level will be computed for each intermediate classification (for example on the basis of multi-resolution segmentation, Fuzzy

fusion and/or Markov Random Field approach [43]). The corresponding information could be effectively adopted at Lesion Diagnosis stage for improving sensitivity and specificity of the software system as whole.

Finally, an intensive measurement campaign will be carried out aiming to a double goal. The control of the image acquisition and the availability of a large image database will allow to deeply investigate the influence of the color calibration on the proposed processing algorithms. Moreover, the diagnosis from the automatic system will be compared with the results from the interactive adoption of the software tool by two groups of physicians (respectively expert and not acquainted with the 7-Point Check List) in order to estimate the improvements in the daily clinical practice of dermatologists.

This system will help dermatologists to deliver a fast and non-invasive diagnosis. Thus, it will help prevent skin cancer and treat it in its early stages, because patients will be more comfortable when tracking their lesions. However, because such instrumentation will never achieve 100 % diagnostic accuracy, and because the gold standard of histopathologic diagnosis suffers from significant interobserver disagreement, the diagnosis can't be only performed by computer, but a semiautomatic computer diagnosis can help the clinicians to achieve the best diagnostic accuracy. These technologies can be used in accordance with the patient history and clinical examination to enhance the ability to diagnose melanoma while avoiding unnecessary biopsies.

References

1. De Vries, W., Coebergh, J.W.: Cutaneous melanoma in Europe. Eur. J. Cancer **40**, 2355–2366 (2004)
2. Parkin, D.M., Whelan, S.L. et al.: Cancer Incidence in Five Continents, vol. I–VIII, no. 7. IARC Cancer Base, Lyon (2007)
3. Curado, M.P., Edwards, B., et al.: Cancer Incidence in Five Continents IX, no. 160. IARC Scientific Publications (IARC), Lyon (2007)
4. Leiter, U., Buttner, P.G., Eigentler, T.K., Garbe, C.: Prognostic factors of thin cutaneous melanoma: an analysis of the central malignant melanoma registry of the German Dermatological Society. J. Clin. Oncol. **22**(18), 3660–3667 (2004)
5. Mayer, J.: Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma. Med. J. Aust. **167**, 206–210 (1997)
6. Argenziano, G., Fabbrocini, G., Carli, P. et al.: 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. Arch. Dermatol. **134**(12), 1563–1570 (1998)
7. Haenssle, B. Korpas, Hansen-Hagg, C. et al.: Seven-point checklist for dermatoscopy: performance during 10 years of prospective surveillance of patients at increased melanoma risk. J. Am. Acad. Dermatol. **62**(5), 785–793 (2010)
8. Soyer, H.P., Argenziano, G., Zalaudek, I. et al.: Three-point checklist of dermoscopy. A new screening method for early detection of melanoma. Dermatology **208**(1), 27–31 (2004)
9. Zalaudek, I., Argenziano, G., Soyer, H.P.G. et al.: Three-point checklist of dermoscopy: an open internet study. Br. J. Dermatol. **154**(3), 431–437 (2006)

10. Gereli, M.C., Onsun, N., Atilganoglu, U. et al.: Comparison of two dermoscopic techniques in the diagnosis of clinically atypical pigmented skin lesions and melanoma: seven-point and three-point checklists. Int. J. Dermatol. **49**(1), 33–38 (2010)
11. Tromme, I., Sacré, L., Hammouch, F. et al.: Availability of digital dermoscopy in daily practice dramatically reduces the number of excised melanocytic lesions: results from an observational study. Br. J. Dermatol. **167**(4), 778–786 (2012)
12. Binder, M., Schwarz, M., Winkler, A., et al.: Epiluminescence microscopy a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch. Dermatol. **131**, 286–291 (1995)
13. Blum, A., Luedtke, A.H., Ellwanger, U., Schwabe, R., Rassner, G., Garbe, C.: Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. Br. J. Dermatol. **151**(5), 1029–1038 (2004)
14. Burroni, M., Corona, R., Dell'Eva, G., et al.: Melanoma computer-aided diagnosis: reliability and feasibility study. Clin. Cancer Res. **10**, 1881–1886 (2004)
15. Rubegni, P., Cevenini, G., Burroni, M., et al.: Automated diagnosis of pigmented skin lesions. Int. J. Cancer **101**, 576–580 (2002)
16. Ganster, H., Pinz, A., et al.: Automated melanoma recognition. IEEE Trans. Med. Imag. **20**, 233–239 (2001)
17. Schmid, P.: Segmentation of digitized dermatoscopic images by two-dimensional color clustering. IEEE Trans. Med. Imag. **18**, 164–171 (1999)
18. Grana, C., Pellacani, G., Cucchiara, R., Seidenari, S.: A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions. IEEE Trans. Med. Imag. **22**, 959–964 (2003)
19. Blum, A., Luedtke, H., et al.: Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. Br. J. Dermatol. **151**, 1029–1038 (2004)
20. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. Artif. Intell. Med. **56**(2), 69–90 (2012)
21. Celebi, M. E., Stoecker, W.V., Moss, R. H.: Advances in skin cancer image analysis. Comput. Med. Imaging Graph. **35**(2), 83–84 (2011)
22. Di Leo, G., Fabbrocini, G. et al.: ELM image processing for melanocytic skin lesion based on 7-point checklist: a preliminary discussion. Proceedings of the 13th IMEKO TC-4 Symposium, vol. 2, pp. 474–479. Athens, Greece (2004)
23. Betta, G., Di Leo, G. et al.: Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. Proceedings of Intern Work on Medical Measurement and Applications (MeMeA), pp. 63–67 (2006)
24. Di Leo, G., Liguori, C. et al.: An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin lesion. IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems 2008, pp. 190–194, Istanbul, Turkey, 14–16 July 2008
25. Di Leo, G., Fabbrocini, G. et al.: Automatic diagnosis of melanoma: a software system based on the 7-point check-list. Proceedings of the 43rd Annual Hawaii International Conference on System Sciences, Computer Society Press, 5–8 Jan 2010
26. Schmid, P., Guillodb, J.: Towards a computer-aided diagnosis system for pigmented skin lesions. Comput. Med. Imag. Graphics **27**, 65–78 (2003)
27. Celebi, M.E., Kingravi, H.A. et al.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imaging Graph. **31**(6), 362–373 (2007)
28. De Vita, V., Di Leo, G. et al.: Statistical image processing for the detection of dermoscopic criteria. Proceedings of XVIII IMEKO TC-4 Symposium, Natal, Brazil, 27–30 Sept 2011
29. Lee, T., Ng, V., Gallagher, R. et al.: DullRazor: a software approach to hair removal from images. Comput. Biol. Med. **27**(6), 533–543 (1997)
30. Schmid, P.: Lesion detection in dermatoscopic images using anisotropic diffusion and morphological flooding. IEEE, pp. 449–453 (1999)

31. Abbas, Q., Garcia, I.F., Celebi, M.E., Ahmad, W.: A feature-preserving hair removal algorithm for dermoscopy images. Skin Res. Technol. **19**(1), e27–e36 (2013)
32. Abbas, Q., Celebi, M.E., Garcia, I.F.: Hair removal methods: a comparative study for dermoscopy images. Biomed. Signal Process. Control **6**(4), 395–404 (2011)
33. Celebi, M.E., Iyatomi, H., Schaefer, G., Stoecker, W.V.: Lesion border detection in dermoscopy images. Comput. Med. Imag. Graph. **33**(2), 148–153 (2009)
34. Donadey, T., Serruys, C. et al.: Boundary detection of black skin tumors using an adaptive radial-based approach. SPIE Med. Imag. **3379**, 810–816 (2000)
35. Cucchiara, R., Grana, C. et al.: Exploiting color and topological features for region segmentation with recursive fuzzy C-means. Mach. Graphics Vis. **11**, 169–182 (2002)
36. Melli, R., Grana, C. et al.: Comparison of color clustering algorithms for segmentation of dermatological images. SPIE Med. Imag. **6144**, 351–359 (2006)
37. Zhou, H., Schaefer, G., Sadka, A., Celebi, M.E.: Anisotropic mean shift based fuzzy C-means segmentation of dermoscopy images. IEEE J. Sel. Top. Signal Process. **3**(1), 26–34 (2009)
38. Erkol, B., Moss, R.H. et al.: Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. Skin Res. Technol. **11**, 17–26 (2005)
39. Zhou, H., Schaefer, G., Celebi, M.E., Lin, F., Liu, T.: Gradient vector flow with mean shift for skin lesion segmentation. Comput. Med. Imag. Graph. **35**(2), 121–127 (2011)
40. Abbas, Q., Celebi, M.E., Garcia, I.F.: A novel perceptually-oriented approach for skin tumor segmentation. Int. J. Innovative Comput. Inf. Control **8**(3), 1837–1848 (2012)
41. Celebi, M.E., Aslandogan, Y.A. et al.: Unsupervised border detection in dermoscopy images. Skin Res. Technol. **13**(4), 454–462 (2007)
42. Garnavi, R., Aldeen, M., Celebi, M.E., Varigos, G., Finch, S.: Border detection in dermoscopy images using hybrid thresholding on optimized color channels. Comput. Med. Imag. Graphics **35**(2), 105–115 (2011)
43. Celebi, M.E., Wen, Q., Hwang, S., Iyatomi, H., Schaefer, G.: Lesion border detection in dermoscopy images using ensembles of thresholding methods. Skin Res. Technol. **19**(1), e252–e258 (2013)
44. Celebi, M.E., Kingravi, H.A. et al.: Border detection in dermoscopy images using statistical region merging. Skin Res. Technol. **13**(4), 347–353 (2008)
45. Iyatomi, H., Oka, H., Celebi, M.E., Hashimoto, M., Hagiwara, M., Tanaka, M., Ogawa, K.: An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. Comput. Med. Imag. Graphics **32**(7), 566–579 (2008)
46. Nock, R., Nielsen, F.: Statistical region merging. IEEE Trans. Pattern Anal. Mach. Intell. **26**(11), 1452–1458 (2004)
47. Otsu, N.: A threshold selection method from gray-level histogram. IEEE Trans. Syst. Man Cybern. **SMC-9**(1), 62–66 (1979)
48. Pradhan, R., Kumar, S., et al.: Contour line tracing algorithm for digital topographic maps. Int. J. Image Process. **4**(2), 156–163 (2010)
49. Kurugollu, F., Sankur, B., Harmanci, A.E.: Color image segmentation using histogram multithresholding and fusion. Imag. Vis. Comput. **19**, 915–928 (2001)
50. Levy, A., Lindenbaum, M.: Sequential Karhuen-Loeve basis extraction and its application to images. IEEE Trans. Image Process. **9**, 1371–1374 (2000)
51. Koonty, W., Narenda, P.M., Fukunya, F.: A graph theoretic approach to non-parametric cluster analysis. IEEE Trans. Comput. **25**, 936–944 (1976)
52. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Prentice Hall, New Jersey (2002)
53. Sadeghi, M., Razmara, M., Wighton, P., Lee, T.K., Atkins, M.S.: A novel method for detection of pigment network in dermoscopic images using graphs. Comput. Med. Imag. Graphics **35**(2), 137–143 (2011)
54. Barata, C., Marques, J.S., Rozeira, J.: A system for the detection of pigment network in dermoscopy images using directional filters. IEEE Trans. Biomed. Eng. **59**(10), 2744–2754 (2012)
55. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **SMC-3**(6), 610–621 (1973)

56. Iyatomi, H., Oka, H., Celebi, M.E., Ogawa, K., Argenziano, G., Soyer, H., Koga, H., Saida, T., Ohara, K., Tanaka, M.: Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. J. Invest. Dermatol. **128**(8), 2049–2054 (2008)
57. Wighton, P., Lee, T.K., Lui, H., McLean, D.I., Atkins, M.S.: Generalizing common tasks in automated skin lesion diagnosis. IEEE Trans. Inf. Technol. Biomed. **15**(4), 622–629 (2011)
58. Witten, H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco (2005)
59. Debeir, O., Decaestecker, C.: Computer-assisted analysis of epiluminescence microscopy images of pigmented skin lesions. Cytometry **37**, 255–266 (1999)
60. Celebi, M.E., Iyatomi, H. et al.: Automatic detection of blue-white veil and related structures in dermoscopy images. Comput. Med. Imag. Graphics **32**(8), 670–677 (2008)
61. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA (1993)
62. Sadeghi, M., Lee, T.K., McLean, D.I., Lui, H., Atkins, M.S.: Detection and analysis of irregular streaks in dermoscopic images of skin lesions. IEEE Trans. Med. Imag. **32**(5), 849–861 (2013). doi:10.1109/TMI.2013.2239307
63. Landwehr, N., Hall, M., Frank, E.: Logistic model trees. 14th European Conference on, Machine Learning (2003)
64. Argenziano, G., Soyer, H.P., De Giorgi, V., et al.: Interactive Atlas of Dermoscopy. EDRA Medical Publishing & New Media, Milan, Italy (2002)
65. Hance, G.A. et al.: Unsupervised color image segmentation with application to skin tumor borders. IEEE Engineering in Medicine and Biology Magazine **15**(1), 104–111 (1996)
66. Florkowski, Christopher, M.: Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin. Biochem. Rev. **29**(Supplement (i)), S83–S87 (2008)

Dermoscopy Image Processing for Chinese

Fengying Xie, Yefen Wu, Zhiguo Jiang and Rusong Meng

Abstract Dermoscopy image analysis technology is discussed based on Chinese in this chapter. It includes four aspects: preprocessing, segmentation, feature extraction and classification. Firstly, in preprocessing stage, hair is extracted out according to the elongate of connected region, and then removed from the image by using PDE-based image inpainting technology. Secondly, a novel dermoscopy image segmentation algorithm is introduced using self-generating neural network (SGNN) combined with genetic algorithm (GA). And in the feature description stage, the features including color, texture, shape and border are extracted for the lesion object. Lastly, the model of combined neural network classifier is employed to classify the lesion object successfully with a sensitivity and specificity of 93.3 and 96.7 % respectively. Based on the image analysis method disscussed in this chapter, an automatic analysis system of dermoscopy images of Chinese is successfully developed and has been applied for the clinical diagnosis of skin tumors.

Keywords hair removal · PDE-based inpainting · automatic segmentation · self-generating neural network · feature extraction · lesion object classification · combined neural network classifier

Introduction

The skin cancer is one of the most rapidly increasing cancers in the world, with an estimated annual incidence of 76,330 and 11,980 deaths in the US in 2011 [1]. At

F. Xie (✉) · Y. Wu · Z. Jiang
School of astronautics, Beihang University, 100191 Beijing, China
e-mail: xfy_73@buaa.edu.cn

R. Meng
Department of Dermatology, General Hospital of the Air Force of the Chinese People's Liberation Army, 100142 Beijing, China

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_5,

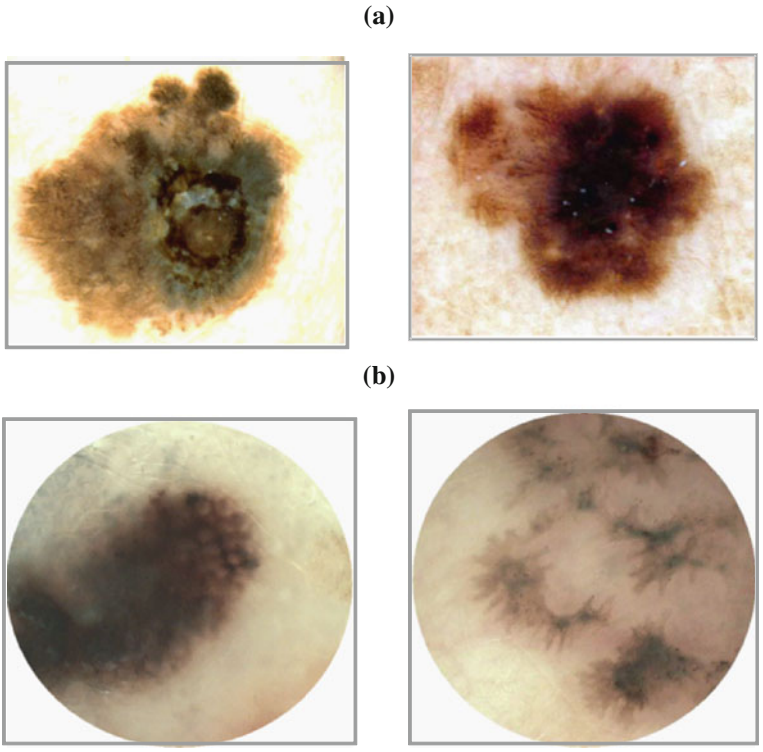


Fig. 1 Dermoscopy images of Caucasians and Chinese. **a** Dermoscopy images of Caucasians (https://b0112-web.k.hosei.ac.jp/DermoPerl/) **b** Dermoscopy images of Chinese

present, the diagnostic accuracy depends on the doctor's experience and computer-aided diagnosis for skin lesions is full of significance. Literature [2] provides a comprehensive review of the research done in the computerized pigmented skin lesion image analysis and literature [3] shows us the latest special issue on skin cancer image analysis.

Although the incidence of skin cancer in China is lower than the U.S., Europe and Australia, it has increased 3–8 % annually and has doubled over the past decade. Because of the different skin color, there exits discrepancy in the color and contrast of dermoscopy image between Caucasians and Chinese (shown in Fig. 1) So the computer-based diagnosis methods used in dermoscopy images of Caucasians do not often work well on Chinese.

Cooperating with the General Hospital of the Air Force of the Chinese People's Liberation Army, the Image Processing Center of Beihang University in China has been studying the analysis and diagnosis technology of dermoscopy images for Chinese since 2007, including removing hair from images in preprocessing phase, image segmentation, feature extraction and recognition. So far, we are the only one

to study the analysis of dermoscopy image of Chinese, and discusses in the chapter is based on our researches.

This study was supported by the Chinese National Natural Science Foundation (Grant Nos. 61027004, 61071138) and the Science Foundation of Beihang University (Grant no. YWF-12- LXGY-013).

Hair Removal from Dermoscopy Image

The factors influencing dermoscopy image quality include defocus blur, poor illumination, weak contrast, hair, air bubble, and other factors. Hair noise is processed in this section. To get the precise segmentation and analysis of the skin malignant melanoma image with hairs, it is very necessary to detect and remove the hair from the image.

In a study conducted in Italy [4], the investigators decided to shave the hairs using a razor before the imaging segmentation. This procedure not only adds extra costs and time to the imaging session, but it is also impractical when applied to total-body nevus imaging [5]. Lee [6] located the hair through morphological closing operator firstly and then removed the hairs from image by replacing the hair pixels by the nearby non-hair pixels, but the algorithm worked well just for the thick and dark hairs. In Literature [7], inpainting performs approximately 33 % better than DullRazor's linear interpolation, and is more stable under heavy occlusion. Literature [8] achieves automatic hair and ruler marking detection using curve fitting and models curve intersection and replaces the artifact pixels using feature guided, exemplar-based inpainting. In [9], Kimia Kiani et al. firstly detects hair pixels by using Radon transform and the Prewitt filters and then replaces the detected hair pixels with the average gray-level of the background. In [10], Qaisar Abbas et al. proposed a fast marching method for hair removal which includes three steps: (a) initial hair detection by derivative of Gaussian, (b) refine hair detection by morphological technique, (c) repairing of hair-occluded information by fast marching. In [11], the fast marching method is developed.

In the section, we introduce an automatic algorithm to remove hair, which includes three steps [12]:

(1) Enhancing the melanoma image with hairs by morphologic closing-based top-hat operator and thresholding the enhanced image; (2) Extracting out hairs based on the elongate of connected region; (3) Repairing the hair-occluded information using the PDE-based image inpainting technology.

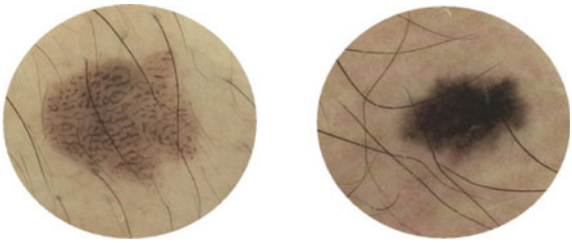


Fig. 2 Two skin melanoma images with hairs

Detecting Hair

The hair objects can be treated as the black structures because the pixels inside the hair are darker than those surrounding it (shown in Fig. 2), thereby hair objects can be enhanced using morphological closing-based top-hat operator.

The morphological top-hat operator for grayscale images is a part of the basic toolbox of mathematical morphology operators [13, 14]. It is often used to detect contrasted objects on non-uniform backgrounds. For grayscale images, there are two versions: the opening-based top-hat operator can extract bright structures in the image, and the closing-based top-hat operator can extract dark structures in the image.

Let the processing image F and the structure element B are respectively:

$$F = \{[x, f(x)] | x \in P, P \subseteq E^2\} \tag{1}$$

$$B = \{[m, b(m)] | m \in S, S \subseteq E^2\} \tag{2}$$

And then for the image $\boldsymbol{F}$, the dilating $\boldsymbol{(f \oplus B)(x)}$ and eroding $\boldsymbol{(f \Theta B)(x)}$ by structure element $\boldsymbol{B}$ at point $\boldsymbol{x}$ can be written as:

$$(F \oplus B)(x) = \sup_{\substack{m \in S \\ x-m \in P}} \{f(x-m) + b(m)\} \tag{3}$$

$$(F \Theta B)(x) = \inf_{\substack{m \in S \\ x+m \in P}} \{f(x+m) - b(m)\} \tag{4}$$

where $f(x \pm m)$ means each point x in the image F moves along the vector m.

The opening $F \circ B$ and the closing $F \bullet B$ by structure element B are respectively defined as:

$$F \circ B = (F \Theta B) \oplus B \tag{5}$$

$$F \bullet B = (F \oplus B) \Theta B \tag{6}$$

(a) **(b)** **(c)**

Fig. 3 Segmentation of melanoma images with hairs. **a** Original image **b** Closing-based top-hat operation **c** binary image

Further the opening-based top-hat $OTH_{F,B}$ and the closing-based top-hat $CTH_{F,B}$ can be respectively written as:

$$OTH_{F,B} = F - F \circ B \tag{7}$$

$$CTH_{F,B} = F \bullet B - F \tag{8}$$

The morphological top-hat operator has the property of high-pass filter. The opening-based top-hat operator, also called peak detector, can extract bright structures in the image. Different with the opening-based top-hat, the closing-based top-hat operator, also called valley detector, can extract black structures in the image.

We enhance hair objects on the grayscale image according to Eq. (8). Figure 3a is a dermoscopy image, and Fig. 3b is the result of carrying out the closing-based top-hat on Fig. 3a. From Fig. 3b, the non-hair regions including lesion are weakened, on the contrary, both weak and strong hair objects are greatly enhanced. Enhanced hairs become the bright regions in image, and the area of these bright regions occupied some ratio in the whole image. According to the experience, the ratio is about 5 % with which Fig. 3c is the threshold segment result for Fig. 3b.

Figure 3 is an example for black hairs. In fact, the blond or gray hairs are darker than the surrounding background, and thereby they can also be enhanced and detected out from the image. But for those brighter hairs than surrounding background, the opening-based top-hat operator can be used to enhance the hair objects.

Extracting Hairs

From Fig. 3c, with the non-hair noises, the hairs are detected, and the connected regions belonging to hair are bigger and longer than those non-hair connected regions

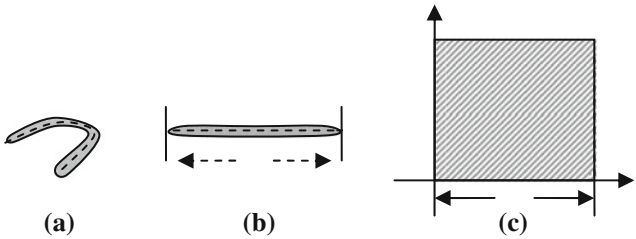


Fig. 4 The illustration of defined elongate function. **a** Connected region **b** Outspreaded region **c** Outspreaded square

which are smaller and shorter. It is simple to take the length or area size of connected region as measure to extract hairs from image, but some short hairs will not be separated from those long non-hair connected regions.

We use the elongate of connected region as measure to extract the hairs from image. For a band-like connected region, the elongate function is defined as:

Definition 1: For a given connected region R, its expanded area, EA, is the area of the square outspreaded by the central axis, the formula is as follows:

$$EA = l^2 \tag{9}$$

where l is the central axis length of R.

The geometrical explanation is illustrated in Fig. 4. The broken line in Fig. 4a is the central axis of the connected region, the area of the square in Fig. 4c is just the expanded area of Fig. 4a. From Fig. 4, for a connected region, the longer the central axis is, the bigger its expanded area is.

Definition 2: For a given connected region R, its elongate function $E(R)$ is the ratio of its expanded area to its real area, the formula is as follows:

$$E(R) = EA/RA = l^2/RA \tag{10}$$

where RA is the real area of R.

From the Eq. (10), the elongate function E measures the size of expanded area in unit area for connected region and it is related to two variables l and RA. Supposing l is invariable, with the decrease of RA, the connected region will be thinner, its elongate becomes more and more strong, and $E(R)$ is increasing; Supposing RA is invariable, with the increasing of l, the connected region will be longer, its elongate becomes more and more strong, and $E(R)$ is increasing. Thereby the thinner and longer a band-like object is, the bigger the function value $E(R)$ is.

The central axis of one band-like connected region can be obtained by skeleton extraction algorithm based on morphological Hit-Miss transform [15]. And the central axis length l, which is one of the parameters of elongate function, can be calculated according to the central axis locus. Taking the pixel number of connected

Fig. 5 Extracted hair regions based on elongate measure

region as the real area RA, and the elongate function value can be obtained according to Eq. (10).

The thin and long hair is a kind of band-like object in Fig. 3c, and its elongate feature function value $E(R)$ is bigger than the non-hair connected region. Thereby an appropriate threshold T can be selected according to real application instance and the hair whose elongate feature function value E is bigger than $T$is determined to be hair object. Figure 5 is the extracted hairs from Fig. 3c with the elongate measure defined by Eq. (10). It can be seen that the elongate feature function proposed in this paper is efficiently measured for extracting hairs.

Repairing the Hair-Occluded Information

The influence of hair to diagnosis analysis can be eliminated through excluding the hair regions from further analysis. But for the image segmentation, it can improve the veracity of segmentation to repair the melanoma texture occluded by hairs. Thus the PDE-based inpainting method is selected to repair the hair-occluded information in this chapter.

Initially proposed by Perona and Mafik, the non-linear diffusion filters have been widely used in the last decade in edge preserving and enhancement filtering. The gray levels of an image (U) are diffused according to [16].

$$\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{t}} = \mathbf{div}(\boldsymbol{c}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t}) \cdot \nabla \boldsymbol{u}) \tag{11}$$

where $\mathbf{div}(\bullet)$ is the divergence operator, ∇ is the gradient operator. the scalar diffusivity $c(x, y, t)$, in a pixel of coordinates (x, y), is chosen as a non-increasing function $g(\bullet)$ of the Gradient ∇U, which governs the behavior of the diffusion process. A typica1 choice for the diffusivity function g is:

$$c(x, y, t) = g(\nabla U) = \frac{1}{1 + (\nabla U/K)^2} \tag{12}$$

Fig. 6 Inpainting result for Fig. 3a based on PDE

where, K is gradient threshold. Practical implementations of the P-M filter are giving impressive results that noise is eliminated and edges are kept or even enhanced provided that their gradient value is greater than K.

Equation (12) is an anisotropic partial differential equation. Let $u_0(x, y)$ be the original image, and the discrete iterating form of Eq. (12) for image inpainting is:

$$u^{t+1}(x, y) = u^t(x, y) + \frac{\lambda}{n} \sum_{p \in D} c(\nabla^t u(x, y)) \nabla u^t(x, y) \tag{13}$$

where, (x, y) is pixel coordinate. D is the neighborhood of (x, y) pixel. n is the number of neighborhood pixels. The positive constant λ denotes smooth degree and t is the iteration times.

For color image of skin melanoma, the hair regions extracted from image with elongate measure are taken as mask, and the inpainting result can be obtained through iterating the Eq. (13) repeatedly within the region of the mask respectively in three color bands on the original image. Figure 6 is the inpainting result for Fig. 3a, and it can be seen that the repaired texture of melanoma is consistent with the human vision.

The automatic repair problem of the hair-occluded information can be resolved very well through combining hair object extraction with image inpainting technology in this section. Figure 7 is the local effect repaired by the method, it can be seen that there is no any blur in the region without hairs and the repaired texture in the hair region is logical. Figure 8 is the segmentation effect by Otsu's threshold after or before repairing the hair-occluded information, and it can be seen that the segmentation is improved after repairing the hair-occluded information. This introduced method is practical and robust, and the segmentation error of the skin melanoma image with hairs is effectively reduced after repairing the occluded information.

Segmentation Based on SGNN

The segmentation stage is quite important since it affects the accuracy of the subsequent feature extraction and analysis. However, segmentation is quite difficult because [17]: (i) the transition between the lesion and the surrounding skin is usually

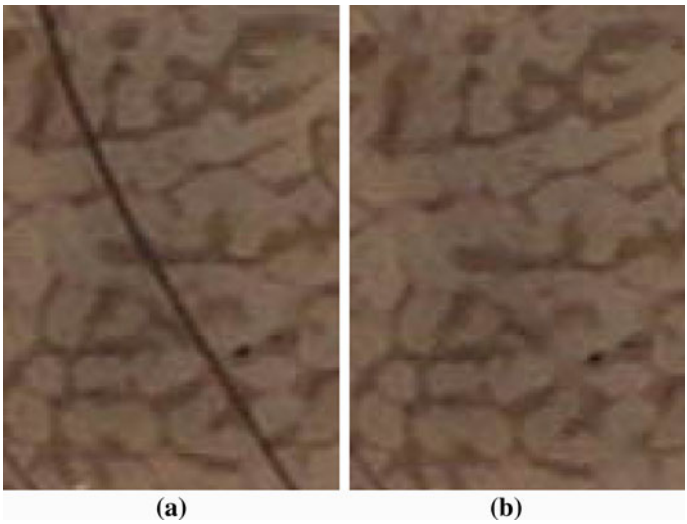


Fig. 7 Local repaired effect. **a** Part cut out from Fig. 2a **b** Repaired result by our method

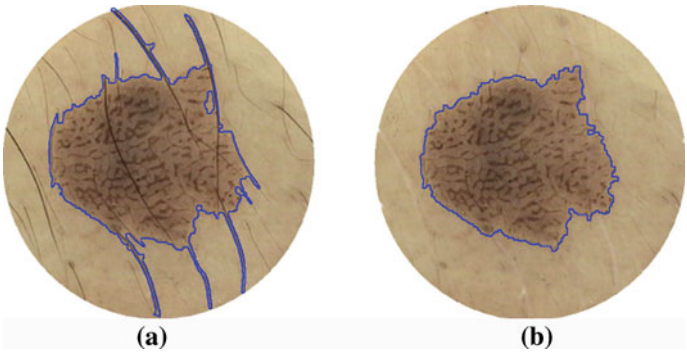


Fig. 8 Segmented result after or before inpainting. **a** Segmented result before inpainting **b** Segmented result after inpainting

of low contrast; (ii) the lesion borders are usually irregular and fuzzy; (iii) complicating artifacts are often present such as skin texture, air bubbles and hairs; and (iv) the interior of the lesion may exhibit variegated coloring.

To address these problems, a number of dermoscopic segmentation algorithms have been developed [18]. For convenience, we broadly classify these into three categories: thresholding, edge/contour-based and region-based. An effective thresholding method proposed by Grana [19] uses Otsu's threshold to automatically segment the melanoma image, then selects k points for spline-based interpolation, yielding a smoothed lesion border. Thresholding methods such as this can achieve good results when there is good contrast between lesion and skin, but encounter prob-

lems when the modes of the two regions overlap. Edge/contour-based approaches were used in [20, 21]. Rubegni [20] segmented dermoscopy images using the zero-crossings of a LoG edge operator, while Zhou [21] used an improved snake model to detect lesion borders. Edge and contour-based approaches perform poorly when the boundaries are not well defined, for instance, when the transition between skin and lesion is smooth. In such situations, the edges have gaps and the contour may leak through them. Region-based approaches have also been used. Some examples include multi-scale region growing [22], fuzzy *c*-means based on anisotropic mean shift [23], multi-resolution Markov random fields [24] and statistical region merging (SRM) [17]. Region-based approaches have difficulties when the lesion or the skin region is textured, or has different colors presentation, which can lead to over-segmentation.

Color is a significant feature for image segmentation and unsupervised color clustering [25] has been successfully used for region-based segmentation. Such data-driven methods have great potential of dealing with varied imaging situations, provided that an accurate model that is flexible enough to span the space of possible lesion image environments can be found. Since modeling such a high-dimensional complex space of possibilities is quite difficult, learning-based methods that can be trained on large datasets are of interest. Towards this end, we introduce a color clustering model for dermoscopy images [26] that combines the technique of the self-generating neural network (SGNN) with genetic algorithms (GA). Using a measure of cluster validity, the clustering algorithm that we develop automatically determines an appropriate number of clusters. By merging the clustering regions into lesion and background skin, segmentation of dermoscopy images is achieved.

Self-generating Neural Networks

SGNN is proposed as one of the competitive learning neural networks firstly by Wen [27], and then is in-depth researched by Inoue [28, 30], which is characterized by simplicity in network design, fast in learning and self-organizing ability, thus can be competent for the clustering or classification with high performance [27–29].

As shown in Fig. 9, the SGNN has been implemented as a self-generating neural tree (SGNT) which is a tree structure for hierarchical classification. Figure 9a is a clustering sample set, where $e_i, i = A, B, \ldots E$ is the sample attribute. Figure 9b is the generated SGNT for Fig. 9a according to the SGNT generating rules [27, 28], where w_i notes the neuron weight. Each leaf neuron corresponds to one or multiple samples, and its weight is the average attribute of the corresponding samples. For the SGNT, every leaf neuron is corresponding to one or multiple samples, and its weight is the attribute average of the corresponding samples, and the weight of every node neuron (non-leaf neuron) is the attribute average of all the leaf neurons covered by this node neuron. Taking every child of the root neuron as a cluster center, all of the leaf neurons in the sub-network rooted by this child belong to the same cluster, and the number of the clusters is just the number of the root neuron's children. Obviously, in the Fig. 9b, A is in the same cluster with B; C is in the same cluster with D and E,

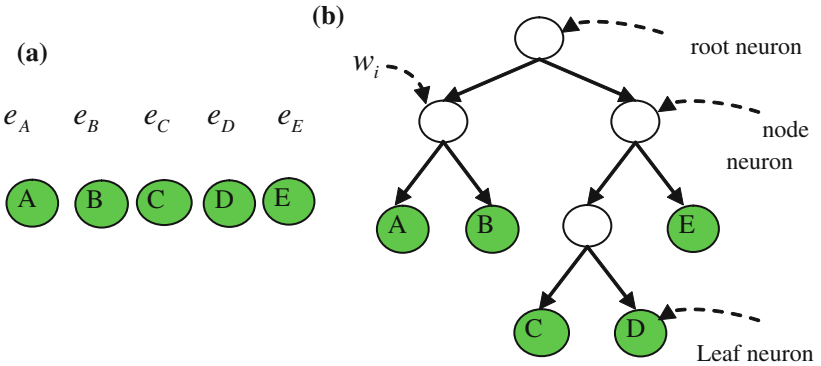


Fig. 9 The structure of the SGNT. **a** 5 samples **b** Generated SGNT for (**a**)

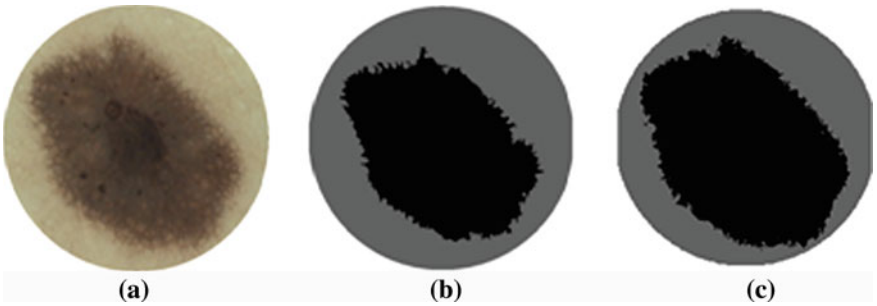


Fig. 10 Different clustering results using SGNN with different input order of samples. **a** Original image. **b** Clustering result 1. **c** Clustering result 2

and the number of the clusters is 2. It may refer to references [27–30] for more details about SGNN and pass over the further discussion here. Taking the image pixels as clustering samples and the color or coordinates information as sample attributes, SGNN can be used for image clustering.

In spite of its fitting capacity for clustering, the SGNN algorithm is influenced by the input order of the training samples, which can cause inconsistent clustering results, as depicted in Fig. 10, where Fig. 10a is the original image, and Fig. 10b, c are the results that are achieved when different samples are selected as the first input into the SGNT.

In Fig. 10b, the area of the lesion is under-segmented, whereas a more accurate result is obtained in Fig. 10c. To ameliorate this in [26], the SGNT is generalized to a Self-Generating Neural Forest (SGNF), and GA is subsequently employed to consistently select an appropriate group of seed samples as the first input into the SGNF, thereby yielding optimized clustering results.

Self-generating Neural Forest

The SGNT can be generalized to a Self-Generating Neural Forest (SGNF) as follows. Suppose that a given sample set has c cluster centers. Then the SGNF generating algorithm can be described as follows:

Step 1: Remove c seed samples randomly from the sample set, treating these seeds as initial neural trees to form an initial forest;

Step 2: Generate neuron n_j for sample i, then search each SGNT in the SGNF to find the neuron n_{win} at the shortest distance from neuron n_j;

Step 3: Connect the n_j into the SGNT covering the neuron n_{win};

Step 4: Repeat steps 2 and 3 until all samples are input into the SGNF;

The generated SGNF includes c SGNTs, each SGNT corresponding to a cluster, and all the leaf neurons in a SGNT belong to the same cluster. Whereas, the number of clusters based on SGNF is user-specified.

Optimal Clustering by Genetic algorithm

GAs [31–33] are efficient searching algorithms as well as a stochastic search technology based on natural selecting and genetics principle, and they are capable of adaptively searching solution space for optimal solution and thus successfully used for pattern classification including clustering of data [34–36]. The relevant steps of GA are:

Step 1: Randomly generate an initial population $G(0)$;

Step 2: Evaluate the fitness $f(m)$ of each individual m in the current population $G(t)$;

Step 3: Execute genetic operators including selection, crossover and mutation;

Step 4: Generate the next population $G(t+1)$ using genetic operators;

Step 5: Return to Step 2 until the maximum of the fitness function is obtained.

The clustering results derived from the SGNF generating algorithm are influenced by the c seed samples used to generate the initial neuron forest. Suppose we partition the sample set X into c clusters $X_1, X_2, \ldots, X_c$; then their cluster centers are $m_1, m_2, \ldots, m_c$ respectively, and the between-class variance can be estimated:

$$\sigma_B^2(X_1, X_2, \ldots, X_c) = \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} p_i p_j (m_i - m_j)^2 \tag{14}$$

where $p_i = n_i/n$, n_i is the number of pixels for cluster X_i, and n is the total number of pixels. For an RGB image, $(m_i - m_j)^2$ is given by:

$$(m_i - m_j)^2 = \sum_{q=1}^{3} (m_i^q - m_j^q)^2 \tag{15}$$

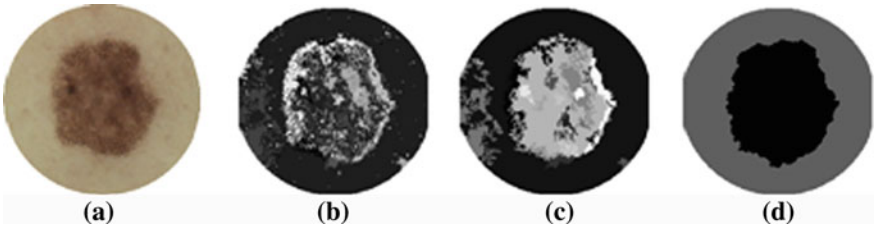


Fig. 11 Clustering based on SGNN seeded by GA. **a** Original image **b** Region growing **c** Trimming by filtering **d** Clustering

where m_i^q is the value of the ith cluster center in the qth color band.

According to the idea behind Otsu's thresholding method [37], the higher the between-class variance σ_B^2, the more accurate the partitioning of the sample set. Clustering with maximum σ_B^2 should subsequently yield an optimal partitioning of the sample set. Therefore the selection of the c seed samples can be modeled as an optimization problem. By taking σ_B^2 as the fitness function, the GA can be used to search for the c seed samples to optimize the clustering solution.

Figure 11 is an example of optimal clustering by GA. In order to enhance the computational speed, the coarse-to-fine segmentation strategy is adopted here. The image is coarsely segmented into some small sub-regions by using the region growing algorithm shown as Fig. 11b. And then some little regions are filtered out to obtain a trimmer image shown as Fig. 11c for the purpose of saving more computational time and reducing the influence of noises. Further treating various sub-regions in Fig. 11c as training samples, two samples are selected to be seeds via GA, and the clustering result is shown in Fig. 11d.

Adaptive Clustering Based on GA and SGNN (ACluster-GA-SGNN)

In Fig. 11, the number of clusters is specified before clustering. In practical applications, the number of clusters is often unknown because of the complexity of dermoscopy image. How to determine this value adaptively? This problem can be addressed using indices of *cluster validity.* Cluster validity indices are commonly used for clustering evaluation and selection of optimal clustering schemes. A number of validity indices have been introduced, e.g., the Davies-Bouldin (DB) index [38], Dunn's index [39], and the SD index [40]. Most validity indices are based on two criteria: *Compactness*, where the idea is that the members of each cluster should be as close to each other as possible, and *Separation,* meaning, the clusters should be widely spaced. We utilize the SD index to determine the proper number of clusters here.

The SD index is based on measurement of the scattering of clusters and the separation between clusters. Let $\sigma(X)$ be the variance of data set X and $\sigma(m_i)$ be

the variance of cluster *i*. Then the scattering of clusters and the separation between clusters are:

$$Scat(c) = \frac{1}{c}\sum_{i=1}^{c} \|\sigma(m_i)\| / \|\sigma(X)\| \tag{16}$$

$$Dis(c) = \frac{D_{\max}}{D_{\min}} \sum_{i=1}^{c} \left(\sum_{j=1}^{c} \left\| m_i - m_j \right\| \right)^{-1} \tag{17}$$

where c is the number of clusters, $D_{\max} = \mathbf{max}(\left\| m_i - m_j \right\|)$ and $D_{\min} = \mathbf{min}(\left\| m_i - m_j \right\|)$ are the maximum and minimum distances between cluster centers, respectively. The SD index is then:

$$SD(c) = \alpha Scat(c) + Dis(c) \tag{18}$$

where $\alpha = \mathrm{Dis}(c_{\max})$ is a weighting factor, where $c_{\max}$ is the maximum number of input clusters.

Suppose the number of clusters for a given dermoscopy image lies between 2 and $c_{\max}$, then calculate the SD index for every c using (18). Then the optimal number of clusters can be determined:

$$c^* = \mathbf{arg}\ \mathbf{min}_{c} SD(c) \tag{19}$$

Suppose the number of clusters for a given dermoscopy image lies between 2 and $c_{\max}$, the adaptive clustering algorithm includes following two steps:

Step 1: For $c = 2$ to $c_{\max}$do

(i) Specify the number of clusters to be $c$and run the SGNN algorithm seeded by GA

(ii) Calculate $SD(c)$ using Eq. (18)

Step 2: determining the optimal number of clusters using Eq. (19)

Using ACluster-GA-SGNN, the number of clusters can be determined adaptively, and the clustering task completed automatically. Figure 12 shows an instance of adaptive clustering by ACluster-GA-SGNN on a malignant skin tumor image.

The ACluster-GA-SGNN model is compared against two widely used color clustering algorithms: k-means and FCM. For k-means and FCM, the number of clusters is specified by the user, and a filtering operator is carried out to remove noise. For the sake of fairness, the number of clusters for ACluster-GA-SGNN is specified by the user also. Figure 13 is a group of clustering results. It may be seen that when the contrast is large between the lesion object and its surrounding skin, and the lesion has a homogeneous texture and clear edge, then all three methods deliver a reasonable clustering result. While the object has no regular edges, shape or uniform color, and the contrast is reduced between the lesion object and its surrounding skin, then our model yields a more satisfying result.

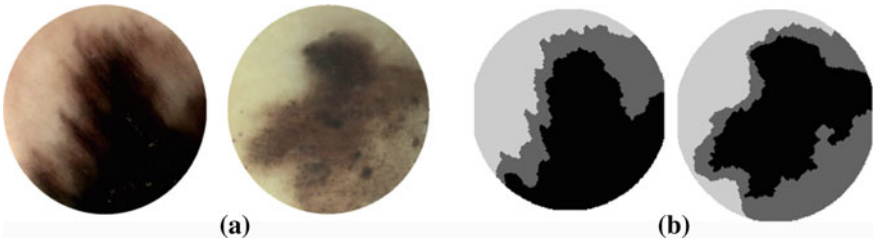


Fig. 12 Clustering results using ACluster-GA-SGNN. **a** Original image **b** Clustering result for (**a**)

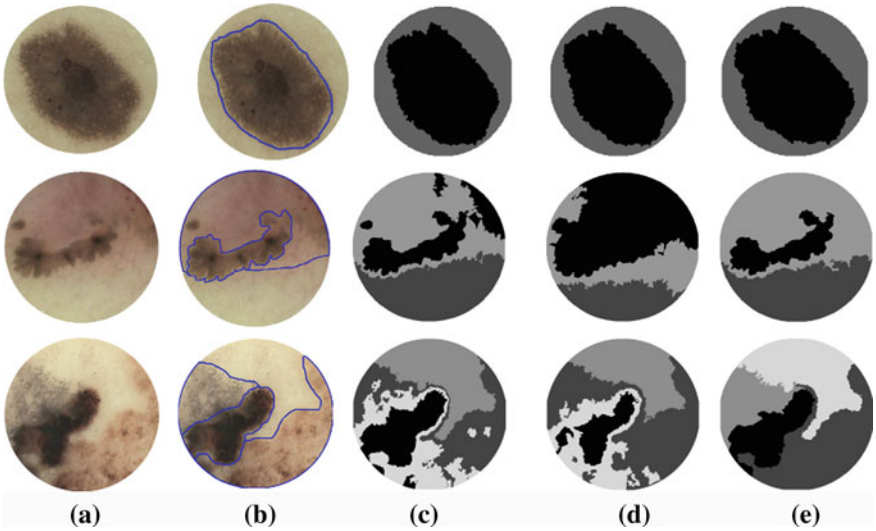


Fig. 13 Clustering results by three methods. **a** Original image **b** Manual border **c** k-means **d** FCM **e** ACluster-GA-SGNN

Clustering Region Merge

The number of clusters varies across dermoscopy images, and the clustering image often includes more than two sub-regions that can be used for subsequent feature extraction and lesion segmentation. Features, such as the number of clusters, the color and texture in each sub-region, are important information for lesion classification. Actually, segmentation must be done before feature extraction, since the purpose is to determine those regions belonging to the lesion object and to detect the lesion border, because [18] firstly, the resulting border structure provides a basis for the calculation of important clinical features, such as lesion size and symmetry axes; and secondly, it is crucial for the extraction of some of the most discriminating dermoscopy features, such as radial streaming and pseudopods.

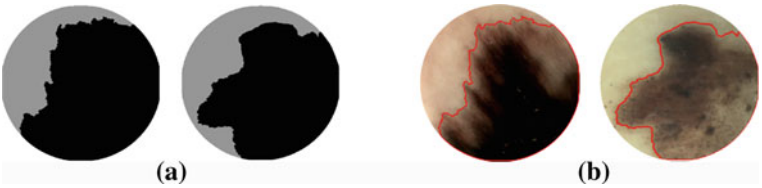


Fig. 14 Segmentation results using ASegment-GA-SGNN. **a** Segmentation result **b** Extracted border

We place the focus on the segmentation of a lesion from its surrounding skin. Generally, the background skin region is brighter than the lesion region and will touch the image frame. According to this feature, we merge these clustering subregions into two parts: background skin and lesion object, and then lesion border can be extracted, as shown in Fig. 14, where the red line in Fig. 14b is the extracted border from Fig. 12b. Actually, this merge step is the post-processing for the ACluster-GA-SGNN, and this segmentation algorithm is named as ASegment-GA-SGNN.

Generally, when the image has strong contrast and homogeneous texture, most of the automatic methods can deliver satisfactory accuracy. While for the images with weak contrast, heterogeneous texture and variational colors, ASegment-GA-SGNN can have better performance, for example Fig. 15. Figure 15 is a group of segmentation instances using the automatic methods including Otsu's threshold, k-means, FCM, and ASegment-GA-SGNN, where the blue line is a manually inscribed border and the red line is the automatically determined border. It can be seen, for the complex image, ASegment-GA-SGNN yields more consistent segmentation result with manual border.

Classification of Lesion Objects

The feature extraction and classification of lesion object is one of the key steps for the automatic analysis of dermoscopy images which have been addressed by many researchers. In 2004, Tanaka et al. [41] used the statistical method based on the ABCD rule to get 105 features and then employed recursive discriminant method to classify lesion object with the accuracy rate of 96 %. In the same year, Motoyarna et al. [42] divided each channel of the RGB color space into 16 equal parts, and the RGB color space was quantized into $16 \times 16 \times 16$ bins. By using these spaces to analyze features of the malignant melanoma, about 26 % of melanoma images could be successfully classified through the color information. Literature [43] classified skin lesion objects by using the neural network technology on dermoscopy images, and by using fuzzy membership function and adaptive wavelet transform on multispectral images. And it was concluded that the combination of dermoscopy and multispectral technologies can improve the diagnostic accuracy of lesions object. In 2007, Celebi et al. [44] extracted 437 features including color, texture and shape informa-

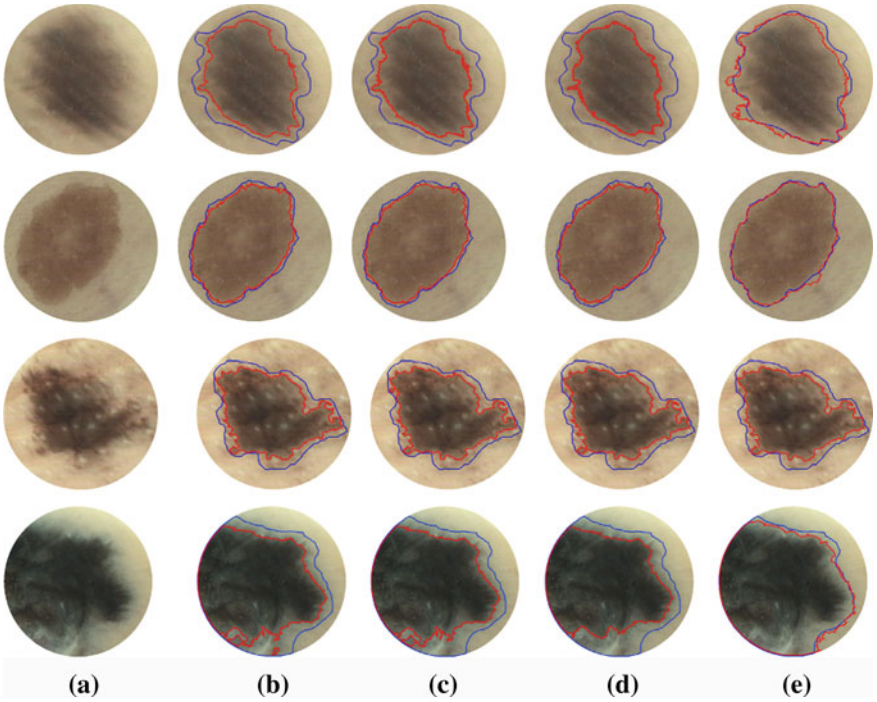


Fig. 15 Group of segmentation instances by automatic methods.The column from (**a**) to (**e**) is the original image, Otsu's threshold, k-means, FCM, and ASegment-GA-SGNN respectively

tion according to the ABCD rule, and then employed Weka data mining to select the most important 18 features. Finally he used SVM classifier to achieve a result with the sensitivity and specificity of 93.3 and 92.3 % respectively. In 2012, based on the color of asymmetric and multi-scale texture features, literature [45] used AdaBoost.MC (adaptive boosting multi-label learning algorithm) classifier model to divide lesions into 7 kinds of model: reticular, globular, cobblestone, homogeneous, parallel ridge, star-burst and multicomponent, with obtaining the sensitivity and specificity of 89.28 and 93.75 % respectively. Some important achievements of image classification for skin cancer since 2001 are shown in Table 1.

The research above is about the dermoscopy image from Caucasians, and next we introduce a classification method aiming at Chinese dermoscopy images. In this method, the features including color, texture and shape are extracted according to the diagnosis standards in clinic, and combined neural network is employed to classify the skin lesion object.

Table 1 Classification research of dermoscopy image

Paper	Date	Feature description	Classifier		
[46]	2001	Color/shape/ boundary gradient distribution	KNN		
[47]	2001	Color/shape/border	ANN		
[48]	2002	Color/texture/shape	ANN		
[49]	2003	Color/shape	Multiple classifiers		
[50]	2004	Color/texture/shape	Logistic regression		
[51]	2005	Color/entropy/shape	Logistic regression		
[52]	2007	Color/texture/shape	SVM		
[53]	2009	Color/texture/shape	Adaboost/C4.5		
Paper	Number	Malignant (%)	Benign (%)	Sensitivity (%)	Specificity (%)
[46]	5363	2	98	87	92
[47]	58	38	62	77	75
[48]	147	39	61	93	92.8
[49]	152	28	72	81	74
[50]	837	10	90	80.0	82.4
[51]	2430	16	74	91	65
[52]	564	16	84	93.3	92.3
[53]	513	16	84	95	NR

NR not reported, *kNN* k nearest-neighbor
ANN artificial neural network

Feature Description

At present, the artificial diagnosis standards of melanoma mainly include: ABCD rule [53], Menzies scoring [54] and seven-point checklist [55]. Benign melanocytic lesions tend to have few colors, an architectural order, and symmetry of pattern and are homogeneous; and malignant melanomas often have many colors, architectural disorder, and asymmetry of pattern and are heterogeneous. The ABCD rule describes these visual features associated to malignant lesions (Asymmetry, Border irregularity, Color irregularity and different structural components), from which a score is computed. In automated diagnosis of skin lesions, most of feature design is based on the ABCD rules. The same with automatic methods in literature [47, 56, 57], we extract the color, texture and shape features based on ABCD rule in this section.

Before extracting features, we divide the dermoscopy image into three parts according to the characteristics of the skin lesion image: inner lesion region, transition zone and the background region (shown in Fig. 16.)And features are extracted in the RGB color space over these regions [38, 56].

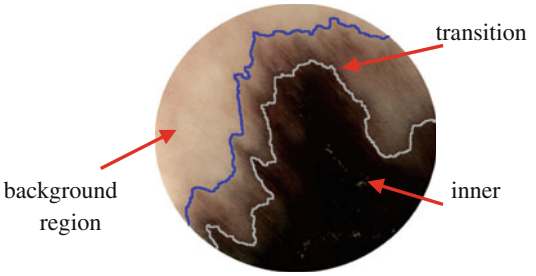


Fig. 16 Three regions of a dermoscopy image

Color Features

(1) Mean and standard deviation of color
The mean and standard deviation values of lesion over each channel of the RGB color space are calculated respectively. The two statistics over three regions are calculated: lesion object (including the inner region and transition region), inner region and transition region. We also calculate the ratios and differences of the two statistics over the inner and transition region for each channel. The total number of color features is n = 30, they are:

$$\begin{aligned} n = (&3\ \textit{channels in RGB color space}) \times (3\ \textit{means} \\ &+ 3\ \textit{standard deviations} + 1\ \textit{difference of mean}, \\ &\quad 1\ \textit{difference of standard deviation} \\ &+ 1\ \textit{ratio of mean} + 1\ \textit{ratio of standard deviation}) \end{aligned} \tag{20}$$

(2) 3D color histogram
Color diversity is an important indicator to judge whether the lesion is malignant or not. The mean and standard deviation of color above is calculated in single channel of color space. Here, the color space was quantized into 16 × 16 × 16 bins and the 3D histogram of the lesion is formulated based on it, for the image color diversity is the result of the interaction of different channels [42]. Let $p(i, j, k)$ be color statistical probability when the value of R, G, B is i, j, k respectively and *ColorNum* be the value of color distribution over the 3D histogram, which is defined as:

$$\boldsymbol{ColorNum} = \sum_{i=1}^{16}\sum_{j=1}^{16}\sum_{k=1}^{16} \boldsymbol{h(i,j,k)} \tag{21}$$

where

$$\boldsymbol{h(i,j,k)} = \begin{cases} \boldsymbol{1,\ p(i,j,k) \neq 0} \\ \boldsymbol{0,\ p(i,j,k) = 0} \end{cases} \tag{22}$$

Compaired with benign lesion, the malignant lesion is full of various color and the value of $ColorNum$ is bigger.

(3) LUV histogram distances
In order to determine the color difference of two regions (the inner lesion and transition area), the histogram distance in the CIE L*u*v space is employed. First, the color space is quantized into $4 \times 8 \times 8$ bins, which resulting in a 256-dimensional vector of color features for a region. Then, the distance of two histogram is employed to calculate the similarity of two regions:

$$\boldsymbol{d_1(h,g)} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} |\boldsymbol{h(i)} - \boldsymbol{g(i)}| \tag{23}$$

$$\boldsymbol{d_2(h,g)} = \left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{256}} (\boldsymbol{h(i)} - \boldsymbol{g(i)})^2 \right)^{\boldsymbol{1/2}} \tag{24}$$

where, $i = 1, 2, \ldots, 256$, is the ith bin of color, $\boldsymbol{h(i)}$ and $\boldsymbol{g(i)}$ are the color histograms of the two regions (the inner lesion region and the transition region) respectively. Thus we obtain two color features d_1 and d_2.

Texture Features

The Grey Level Co-occurrence Matrix (GLCM) [58] is one of the methods most commonly used in texture statistical analysis. Given an image composed of pixels each with an intensity (a specific gray level), the GLCM is a tabulation of how often different combinations of gray levels in an image co-occur in the certain direction. Texture feature calculations [40] use the contents of the GLCM to give a measure of the variation in intensity at the pixel of interest. Given an image $f(x, y)$ with size $N \times N$, the GLCM is defined as:

$$P(i, j|d, \theta) = \{(x, y)|f(x, y) = i, f(x + dx, y + dy) = j\} \tag{25}$$

where $d = (dx, dy)$ is an offset vector in x and y directions.

Five textrue features, which include contrast, energy, correlation, entropy and inverse difference, are calculated from GLCM and employed to quantify the texture of dermoscopy image here. As the calculation of mean and standard deviation, the texture statistics are calculated in three regions: the lesion object (including the inner and the transition), the inner lesion region and the transition region, and then the ratios and differences of the latter two regions are calculated. The number of texture features is 25.

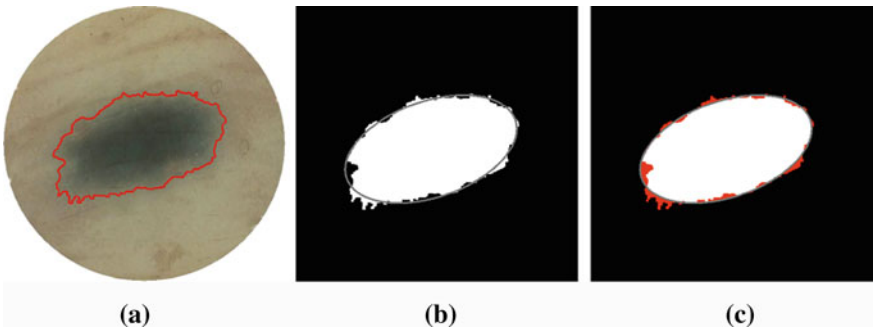


Fig. 17 Lesion object and its ellipse. **a** Lesion object **b** Drawing an ellipse **c** Non-overlaping regions

Shape Features

Compared with the malignant tumors, the shape of benign tumors is closer to a circle or ellipse and is symmetric, and the object boundary is more regular with less prominent twigs. The actual size of the benign tumors is closer to its convex hull area. Thus the common shape features for lesions include ellipseness, aspect ratio, symmetric ratio, and solidity.

(1) Ellipseness
For an object, the major axis L_{long}, minor axis L_{short} and the major axis orientation (θ) can be calculated respectively as:

$$L_{long} = \sqrt{2(\mu_{02} + \mu_{20} + (4\mu_{11}^2 + (\mu_{02} - \mu_{20})^2)^{1/2})} \tag{26}$$

$$L_{short} = \sqrt{2(\mu_{02} + \mu_{20} - (4\mu_{11}^2 + (\mu_{02} - \mu_{20})^2)^{1/2})} \tag{27}$$

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{28}$$

An ellipse for the lesion object is drawn in Fig. 17b. There are some black regions in the ellipse and some white regions out the ellipse near the boundary line. Let $A_{difference}$ be the sum area of the black and white regions (the red regions in Fig. 17c), and $A_{ellipse}$ be the area of the ellipse, then the ellipseness is defined as:

$$Ellipseness = 1 - \frac{A_{difference}}{A_{ellipse}} \tag{29}$$

(2) Aspect Ratio
Aspect Ratio is the ratio of the length of the major axis to the length of the minor axis:

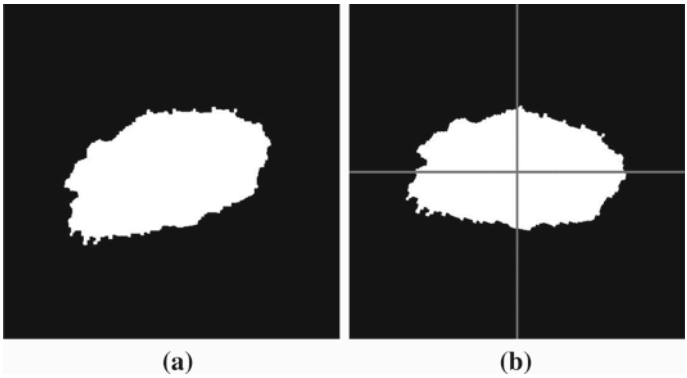


Fig. 18 Rotating object according to the major axis orientation. **a** Lesion object in Fig. 17a **b** Object is rotated θ degree clock-wise

$$Aspect\_Ratio = \frac{L_{short}}{L_{long}} \tag{30}$$

(3) Symmetry
In order to calculate the symmetry, we rotate the object θ degrees clockwise (shown in Fig. 18) firstly. And then, taking the major axis as x-axis, the object is folded about the x-axis hypothetically. The area A_y of the same region between the overlapping folds is calculated as the amount of symmetry about the x-axis. Taking the minor axis as image y-axis, the same area A_x is obtained by the same procedure. As shown in Fig. 19, the gray region is the same region between the overlapping folds.
Let A be the total area of the object, and the symmetry rates is defined as:

$$Symmetry\_Rate = \frac{A_x + A_y}{2A} \tag{31}$$

(4) Solidity
The convex hull for the lesion is defined as the convex polygon region with the minimum area including the lesion object (shown in Fig. 20).
Let A_{convex_hull} be the area of convex hull, the solidity is defined as [59]:

$$Solidity = \frac{A}{A_{convex\_hull}} \tag{32}$$

According to the features of lesion object, for the four shape indices, the lower the value, the more the lesion tends to be malignant.

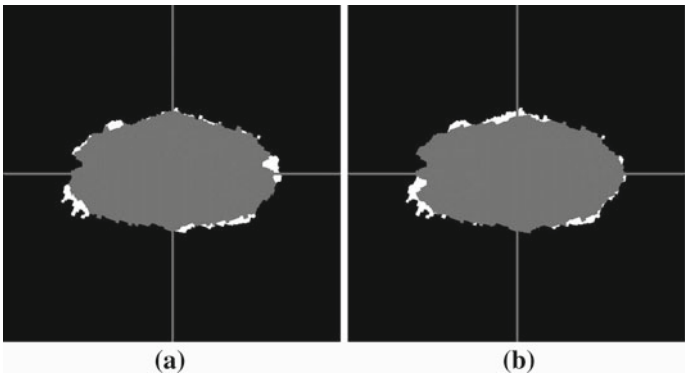


Fig. 19 Symmetry region about the symmetry axis. **a** Symmetry region about the y-axis **b** Symmetry region about the x-axis

Fig. 20 The convex hull for the lesion

Border Features

In practical applications, there are often big objects which can not be acquired completely. To overcome this shortcoming, two border features including the border depression degree and the non-similar degree between the inner and outer border are introduced in [57].

(1) Border depressed degree
For a lesion object, the border depressed degree $\boldsymbol{R}_{depressed}$ is defined as:

$$\boldsymbol{R}_{depressed} = \frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \frac{\boldsymbol{RA_i}}{\boldsymbol{l_i}} \tag{33}$$

where n denotes the number of depressed area of one object, l_i is the span of the ith depressed area, and RA denotes the area of depressed border. Figure 21 shows the intuitive explanation. $R_{depressed}$ represents the average depth of all depressed area of one object.

(2) Non-similar degree between the inner and outer borders
The outer and inner border of lesion object are shown in Fig. 22. Let Γ_{outer} and Γ_{inner} be the pixel sets of outer and inner border respectively; $D(p_i, p_j)$

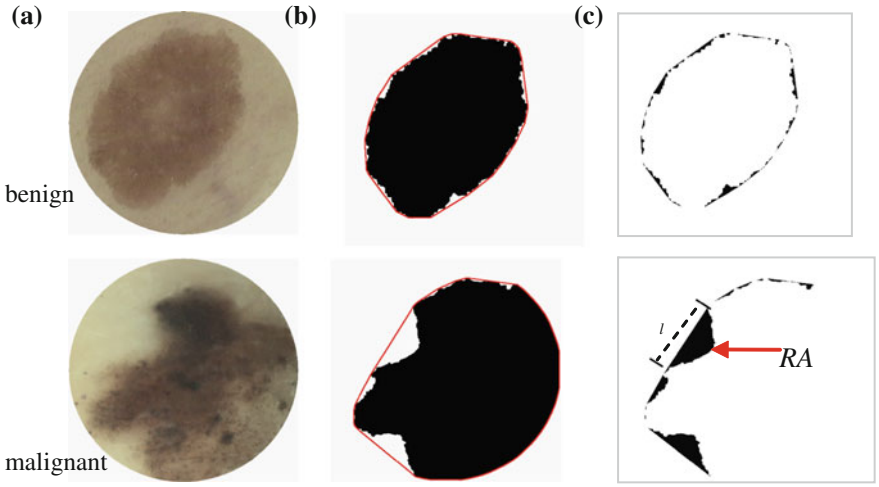


Fig. 21 BDR of the lesion. **a** Lesion **b** Convex hull **c** Concave region

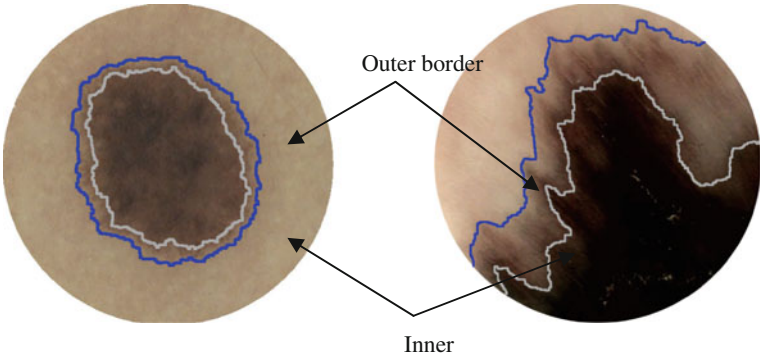


Fig. 22 The outer and inner border of the lesion. **a** Inner and outer border of the benign **b** Inner and outer border of the malignant

denotes the distance between pixel p_i to p_j. Thus, the minimum distance from outer border pixel p_i to inner border pixel p_j is defined as:

$$d_i = \min_j(D(p_i, p_j)),\ p_i \in \Gamma_{outer},\ p_j \in \Gamma_{inner} \tag{34}$$

The non-similar degree between the inner and outer borders can be represented by the variance of the distance from the outer border pixel to the inner border:

$$\boldsymbol{BorderUnsimilar} = \sqrt{\frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{d_i} - \boldsymbol{m})^{\boldsymbol{2}}} \tag{35}$$

Where n denotes the number of pixels of outer border, and $m = \frac{1}{n}\sum_{i=1}^{n} d_i$ is the mean of $\boldsymbol{d}_i$.

Compared with benign, the boundary of malignant lesion object has more prominent twigs, and the inner border is very different with the outer border. Thus the $\boldsymbol{R}_{depressed}$ and ***BorderUnsimilar*** for malignant lesion have bigger values.

Feature Selection

The total number of features is 64 and too many features will introduce irrelevant noisy features, which reduce the recognition rate. Good feature set would be of good discernibility, high reliability, strong independence, high stability and includes fewer members. GA is a kind of optimizing search method based on biological selection and evolution process. For its nice adaptation, good parallel search property and satisfied performance in many complex optimization decision and design, GA is employed to find out the most optimal combination of features.

The Model of Combined Neural Network Classifier

Combined neural network consists of multiple sub-networks. Compared with the single neural network, the performance and classification accuracy of combined neural network classifier is much better [60, 61].

A combined neural network topology diagram of parallel structure is shown in Fig. 5. The combined neural network is a combination of relatively independent sub-neural networks, which can be trained parallel and independently. The final output of the combined neural network is the weighted sum of the output of every single sub-network. The function is:

$$\vec{\mathbf{y}}(\vec{\mathbf{x}}, \vec{\mathbf{a}}) = \sum_{j=1}^{p} \boldsymbol{a_j y_j}(\vec{\mathbf{x}}) \tag{36}$$

where $\mathbf{x}$ is the feature input vector and a_i is the weight of the ith single network.

From above topology structure in Fig. 23, it can be known that the combined neural network will have no any improvement if generalization ability of every sub network is the same. So sub-neural network with individual differences need to be trained and that's the way how individual generates. At the same time, the combination of output weights need to be optimized to obtain a high-quality combined neural network classifier (conclusion generating), taking the BP neural network as a basic learning classifier, we employ Adaboost [62] method to generate individual subnets,

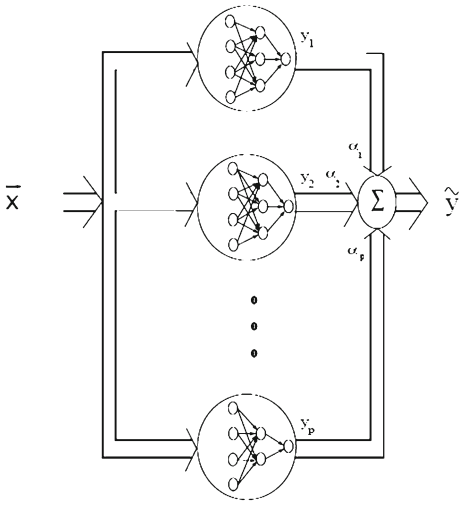


Fig. 23 Topology structure of combined neural nework

and use GA to optimize the output weights [63], and then a combined BP neural network classifier is generated.

Experiment analysis is carried out on the data set with 70 samples from the General Hospital of the Air Force of the Chinese People's Liberation Army. The optimal features selected by GA from the 64 original features include: the two border features, correlation and entropy of the lesion object, 3D color histogram, mean of red and green of the transition region and deviation of all color channels of both lesion object and the inner lesion. Then the features are input the combined BP neural network classifier. seven fold cross-validated is adopted in the section. Finally we achieve sensitivity and specificity of 93.3 and 96.7 % respectively and this is very useful to the research of the automatic aided diagnosis system for Chinese dermoscopy image.

Conclusion

Aiming at Chinese dermoscopy images, the methods including hair removal, segmentation, feature extraction and classification are introduced in this chapter. Image processing center of Beihang is the firstand the only one group by far, to research computational analysis technology for Chinese skin images. At present, an automatic analysis system of dermoscopy images for Chinese has been developed successfully and has been applied for the clinical diagnosis of skin tumors in the General Hospital of the Air Force of the Chinese People's Liberation Army. The analysis result from this system affords an important reference for the doctors, which contributes to the development of dermatology in China.

References

1. Siegel, R., Ward, E., Brawley, O., Jemal, A.: Cancer statistics. CA Cancer J. Clin. **61**(4), 212–236 (2011)
2. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. Artif. Intell. Med. **56**(2), 69–90 (2012)
3. Celebi, M.E., Stoecker, W.V., Moss, R.H.: advances in skin cancer image analysis. Comput. Med. Imaging Graph. **35**(2), 83–84 (2011)
4. Andreassi, L., Pemtti, R., Burroni, M.: Computerized image analysis of pigmented lesions. Chronica Dermatol **1**, 11–24 (1995)
5. Slue, W., Kopf, A.W., Rivers, J.K.: Total body photographs of dysplastic nevi. Archives of Dermatology **124**, 1239–1243 (1988)
6. Lee, T.K., Ng, V., Gallagher, R., et al.: Dullrazor: a software approach to hair removal from images. Comput. Biol. Med. **27**(6), 533–543 (1997)
7. Wighton, P., Lee, T.K., AtkinsaM.S.: Dermoscopic hair disocclusion using inpainting. In: Proceedings of the SPIE medical imaging, vol. 6914, pp. 691427–691427-8 (2008)
8. Zhou, H., Chen, M., Gass, R., Rehg, J.M., Ferris, L., Ho J., et al.: Feature-preserving artifact removal from dermoscopy images. In: Proceedings of the SPIE medical imaging, vol. 6914, pp. 69141B–69141B-9 (2008)
9. Kiani, K., Sharafat, A.R., Shaver, E., etal.: An improved dullrazor for digitally removing dark and light-colored hairs in dermoscopic images. Comput. Biol. Med. **41**(3), 139–145 (2011)
10. Abbas, Q., Celebi, M.E., Garcia, I.F.: Hair removal methods: a comparative study for dermoscopy images. Biomed. Signal Process. Control **6**(4), 395–404 (2011)
11. Abbas, Q., Garcia, I.F., Celebi, M.E., Ahmad, W.: A feature-preserving hair removal algorithm for dermoscopy images. Skin Res. Techhnol. **19**(1), e27–e36 (2013)
12. Xie, F.Y., Qin, S.Y., Jiang, Z.G., Meng, R.S.: PDE-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma. Comput. Med. Imaging Graph. **33**(4), 275–282 (2009)
13. Zeng, M., Li, J.X.: Optimized design of morphological improved top-hat filter based on improved genetic algorithms. Acta Optica Sinica **26**(4), 510–515 (2006)
14. Soille, P.: Morphological Image Analysis: Principles and Applications, Springer, Berlin (1999)
15. Cui, Y.: Mathematics morphological algorithms and its application. Science Press, Beijing, China (2000)
16. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990)
17. Celebi, M.E., Kingravi, H., Iyatomi, H., Aslandogan, A., Stoecker, W.V., Moss, R.H.: Border detection in dermoscopy images using statistical region merging. Skin Res. and Technol. **14**(3), 347–353 (2008)
18. Celebi, M.E., Iyatomi, H., Schaefer, G., Stoecker, W.V.: Lesion border detection in dermoscopy images. Comput. Med. Imaging Graph. **33**(2), 148–153 (2009)
19. Grana, C., Pellacani, G., Cucchiara, R., Seidenari, S.: A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions. IEEE Trans. Med. Imaging **22**(8), 959–964 (2003)
20. Rubegni, P., Ferrari, A., Cevenini, G., Piccolo, D., Burron, M., et al.: Differentiation between pigmented spitz naevus and melanoma by digital dermoscopy and stepwise logistic discriminant analysis. Melanoma Res. **11**(1), 37–44 (2011)
21. Zhou, H., Schaefer, G., Celebi, M.E., Lin, F., Liu, T.: Gradient Vector Flow with Mean Shift for Skin Lesion Segmentation. Comput. Med. Imaging Graph. **35**(2), 121–127 (2011)
22. Celebi, M.E., Aslandogan, A., Stoecker, W.V.: Unsupervised border detection in dermoscopy images. Skin Res. Technol. **13**(4), 454–462 (2007)
23. Zhou, H., Schaefer, G., Sadka, A.H., Celebi, M.E.: Anisotropic mean shift based fuzzy c-means segmentation of dermoscopy images. IEEE J. Sel. Top. Sign. Proces. **3**(1), 26–34 (2009)

24. Gao, J., Zhang, J., Fleming, M.G.: A novel multiresolution color image segmentation technique and its application to dermatoscopic image segmentation. In: Proceeding of IEEE International Conference on Image Process, Vancouver, BC, Canada (2000)
25. Cucchiara, R., Grana, C., Seidenari, S., Pellacani, G.: Exploiting color and topological features for region segmentation with recursive fuzzy C-means. Machine Graphics Vision **11**(2/3), 169–182 (2002)
26. Xie, F.,Bovik, Al.: Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. Pattern Recognit. **46**(3), 1012–1019 (2013)
27. Wen, W.X., Jennings, A., Liu, H.: Learning a neural tree. In: Proceeding of International Joint Conference on Neural Networks, vol. 2, pp. 751–756, Beijing, China (1992)
28. Inoue, H., Narihisa, H.: Efficiency of self-generating neural networks applied to pattern recognition. Math. Comput. Model. **38**(11–13), 1225–1232 (2003)
29. Feng, S., Chaudhari, N.S.: A chaotic behavior decision algorithm based on self-generating neural network for computer games. In: Proceedings of IEEE Conference on 3rd Industrial Electronics and Applications, pp. 1912–1915, Singapore (2008)
30. Inoue, H., Narihisa, H.: Efficient pruning method for ensemble self-generating nerual networks. J. Syst. Cybern. Inf. **1**(6), 72–77 (2003)
31. Mukhopadhyay, A., Bandyopadhyay, S., Maulik U.l.: Clustering using Multi-objective genetic algorithm and its application to image segmentation. In: Proceeding of IEEE International Conference on Systems, Man, and Cybernetics, vol. 10, pp. 2678–2683 (2006)
32. Zhu, Y., Jiang, L.J.: Image Segmentation Based on GA-FCM Clustering and Probability Relaxation. Laser Infrared **38**(4), 292–295 (2008)
33. Awad, M., Chehdi, K., Nasri, A.: Multicomponent image segmentation using a genetic algorithm and artificial neural network. IEEE Geosci. Remote Sens. Lett. **4**(4), 571–575 (2007)
34. Hall, L.O., Ozyurt, I.B., Bezdek, J.C.: Clustering with a genetically optimized approach. IEEE Trans. Evolut Ccmput. **3**(2), 103–112 (1999)
35. Maulik, U., Bandyopadhyay, S.: Genetic algorithm based clustering technique. Pattern Recogn. **33**, 1455–1465 (2000)
36. Maulik, U., Bandyopadhyay, S.: Fuzzy partitioning using a real-coded variable-length genetic algorithm for pixel classification. IEEE Trans. Geosci. Remote Sens. **41**(5), 1075–1081 (2003)
37. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. B Cybern. **9**(1), 62–66 (1979)
38. Davies, D.L., Bouldin, D.W.: A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. **1**, 224–227 (1979)
39. Dunn, J.C.: Well separated clusters and optimal fuzzy partitions. J. Cybern. **4**, 95–104 (1974)
40. Halkidi, M., Vazirgiannis, M., Batistakis, Y.: Quality scheme assessment in the clustering process. In: Procedding of European Conference on Principles and Practice of Knowledge Discovery in Databases, pp. 265–276, Lyon, France (2000)
41. Tanaka, T., Yamada, R., Tanaka, M., et al.: A study on the image diagnosis of melanoma. Proceedings of the 26th Annual International Conference of the IEEE EMBS, vol. 9, pp. 1597–1600 (2004)
42. Motoyarna, H., Tanaka, T., Tanka, M., et al.: Feature of malignant melanoma based on color information. SICE Annual Conf. Sapporo **1**, 230–233 (2004)
43. Sachin, V., Atam, P.: Multi-spectral imaging and analysis for classification of melanoma. In: Proceedings of the 26th Annual International Conference of the IEEE EMBS vol. 9, pp. 503–506 (2004)
44. Celebi, M.E., Kingravi, H.A., Uddin, B., et al.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imaging Graph. **31**, 362–373 (2007)
45. Abbas, Q., Celebi, M.E., Serrano, C., García, I.F.: Pattern classification of dermoscopy images: A perceptually uniform model. Pattern recogn. **46**(1), 86–97 (2013)
46. Ganster, H., Pinz, A., Rohrer, R., et al.: Automated melanoma recognition. IEEE Trans. Med. Imaging **20**(3), 233–239 (2001)
47. Hintz-Madsen M., Hansen L. K., Larsen J., et al. A probabilistic neural network framework for detection of malignant melanoma. In: Naguib RNG, Sherbet GV(ed) Artificial Neural Networks

in Cancer Diagnosis, Prognosis, and Patient Management, CRC Press, Boca Raton, 141–183 (2001)
48. Rubegni, P., Burroni, M., Cevenini, G., et al.: Digital dermoscopy. Analysis and artificial neural network for the differentiation of clinically atypical pigmented skin lesions: A retrospective study. J. Invest. Dermatol. **119**(2), 471–474 (2002)
49. Sboner, A., Eccher, C., Blanzieri, E., et al.: A multiple classifier system for early melanoma diagnosis. Artif. Intell. Med. **27**(1), 29–44 (2003)
50. Blum, A., Luedtke, H., Ellwanger, U., et al.: Digital image analysis for diagnosis of cutaneous melanoma development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. Br. J. Dermatol. **151**(5), 1029–1038 (2004)
51. Menzies, S.W., Bischof, L., Talbot, H., et al.: The performance of SolarScan: An automated dermoscopy image analysis instrument for the diagnosis of primary melanoma. Arch. Dermatol. **141**(11), 1388–1396 (2005)
52. Capdehourat, G., Corez, A., Bazzano, A., et al.: pigmented skin lesions classification using dermatoscopic images. Progress Pattern Recognition, Image Analysis, Computer Vision, Applications **5856**(11), 537–544 (2009)
53. Stolz, W., Riemann, A., Cognetta, A.B., et al.: ABCD rules of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur. J. Dermatol. **4**(7), 521–527 (1994)
54. Menzies, S., Crook, B., McCarthy, W., et al.: Automated instrumentation and diagnosis of invasive melanoma. Melanoma Res. **7**(Suppl. 1), s13 (1997)
55. McGovern, T.W., Litaker, M.S.: Clinical predictors of malignant pigmented lesions: a comparson of the Glasgow seven-point checklist and the American Cancer Society's ABCDs of pigmented lesions. J. Dermatol. Surg. Onc **18**, 22–26 (1992)
56. Celebi, M.E., Kingravi, H.A., Uddin, B., et al.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imaging Graph. **31**, 362–373 (2007)
57. Fengying Xie. (2009) Segmentation and Classification of Dermoscopy Images Based on Computational Intelligence.doctor dissertation at Beihang University in China.
58. Clausi, D.A.: An analysis of co-occurrence texture statistics as a function of gray level quantization. Can. J. Remote Sens. **28**(1), 45–62 (2002)
59. Andrew, A.M.: Another Efficient Algorithm for Convex Hulls in Two Dimensions. Inf. Process. Lett. **9**(5), 216–219 (1979)
60. Hansen, L.K., Salamon, P.: Neural network ensembles. IEEE Trans. Pattern Anal. Mach. Intell. **12**(10), 993–1001 (1990)
61. Hashem, S., Schmeiser, B., Yih Y.:Optimal linear combinations of neural networks: An Overview. In: IEEE International Conference on Neyral Networks pp. 93–19 (1994)
62. Freund, Y., Schapire, R.E.: A Decision-TheoreticGeneralization of On-Line Learning and an Application to Boosting. J. Comput. Syst. Sci. **55**(1), 119–139 (1997)
63. Shen, Z.Q., Kong, F.S.: Optimizing weights by genetic algorithm for neural network ensemble. Lect. Notes Comput. Sci. 323–331 (2004)

Automated Detection of Melanoma in Dermoscopic Images

Jose Luis García Arroyo and Begoña García Zapirain

Abstract In this chapter a software system for the automated detection of melanoma over dermoscopic images is presented. The analysis is carried out by supporting in the "ABCD Rule" medical algorithm, undertaking the automated detection and characterization of the corresponding indicators. For this purpose the system uses different image processing techniques and three supervised machine learning tasks. To test the robustness of the system the different indicators of the algorithm are tested, obtaining good results of accuracy in all of them, and there also was determined in a direct way the diagnostic capacity of the system, obtaining results of 81.25 % of sensitivity and 77.14 % of specificity. Moreover, the system is also capable of analyzing macroscopic images, having been designd with multiplatform architecture and being firmly oriented to teledermatology, which is increasingly used.

Keywords Melanoma · Segmentation · Biosignal processing techniques · Pattern recognition · Machine learning

1 Introduction

In this chapter a software system for the automated detection of melanoma over dermoscopic images designed and developed by our development team is presented.

Melanoma is a type of skin cancer which represents approximately 1.5 % of cancer worldwide, being diagnosed approximately 1,60,000 new cases every year. The phototype of a person, the ability of the skin to assimilate the solar radiation, is important to determine the melanoma risk and, therefore, the highest incidence is

J. L. G. Arroyo · B. G. Zapirain(✉)
DeustoTech-LIFE Laboratory, Faculty of Engineering, University of Deusto, Bilbao, Spain
e-mail: jlgarcia@deusto.es

B.G. Zapirain
e-mail: mbgarciazapi@deusto.es

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_6,

registered in developed countries, with white population predominance, where 81 % of the worldwide cases are given, increasing every year to the extent that, in some countries, it has become into a serious health problem in citizens [1, 2].

In the fight against melanoma it is fundamental the early detection. To get a better idea of the importance of it, the following fact: the survival rate for those patients whose melanoma is detected in good time, before the tumor has penetrated the skin-non-invasive melanoma or melanoma in situ-is 99 %, falling by 15 % in advanced cases-invasive melanoma-when the metastases has entered [1].

Historically, the main method used for the early detection of melanoma was the "naked-eye" vision of the doctor, in pursuit of suspicious lesions, and for these the biopsy, which consists on removing the skin lesion for its subsequent analysis, with 100 % reliability. The problem of this method is that the accuracy of the "naked eye" method is manifestly improvable, especially if it is not carried out by a medical dermatologist and, conversely, the biopsy is the most secure diagnostic procedure, but its utilization presents the inconveniences of being an invasive method and the needing of many resources, either human or material ones. It therefore becomes evident the necessity of using non-invasive techniques, with good precision.

The most used is, by far, the dermatoscopy, or epiluminescence microscopy, which consists on the exam of the skin through the utilization of an optical system and a light source designed to reduce the irregular refraction and the light reflection in its surface. Thus, it is possible to visualize in depth structures, forms and colors which are not accessible to the visual inspection. The dermatoscopy has been demonstrated especially useful to define the benignity or malignancy of the pigmented lesions. The device used is called dermatoscope which, as it can be seen in Fig. 1, connected to a computer captures the images, existing two types of taken images, those of the dermatoscopy with contact illumination type and non-polarized light, using a liquid in the interphase, and those of the dermatoscopy with cross-polarized light, for which it is needed the dermatoscope to be equipped with polarized light lenses, with a cross polarization filter [3]. Although there are many differences, most of the structures

Fig. 1 Using dermatoscope. *Source* Ref. [13]

and observed colors with contact dermatoscopy and polarized light dermatoscopy are equivalent.

This method presents better results than the "naked eye" analysis. In expert hands a sensitivity of 89 % is achieved (against the 70–85 % obtained through the "naked eye") [4]. This is an indicative of the robustness of this technique and means that most melanomas can be diagnosed, although it should always be taken into account that there is a percentage of 11 % of the cases in which it is not applicable. This system also allows the reproducibility in the diagnosis, enabling second opinions and allowing the use of image processing technologies.

There exist new promising techniques which are alternative to dermatoscopy [5–8], notwithstanding the easiness of image acquisition, its good results and its utilization level between medical experts ensures its utilization over a long period of time. In fact, dermoscopy has been recognized as the "gold standard" in the screening phase [8].

In the diagnosis of the melanoma over dermoscopic images the most used method is the 2-stage method. As it can be seen in Fig. 2, this method consists on, in a first stage, the dermatologist has to discern whether the lesion is melanocytic or not, with a series of criteria and, in a second stage, if so, to use a diagnostic method to determine the level of malignancy, with the aim of deciding if the skin lesion should be removed [3]. The 4 main methods used by the dermatologists are: "Pattern Analysis" [9], "ABCD Rule" [10], "Menzies Method" [11] and "7-Point Checklist" [12]. It is, in all of them, about to detect and to characterize a series of indicators present in the skin lesion image, determining the diagnosis in line with pre-established criteria. Notwithstanding, it should be noted that its objectivation is difficult, since, in many cases, the detection and characterization of these indicators are highly biased by the subjectivity of the dermatologists.

It is clear that the automation of this process through computer aided diagnostic software is of great importance, since, on the one hand, it can accelerate the medical task, and on the other hand to provide a qualitative and objective evaluation of the

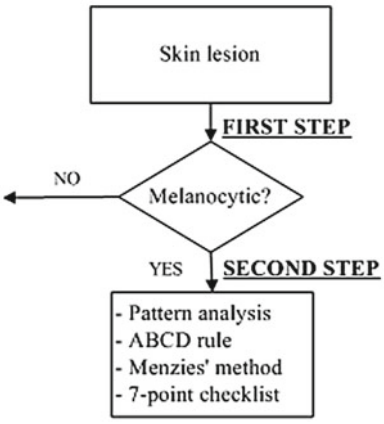


Fig. 2 Two steps procedure

skin lesion, decreasing the inter-observer and intra-observer variability which can be found in the diagnoses made by human experts.

Facing to the automation of the diagnosis, there exist a great number of algorithms which address different parts of this process, like certain processing or segmentation tasks, or the detection and characterization of indicators. However, there exist few software systems that address the automation of the whole process, and even less commercial tools which offer a certain degree of robustness in the diagnosis.

In this context, our research team has designed and developed a complete family of digital image processing algorithms which have been integrated in an automated system in support of the detection of melanoma, carrying out the help in the diagnosis over the quantitative base of the medical algorithm "ABCD rule", also supporting the decision in the "Pattern Analysis".

The results that are being obtained are very promising in the different stages of the diagnostic system: preprocessing, segmentation, detection, characterization of indicators and diagnosis.

Currently, this system is in piloting stage, and the aim is that the system can be applied in different areas. For instance, it can be used as a guide to medical professionals which are not experts in dermatology. The case of primary care physicians, or even if it is an expert dermatologist the task will be make it easier, either because of being great the number of moles to analyze or, simply, to give a second opinion. Another possible application is the execution of the system in massive analysis processes, over repositories of dermoscopic images, with the purpose of searching possible malign lesions. Or even in screening systems directed to the population in order to detect melanomas in early stages, over telemedicine platforms or from another type

In fact, the system has been designed in such a way that it is multiplatform, likely to be accessed by the doctor, apart from the desktop application, via web or mobile, to a server system centralized to the diagnoses. Furthermore, the system can be accessed by a normal user, via web or mobile, for screening objectives withstanding for this the analysis of macroscopic images: in fact, the system has a determined vocation towards the teledermatology, which is increasingly used.

This chapter has been structured as follows: in section "State of the Art", a detailed review of the state of the art in the automated detection of melanoma is made, in section "System Design" the design of the proposed system is explained, in section "Results" the results of the system are presented, and, finally, in section "Conclusions and Future Work" the conclusions and the future work are showed.

2 State of the Art

In this section the state of the art in the automated detection of melanoma is presented. First of all, the aforementioned medical diagnostic methods are commented briefly, deepening in the "ABCD rule" and the "Pattern Analysis". Secondly, the state of the art techniques for the automated detection of melanoma in dermoscopic images are

explained in detail. Finally, a review of some of the most relevant commercial tools which currently exist for the early detection of melanoma is carried out.

2.1 Medical Methods in Dermoscopy: "Pattern Analysis" and "ABCD Rule"

As stated previously, in the diagnosis of melanoma over dermoscopic images the most used method is the 2 stage method, being these the most important 4 diagnostic methods used: "Pattern analysis", "ABCD Rule", "Menzies Method" and "7-Point Checklist". It is about, in all of them, to detect and characterize a series of indicators, present in the image of the skin lesion, determining the diagnosis according to some established criteria.

The "Pattern Analysis" is the method by which the best results are obtained, since it allows a more complete approach of the pigmented lesions although, conversely, it is the one which requires more experience [3]. The other 3 methods are named medical algorithms, since in all of them it is about to detect and characterize quantitatively a series of indicators, present in the skin-lesion image, and thus assign them numerical scores, being the final value obtained the malignancy measurement result of the algorithm, determining the diagnosis according to pre-established value ranges. Obviously, owing to manage with characterizations and numerical punctuations, when approaching an automation of the diagnosis, the most appropriate is to implement a medical algorithm.

2.1.1 Pattern Analysis

It consists of finding the existing patterns in the lesion and taking a diagnostic decision. These patterns are visual structures with forms, colors and characteristic textures, as it can be seen in Fig. 3. It allows the recognition and valuation of a series of parameters and patterns which allow the dermoscopic distinction between a benign and a malignant lesion [9].

2.1.2 ABCD Rule

"ABCD Rule" is a medical algorithm that intends to facilitate diagnosis of melanoma to those observers less experienced in the technique. It is based on the valuation of 4 criteria [10], as can be seen in the Table 1:

Asymmetry: The pigmented lesion is divided in two perpendicular axes, in such a way that we achieve as much symmetry as possible, and the asymmetry is evaluated with respect to the shape, the color and the structure in both sides of each axis. A punctuation of 0 is granted if there is no asymmetry in any axis, of 1 if there exists asymmetry in one axis and of 2 if it presents it in both axes, in such a way that the lesion can have a score of 0–2.

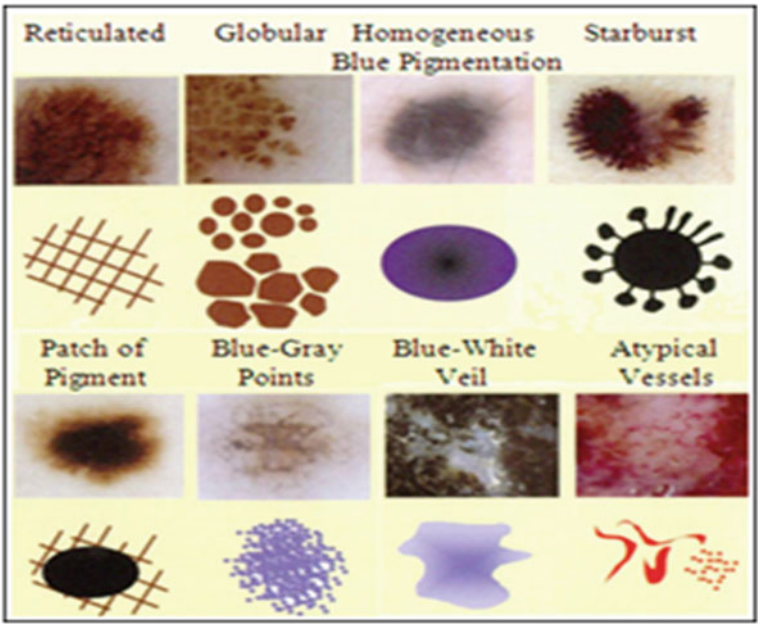


Fig. 3 Dermoscopic patterns. *Source* Ref. [14]

Table 1 Criteria and punctuation of "ABCD rule"

Dermoscopic criteria	Score	Weighting/ correction factor
Asymmetry	0 to 2	X 1, 3
Border	0 to 8	X 0, 1
Color	1 to 6	X 0, 5
Dermoscopic/ dermoscopic structures	1 to 5	X 0, 5

Border: The lesion is divided in eight segments and it is scored with 1 each portion which might present an abrupt completion of the border. In such a way that the minimum score that it can be obtained is 0 and the maximum is 8.

Colors: It evaluates the presence of six colors (white, light brown, dark brown, blue-gray, red and black). Consequently, the maximal score will be 6 and the minimum 1. The white color will only be scored if it is lighter than the adjacent skin; that is, when it belongs to white areas of regression.

Dermoscopic Structures:

In the clinical ABCD, applicable to macroscopic images, the "D" is corresponding to the Diameter and it is considered a symptom of malignancy that it drops from 6 mm.

Here, we are describing the dermoscopic ABCD, in which the "D" is corresponding to dermoscopic structures, a similar concept to the "pattern" which it was

Table 2 Diagnosis with "ABCD rule"

TDS		
<4.75	4.8–5.45	>5.45
Benign	Suspicious	Malignant

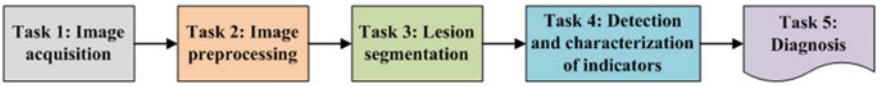


Fig. 4 Tasks of the life cycle of an automated system for the detection of melanoma

considered in the "Pattern Analysis". It is considered five structures, each of them scored with 1. Consequently the maximum score will be of 5 and the minimum of 1. The dermoscopic patterns are:

- Pigmented network (reticular): either if it is typical of atypical.
- Homogeneous and unstructured areas: should cover more than 10 % of the lesion.
- Points: should be more than two.
- Globule: should be at least two.
- Linear ramifications/pseudopods: should be more than two.

Subsequently each value is multiplied by each weighting factor, the partial scores are added, and the total dermoscopic score (TDS) is obtained.

Once calculated the TDS, as can be seen in the Table 2, the diagnosis is determined, with the following ranges:

TDS < 4.75: BENIGN
4.8 <= TDS <= 5.45: SUSPICIOUS
5.45 < TDS: MALIGNANT

2.2 Techniques for the Automated Detection of Melanoma in Dermoscopic Images

Here a state of the art of the techniques for the automated detection of melanoma in dermoscopic images used in the most relevant works is presented. To do so, first of all, the complete life cycle of an automated system in support of the detection of melanoma is going to be undertaken, and secondly, the techniques used in some of the most relevant works will be commented.

2.2.1 Life Cycle of an Automated System for the Detection of Melanoma

The complete life cycle of an automated system for the detection of melanoma, using the dermoscopic technique, as can be seen in Fig. 4, consists on the following stages:

1. Image acquisition: The acquisition of the dermoscopic image is carried out.

2. Image preprocessing: The preprocessing of the image is carried out. First of all, the problems of the improvement in the image quality are covered, with the aim of these representing as closely as possible the original one. Secondly, and it is here where the automated process of a software tool really begins, the problem in the detection and treatment of the "noise" is covered, which used to appear in this kind of images, such as hair, bubbles, flashes, shadows, ink marks in the skin, electronic marks (usually digital identifiers or copyright information), black frames and devices and rulers to measure.

3. Skin lesion segmentation: The skin lesion segmentation is carried out. Habitually, it is undertaken in an automated way. Notwithstanding, there exist semi-automated systems in which the expert is allowed to intervene in this stage, introducing information for the segmentation improvement.

4. Detection and characterization of indicators: The automated detection and characterization of the chosen indicators to undertake the diagnosis.

5. Diagnosis: On the basis of the medical algorithm, the quantitative calculation of the degree of malignancy is made to determine the diagnosis according to the ranges of the pre-established values.

2.2.2 Techniques Used in Relevant Works

Here the techniques used in the 2–5 of the life cycle of a system of automated detection of melanoma will be explained.

2.2.3 Image Preprocessing

Here the problem in the detection and the treatment of the "noise" which sometimes appear in these type of images, detecting the devices existing in the image, and substituting the existing pixels in an appropriate manner.

Due to its inevitability in the acquisition of images, the most relevant problem is, by far, the detection and treatment of hair, and in fact, it is the one which is treated more in the medical literature.

The majority of works are based on the DullRazor® algorithm, presented in [15] by Lee et al., using morphological filters for each of the RGB bands, and bilinear interpolation for replacing the pixels. In [16], Zhou et al. present a novel system which automatically detects the artifacts with curvilinear property, i.e. hair and ruler markings, and removes them from dermoscopy images using exemplar-based inpainting. In [17], Wighton et al. carry out an inpainting study as an alternative to DullRazor's bilinear interpolation.

In [18], Kiani et al. present a method for hair removal, better and more efficient than DullRazor® algorithm, whose principal contribution is in the hair detection, for which an initial task of determining the hairs orientation is performed, using the Radon Transform, followed by a second task that subsequently carries out a border detection, made by Prewitt filters.

In [19], Abbas et al. present a novel algorithm for detecting and repairing of hair-occluded information from dermoscopy images, which has a capability to preserve the skin lesion features such as color and texture. For the hair detection the method uses Gaussian and morphological filters, and for the restoration it uses a fast marching inpainting method.

In [20], Abbas et al. undertake a comparative study between the previously presented method and the three main method of the state of art: linear interpolation, inpainting by nonlinear partial differential equation (PDE) and inpainting by exemplar-based repairing techniques.

2.2.4 Lesion Segmentation

For the segmentation of the skin lesions within the dermoscopic lesion numerous methods have been proposed, corresponding to different approaches which the segmentation can be addressed with.

In [21], Celebi et al. analyze the state of the art and the issue existing in the segmentation of the skin lesions in dermoscopic images. In this work the different segmentation methods are classified roughly in the following categories: 1. Histogram thresholding: obtaining an histogram threshold to make the segmentation; 2. Clustering: obtaining homogenous regions using color clustering algorithms; 3. Edge-based: using edge operators; 4. Region-based: obtaining homogenous regions using region merging, region splitting, or both; 5. Morphological: using the watershed transform; 6. Model-based: modeling of images as random fields; 7. Active contours (snakes and their variants): using active contours techniques; 8. Soft computing: using soft-computing techniques.

In [22], Delgado et al. carry out the segmentation, using the IHP method, a method based on the accomplishment of a linear transformation over another color space in which the segmentation is made using techniques through histogram thresholding. With an idea similar to the previous one, in [23], Garnavi et al. undertake the projection of the dermoscopic image to different color channels, making subsequently the segmentation in two regions, through histogram thresholding, using the Otsu method, with the aim of obtaining the optimal projection. In [24], Celebi et al. propose a new method based on the fusion of some thresholding methods, using a Markov random field (MRF) framework, and subsequently this method is compared with nine methods of the state of art.

In [25], Schmid et al. carry out the segmentation on the basis of the fuzzy c-means (FCM). In [26], Cucchiara et al. also use that method, paying the attention in issues of color, as well as topological ones, guiding the recursion of the algorithm by evaluation of adjacency and mutual inclusion properties of extracted regions. In [27], Zhou et al. present a new mean shift based FCM algorithm improving the computing cost. In [28], Melli et al. compare four color clusterization methods: median-cut, k-means, FCM and mean shift, the latter having obtained the best results. In [29], Celebi et al. carry out the segmentation using the JSEG algorithm based on the color quantization and spatial segmentation. In [30], Celebi et al. use the statistical region

merging algorithm method comparing the results with six methods of the state of the art, with very good results.

The use of the watershed technique has been proposed in some works; one of the most relevance is presented in [31], where Wang et al. use the watershed approach, and adding for the improvement of the segmentation a postprocess by a correction based on neural networks.

The technique of the active contours/snakes is described in the methods section. This method is normally used in semiautomatic systems, which require some interaction by the user, although there also exist methods in which this interaction is not undertaken, like [32], in which Erkol et al. use a method based on gradient vector flow (GVF) snakes with an automatic initialization. More recently, in [33], Zhou et al. propose an improved algorithm based on GVF and, in [34], Abbas et al. present a work that uses the active contour model in the CIECAM02 based uniform color space, to achieve an adaptation to human perception.

2.2.5 Detection and Characterization of Indicators and Diagnosis

Previously the most commonly used medical methods for the diagnosis were commented. But examining the indicators present in those methods, it can be observed that the objectivization of some of the present indicators, the C of the "Color" of the "ABCD Rule", can be easily quantified and objectifiable. However, there exist other indicators, as it is the case of many of the patterns or dermoscopic structures, which are very difficult to quantify and objectify, in fact, its determination by human experts is largely subjective. Thus, many methods of the state of the art have other approach.

Figures 5 and 6 show the two approaches, firstly the classical approach, that maps directly the detection and characterization of indicators from the medical algorithm, and secondly the other approach, where a process of extraction of features is undertaken and, on the basis of a classification procedure, the diagnosis is determined.

In the second approach should be noted that, really, these features behave themselves like the indicators of the algorithm. Normally, this approach uses supervised machine learning techniques: based on a set of features, a sample of image values is extracted, and its numerical values are studied in relation with its benignity or malignancy. On the basis of this, some classification rules are generated, which determine the diagnosis, and which are applied over the entirety of images.

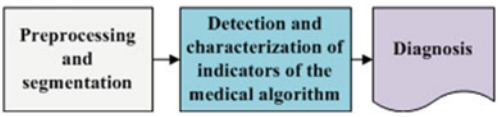


Fig. 5 Diagnosis made with the classical approach mapping directly from the medical algorithm

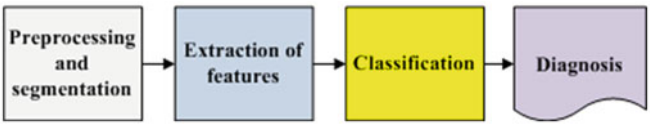


Fig. 6 Diagnosis made with the approach based on the extraction of features and classification algorithms

The problem this approach has is that, these being habitually automations are a support for the diagnosis of the doctor and, since he is the one who takes the decision at the end, he usually prefers to do it supported by algorithms used in the medical practice, and that is why he feels more comfortable if the guide to the diagnostic shows him the estimated value of the medical indicators.

Hence, the classical approach which is to be followed is to attempt to reproduce as closely as possible the medical method, making to do so a detection and characterization of the indicators of the medical algorithms. The system presented here would be located in this classical approach.

We are watching some of the most relevant works, using both approaches.

In [35], Iyatomi et al. carry out an implementation of the ABCD algorithm by selecting 80 features for the A, 32 for the B, 140 for the C and for the D, instead of undertaking the detection of different dermoscopic structures, it selects 176 texture features, and the process of machine learning and rule generations is made with ANN (Artificial Neural Network).

In [36], Alcon et al. present a detection system based on the ABCD algorithm, with an approach similar to the previous one, selecting features for A, B, C and D, and taking texture features for the D. For the rule generations various classifiers are used.

In [37], Di Leo et al. present a detection system based on the 7-Point Checklist Method. For the detection of the different structures they use color, texture, spectral and structural analysis features. The analysis and classification are made using statistical methods and machine learning methods.

In [38], Ganster et al. select a set of color, geometric and texture features to describe the malignancy of a lesion, delivering the final kNN classification a sensitivity of 87 % with a specificity of 92 %.

In [39], Tomassi et al. define a series of color and texture features, then a classification with SVM and spin glass-Markov random fields. The results are compared, over 270 images, with those of the previous method exposed by Ganster et al.

In [40], Celebi et al. present a work in which a series of geometric features are defined, color and texture ones, which are subsequently studied with SVM, using kernel RBF. Out of a total of 564 images a sensitivity of 84.33 % and a specificity of 96.19 % are obtained.

In [41], Rahman et al. present a system of detection of melanoma. 128 features of color and texture are presented, the classification is made with Bayes and 358 images are tested, with an accuracy of 75.69 %.

In [42], Surowka presents a work in which supervised learning with MLP and SVM is used. The feature set is composed of wavelet-based multi-resolution filters of dermoscopy images, and feature selection is done by the Ridge linear models. A sensitivity of 89.2–94.7 % and a specificity of 85–95 % are obtained.

In [43], Situ et al. use the technique of Bag-of-Features (BoF), widely used in artificial vision and equivalent to the Bag-of-Words used in the documents. The classification results are made using Naive Bayes classification and Support Vector Machines. The best performance obtained is 82.21 % on a dataset of 100 skin lesion images.

In [44], Zhou et al. present a new technique for the detection and description of invariant rotation and scaling features. Based on the interest point detection model, the concept of DIP (Dermoscopic Interest Points) is introduced for the first time, and a detector and descriptor of those points is proposed, comparing the information taken with others of the state of the art like SIFT and SURF over 150 images, obtaining good results.

In [45], Situ et al. present a work in which, using the same BoF framework designed in [43], implement lesion classification using DIPs. SVM classifier is used, and over 1505 images 86.17 % of accuracy and 84.68 % of specificity are obtained.

Regarding the detection of the individual patterns, it will be commented hereafter.

In [46], Anantha et al. use for the detection of the pigment network texture analysis algorithms.

In [47], Betta et al. for the detection of the atypical pigment network use two techniques, one structural, in which morphological methods are used, and another one spectral, in which the FFT is used, high pass filters, inverse FFT and finally thresholding techniques. Furthermore, in this work, a pixel classification method is used to find the pixels belonging to a vascular pattern.

In [48], Di Leo et al. from the same research group, improve the previous work, defining 9 chromatic characteristics and 4 spatial ones, and use decision tree classifiers, in the categories of “Absent”, “Typical” and “Atypical”, using the C4.5 algorithm. The process is carried out over 173 images with more than 85 % of sensitivity and specificity (the exact values are not specified).

In [49], Barata et al. undertake the detection of the pigment network using a bank of directional filters and morphological operations, followed by a feature extraction and an AdaBoost algorithm for the classification, and then obtaining over a database of 200 images a sensitivity of 91.1 % and a specificity of 82.1 %.

In [50], Shrestha et al. use different texture metrics for the analysis of the del atypical pigment network, with an accuracy of 95 %. In [51], Sadeghi et al. carry out the detection of the pigment network using the Laplacian of Gaussian (LoG) filter in the first place to capture correctly the “Light-Dark-Light” and then, over the generated binary image, seeks cyclical sub-networks, using the ILCA algorithm (Iterative Loop Counting Algorithm). 500 images are tested with an accuracy of 94.3 %.

In [52], Sadeghi et al. from the same research group, improve the previous work and present a method for the classification between “absence of pigment network”, “typical pigment network” and “atypical pigment network”. To do so they propose

an algorithm based on the previous work to find the mesh and the extraction of structural, geometric, chromatic and texture characteristics, treated with the Login-Boost algorithm, thereby generating a classification between "Absent", "Typical" and "Atypical", over 436 images. 82.3 % of accuracy is obtained. If it is made over the "Absent" and "Present" categories, 93.30 % of accuracy is obtained.

In [53], Skrovseth et al. seek different texture parameters for the detection of the pigment network and the globular pattern.

In [54],Yoshino et al. present a work in which morphological filters are used for the detection of the globular pattern.

In [14], Gola et al. undertake the detection of globules combining morphological techniques with an edge detection algorithm.

In [55], Celebi et al. study the presence of the pattern blue-white veil in the melanocytic lesions. To do so a series of features corresponding to the veil pixels are captured, and a tree for the classification is generated, using the C4.5 algorithm, with the rules obtained for the classification of the pixels of a skin lesion between veil and not veil. After that, the study about if the region with the pixels might have the blue-white pattern, on the basis of geometric features, using the C4.5, and generating a tree with the corresponding rules.

In [56], Di Leo et al. carry out a partition of the lesion into different colors using PCA. Over the regions some features are defined, to determine those of blue-white veil and those of regression. LMT are used, generating a tree of rules, achieving sensitivity results of 87 % and specificity results of 85 %.

In [57], Garcia et al. study the blue-white veil and the dark-red patch of pigment patterns in melanocytic lesions, using machine learning techniques, obtaining color features for the pixels detection and calculating morphological and geometric features for the regions classification, and also incorporating in the generated rules for the patterns detection the relation between both types of structures.

For the streaks detection, in [58], Mirzaalian et al. develop a machine-learning approach, using for the identification of the streaks color, spectral and geometric features, and for the classification a SVM model, validating the method over 99 images, differenced in "absence", "presence of regular" and "presence of irregular", with an AUC of 91 %.

In [59] Sadeghi et al. present a novel approach for the detection of the streaks, applying techniques based on ridge and valley detection used in fingerprint image recognition. It is tested in 300 images of melanocytic lesions, with an AUC of 90.5 %. In [60], Sadeghi et al. from the same research group, improve the previous work, and present a method for the detection of the streaks. They classify the dermoscopy images into streaks "Absent", "Regular" and "Irregular", testing in 945 images, and obtaining an AUC of 85 %. This method also includes the Absent/Present and Regular/Irregular classifications.

For the detection of the hypopigmentation pattern, in [61], Dalal et al. study the features of the hypopigmentation areas from different point of views (color, geometry, position inside the lesion, etc.), and using neuronal networks arrive to a series of rules.

For the detection of blotches, in [62], Madasu et al. present a work in which blotches are identified using the texture analysis.

In [63], Stoecker et al. design a blotches extraction algorithm divided into two parts. First of all the thresholding is made to find the regions and, secondly, they are analyzed through a series of features, using machine learning techniques. Results with an accuracy of 77 % are obtained.

For the detection of blotches, in [64], Pellacani et al. present two methods for the automated extraction of 'Absolute' (ADAS) and "Relative" dark areas (RDA) and a set of parameters for its description. It is tested in 339 images of melanocytic lesions, sifting from melanomas and nevus, obtaining a result of 74.2 and 71.2 % respectively.

In [65], Khan et al. use fuzzy techniques for the detection of blotches. A machine learning process is carried out, obtaining some rules, using neuronal networks. An accuracy of 81.2 % is obtained.

For the detection of parallel pattern, typical in acral volar melanomas, in [66], Iyatomi et al. propose a novel method, obtaining over a dataset of 213 images a sensitivity of 81.1 % and a specificity of 92.1 %, consisting in three subpattern detectors for the three typical parallel structures such as parallel ridge, parallel furrow and fibrillar pattern. They use color, texture and structural features, reduced by using PCA, and classified by a linear model.

If wanting to enlarge more, it can be consulted in [67], from Korotkov et al., that present a comprehensive survey in the field of computerized pigmented skin lesion image analysis. Finally, in [68], Celebi et al. present the editorial of the latest special issue on skin cancer image analysis.

2.3 Most Relevant Tools for the Automated Detection of Melanoma in Dermoscopic Images

A state of the art of some of the most relevant software systems for the automated detection of melanoma in dermoscopic images will be carried out.

To do so, first of all, a characterization of the automated systems in dermoscopic images will be undertaken, describing some of the criteria to be taken into account in those systems, secondly a brief description of the most relevant systems in the market will be made, and finally, thirdly, a comparison between them is carried out.

2.3.1 Characterization of the Software Systems for the Automated Detection of Melanoma in Dermoscopic Images

Some of the characteristics of interest which are used to be taken into account when characterizing a software system for the automated detection of melanoma in dermoscopic images will be carried out.

It should be noted, in any case, that the approach taken in this review is to place especially the standpoint in the analysis of the software systems in support to the

diagnosis used and it shall not be assessed the technical specifications of the hardware which in most cases are integrated with the software.

Integrated with proprietary hardware: Most of the relevant systems work with proprietary hardware, having been developed by the manufacturers themselves, which complement the acquisition of images with a software to support to the diagnosis. The advantage of this is that it assures a stronger integration between the software system and the acquired images. The disadvantage is the coupling this may induce to, and the worse flexibility when analyzing the not acquired with proprietary technology.

Teledermatoscopy: A system with teledermatoscopy support is capable of undertaking the distance transmission of dermoscopic images, for its remote analysis. The teledermatoscopy can be used with different purposes: for the medical user it allows to carry out the diagnosis from a remote center, or else a second diagnosis to review the diagnosis already carried out; for the non-medical user it allows the possibility of making the screening directly, via mobile or web, or else, it should be integrated in a telemedicine platform.

Multiplatform: Quite related to the previous one, this characteristic has to do with the different platforms from which it can be accessed to the diagnostic system, via web, mobile devices, via web services, via desktop application, etc.

Multiaccess: There exist systems which can be accessed by the public in general, other ones which can only be accessed by health professionals and other ones which allow both types of users.

Bodymapping: Currently, most of the systems have access to the bodymapping, that is to say, they are capable of storing and processing macroscopic images corresponding to parts of the body, with sets of lesions, mapped to dermoscopic images corresponding to the different lesions. Notwithstanding, for the accomplishment of this task there exist different degrees of sophistication, which allow a greater agility in the procurement and processing of this information.

Macroscopic images: Most of the systems of automated detection of melanoma are capable of analyzing macroscopic images of the lesion, carried out with a normal photo camera, generally without any magnification. Obviously, this analysis always has a worse accuracy, but in some occasions it can be convenient.

Classic dermoscopic images: By "classic dermoscopic" images we mean the ones obtained through the utilization of a dermatoscope. Usually, these images are defined as just dermoscopic, however, here we have established the convention of encompassing the multispectral ones within the dermoscopic ones, and to differentiate them they will be named as "classic dermoscopic" .

Multispectral dermoscopic images: Images obtained by a device using different wavelenghts of light, which penetrate the skin to different depths, which allow the visualization of invisible criteria to macroscopic images and classic dermoscopic ones. Various images in gray are obtained corresponding to the different bands on the spectrum which are handled.

Frequently quoted and positively evaluated in relevant publications in recent years: When a system is quoted in a relevant publication is regarded as a system of greater acceptance by the scientific and medical community, being also important

the assessment given to it. Obviously, there are always questionable systems of analyzing, depending on who carries out the experiment and how this is undertaken, the database tested, and so on.

High degree of penetration on the market of the health professionals: It is an important characteristic the analysis of the degree of penetration on the market, among the health professionals, since if this is high it can be an indicator that the experts trust in it. In any case, it always has to be taken into account that, in many instances the hardware and the software are sold integrated and, consequently, the sale of the software might be due to the hardware. And, of course, many other commercial considerations can be taken.

Approval by institutions: A key feature when it comes to assess the reliability of a diagnostic is its approval by the institutions.

Algorithm for the diagnosis: Obviously, all the software systems implement proprietary algorithms. Notwithstanding, this automations are really a guide to the diagnosis of the doctor, and since he is the one who takes the decision at the end, in many instances he prefers to do so supporting in algorithms accepted by the medicine. Consequently, normally, the diagnosis is accompanied by the quantitative information corresponding to a medical algorithm.

Accuracy: It is obvious that the most important feature of an automated system is its level of accuracy, and more specifically its sensitivity and specificity. Notwithstanding, it deals with a sensitive issue. Many corporations do not provide these data explicitly and, in any case, as it has been commented previously, these data need to be taken with great caution, since the accuracy data are always questionable data to analyze, depending on who does the experiment and how this was carried out, the database tested, whether the difficulty is higher or lower in the discernment of images, the factors to be taken into account, and so on.

2.3.2 Most Relevant Automated Systems for the Detection of Melanoma on the Market

Some of the, in our opinion, most relevant software systems for the automated detection of melanoma in dermoscopic images will be commented here, based on the previously mentioned criteria and always from the point of view of the software in support of the diagnosis.

Additionally, it should be also noted that, presently, due to the importance of the early detection of melanoma, the applications in macroscopic images aimed at the public in general are proliferating. To illustrate this type of tools, which are obviously less accurate, three systems which work exclusively with macroscopic images will be commented, two of them are mobile applications.

Finally, the Molemap, the most important screening worldwide program for the detection of melanoma, and Melanoscan, a screening program which is having good results lately, will be commented. In both cases the diagnostic software is not commercialized per se but, nevertheless, due to its importance, it has been considered appropriate to include them here.

As it has been stated previously, it should be noted that most of the relevant systems work with proprietary hardware, having been developed by the manufacturers themselves, which complement the acquisition of images with a software supporting the diagnosis though, in many cases, this was developed in principle by a research team external to the manufacturing company.

MoleMax™

It is a property of the Derma Medical Systems company (Austria) and it was developed in cooperation with the medical faculty of the University of Vienna [69]. It is a system which lasts present on the market for many years, and it works in classic dermoscopic images obtained with proprietary hardware. In fact, today MoleMax™ is a worldwide accepted clinical standard in digital epiluminescence microscopy. It also allows the analysis of macroscopic images, captured with digital cameras, and giving support for webcam. It is sold only to health professionals, although the interaction with the non-medical user system is allowed, for the accomplishment of screening and, in fact it is focused to teledermoscopy. It is multiplatform to a great extent, and it has been quoted and positively evaluated in various relevant publications. Its degree of penetration among the health professionals is very high. The diagnosis is undertaken in an automated way, using as diagnostic support the ABCD Rule and the 7-Point Checklist. Recent data about the accuracy of the Molemax diagnostic software are not available. It can be seen in Fig. 7.

Fig. 7 Molemax software tool, integrated with proprietary system. *Source* in [69]

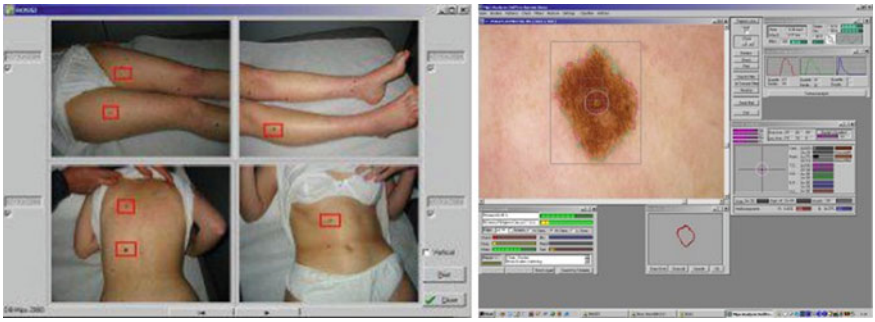


Fig. 8 DBDermo-Mips/DDAX the software tool. *Source* in [70]

DBDermo-Mips/DDAX

It is a property of the Biomips company (Italy) [70]. It is a system for the aided diagnosis of melanoma which works in classic dermoscopic images, on the basis of the DBDermo-Mips system which remains for many years on the market. The main addendum is its great speed, which provides the possibility of performing in real-time analysis of the patient lesions, without the necessity of storing previously the images. It is not integrated with proprietary hardware, but there are dermatoscopes which it is more integrated with. It is only directed to health professionals, being its level of penetration between them high and having been quoted and positively evaluated in some relevant publications. The diagnosis is carried out using as medical algorithm the ABCD rule, although it analyses a large number of indicators using the same similarity pattern and the company itself reports an accuracy with a mean sensitivity of 92 % and a mean specificity higher than 80 % [70]. It can be seen in Fig. 8.

Fotofinder

It is a property of the FotoFinder Systems Inc. company (Germany) [13]. It is a system which works in classic dermoscopic images obtained with proprietary hardware. It is widespread among the health professionals being aimed only at them, offering a large number of performances, among them an advanced capture software for the bodymapping. It is multiplatform, allowing use even mobile devices integrating an optical device. It has been quoted and positively evaluated in various relevant publications. The diagnosis is made in an automated way using the diagnosis Mole Analyzer tool, having as diagnostic support the ABCD Rule. No recent data about the accuracy of the diagnostic Mole Analyzer software are available. It can be seen in Fig. 9.

Fig. 9 Mole Analyzer tool, integrated with the Fotofinder system. *Source* in [13]

DermaGraphix/Mirror

It is a property of the Canfield Imaging Systems company (US) [71]. It is a system which works in classic dermoscopic images with proprietary hardware. It is quite established among health professionals, mainly in UK and US, and it has been quoted and positively evaluated in some relevant publications. The diagnosis is carried out in an automated way monitoring the evolution of the lesions on the skin and using for the diagnosis the ABCD rule. No relevant data about the accuracy of the diagnostic software is available. It can be seen in Fig. 10.

SIAscopy™-Molemate

It is a property of the MedX Health Corporation company (Canada) [72]. It is a system which works in multispectral dermoscopic images obtained with proprietary hardware, which are generated using the SIAscopy™ technology, and they correspond to eight consecutive discrete wavebands between 400 and 1,000 nm penetrating to a depth of 2 mm. It is directed only to the health professionals, being high its level of penetration and having been quoted and positively evaluated in some relevant

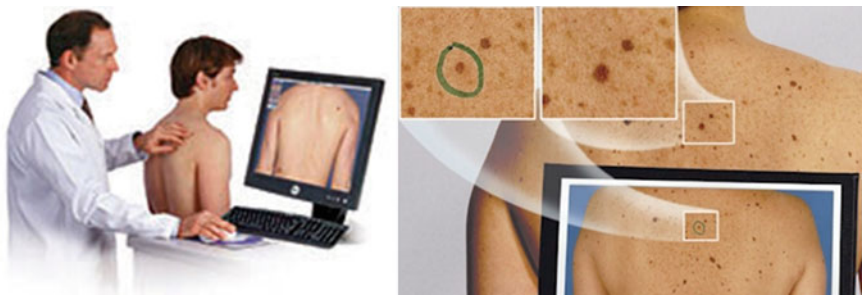


Fig. 10 Mirror software tool, Canfield dermatoscope. *Source* Ref. [71]

Fig. 11 Molemate system. *Source* Ref. [72]

publications. The diagnosis is made in an automated way, using the Molemate tool, on the basis of proprietary algorithms. On July 2012, the British Medical Journal published a study in which Molemate obtained good results, though no evidence that the MoleMate system improved appropriateness of referral was found and, neither if the systematic application of best practice guidelines alone was more accurate than the MoleMate system [73]. It can be seen in Fig. 11.

MelaFind®

It is a property of the MELA Sciences Inc. (U.S.) company [74]. It is a system which works in multispectral dermoscopic images obtained with proprietary hardware, corresponding to 10 distinct spectral bands from 430 to 950 nm, from near infrared through visible light spectrum, with 20 micron resolution. MelaFind analyzes the data using proprietary algorithms which work in a database containing multi-spectral data and histological diagnoses for more than 10,000 biopsied lesions from over 7,000 patients. It is directed only to dermatologists providing them with an output of yes or no to do the biopsy. It is considered presently as the most reliable diagnostic system on the market. In fact it has been the first software system accepted by the FDA of U.S. On February 2011, the Archives of Dermatology published a study on the diagnostic accuracy of MelaFind vs dermatologists. MelaFind achieved a sensitivity of more than 98 %, and the average specificity was 91.5 %. Both of these parameters were better than the clinical diagnoses made by the investigators [75]. It can be seen in Fig. 12.

Dermoscopix

It is a property of the Dermoscopix company (Germany) [76]. It is an automated system in support of the diagnosis which admits classic macroscopic and dermoscopic images, from different sources, not being integrated with any particular device of image capturing, allowing to be used with any camera-dermatoscope combination. It is multiplatform, being aimed at health professionals. The diagnosis is carried out

Fig. 12 Melafind. *Source* Ref. [74]

in an automated way, using as medical algorithm the Menzies Method. No recent data about the accuracy of the diagnostic software are available. It can be seen in Fig. 13.

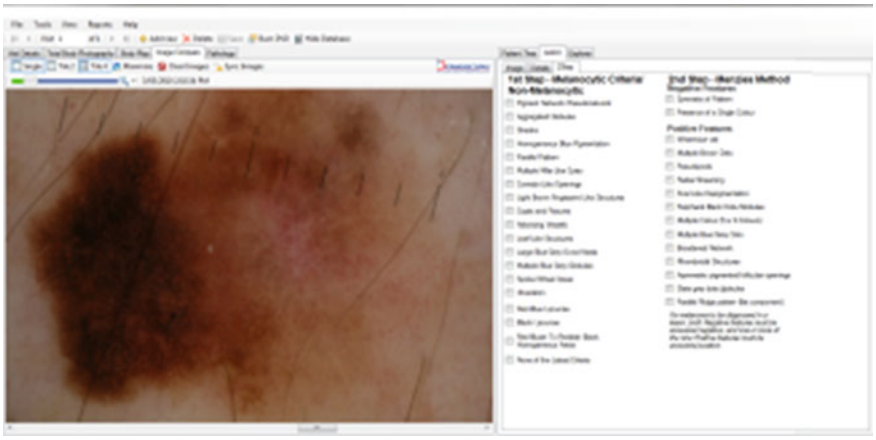


Fig. 13 Dermoscopix software. *Source* Ref. [76]

Fig. 14 Microderm. *Source* Ref. [77]

MicroDerm

It is a property of the VisioMed (Germany) company [77]. It is an automated system in support of the diagnosis in classic dermoscopic images, integrated with proprietary hardware. It is multiplatform, being aimed at health professionals. The diagnosis is carried out in an automated way, using as medical algorithm the ABCD rule. No recent data about the accuracy are available. It can be seen in Fig. 14.

MoleExpert

It is a property of the DatInf GmbH (Germany) company [78]. It has a support for macroscopic and dermoscopic images, allowing to monitor the evolution between images taken in different dates. It uses for the diagnosis the ABCD Rule, not reporting accuracy data. It can be seen in Fig. 15.

DermAlert®

It is a property of the Western Research Company Inc. (US) [79]. It works in macroscopic images, allowing the comparison between images of parts of the body taken

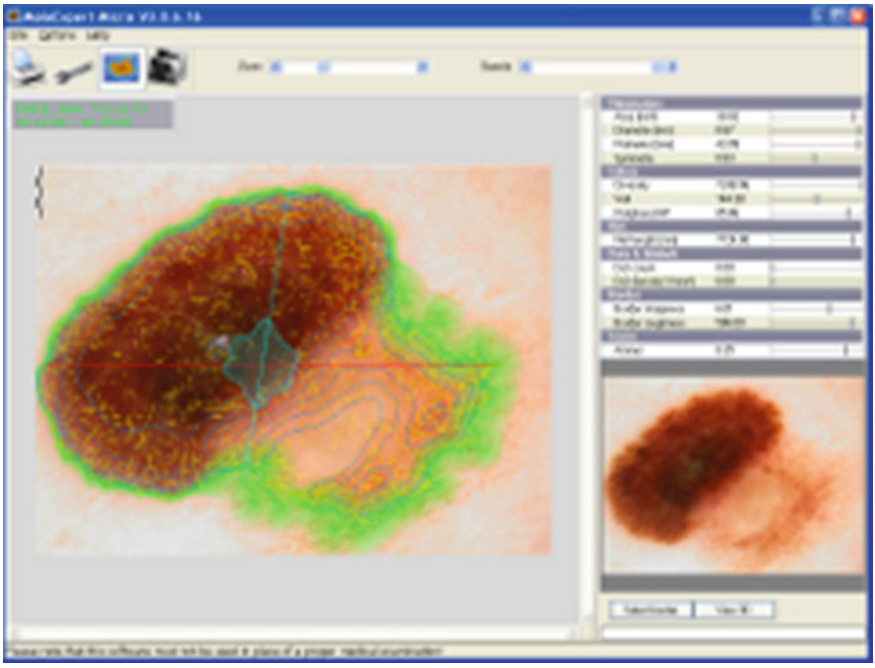


Fig. 15 Moleexpert. *Source* Ref. [78]

in different periods and the visualization of the existing differences. It uses for the diagnosis the ABCD Rule, not reporting the accuracy data. It can be seen in Fig. 16.

MelApp

It is a property of the Health Discovery Corporation company (US) [80]. It works in dermoscopic images and it is available for Iphone and Android. It is supported by a database of images over which it uses recognition of patterns using the machine learning Support Vector Machines (SVM) technique, carrying out using the ABCD rule. It can be seen in Fig. 17.

UMSkinCheck

It is a property of the University of Michigan (US) [81]. It works in macroscopic images and it is available for Iphone. It proposes a protocol based on the taking of 23 images, covering the whole body, whose aim is to monitor the evolution, being this its main innovation. It can be seen in Fig. 18.

Molemap

It is a property of the Molemap company (Australia) [82]. It is the world leading melanoma detection program for screening tasks with the aim of helping to identify melanoma skin cancer at an early stage. It comprises other programs such as the

Fig. 16 DermAlert®. *Source* Ref. [79]

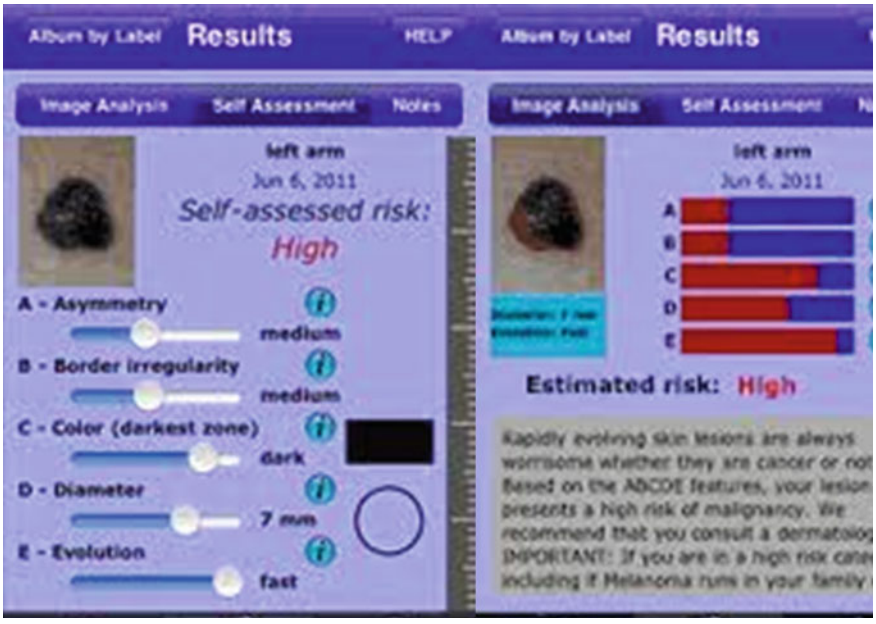


Fig. 17 MelApp. *Source* Ref. [80]

Molesafe [83], existing a network of centers worldwide, in which the specialists take the images, consisting on a full scan of the body of the patient, and classic dermoscopic ones corresponding to the lesions which are seen irregular or have changed from a previous review. The images are sent to a diagnostic center, where

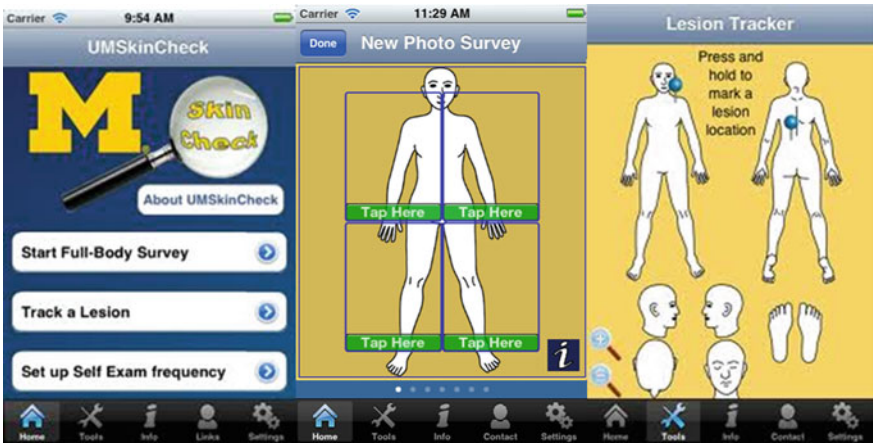


Fig. 18 UMSkinCheck tool. *Source* Ref. [81]

expert dermatologists analyze the images and draw a report. They report a high accuracy, though no quantitative data are specified. It is really a screening program and it is aimed at the patients directly, it does not commercialize diagnostic software as such, nevertheless, due to the importance of this program it has been considered adequate to include it here. It can be seen in Fig. 19.

Melanoscan®

It is a property of the Melanoscan Inc. company (US) [84]. It is a screening system developed recently, with very encouraging results. It undertakes a rapid bodyscanning seeking suspicious lesions, from which it obtains more detailed images. If it has previous studies, it carries out comparisons in real time. It is really a screening program and it is aimed at the patients directly, it does not commercialize diagnostic software as such, nevertheless, due to the importance of this program it has been considered adequate to include it here, since currently it is being used increasingly in US. For the diagnosis it uses the ABCD algorithm. In the study [85] a sensitivity of 75 % is reported and a specificity of 73.70 % is reported. It can be seen in Fig. 20.

2.3.3 Comparative Among Systems

Hereafter the features which are, as far as we are concerned, the most relevant ones in each system of the previously described ones are to be shown, together with the algorithms which are used in the diagnosis, the accuracy reported for these (in the event of existing) and, in those cases where such information is reported by the company, the information of the used datasets. It can be seen in Table 3.

Fig. 19 Molemap. *Source* Ref. [82]

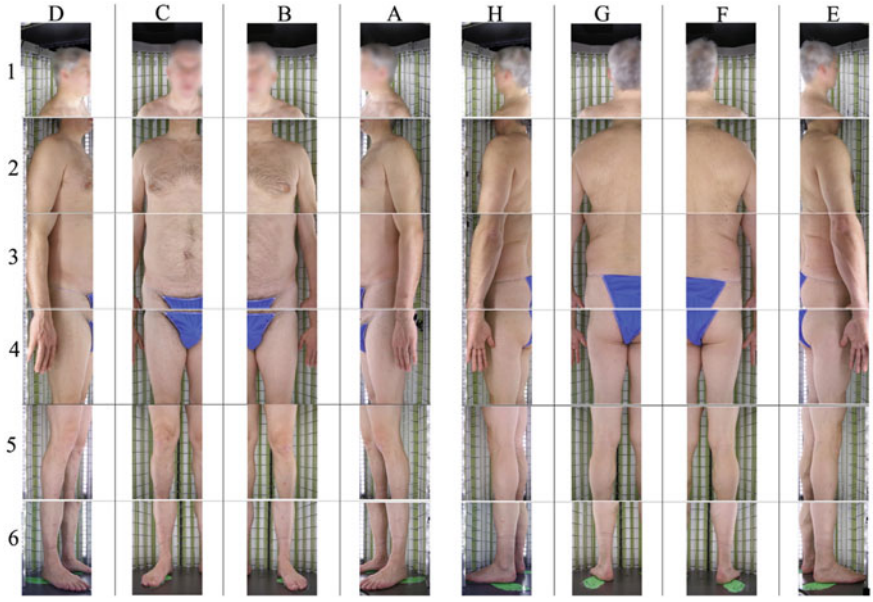


Fig. 20 Melanoscan. *Source* Ref. [84]

Table 3 Comparative of the systems

	proprietary hardware	teledermoscopy	multiplatform	bodymapping	macroscopic	classic. dermoscopic	multispectral	approval	medical algorithm	accuracy	database
MoleMax™	☑	☑	☑	☑	☑	☑			ABCD/ 7-point checklist		
DBDermo-Mips/DDAX			☑			☑			ABCD	Sensit. 92% Specif. 80%	Not reported
Fotofinder	☑		☑	☑		☑			ABCD		
DermaGraphix-Mirror	☑			☑		☑			ABCD		
SIAscopy™-Molemate	☑						☑		Proprietary algorithms		
MelaFind®	☑						☑	☑	Proprietary algorithms	Sensit. 98% Specif. 91.5%	1632 images / 127 melanomas
Dermoscopix		☑	☑		☑	☑			Menzies' method		
MicroDerm		☑	☑			☑			ABCD		
MoleExpert		☑	☑		☑	☑			ABCD		
DermAlert®		☑	☑		☑				ABCD		
MelApp		☑	☑		☑				ABCD		
UMSkinCheck		☑	☑	☑	☑				ABCD		
Molemap	☑	☑		☑		☑			Experts		
Melanoscan®	☑			☑		☑	☑		ABCD	Sensit. 75% Specif. 73.7%	394 images

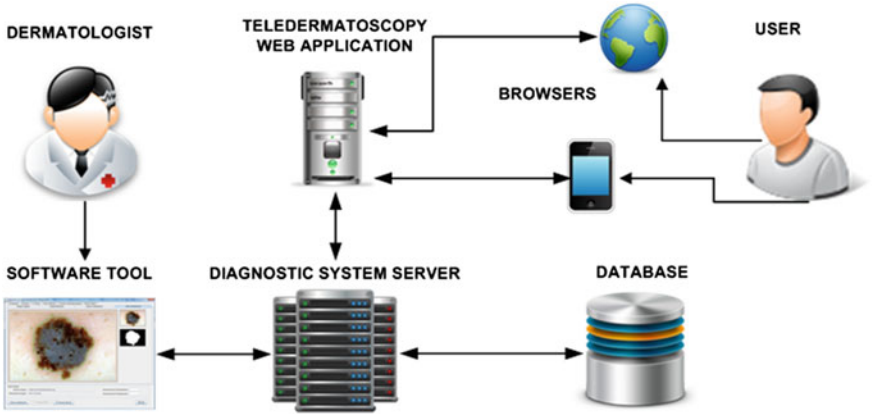


Fig. 21 Architecture of the system

3 System Design

Our research team designed and developed a complete family of digital image processing algorithms which have been integrated in an automated system in support of the detection of melanoma, which is to be presented in this section. It will start with a high level to thereupon explain the design of the system in its different parts.

3.1 High Level View

It will start commenting the architecture and the design of the system.

As it can be seen in Fig. 21, the system is designed multiplatform and multiaccess, offering the possibility of analysing macroscopic and classic dermoscopic images. The doctor can access to it via software tool to the diagnostic system, which is integrated with the database. The non-medical users, for screening works, can access via mobile phone or via web, to the application of teledermoscopy, which is integrated with the own diagnostic system.

The system core is the software tool and the diagnostic system, and its analysis of the dermoscopic images, taken from different sources, from different manufacturers, not being integrated with any particular hardware. As it can be seen in Fig. 22, four modules are taken, corresponding to the four last tasks of the life cycle of an automated system of detection of melanoma, explained in the previous section.

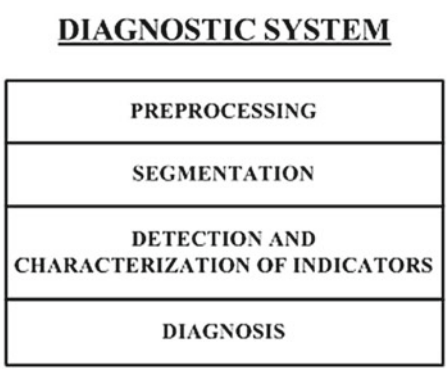


Fig. 22 Design of the diagnostic system

3.2 Preprocessing

For the undertaking of the preprocessing in the original image IMG_SOURCE the following steps are carried out:

P1—Generation of the color map of the complete image IMG_SOURCE. This color map takes advantage for the preprocessing and the subsequent C of the ABCD.

P1.1—The color map is carried out through a machine learning supervised process. In 50 images doctors were required to catalogue the pixels of an image sample of each of the colors: white, red, light-brown, dark-brown, blue-gray and black.
P1.2—From the selected pixels features from different color spaces are extracted: RGB, rgb (normalization of RGB), HSV, CIEXYZ and CIELab, classifying each one of the pixels in the 6 described colors.
P1.3—Through the obtained data a tree classifier was used, the C4.5 algorithm being chosen to generate the decision tree, which possesses the classification rules [86]
P1.4—Through the generated rules, the partition of the lesion is made in regions corresponding to the 6 colors of ABCD Rule named R_WHITE, R_RED, R_LIGHT_BROWN, R_DARK_BROWN, R_BLUE_GRAY and R_BLACK.

P2—Detection of the black frames:

P2.1—From each one of the edges of the image (up, right, down and left) it starts to examine the parallel lines
P2.2—For each parallel line:
P2.2.1—The pixels are gone through. For each pixel:
- *P2.2.1.1—It is calculated as darkness measure: darknessMeasure=* $(max(R, G, B) + min(R.G, B))/2$.
- *P2.2.1.2—The condition is evaluated* $(darknessMeasure < THR\_DARKNESS\_PIXEL)$, *which is a threshold darkness value.*

P2.2.2—If the THR_PERCENT_PIXELS_LINE% of the pixels fulfills the previous condition, the line as part of the frame is preselected, where that percentage is a threshold value.

P2.2.3—The median of the red and blue of the pixels of the line are calculated, medianRed and medianBlue, and the variance of the Red varianceRed.

P2.2.4—The condition is evaluated $((medianRed - medianBlue) < THR\_MEDIANS\_RED\_BLUE) AND (varianceRed < THR\_VARIANCE\_RED)$*, where both of them are threshold values. If the line is fulfilled it is from the black frame.*

P2.3—The previous process ends when one of the parallel lines does not fulfill the conditions to be selected.

P2.4—The mask of the black frames is the attachment of all the selected parallel lines, in each one of the directions, naming the region R_MASK_MARKS.

P3—Detection of the flashes and bubbles:

P3.1—All the pixels are gone through and it is compared $(R >= THR\_FLASH\_BUBBLES) OR (G >= THR\_FLASH\_BUBBLES) OR (B >= THR\_FLASH\_BUBBLES)$*, where* $THR\_FLASH\_BUBBLES$ *is a threshold value.*

P3.2—The mask with the flash pixels and bubbles will be named R_MASK_FLASH_BUBBLES.

P4—Detection of rules:

P4.1—The dark side of the image is sought R_MASK_DARK through the threshold value THR_DARKNESS.

P4.2—The dark side of the R image which is not frame is sought R_MASK_DARK_2= R_MASK_DARK- R_MASK_MARKS.

P4.3—To the previous one, the regions corresponding to white, red and light brown from the color map is subtracted. R_MASK_DARK_3 = R_MASK_DARK_2- R_WHITE- R_RED- R_LIGHT_BROWN.

P4.4—The components of 8-connected from R_MASK_DARK_3 higher than an amount of threshold pixels THR_NUM_MIN_PIXELS_MAJOR, with the aim of taking large pieces of the rule.

P4.5—With each of the founded components C_MAJOR:

P4.5.1—It is preselected if it does not intersect a circle with center and radius a value THR_RADIO_CIRC_CENTER, to avoid selecting part of the rule to the lesion.

P4.5.2—The median of the color of the values of C_MAJOR, is calculated, naming colorMedian_M.

P4.5.3—C_MAJOR is selected if the conditions are fulfilled:

- *P4.5.3.1—If it has big holes in the middle the proportion of C_MAJOR_H is calculated, which is the C_MAJOR with holes bigger than a threshold value THR_SIZE_BIG_HOLES filled: percent=Area(C_MAJOR_H)/ Area(C_MAJOR) and is selected if* $(percent > THR\_COMPARE\_ORIG\_FILLED)$*, which is a threshold value.*

- *P4.5.3.2—If the color is similar to THR_COLOR_M: if* $||colorMedian\_M\text{-}THR\_COLOR\_M|| < THR\_DIFF\_MEDIANS$*, where THR_COLOR_M is a threshold color and THR_DIFF_MEDIANS is a threshold value.*

P4.5.4—If C_MAJOR has been selected C_MAJOR_NO_NOISE is calculated, with C_MAJOR filling of the holes with a smaller size than a threshold value THR_SIZE_SMALL_HOLES, with the aim of removing noise pixels which may be like holes of the mask .

P4.6—The union of all of the C_MAJOR_NO_NOISE obtained in R_MASK_UNION_MAJORS.

P4.7—If there is any C_MAJOR selected, that is to say, if there are big pieces of rule, the components 8-connected of R_MASK_DARK_3 are selected, lower than the amount of pixels THR_NUM_MIN_PIXELS_ MAJOR already defined and higher than a threshold value THR_NUM_MIN_PIXELS_MINOR, with the aim of taking small pieces of the rule.

P4.8—With each one of the components found C_MINOR:

P4.8.1—C_MINOR is selected if it fulfills the conditions:

- *P4.8.1.1—If* $||colorMedian\_M - THR\_COLOR\_M|| < THR\_DIFF\_MEDIANS$*, where THR_COLOR_M and THR _DIFF_MEDIANS are thresholds already defined.*
- *P4.8.1.2—If the distance of C_MINOR to R_MASK_ UNION_MAJORS is lower than a threshold value THR_DIST_MAJOR_MINOR.*

P4.8.2—If C_MINOR has been selected C_MINOR_NO_NOISE is calculated , with C_MINOR filling of the holes of a lower size than a threshold value THR_SIZE_SMALL_HOLES.

P4.9—The union of all of the C_MINOR_NO_NOISE are calculated, obtained in R_MASK_UNION_MINORS.

P4.10—The mask with the pixels of the rule will be named R_MASK_ RULE and it is the result of the union of R_MASK_UNION_MAJORS and R_MASK_UNION_MINORS.

P5—Calculation of the color of the skin:

P5.1—The region of the skin pixels is calculated R_MASK_COLORS_SKIN as the ones which fulfill the condition of $(R > THR\_RED)AND(R > G)AND(R > G)$ *where THR_RED is a threshold value.*

P5.2—From that region 3 masks already defined are subtracted R_MASK_SKIN =R_MASK_COLORS_SKIN-R_MASK_MARKS-R_MASK_FLASH_BUBBLES-R_MASK_DARK.

P5.3—The color of the skin skinColor is calculated as the value of the median of the pixels of the region R_MASK_SKIN.

P6—Detection of the mask with the hairs. The idea is similar to that one exposed in [15] *by Lee et al., with some modifications:*

P6.1—A morphological closing is made in the 3 RGB bands.

P6.2—The 3 bands RGB are subtracted to the 3 bands with the closing accomplished.
P6.3—A grey image is created in which the value of each pixel is the minimum of the 3 gray images created in the previous stage.
P6.4—Through the previous stage a mask R_MASK_HAIR_1 is created through a threshold value THR_HAIR.
P6.5—The 8-connected components of the R_MASK_HAIR_1 are selected, higher than an amount of threshold pixels THR_NUM_ MIN_PIXELS_HAIR, and the union of all of them is made, creating R_MASK_HAIR_2, with the aim of removing loose pixels.
P6.6—R_MASK_HAIR=R_MASK_HAIR_2-R_WHITE-R_RED-R_LIGHT _ BROWN-R_DARK_BROWN is calculated, that is to say, 4 of the values of the initial map of colors are removed with this, the trusses are removed or from other type which should not be confused with hairs R_MASK_HAIR is the mask with the hairs.

P7—Substitution of the detected artifacts pixels:

P7.1—The pixels of the mask R_MASK_MARKS and R_MASK_RULE are substituted by the skin color skinColor.
P7.2—The pixels of the masks R_MASK_FLASH_BUBBLES and R_MASK_ HAIR are substituted by a bilinear interpolation.

P8—For its use in the C of ABCD a mask is kept with the union of all the substituted artifacts R_MASK_ARTIFACTS, since the pixels of that mask should be classified once again for the map of colors, since the initial classification is made from the original one and not from the preprocessed one.
P9—As the output of the process it is taken that through the original image IMG_SOURCE a preprocessed image has been generated IMG_PREPROC.

3.3 Segmentation

Here the automated segmentation accomplished by the application will be commented in detail:

S_SEG_1—Determination of coefficients (COEF_RED,COEF_GREEN, COEF _BLUE) for a linear transformation of RGB to a grayscale which allows the optimization of the segmentation using the Otsu Method. The idea is similar to the one exposed in [22] *of Delgado et al., with some modifications. Those coefficients will fulfill the condition of for all of the pixels of IMG_PREPROC, the image in gray IMG_PREPROC_GRAY in which IMG_PREPROC_GRAY(x,y)= COEF_RED*IMG_PREPROC(x,y,1)+COEF_GREEN*IMG_PREPROC(x,y,2) +COEF_BLUE* IMG_PREPROC(x,y,3) will be segmented by the Otsu Method* [87]*in an optimal way.*

S_SEG_1.1—In 110 images the doctors were required to carry out the segmentation of the lesions, generating 110 masks R_MASK_SEG_MED.
S_SEG_1.2—An experiment is carried out with different combinations of coefficients (COEF_R,COEF_G, COEF_B), positive numbers which fulfill the condition (COEF_R+COEF_G+COEF_B)=1.
S_SEG_1.3—With each (COEF_R,COEF_G,COEF_B):
S_SEG_1.3.1—With each one of those 110 images, through the preprocessed images IMG_PREPROC, the linear transformation is undertaken with the coefficients (COEF_R,COEF_G,COEF_B), a blurring is carried out applying to the low pass filter, a Gaussian filter with 3 pixels of radius, and subsequently the segmentation is made using the Otsu Method, generating a mask. The 8-connected component of higher size is selected, which is finally smoothed by using a morphological opening, having a result R_MASK_TEST.
S_SEG_1.3.2—A measure to determine whether the segmentation was correct is defined: measureSeg=Area(Intersection(R_MASK_TEST,R_MASK_SEG_MED))/Area(Union(R_MASK_TEST, R_MASK_SEG_MED)).
S_SEG_1.3.3—The segmentation is considered good if $(measureSeg > THR\_SEG)$*.*
S_SEG_1.3.4—The number of good segmentations are counted numSegOk.
S_SEG_1.4—The coefficients with a greater number of good segmentations are selected numSegOk. Those coefficients are: (COEF_RED, COEF_GREEN, COEF_BLUE).

S_SEG_2—To segment an image IMG_PREPROC:

S_SEG_2.1—The image in gray is generated IMG_PREPROC_GRAY through the coefficients (COEF_RED,COEF_GREEN,COEF_BLUE).
S_SEG_2.2—A blurring applying to the low pass filter and a Gaussian filter with 3 radius pixels is carried out.
S_SEG_2.3—The segmentation using Otsu Method is carried out.
S_SEG_2.4—The 8-connected component of larger size is selected.
S_SEG_2.5—It is smoothed using a morphological opening.

S_SEG_3—Through the previous segmentation the mask of the lesion is generated R_MASK_SEG.

3.4 Detection and Characterization of Indicators

In the state of art section it was already commented that in the works of computer aided diagnosis of melanoma there exist two approaches: on the one hand, those which are based on a medical algorithm and on the other hand, those which are not based on any medical algorithm, and which consists on the extraction of features followed by a classification algorithm in the values of those features, based on a set of rules generated in a machine learning process.

Here the first approach was chosen. The choice of a medical algorithm as a basis of the diagnosis enables the analysis to have their basis on the existing extensive literature, which provides the experts with more confidence.

In the system indicators corresponding to the medical algorithm of the "ABCD rule" are sought, having selected as diagnostic algorithm the "ABCD Rule" for being easier to objectify and quantify.

A: Asymmetry

The algorithm for the calculation of the asymmetry is decomposed into the following steps:

A1—Calculation of the axes: higher and lower. The major axis will be an axis which will cut the lesion by the middle in two halves, more or less equal and symmetrical, and the lower axis will be the perpendicular to this by the middle.

A1.1—The centroid of the lesion of the mask R_MASK_SEG is calculated, which is assumed to be the origin of the coordinates.
A1.2—A threshold value has been defined THR_DIFF_PIXELS. Through the calculated point, and assuming that point a coordinate axis with parallel axes to the border of the images some straight line segments of the first quadrant which pass through that point and cut the mask of the lesions in two halves which have as maximum an area difference of THR_DIFF_PIXELS. The one with longer length is selected.
A1.3—The same thing as in the previous point is made, with the straight segments of the second.
A1.4—The segment of longer length is selected among the ones of the first and the second quadrant: the major axis is already available.
A1.5—The minor axis will be the segment perpendicular to the major axis which passes through the center point.

A2—The calculation of the asymmetry in shape with respect to the major and minor axis:

A2.1—With respect to the major axis:
A2.1.1—The left and right regions of R_MASK_SEG are calculated with respect to the major axis: R1 and R2.
A2.1.2—The symmetric region with respect to the major axis of one of the two halves, of the R2, for instance, is calculated: It is obtained S2=Sim(R2).
A2.1.3—The intersection of R1 and S2 is calculated: I= Intersection (R1,Sim (R2)).
A2.1.4—The union of R1 and S2: U=Union(R1, Sim(R2)).
A2.1.5—Hereafter the number of pixels of the intersection I is divided by the number of pixels the U union, which we will name intersectionPercent.
A2.1.5—The next condition is evaluated: (intersectionPercent<THR_INTERSECTION_PERCENT),*being this a threshold value, in positive case it is asymmetric in shape with respect to the major axis.*
A2.2—The same with respect to the minor axis.

A3—The calculation of the asymmetry in color with respect to the major and minor axis:

A3.1—With respect to the major axis:

A3.1.1—The RGB histograms are calculated with respect to both halves.

A3.1.2—The means of the RGB histograms of both halves are obtained, finding meanHistMajor1(meanRedHistMajor1,meanGreenHistMajor1,meanBlue HistMajor1)and meanHistMajor2(meanRedHisMjor2,meanGreenHist-Major2,meanBlueHistMajor2).

A3.1.3—The Euclidean distance is found: distMeansHistMajor=||meanHist-Major1-meanHistMajor2||

A3.1.4—The condition is evaluated: ($distMeansHistMajor$ > THR_ MEAN_HIST), *in positive case, it is asymmetric in color with respect to the major axis.*

A3.2—The same with respect to the minor axis.

A4—The asymmetry is calculated in each axis: if with respect to the axis it is asymmetric in shape or color, then it is asymmetric.

B: Borders

The algorithm for the calculation of borders is decomposed in the following stages:

B1—Calculation of bisectors and determination of the octants:

B1.1—The 2 bisectors are calculated.

B1.2—The skin lesion is divided in 8 octants.

B2—Calculation of the crown for the analysis of the borders. Through THR_WIDTH_INTERIOR_BORDER and THR_WIDTH_EXTERIOR _BORDER, threshold values, the crown is calculated with the origin the border of the lesion, which has a width on the inside THR_WIDTH_INTERIOR_ BORDER and on the outside THR_WIDTH_EXTERIOR_BORDER.

B3—Calculation of radii: The 7 radii which divide the angle of the octant in 8 pieces of PI/32 radians are calculated.

B4—The study of the graphics of the radii in each octant:

B4.1—The graphic with the values of each one of the radii in the grayscale in (WIDTH_INTERIOR + WIDTH_EXTERIOR + 1) values corresponding to the interval of the segment corresponding to the radius inside the crown.

B4.2—Given a threshold constant THR_DIFFERENCES_RANGE in all the values of the interval the wider of the differences which are given in each one of the subintervals of the size THR_DIFFERENCES_ RANGE is sought, satisfying the fact of being more or less monotonous the increasing, which is calculated through a threshold value THR_CHANGES_IN_SUCCESSIVE_POINTS, controlling the fact that in that interval would not vary in more than that value. For each one of the radii it is obtained a maxDiffRadio value.

B5—For each octant the degree of abruption of the border is the maximum of the maxDiffRadio of each of the radii found in the previous point, this value is called abruptionLevel.
B6—For each octant the degree of abruption of the border is calculated: the found value is compared with a threshold constant THR_ABRUPT_BORDER, which determines if this degree of increasing can be considered as an abrupt border or not, if (abruptionLevel > THR_ABRUPT_BORDER)*then it is abrupt.*

C: Colors
The determination of the colors in the application is shown thereupon:

C1—Generation of the map of colors:

C1.1—Through the supervised machine learning process described in the image processing section, the masks through the original image IMG_SOURCE were generated, corresponding to the 6 colors, named R_WHITE, R_RED, R_LIGHT_BROWN, R_DARK_BROWN, R_BLUE _GRAY and R_BLACK.
C1.2—The intersection of these masks with the mask of the lesion R_MASK_SEG is calculated.
C1.3—To generate the definitive masks R_WHITE, R_RED, R_LIGHT _BROWN, R_DARK_BROWN, R_BLUE_GRAY and R_BLACK corresponding to the 6 colors of the lesion, it only needs to be determined the pixels of the masks R_MASK_ARTIFACTS of the preprocessed image IMG_PREPROC. Through the rules generated in this process of machine learning the different pixels of this mask are made to correspond to each one of the masks of the 6 colors.

C2—The percentages of each one of the regions with respect to the total of the lesion, selecting the colors with the regions which pass through a percentage THR_PORCENTAJE_COLOR.

D: Dermoscopic Structures
The detection of the dermoscopic structures is shown thereupon:
Homogeneous and unstructured areas

D_HOM_UNSTR1—Determination of the blue homogeneous, hiperpigmentation and hypopigmentation: they are calculated through the color map generated in the previous stage. The blue homogeneous region coincides with the blue-gray region, the hiperpigmentation region coincides with the union of the dark brown and black regions and the hypopigmentation region coincides with the white region. Thus, 3 regions are generated: R1, R2 and R3.
D_HOM_UNSTR2—For each one of the regions the union of the 8-connected components of large size to a threshold value are selected of size bigger than a threshold THR_SIZE_HOMOG_UNSTR, obtaining the regions C1, C2 and C3.
D_HOM_UNSTR3—The percentages of each one of the previous regions are calculated with respect to the total lesion. We shall obtain the percentages, percent1, percent2 and percent3.

D_HOM_UNSTR4—It is seen which of the previous ones are superior to a threshold THR_PERCENT. If even one of the percentages will be superior it would already be punctuation in this criterion.

MACHINE LEARNING PROCESSES FOR: 1—PIGMENT NETWORK; 2—GLOBULES, POINTS AND PSEUDOPODS

For the detection and characterization of, in the first place the reticular pattern, and in the second place the globular, points and pseudopods patterns two supervised machine learning processes have been carried out, with similar features.

For the first process the experienced doctors were required to catalogue in a sample of images the reticular and non-reticular pixels. For the second process, the same with the globular, points and pseudopods pixels and the ones not satisfying that condition.

For both processes color features have been extracted, DoG (Difference Of Gaussians), Laplacian of Gaussians (LoG), Determinant of Hessian (DoH), gradient, and also typical texture variables based on the Gray Level Co-occurrence Matrix (GLCM): entropy, contrast and correlation.

According to the data obtained, in both cases tree classifiers were used, for efficiency reasons, RandomForest algorithm [88] *being chosen to generate the decision trees, with the classification rules.*

Reticular

D_RETIC1—Determination of the region with the pixels which fulfill the rules generated in the machine learning process for the determination of the pigment network pixels: R1

D_RETIC2—The intersection of the previous rule is carried out with the color R_LIGHT_BROWN, R_DARK_BROWN and R_BLACK: R2.

D_RETIC3—The 8-connected components higher than an amount of pixels are selected THR_NUM_PIXELS_RET_REGIONS.

D_RETIC4—With each one of the found C components the holes which it has been calculated (since the reticular pattern has network shape):

D_RETIC4.1—To do so first of all the presence of holes is calculated, calculating the percentage of the C_H area, which is the mask with all holes of more than one threshold value THR_NUM_PIXELS_HOLE filled, with respect to the area of C: percent=Area(C_H)/Area(C).

D_RETIC4.2—C The condition is evaluated with respect to one threshold value (percent> THR_COMPARE_ORIG_FILLED).

D_RETIC4.3—In positive case, the possessing holes are calculated, with an area superior to the THR_NUM_PIXELS_HOLE value, focusing this value in numHoles.

D_RETIC4.4—C is selected if numHoles has more than one threshold value THR_NUM_HOLES holes, that is to say, if (numHoles>THR_NUM_HOLES).

D_RETIC5—The union of all the C regions which fulfill the previous conditions is made, generating the R_MASK_RET region.

D_RETIC6—The addition of all of the numHoles corresponding to the C regions which fulfill the previous conditions is calculated, sumNumHoles=sum(numHoles).
D_RETIC7—If the number of sumNumHoles is higher than a threshold value THR_NUM_MIN_HOLES, that is to say if (sumNumHoles> THR_NUM_ MIN_ HOLES)*then it has the reticular pattern, with R_MASK_RET mask.*

Previous to Globular, Points and Pseudopods

D_GLOB_POINT_PSEUDO1—Determination of the region with the pixels which fulfill the rules generated in the machine learning process, for the determination of the globule, points and pseudopod pixels: R1.
D_GLOB_POINT_PSEUDO2—The intersection of the previous region is undertaken with the color masks R_LIGHT_BROWN, R_DARK_BROWN, R_BLUE_ GRAY and R_BLACK: R2.
D_GLOB_POINT_PSEUDO3—A subtraction of the reticular mask is made R_MASK_RET: R3.
D_GLOB_POINT_PSEUDO4—For each 8-connected C component of R3, those which are higher than the value THR_PIXELS_REGIONS_GLOB_ PSEUDO pixels are considered as potential globule or pseudopods, and those which are lower are considered as potential points, though higher than a value THR_ PIXELS_REGIONS_POINTS.

Previous to Globular and Pseudopods

D_GLOB_PSEUDO1—A crown similar to the one created in the borders section is created through THR_WIDTH_INTERIOR_GLOB_PSEUDO and THR_WIDTH_ EXTERIOR_ GLOB_PSEUDO.
D_GLOB_PSEUDO2—For each 8-connected C component (already described previously):

D_GLOB_PSEUDO2.1—The maximum diameter is calculated maxDiameter and the elongation, where the formulation is elongation=maxDiameter /area(C).
D_GLOB_PSEUDO2.2—If C satisfies the condition (maxDiameter>THR_ MIN_DIAM_GL) AND (maxDiameter<THR_MAX_DIAM_GL) AND (elongation< THR_ELONGATION) *C is globule, being thresholds the described values.*
D_GLOB_PSEUDO2.3—If C is contained in the crown by the border of the lesion and (maxDiameter>THR_MIN_DIAM_PSEU) AND (maxDiameter <THR_MAX_DIAM_PSEU) *C is a pseudopod, being thresholds the described values.*

Globular

D_GLOB1—The regions C which fulfill the condition described of globule are counted: numGlob, and the union of these in R_MASK_GLOB mask.
D_GLOB2—If (numGlob>THR_MIN_GLOB)*is fulfilled, then it has the globular pattern, with R_MASK_GLOB mask.*

Points

D_POINTS1—The C regions which fulfill the condition described of point: numPoints, and the union of these in R_MASK_POINTS mask.
D_POINTS2—If (numPoints>THR_MIN_POINTS)*is fulfilled, then it has the points pattern, with R_MASK_POINTS mask .*

Pseudopods

D_PSEUDO1—The C regions which fulfill the condition described of pseudopod: numPseudo, and the union of these in R_MASK_ PSEUDO mask.
D_ PSEUDO 2—If (numPseudo>THR_MIN_PSEUDO) *is fulfilled, then it has the pseudopods pattern, with R_MASK_ PSEUDO mask.*

3.5 Diagnosis

The diagnosis is carried out with the "ABCD Rule", from the calculated values of the indicators. Concerning the described methods, an observation, in the algorithms intervene some numerical values which start by "THR", which correspond to determined thresholds. The values of these thresholds were obtained by a statistic study as well as by an empirical study.

Hereafter, in Fig. 23 some images of the complete process can be seen, captured using the software tool, with the diagnosis at the end.

4 Results

In this section the results of the system will be presented. The image database used to carry out the research, was created with the collaboration of J.L. Díaz and J. Gardeazabal, dermatologists from IMQ clinic company, in Bilbao, Spain. It consists of 1179 images, with a resolution of 768×512 pixels. All of the images are cataloged by experts, having 287 images of melanoma.

4.1 Preprocessing

When applying the preprocessing in the different images satisfactory results were obtained. The results of the preprocessing in a visual way will be showed thereupon in Fig. 24, with various images, to illustrate it.

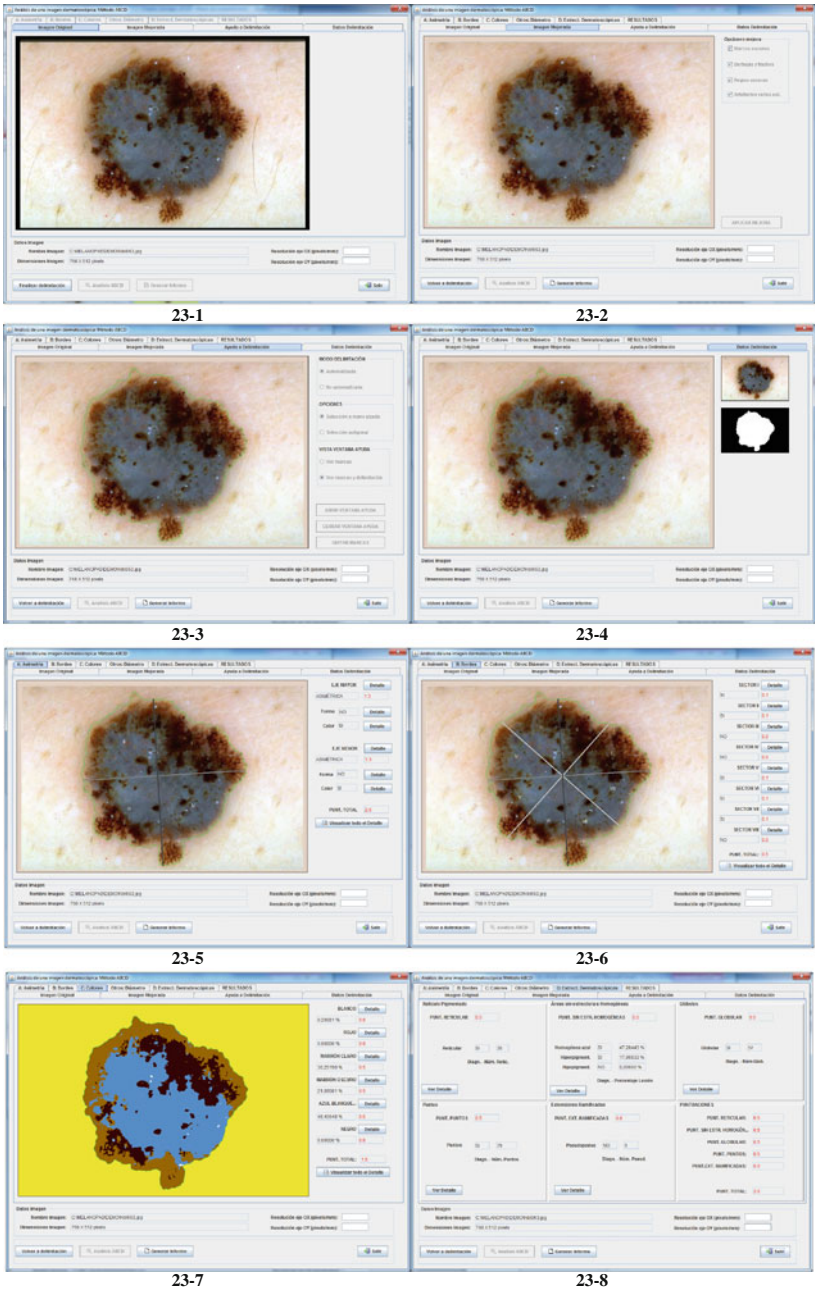
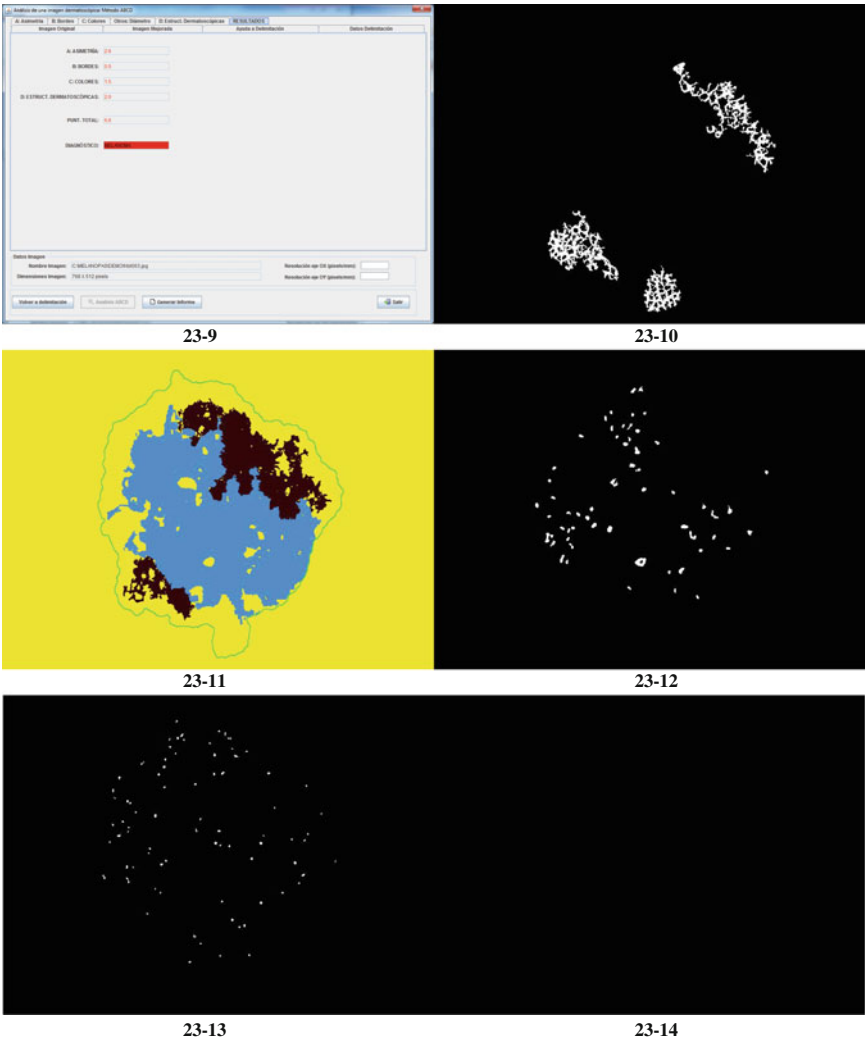


Fig. 23 Some windows of the software tool: *1* Original image; *2* Preprocessed; *3* Segmented with help to segmentation; *4* Segmented; *5* A: Assimetry; *6* B: Borders; *7* C: Colors; *8* D: Dermoscopic structures; *9* Results; *10* Reticular Mask; *11* Homogeous and unstructured areas; *12* Globular mask; *13* Points Mask; *14* Pseudopods mask; *15* Diagnosis in the generated report



23-9 23-10

23-11 23-12

23-13 23-14

Fig. 23 (continued)

DIAGNOSIS

A: ASIMMETRY: **2,6 p.**
MAJOR AXIS: YES MINOR AXIS: YES
SHAPE: NO SHAPE: NO
COLOR: YES COLOR: YES
B: BORDER: **0,5 p.**
ABRUPT: I, II, V, VI, VII
NO ABRUPT: III, IV, VIII
C: COLOR: **1,5 p.**
COLORS: LIGHT BROWN, DARK BROWN, BLUE-GRAY
D: DERMOSCOPIC PATTERNS: 2 p.
RETICULAR: YES GLOBULAR: YES
POINTS: YES PSEUDOPODS: NO
HOMOG. AND UNSTRUCT.: YES

RESULTS: **6.6 p.**
DIAGNOSTIC: MELANOMA

Fig. 23 (continued)

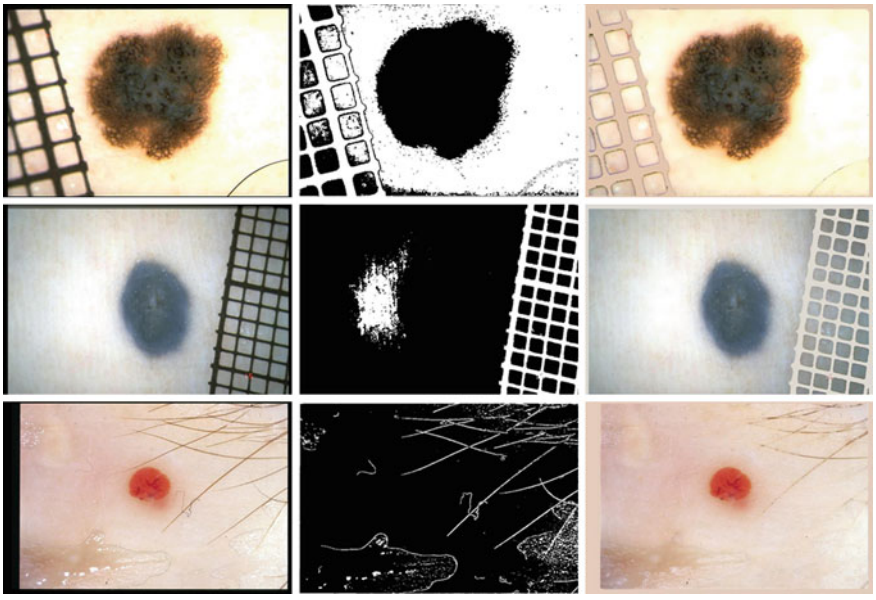


Fig. 24 Processing process in three images: for each image, first the original image, second the artifacts mask image (the fact that it is detecting the *black frames* in the edge of the images should be noted) and third the preprocessed image

4.2 Segmentation

The testing to the images to test the segmentation is intrinsically united to the process described in the system design for the determination of the coefficients (COEF_RED,COEF_GREEN,COEF_BLUE).

In 110 images experienced doctors were required to carry out the segmentation of the lesions, generating 110 masks R_MASK_SEG_MED. In those own images the segmentation has been made generating 110 masks R_MASK_SEG. As a segmentation measure measureSegment=Area(Intersection(R_MASK_SEG, R_MASK_SEG_MED))/Area(Union(R_MASK_SEG,R_MASK_SEG_MED)) has been defined.

Good segmentations are considered those which fulfill (measureSegment>THR_SEG), where THR_SEG is a threshold value.

Accuracy results of 87.27 % were obtained. On the whole, the segmentation method works correctly. The vast majority of the misfires are in the images in which exist hypopigmented lesions not contained within the surface of the region. In these cases it is used in the system the non-automated segmentation.

Some segmentation examples can be seen thereupon, in Fig. 25.

4.3 Detection and Characterization of Indicators

For the determination of the results in the detection and characterization of indicators the following indicators of the "ABCD Rule" were studied, separately.

A: Asymmetry

40 images were taken, which have been diagnosed by expert dermatologists about this indicator and comparing with which was generated by the system. In the asymmetry the evaluations of the experts were computed, for each one of the images, in the minor and major axes. Thus, 80 evaluations of asymmetry are obtained. Obviously, before the own evaluations, a coordinate axis was specified, being in general the ones generated by the system satisfactory for the experts.

88.75 % of accuracy is obtained in the evaluation of asymmetry. In most cases, the errors have been product of the own subjectivity in the appreciation. Some examples can be seen thereupon, in Figs. 26 and 27.

B: Border

40 images were taken, which have been diagnosed by expert dermatologists concerning that indicator and comparing it with which it was generated by the system. In the borders the evaluation of the experts were computed, for each one of the images, in the 8 octants. Thus, 320 evaluations of borders are obtained (Figs. 28, 29, 30 and 31).

86.88 % of accuracy in the evaluation of the borders is obtained. In most cases, as it happened in the asymmetry, the errors have been product of the own subjectivity in the appreciation. Some examples can be seen thereupon:

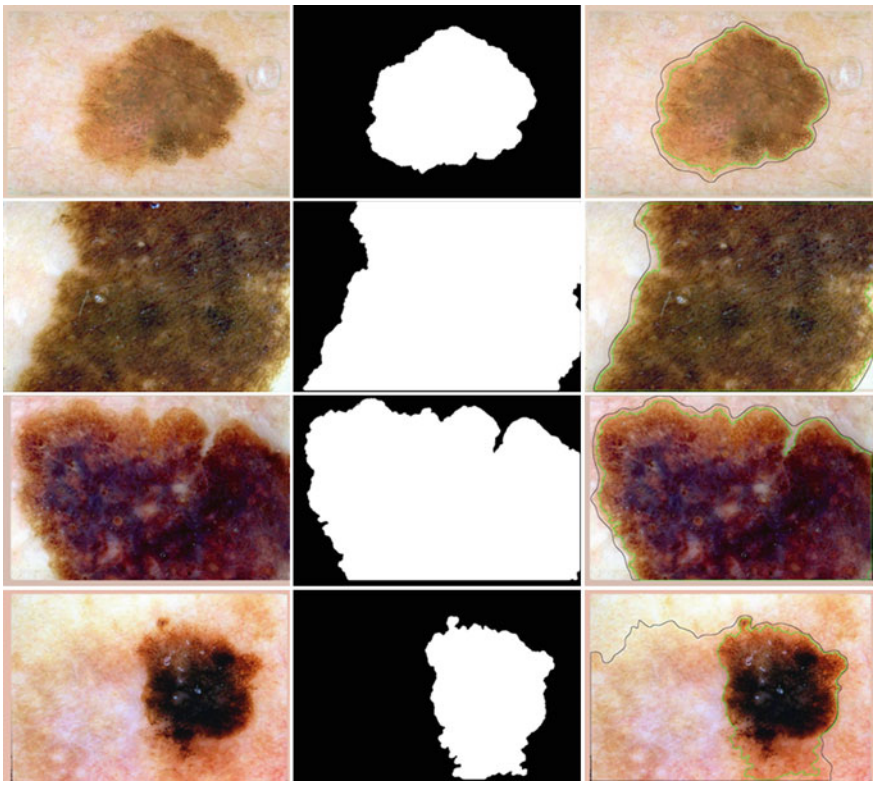


Fig. 25 Segmentation in four images: for each image, first the image, second the mask of the automated segmentation of the lesion and third the image with the segmentation by experts (*black line*) and automated (*green line*). In the last image there is an error in the automated segmentation, which it does not select correctly the part of the hypopigmented lesion

C: Colors

40 images were taken, which have been diagnosed by expert dermatologists concerning that indicator and compared with which was generated in the system. In the colors the evaluations of the experts were computed, determining the number of colors which were seen, and comparing with the generated color maps. Thus 40 evaluations of colors are obtained. It should be noted here that in this case before this test a supervised machine learning process was carried out to the own doctors, requiring them to report about the color of various pixels in each one of 50 images, as it has been described previously in the System Design.

62.5 % of accuracy is obtained in the evaluation of colors. It should be noted here that in these results it is considered as a hit the fact of having success in the colors acutely with the criterion of the expert. This does not mean that in many lesions which are not fully correct has not coincided with the experts partly the colors of the generated color map, in some cases it was considered to be some further one, in

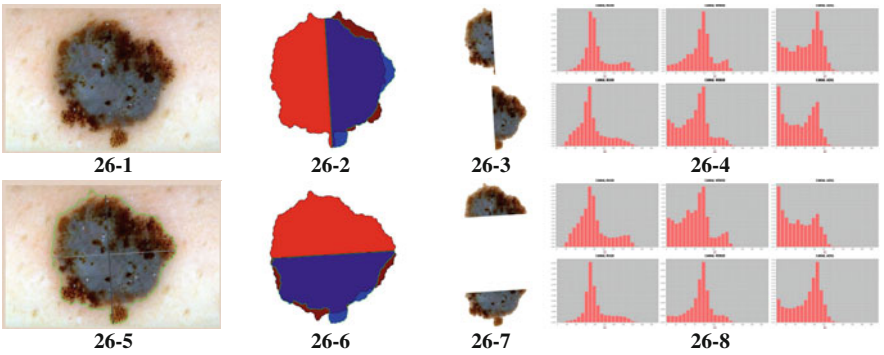


Fig. 26 Computing asymmetry in an image: *1* Preprocessed, *2* Comparison of shape with respect to major axis; *3* Halves with respect to the major axis; *4* Comparison of histograms with respect of the major axis; *5* Image with the axes; *6* Comparison of shape with respect to minor axis; *7* Halves with respect to the minor axis; *8* Comparison of histograms with respect to minor axis. The image is symmetric in shape with respect to both axes, and asymmetric in color with respect to both axes. Hence, it is asymmetric in both axes. Punctuation: 2.6

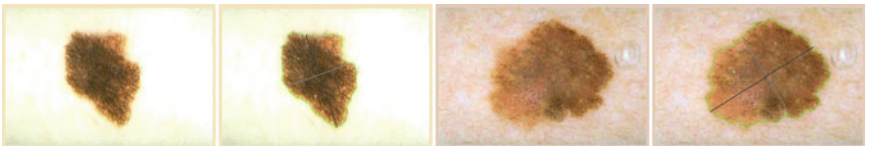


Fig. 27 Computing asymmetry in two images, from each one the preprocessed is taken and the one containing axes. The first one is symmetric with respect to both axes. Punctuation: 0. Second one is symmetric with respect to the major axis and asymmetric with respect to the minor axis. Punctuation: 1.3

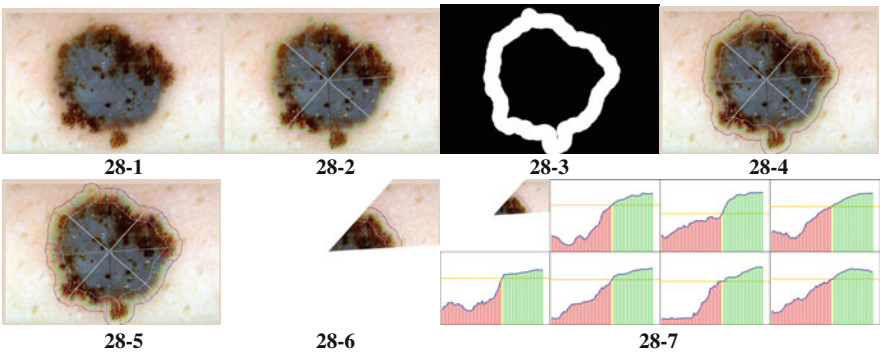


Fig. 28 Computing borders in an image: *1* Preprocessed; *2* Image with the generated sectors; *3* Crown surrounding the border of the lesion where the abruption of the lesion is examined; *4* Image with the limits drawn; *5* Image with the radius for the generated analyses; *6* Sector I; *7* Analysis of the sector I. The image has abrupt sectors: I, II, V, VI and VII. Punctuation: 0.5

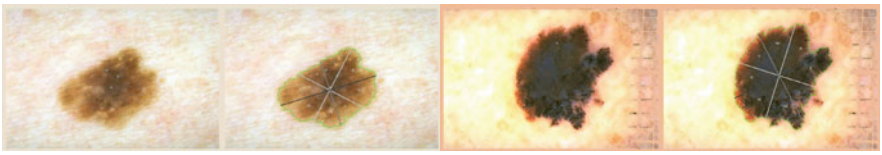


Fig. 29 Computing borders in two images, from each one the preprocessed and the one which has the two sectors is taken. The first one has no abrupt sectors. Punctuation: 0.0. The second one has the abrupt sectors: I, II, III, IV, VI, VII and VIII. Punctuation: 0.7

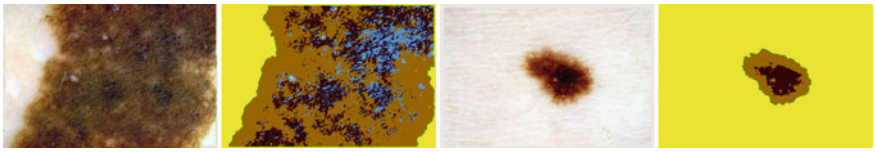


Fig. 30 Computing colors in two images. First it has three colors. Punctuation: 1.5. Second it has 2 colors. Punctuation: 1.0

other some pieces of regions catalogued as a certain color there were considered as another one. Some examples can be seen thereupon:

D: Dermoscopic Structures

Homogeneous and Unstructured areas

40 images were taken, which have been diagnosed by expert dermatologists concerning that indicator and compared with which was generated by the system. 92.5 % of accuracy is obtained in the evaluation of whether there are homogeneous and unstructured areas or not.

Reticular

40 images were taken, which have been diagnosed by expert dermatologists concerning that indicator and compared with which was generated by the system. 90 % of accuracy is obtained in the evaluation of whether there is a reticular pattern or not. The proposed algorithm presents a very robust behaviour.

Globular

40 images were selected, which have been diagnosed by expert dermatologists concerning which was generated by the system. 77.5 % of accuracy is obtained in the evaluation of whether there is a globular pattern or not. In most cases the errors had a great deal to do with the subjectivity.

Points

40 images were taken, which have been diagnosed by expert dermatologists concerning that indicator and compared with which was generated in the system. An accuracy of 77.5 % is obtained in the evaluation of whether there is a points pattern or not. In most cases, as it happens in the previous case, errors had a great deal to do with the subjectivity.

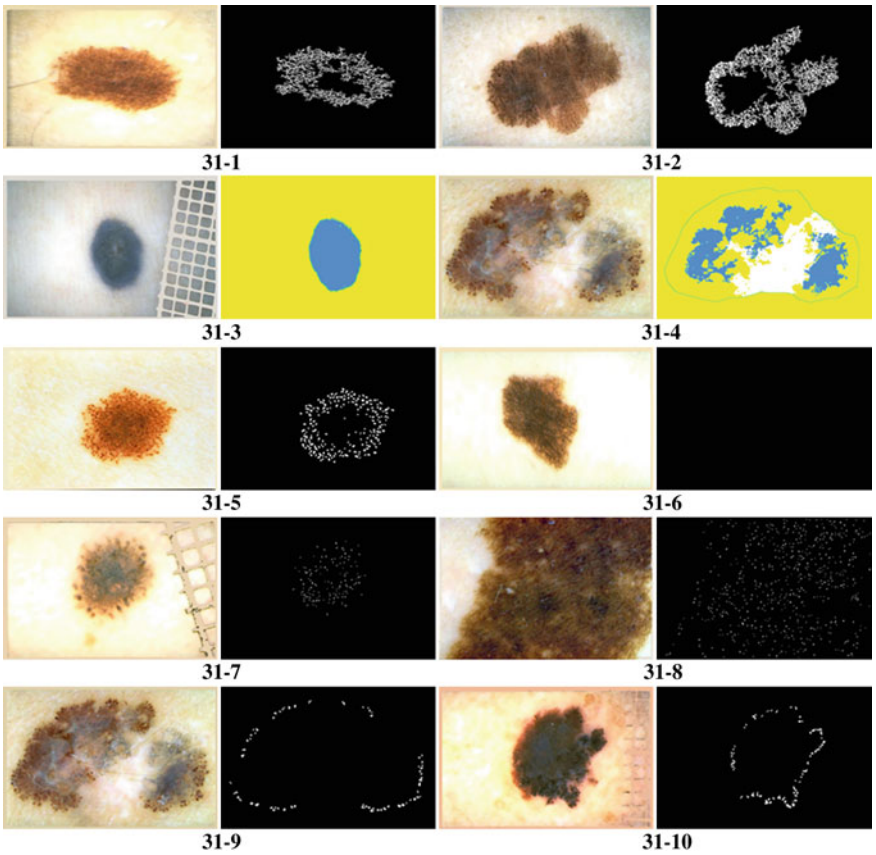


Fig. 31 Computing dermoscopic patterns: first the preprocessed image, second the mask of the pattern (map in the case of homog./unstructured). *1–2* Reticular, punctuation: 0.5; *3–4* Homogeneus and unstructured, punctuation: 0.5. *5–6* Globular, punctuation of first: 0.5, punctuation of second: 0.0; *7–8* Points, punctuation: 0.5; *9–10* Pseudopods, punctuation: 0.5

Pseudopods

40 images were taken, which have been diagnosed by expert dermatologists concerning that indicator and compared with which was generated by the system. An accuracy of 70 % is obtained in the evaluation of whether there is a pseudopods pattern or not. In this case the proposed algorithm does not present a very robust behaviour, and should be improved in the future.

Some examples can be seen thereupon:

Table 4 Diagnosis results

	Positive	Negative
Melanoma	13 (TP)	3 (FN)
Not melanoma	7 (FP)	27 (TN)

4.4 Diagnosis

For testing the robustness of the whole system in the diagnosis support, in distinguishing between malignant and benign lesions, 50 images were taken, with a total amount of 16 melanomas.

The results are presented thereupon, in the Table 4:

The results are therefore of 81.25 % of sensitivity and 77.14 % of specificity, in the tested set of images.

5 Conclusions and Future Work

The design of a system of automated detection of melanoma created by our research team is presented. Being still at the pilot scheme stage, it already presents very encouraging results.

The described system offers the possibility to analyze images of two types: macroscopic and dermoscopic, which can be taken from different sources and different manufacturers, since it is not integrated with any particular hardware. Moreover, it has been designed with multiplatform and multiaccess architecture, having a firm orientation to teledermatology which is increasingly using.

The analysis of the dermoscopic images is carried out by supporting in the "ABCD Rule" medical algorithm, undertaking the automated detection and characterization of the corresponding indicators. That enables the analysis to have a basis on the existing extensive medical literature, which gives the expert dermatologists a greater confidence.

To measure the degree of robustness of the system a set of tests was carried out, obtaining good results in overall.

In the case of preprocessing, the detection of artifacts (black frames, flashes, bubbles, rulers and hairs) is undertaken satisfactorily, although it is working in its improvement. In particular, for the detection of hairs, it has been thinking on incorporating a supervised machine learning process, extracting similar features to the ones used in the detection of the reticular pattern, combining it with curvilinear detection functions, and for the elimination of hairs inpainting techniques will be tested.

In the case of the segmentation, 87.27 % of accuracy is obtained. Those are good results, but they can be improved, especially in the case of having hypopigmented zones not contained absolutely within the lesion. To do so the design of the algorithm

should be improved making it scale to a better behaviour in cases of lesions with infrequent shape and distribution, as well as to test other techniques.

In the case of the asymmetry, borders and colors, the results are also good, with an accuracy of 88.75, 86.88 and 62.5 % respectively; in the case of the colors, as it was commented in the previous section, this result is much higher than might be assumed, since it is considered a merit the fact of being fully successful in the colors, with the criterion of the expert, which makes it much difficult the success. The results for these three indicators are good, though it can be improved, having to take into consideration the improvement process of the algorithms which, in most cases, the errors have been product of the own subjectivity in the appreciation of the expert dermatologists, which were taken as a reference to validate the accuracy of the algorithm.

In the case of dermoscopic patterns, given the good results in the case of the reticular pattern and the homogeneous areas, with an accuracy of 92.5 and 90 % respectively, though they might be improved by refining the algorithm. In the case of the globular and points patterns the accuracy is 77.5 and 77.5 % respectively, reaching a considerably improvement margin, although it has to be born in mind that in this case there is also a great degree of subjectivity in the appreciation. In the case of the pseudopods pattern, the accuracy is 70 %, given nonetheless in this case the fact that the proposed algorithm does not present a very robust behaviour, thus, it should be improved in the future. Besides improving the design of the algorithms, for the reticular, globular, points and pseudopods patterns it is thought to test new techniques to improve the results.

There also carried out tests to determine in a direct way the diagnostic capacity of the system, obtaining results of 81.25 % of sensitivity and 77.14 % of specificity, applying to do so the diagnostic rules of the "ABCD Rule" medical algorithm through the undertaken detection and automated characterization of the indicators. The results are considered good, though they also should be improved, through the improvement of each one of the indicators.

In the evolution of the system, which has been designed in a modular and scalable way, a multidisciplinary team, composed by expert dermatologists, researchers specialized in image processing and computer programmers continues working, adding new images to the database, improving the existing algorithms and adding new functionalities. The research work is going on, in a specialized way in the improvement of the existing diagnostic algorithm, to obtain better results and, as it has been commented before, in the incorporation to that algorithm of the analysis of new dermoscopic patterns, like indicators, in such a way that the system gives a more robust diagnosis. To do so, new digital image processing techniques will be also included.

Once the pilot scheme stage is completed, the system may be used as a guide to medical professionals who are not expert in dermatology, which is for instance the case of the primary doctors, or even if it is the case of an expert dermatologist the task will be facilitated to this medical professional, either for being high the number of moles to be analyzed, or simply for giving the doctor a second opinion. Another possible application will be the execution of the system in massive analysis processes, in repositories of dermoscopic images, with the aim of seeking possible

malignant lesions, or even in screening systems directed to the population with the aim of detecting melanomas in early stages, in telemedicine platforms or from another type.

Acknowledgments The authors wish to acknowledge the help of the companies GAIA, IMQ and Maser, as well as the support provided by the Basque Country Government Departments of Education and Industry, Commerce and Tourism.

References

1. Skin Cancer Foundation, "Skin Cancer Foundation–Skin cancer facts". http://www.skincancer.org/Skin-Cancer-Facts/ (2013)
2. Asociación Española Contra el Cáncer, "Melanoma". http://www.aecc.es/SobreElCancer/CancerPorLocalizacion/melanoma/ (2013)
3. International Dermoscopy Society, "Consensus Net Meeting on Dermoscopy". www.dermoscopy.org/consensus/default.asp (2013).
4. Zaballos, P., Carrera, C., Puig, S., Malvehy, J.: Educación Médica Continuada Criterios dermatoscópicos para el diagnóstico del melanoma, Medicina Cutánea Ibero-Latino-Americana, vol. 32 (2004).
5. Psaty, E.L., Halpern, A.C.: Current and emerging technologies in melanoma diagnosis: the state of the art. Clin. dermatol. **27**(1), 35–45 (2009)
6. Smith, L., Macneil, S.: State of the art in non-invasive imaging of cutaneous melanoma. Skin res. technol. **17**(3), 257–269 (2011)
7. Goodson, A.G., Grossman, D.: Strategies for early melanoma detection: approaches to the patient with nevi. J. Am. Acad. Dermatol. **60**(5), 719–738 (2009)
8. Guitera, P., Menzies, S.W.: State of the art of diagnostic technology for early-stage melanoma. Expert Rev. Anticancer Ther. **11**(5), 715–723 (2011)
9. Pehamberger, H., Steiner, A., Wolff, K.: In vivo epiluminescence microscopy of pigmented skin lesions. I: Pattern analysis of pigmented skin lesions. J. Am. Acad. Dermatol. **17**(4), 571–583 (1987)
10. Stolz, W., Riemann, A., Cognetta, A.B.: ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur. J. Dermatol. **4**, 521–527 (1994)
11. Menzies, S.W., Ingvar, C., Crotty, K.A., McCarthy, W.H.: Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. Arch. dermatol. **132**(10), 1178–1182 (1996)
12. Argenziano, G., Fabbrocini, G., Carli, P., De Giorgi, V., Sammarco, E., Delfino, M.: 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. Arch. dermatol. **134**(12), 1563–1570 (1998)
13. FotoFinder Systems, "Fotofinder". http://www.fotofinder-systems.com/ (2013)
14. Isasi, Gola: A., García Zapirain, B., Méndez Zorrilla, A.: Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. Comput. Biol. Med. **41**(9), 742–755 (2011)
15. Lee, T., Ng, V., Gallagher, R., Coldman, A., McLean, D.: Dullrazor®: a software approach to hair removal from images. Comput. Biol. Med. **27**(6), 533–543 (1997)
16. Zhou, H., Chen, M., Gass, R., Rehg, J.M., Ferris, L., Ho, J., Drogowski, L.: Feature-preserving artifact removal from dermoscopy images. In: Proceedings of the SPIE Medical Imaging Conference, vol. 6914, pp. 69141B–69141B-9 (2008).
17. Wighton, P., Lee, T.K., Atkins, M.S.: Dermascopic hair disocclusion using inpainting. In: Proceedings of the SPIE Medical Imaging Conference, vol. 6914, pp. 691427–691427-8 (2008).

18. Kiani, K., Sharafat, A.R.: E-shaver: an improved DullRazor(®) for digitally removing dark and light-colored hairs in dermoscopic images. Comput. biol. med. **41**(3), 139–145 (2011)
19. Abbas, Q., Garcia, I.F.: Emre Celebi, M., Ahmad, W.: A feature-preserving hair removal algorithm for dermoscopy images. Skin Res. technol. **19**(1), e27–e36 (2013)
20. Abbas, Q., Celebi, M.E., García, I.F.: Hair removal methods: a comparative study for dermoscopy images. Biomed. Signal Process. Control **6**(4), 395–404 (2011)
21. Celebi, M.E., Iyatomi, H., Schaefer, G., Stoecker, W.V.: Lesion border detection in dermoscopy images. Comput. Med. Imaging Graph. **33**(2), 148–153 (2009)
22. Gómez, Delgado: D., Butakoff, C., Ersbøll, B.K., Stoecker, W.: Independent histogram pursuit for segmentation of skin lesions. IEEE Trans. Biomed. Eng. **55**(1), 157–161 (2008)
23. Garnavi, R., Aldeen, M., Celebi, M.E., Varigos, G., Finch, S.: Border detection in dermoscopy images using hybrid thresholding on optimized color channels. Comput. Med. Imaging Graph. **35**(2), 105–115 (2011)
24. Celebi, Emre: M., Wen, Q., Hwang, S., Iyatomi, H., Schaefer, G.: Lesion border detection in dermoscopy images using ensembles of thresholding methods. Skin Res. Technol. **19**(1), e252–e258 (2013)
25. Schmid, P.: Segmentation of digitized dermatoscopic images by two-dimensional color clustering. IEEE Trans. Med. Imaging **18**(2), 164–171 (1999)
26. Cucchiara, R., Grana, C., Seidenari, S., Pellacani, G.: Exploiting color and topological features for region segmentation with recursive fuzzy C-means. Mach. Graph. Vis. Int. J. 11(2/3), 169–182 (2002).
27. Zhou, H., Schaefer, G., Sadka, A.H., Celebi, M.E.: Anisotropic mean shift based fuzzy C-means segmentation of dermoscopy images. IEEE J. Sel. Top. Signal Process. **3**(1), 26–34 (2009)
28. Melli, R., Grana, C., Cucchiara, R.: Comparison of color clustering algorithms for segmentation of dermatological images. In: Proceedings of the SPIE Medical Imaging Conference, vol. 6144, pp. 61443S–61443S-9 (2006).
29. Celebi, Emre: M., Alp Aslandogan, Y., Stoecker, W.V., Iyatomi, H., Oka, H., Chen, X.: Unsupervised border detection in dermoscopy images. Skin Res. Technol. **13**(4), 454–462 (2007)
30. Celebi, M.E., Kingravi, H.A., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H., Malters, J.M., Grichnik, J.M., Marghoob, A.A., Rabinovitz, H.S., Menzies, S.W.: Border detection in dermoscopy images using statistical region merging. Skin Res. Technol. **14**(3), 347–353 (2008)
31. Wang, H., Chen, X., Moss, R.H., Stanley, R.J., Stoecker, W.V., Celebi, M.E., Szalapski, T.M., Malters, J.M., Grichnik, J.M., Marghoob, A.A., Rabinovitz, H.S., Menzies, S.W.: Watershed segmentation of dermoscopy images using a watershed technique. Skin Res. Technol. **16**(3), 378–384 (2010)
32. Erkol, B.H.E., Moss, R., Stanley, R., Stoecker, W.: Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. Skin Res. Technol. **11**, 17–26 (2005)
33. Zhou, H., Schaefer, G., Celebi, M.E., Lin, F., Liu, T.: Gradient vector flow with mean shift for skin lesion segmentation. Comput. Med. Imaging Graph. **35**(2), 121–127 (2011)
34. Abbas, Q., Celebi, M.E.: Fondón García, I.: A novel perceptually-oriented approach for skin tumor segmentation. Int. J. Innovative Comput. Inf. Control **8**(3), 1837–1848 (2012)
35. Iyatomi, H., Oka, H., Celebi, M.E., Hashimoto, M., Hagiwara, M., Tanaka, M., Ogawa, K.: An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. Comput. Med. Imaging Graph. **32**(7), 566–579 (2008)
36. Alcon, Fernandez: J., Ciuhu, C., Ten Kate, W., Heinrich, A., Uzunbajakava, N., Krekels, G., Siem, D., De Haan, G.: Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J. Sel. Top. Signal Process. **3**(1), 14–25 (2009)
37. Di Leo, G., Paolillo, A., Sommella, P., Fabbrocini, G., Rescigno, O.: A software tool for the diagnosis of melanomas. In: IEEE Instrumentation & Measurement Technology Conference Proceedings, vol. 2010, pp. 886–891 (2010).
38. Ganster, H., Pinz, A., Röhrer, R., Wildling, E., Binder, M., Kittler, H.: Automated melanoma recognition. IEEE Trans. Med. Imaging **20**(3), 233–239 (2001)

39. Tommasi, T., La Torre, E., Caputo, B., Beichel, R., Sonka, M.: Melanoma Recognition Using Representative and Discriminative Kernel Classifiers. Computer Vision Approaches to Medical Image Analysis Lecture Notes in Computer Science **4241**, 1–12 (2006)
40. Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imaging Graph. **31**(6), 362–373 (2007)
41. Rahman, M.M., Bhattacharya, P., Desai, B.C.: A multiple expert-based melanoma recognition system for dermoscopic images of pigmented skin lesions. In: 8th IEEE International Conference on BioInformatics and. BioEngineering **2008**, 1–6 (2008)
42. Surowka, G.: Supervised learning of melanocytic skin lesion images. In: Conference on Human System. Interactions **2008**, 121–125 (2008)
43. Situ, N., Yuan, X., Chen, J., Zouridakis, G.: Malignant melanoma detection by Bag-of-Features classification. In: 30th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, EMBS 2008, pp. 3110–3113 (2008).
44. Zhou, H., Rehg, J.M.: Dermoscopic interest point detector and descriptor. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, vol. 2009, pp. 1318–1321 (2009).
45. Situ, N., Wadhawan, T., Hu, R., Lancaster, K., Yuan, X., Zouridakis, G.: Evaluating sampling strategies of dermoscopic interest points. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, vol. 2011, pp. 109–112 (2011).
46. Anantha, M., Moss, R.H., Stoecker, W.V.: Detection of pigment network in dermatoscopy images using texture analysis. Comput. Med. Imaging Graph. **28**(5), 225–234 (2004)
47. Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., Sommella, P.: Dermoscopic image-analysis system: estimation of atypical pigment network and atypical vascular pattern. In: IEEE International Workshop on Medical Measurement and Applications, MeMea 2006, vol. 2006, pp. 63–67 (2006).
48. Di Leo, G., Liguori, C., Paolillo, A., Sommella, P.: An improved procedure for the automatic detection of dermoscopic structures in digital ELM images of skin lesions. In: IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems, vol. 2008, pp. 190–194 (2008).
49. Barata, C., Marques, J.S., Rozeira, J.: A system for the detection of pigment network in dermoscopy images using directional filters. IEEE Trans. Biomed. Eng. **59**(10), 2744–2754 (2012)
50. 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., Argenziano, G., Soyer, H.P.: Detection of atypical texture features in early malignant melanoma. Skin Res. Technol. **16**(1), 60–65 (2010)
51. Sadeghi, M., Razmara, M., Lee, T.K., Atkins, M.S.: A novel method for detection of pigment network in dermoscopic images using graphs. Comput. Med. Imaging Graph. **35**(2), 137–143 (2011)
52. Sadeghi, M., Razmara, M., Wighton, P., Lee, T.K., Atkins, M.S.: Modeling the dermoscopic structure pigment network using a clinically inspired feature set. Lect. Notes Comput. Sci. **6326**, 467–474 (2010)
53. Skrovseth, S.O., Schopf, T.R., Thon, K., Zortea, M., Geilhufe, M., Mollersen, K., Kirchesch, H.M., Godtliebsen, F.: A computer aided diagnostic system for malignant melanomas. In: 2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), pp. 1–5 (2010).
54. Yoshino, S., Tanaka, T., Tanaka, M., Oka, H.: Application of morphology for detection of dots in tumor. In: SICE 2004 Annual Conference, vol. 1, pp. 591–594 (2004).
55. Celebi, M.E., Iyatomi, H., Stoecker, W.V., Moss, R.H., Rabinovitz, H.S., Argenziano, G., Soyer, H.P.: Automatic detection of blue-white veil and related structures in dermoscopy images. Comput. Med. Imaging Graph. **32**(8), 670–677 (2008)
56. Di Leo, G., Fabbrocini, G., Paolillo, A., Rescigno, O., Sommella, P.: Towards an automatic diagnosis system for skin lesions: Estimation of blue-whitish veil and regression structures. In: 6th International Multi-Conference on Systems. Signals and Devices, vol. 2009, pp. 1–6 (2009).

57. Arroyo, J.L.G., Zapirain, B.G., Zorrilla, A.M.: 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 (ISSPIT), vol. 2011, pp. 196–201 (2011).
58. Mirzaalian, H., Lee, T.K., Hamarneh, G.: Learning features for streak detection in dermoscopic color images using localized radial flux of principal intensity curvature. In: IEEE Workshop on Mathematical Methods in Biomedical Image. Analysis **2012**, 97–101 (2012)
59. Sadeghi, M., Lee, T.K., Mclean, D., Lui, H., Atkins, M.S.: Oriented pattern analysis for streak detection in dermoscopy images. In: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2012, pp. 298–306 (2012).
60. Sadeghi, M., Lee, T., Lui, H., McLean, D., Atkins, S.: Detection and analysis of irregular streaks in dermoscopic images of skin lesions. IEEE Trans. Med, Imaging (2013)
61. 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.: Concentric decile segmentation of white and hypopigmented areas in dermoscopy images of skin lesions allows discrimination of malignant melanoma. Comput. Med. Imaging Graph. **35**(2), 148–154 (2011)
62. Madasu, V.K., Lovell, B.C.: Blotch detection in pigmented skin lesions using fuzzy co-clustering and texture segmentation. In: Digital Image Computing: Techniques and Applications, vol. 2009, pp. 25–31 (2009).
63. Stoecker, W.V., Gupta, K., Stanley, R.J., Moss, R.H., Shrestha, B.: Detection of asymmetric blotches (asymmetric structureless areas) in dermoscopy images of malignant melanoma using relative color. Skin Res. Technol. **11**(3), 179–184 (2005)
64. Pellacani, G., Grana, C., Cucchiara, R., Seidenari, S.: Automated extraction and description of dark areas in surface microscopy melanocytic lesion images. Dermatology (Basel, Switzerland) 208(1), 21–26 (2004).
65. Khan, A., Gupta, K., Stanley, R.J., Stoecker, W.V., Moss, R.H., Argenziano, G., Soyer, H.P., Rabinovitz, H.S., Cognetta, A.B.: Fuzzy logic techniques for blotch feature evaluation in dermoscopy images. Comput. Med. Imaging Graph. **33**(1), 50–57 (2009)
66. Iyatomi, H., Oka, H., Celebi, M.E., Ogawa, K., Argenziano, G., Soyer, H.P., Koga, H., Saida, T., Ohara, K., Tanaka, M.: Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. J. Invest. Dermatol. **128**(8), 2049–2054 (2008)
67. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. Artif. Intell. Med. **56**(2), 69–90 (2012)
68. Celebi, M.E., Stoecker, W.V., Moss, R.H.: Advances in skin cancer image analysis. Comput. Med. Imaging Graph. **35**(2), 83–84 (2011)
69. Derma Medical Systems, "MoleMax". http://www.dermamedicalsystems.com (2013)
70. Biomips, "DBDermo-Mips/DDAX Software". http://www.ddax3.com/eng/index.html (2013)
71. Canfield Imaging Systems, "Canfield". http://www.canfieldsci.com/ (2013)
72. MedX Health Corporation, "SIAscopy^{TM}-Molemate". http://simsys-molemate.com/ (2013)
73. Walter, F.M., Morris, H.C., Humphrys, E., Hall, P.N., Prevost, A.T., Burrows, N., Bradshaw, L., Wilson, E.C.F., Norris, P., Walls, J., Johnson, M., Kinmonth, A.L., Emery, J.D.: Effect of adding a diagnostic aid to best practice to manage suspicious pigmented lesions in primary care: randomised controlled trial. Br. Med. J. 345, e4110:1–14 (2012).
74. MELA Sciences Inc., "MelaFind". http://www.melafind.com (2013)
75. Monheit, G., Cognetta, A.B., Ferris, L., Rabinovitz, H., Gross, K.: Martini, M., Grichnik, J.M., Mihm, M., Prieto, V.G., Googe, P., King, R., Toledano, A., Kabelev, N., Wojton, M., Gutkowicz-Krusin, D.: The performance of MelaFind: a prospective multicenter study. Arch. Dermatol. **147**(2), 188–194 (2011)
76. Dermoscopix, "Dermoscopix". http://www.dermoscopix.com (2013)
77. Visiomed AG, "MicroDerm". http://www.visiomedag.com/ (2013)
78. DatInf GmbH, "MoleExpert". http://moleexpert.com/ (2013)
79. Western Research Inc., "DermAlert". http://www.dermalert.com/ (2013)
80. Health Discovery Corporation, "MelApp". http://www.melapp.net/ (2013)

81. University of Michigan, “UMSkinCheck”. http://uofmhealth.org/patientandvisitorguide/my-skin-check-app (2013)
82. Molemap Inc., “Molemap”. http://www.molemap.net.au/ (2013)
83. Molesafe Inc., “Molesafe”. http://www.molesafe.com/ (2013)
84. Melanoscan Inc., “Melanoscan”. http://www.melanoscan.com/ (2013)
85. Elizabeth, D.D.P., Rhett, J., Drugge, M.D., Nguyen, C., Luciana, G.: Clinical pathway for melanoma detection using comprehensive cutaneous analysis with Melanoscan®, Dermatol. Online J. http://dermatology.cdlib.org/1608/1_original/1_10-00149/drugge.html (2010). Accessed 08 Oct 2012
86. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann, San Francisco (1993).
87. Otsu, N.: A threshold selection method from gray level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
88. Breiman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)

Melanoma Decision Support Using Lighting-Corrected Intuitive Feature Models

Robert Amelard, Jeffrey Glaister, Alexander Wong and David A. Clausi

Abstract Skin lesion analysis using standard camera images has received limited attention from the scientific community due to its technical complexity and scarcity of data. The images are privy to lighting variations caused by uneven source lighting, and unconstrained differences in resolution, scale, and equipment. In this chapter, we propose a framework that performs illumination correction and feature extraction on photographs of skin lesions acquired using standard consumer-grade cameras. We apply a multi-stage illumination correction algorithm and define a set of high-level intuitive features (HLIF) that quantifies the level of asymmetry and border irregularity about a lesion. This lighting-corrected intuitive feature model framework can be used to classify skin lesion diagnoses with high accuracy. The framework accurately corrects the illumination variations and achieves high and precise sensitivity (95 % confidence interval (CI), 73.1–73.5 %) and specificity (95 % CI, 72.0–72.4 %) using a linear support vector machine classifier with cross-validation trials. It exhibits higher test-retest reliability than the much larger state-of-the-art low-level feature set (95 % CI, 78.1–79.7 % sensitivity, 75.3–76.3 % specificity). Combining our framework with these low-level features attains sensitivity (95 % CI, 83.3–84.8 %) and specificity (95 % CI, 79.7–80.1 %), which is more accurate and reliable than classification using the low-level feature set.

Keywords Melanoma · Decision support system · Pigmented skin lesion · Feature extraction · Illumination correction · Standard camera

R. Amelard (✉) · J. Glaister · A. Wong · D. A. Clausi
Department of Systems Design Engineering, University of Waterloo, Waterloo, Canada
e-mail: ramelard@uwaterloo.ca

J. Glaister
e-mail: jlglaist@uwaterloo.ca

A. Wong
e-mail: a28wong@uwaterloo.ca

D. A. Clausi
e-mail: dclausi@uwaterloo.ca

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_7,

Introduction

Melanoma is the most deadly form of skin cancer and is caused by the development of a malignant tumour of the melanocytes [42]. It was estimated that 76,250 men and women were diagnosed with melanoma in 2012 and 9,180 men and women died of it in the US alone [31]. One in five Americans are expected to be diagnosed with melanoma in their lifetime [54]. One of the most alarming facts about melanoma is that it is the most common form of cancer for young adults [30] and it is one of the few cancers where the incidence rate is increasing for men [34]. Fortunately, the five-year survival rate is 98 % if the lesion is extracted while the cancer is still confined to its primary location. However, the five-year survival rate decreases to 15 % if the cancer has spread to remote parts of the body [31].

Some clinical tools exist that can assist dermatologists diagnose skin lesions. For example, the Asymmetry, Border irregularity, Colour variation, and Diameter (ABCD) rubric serves as a guide for dermatologists to check skin lesions in a systematic manner [41, 50]. However, expert dermatologists using the ABCD rubric with a dermatoscope (a specialized tool that optically magnifies and enhances skin structures) reported a sensitivity of 76.0–87.7 % and a specificity of 61.0–77.8 % [5]. Furthermore, only 48 % of US fellows of the American Academy of Dermatology reported using a dermatoscope [25]. Recent developments include melanoma detection using standard camera images [1, 8, 9], but there is still much room for technological advances.

This chapter presents a systematic framework to analyse and assess the risk of melanoma using dermatological photographs taken with a standard consumer-grade camera. The framework consists of illumination preprocessing and feature extraction, and is validated using a simple malignancy classification scheme. The preprocessing step consists of a multi-stage illumination modeling algorithm. The proposed features that are extracted are high-level intuitive features (HLIF) describing lesion asymmetry and border irregularity. The segmentation is obtained using a manually-drawn ground-truth border and an existing classification algorithm is used. Automatic segmentation is not discussed here.

The first step in the proposed framework is a preprocessing step, where the image is corrected for illumination variation. This preprocessing step serves to improve performance of the subsequent steps. Illumination correction tries to remove illumination variation, such as shadows, so that healthy skin is a more consistent colour throughout the photograph. Since the lesion is typically a darker colour than the surrounding healthy skin, a segmentation algorithm may misclassify shadows as lesion. The illumination correction can improve the classification algorithm by standardizing features reliant on underlying pixel values. For example, features that rely on lesion colour are affected by shadows.

Once an image has been preprocessed, features are extracted and used to classify the image as “malignant” or “benign”. The quality of this feature set is very important in differentiating the two classes. Furthermore, it is generally preferable to project the image into a low-dimensional feature space, since high-dimensional

feature spaces are usually associated with larger computational cost and possibly various classification problems such as overfitting and the curse of dimensionality. In the proposed framework, we extract HLIFs that were designed specifically to model a human-observable characteristic. These features may take more effort to design than low-level features, but we show in the experimental results that classification accuracy improves dramatically by integrating HLIFs with low-level features.

The rest of the chapter is organized as follows. A literature review of existing methods and algorithms is presented in "Background". The multi-stage illumination modeling algorithm is described in "Illumination Correction Preprocessing". A set of high-level intuitive features for describing skin lesions is presented in "Feature Extraction". Experimental results of the proposed unified framework are shown in "Results" and conclusions are drawn in "Conclusion".

Background

The problem of diagnosing cutaneous cancer has received more attention from the technical community in recent years. Unlike classical medical imaging modalities that produce an image using dedicated equipment (e.g., magnetic resonance imaging and X-ray), skin diagnosis is performed using visual information. This workflow poses more difficult computer vision problems as there is no standard modality with which data is captured.

The majority of the melanoma detection methods in the literature analyse images acquired using dermoscopy [21], also known as epiluminescence microscopy (ELM) or the unfortunate similarly-named "dermatoscopy". Dermoscopy is a non-invasive technique whereby the dermatologist uses a handheld dermatoscope to visually analyse the skin. Dermatoscopes optically magnify the area of interest, and most can elucidate sub-surface structures by applying a dermoscopic oil on the skin or, more recently, employing light cross-polarization. These produce images with standardized lighting conditions, and show sub-surface microstructures of the epidermis. However, it has been reported that only 48 % of American Academy of Dermatology fellows use dermatoscopes [25]. There is, therefore, a large demand for methods that analyse images taken without a dermatoscope.

Our framework's workflow is summarized in Fig. 1. We present a literature review of preprocessing, feature extraction, and classification methods in the following sections, as it pertains to our framework.

Illumination Correction Preprocessing

Most illumination correction algorithms are not specific to skin lesion photographs and can be applied to any scene. Histogram equalization adjusts the distribution of pixel intensities, minimizing illumination variation globally [47]. Other algo-

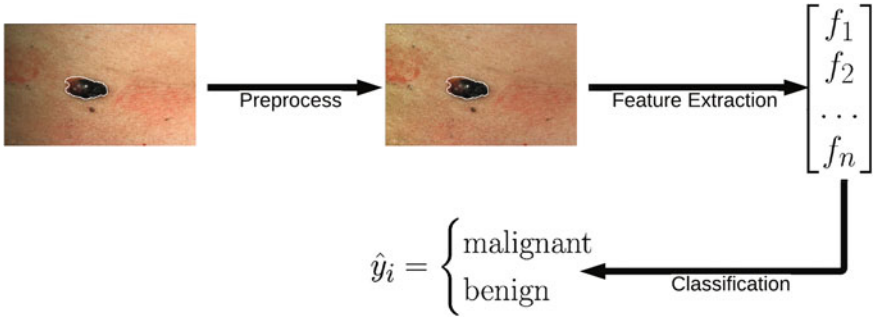


Fig. 1 Workflow of the proposed melanoma decision support framework

rithms correct for local illumination variation. These algorithms typically assume a multiplicative relationship between illumination and reflectance components. The estimated illumination component is estimated and used to find the reflectance component. The illumination component is assumed to be low-frequency, while the high-frequency detail is in the reflectance component. Using this assumption, there are many different algorithms that estimate illumination. One of the earliest is the Retinex algorithm, which uses a set of Gaussian filters of different sizes to remove detail and to estimate illumination [27, 36]. Morphological operators [49], bilateral filters [24], Monte Carlo sampling [53] and total variation [16] approaches have also been used to estimate illumination.

Other methods involve correction algorithms that are specific to images of skin lesions. Earlier algorithms enhance images taken with a dermatoscope to better separate lesion pixels from healthy skin. These algorithms include colour calibration [29] and normalization [33] to improve lesion classification or contrast enhancement [13, 46] to improve segmentation.

Recent work focuses on correcting photographs of skin lesions acquired using standard digital cameras to improve segmentation and classification. Work by Cavalcanti et al. [10] apply morphological operators to estimate the illumination component. The initial estimate of illumination is used to fit a parametric surface using the illumination intensities in the four corners of the photograph. The reflectance component is estimated using the parametric surface. Initial work on the correction algorithm outlined in this chapter was initially presented by Glaister et al. [28].

Feature Extraction and Classification

Most existing feature sets have been designed to model the ABCD criteria using dermoscopic images. Lee and Claridge propose irregularity indices to quantify the amount of border irregularity [37]. Aribisala and Claridge propose another border irregularity metric based on conditional entropy [6]. Celebi et al. propose shape,

colour, and texture features with rationale, and using a filter feature selection method [12]. Colour features are primarily taken either in the RGB space (usually mean and standard deviation of the three channels), or a perceptually-uniform CIE colour space. Good overviews of existing features can be found in [35, 39].

Features designed to analyse dermoscopic images may not necessarily be suitable for the noisy unconstrained environment of standard camera images. Some work has been done to identify suitable features for standard camera images [1, 8, 9], however the focus of these methods has primarily been in the preprocessing and segmentation stages, resulting in large sets of low-level features. For example, Cavalcanti and Scharcanski [9] propose the same low-level feature set as Alcon et al. [1] with a few minor adjustments. Amelard et al. proposed the first set of high-level asymmetry and border irregularity features that were modeled assuming standard camera images [3, 4], which are used in this chapter.

Most of the methods use existing classification schemes, such as support vector machines (SVM), artificial neural networks (ANN), decision trees, and k-nearest neighbour (K-NN) [35]. Ballerini et al. designed a hierarchical classification system based on K-NN using texture and colour features to classify different types of non-melanoma skin lesions with 93 % malignant-versus-benign accuracy and 74 % inter-class accuracy [8]. Piatkowska et al. achieved 96 % classification accuracy using a multi-elitist particle swarm optimization method [44]. Thorough reviews of existing classification schemes can be found in [35, 39].

Some emphasis has been placed on constructing content-based image retrieval (CBIR) frameworks for recalling similar lesions. These methods rely on constructing a representative feature set that can be used to determine the similarity of two images. Ballerini et al. extracted basic colour and texture features such as colour mean, covariance, and texture co-occurrence matrix calculations, and used a weighted sum of Bhattacharyya distance and Euclidean distance to find visually similar lesions [7]. Celebi and Aslandogan incorporated a human response matrix based on psychovisual similarity experiments along with shape features to denote similarity [11]. Aldridge et al. designed CBIR software and experiments which showed with high statistical significance that diagnostic accuracy among laypeople and first-year dermatology students was drastically improved when using a CBIR system [2].

There has been some work done to extract melanin and hemoglobin information from skin images. The melanin and hemoglobin information can be very useful in trying to identify the stage and type of the lesion. All of the proposed methods rely on various physics-based models of the skin to characterize the reflectance under some assumptions about the absorption, reflectance, and transmission of the skin layers. Work primarily led by Claridge explores using multispectral images using spectrophotometric intracutaneous analysis to analyse melanin, hemoglobin, and collagen densities [19, 40]. Claridge used a physics-based forward parameter grid-search to determine the most feasible skin model parameters assuming standardized images [17]. Tsumura et al. used an independent component analysis (ICA) scheme to decompose the image into two independent channels, which they assumed are the melanin and hemoglobin channels [51]. D'Alessandro et al. used multispectral images obtained using a nevoscope and used a genetic algorithm to estimate melanin

and blood volume [20]. Madooei et al. used blind-source separation techniques using a proposed corrected log-chromaticity 2-D colour space to obtain melanin and hemoglobin information [38].

Illumination Correction Preprocessing

The proposed framework first corrects for illumination variation using the multi-stage algorithm outlined in this section. The illumination correction algorithm uses three stages to estimate and correct for illumination variation. First, an initial non-parametric illumination model is estimated using a Monte Carlo sampling algorithm. Second, the final parametric illumination model is acquired using the initial model as a prior. Finally, the parametric model is applied to the reflectance map to correct for illumination variation. The three stages are outlined in this section.

Initial Non-parametric Illumination Modeling

The first stage involves estimating the initial non-parametric illumination model. This stage is required to estimate illumination robustly in the presence of artefacts, such as hair or prominent skin texture. Certain assumptions are made about the illumination in the dermatological images. The images are assumed to have been taken inside a doctor's office, in a controlled environment and beneath overhead lights. This means that the illumination model does not need to account for sudden changes in lighting conditions. Instead, the illumination will change gradually throughout an image. Illumination variation is assumed to be produced by white lights, so the correction algorithm does not need to correct colour variation. Finally, a multiplicative illumination-reflectance model is assumed [36]. In this model, the V (value) channel in the HSV (hue-saturation-value) colour space [48] is modeled as the entry-wise product of illumination i and reflectance r components. After applying the logarithmic operator, this relationship becomes additive (1).

$$\begin{aligned} v(x, y) &= i(x, y) \times r(x, y) \\ v_{log}(x, y) &= i_{log}(x, y) + r_{log}(x, y) \end{aligned} \tag{1}$$

To estimate the illumination map i, the problem can be formulated as Bayesian least squares (2), where $p(i_{log}|v_{log})$ is the posterior distribution.

$$\begin{aligned} \hat{i}_{log} &= \underset{i_{log}}{\text{argmin}} \left\{ E\left((i_{log} - \hat{i}_{log})^2\right) | v_{log} \right\} \\ &= \underset{i_{log}}{\text{argmin}} \left\{ \int (i_{log} - \hat{i}_{log})^2 \, p(i_{log}|v_{log}) di_{log} \right\} \end{aligned} \tag{2}$$

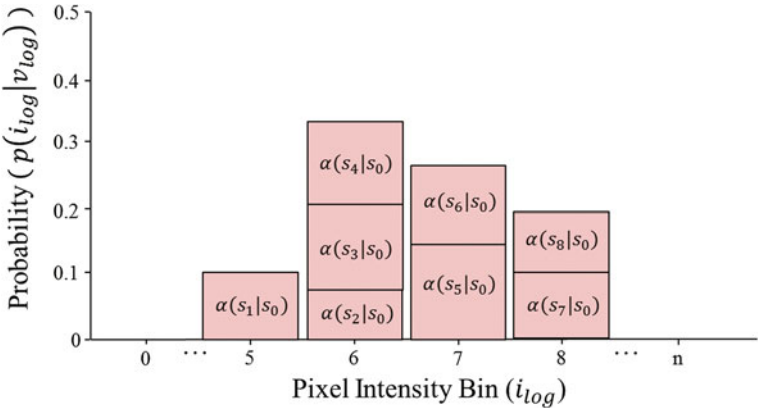


Fig. 2 Sample posterior distribution $\hat{p}\left(i_{log}|v_{log}\right)$, built from pixels accepted in the set Ω. Each stacked element corresponds to a pixel s_k in Ω, where the height is $\alpha(s_k|s_0)$ and bin location is s_k

To estimate the posterior distribution, a Monte Carlo posterior estimation algorithm is used [15]. A Monte Carlo estimation algorithm is used to avoid assuming a parametric model for the posterior distribution. In this Monte Carlo estimation strategy, candidate samples are drawn from a search space surrounding the pixel of interest s_0 using a uniform instrumental distribution. An acceptance probability α is computed based on the neighbourhoods around the candidate sample s_k and pixel of interest s_0. The Gaussian error statistic used in this implementation is shown in (3). The parameter σ controls the shape of the Gaussian function and is based on local variance and h_k and h_0 represent the neighbourhoods around s_k and s_0 respectively. The term λ in the denominator normalizes the acceptance probability, such that $\alpha(s_k|s_0) = 1$ if the neighbourhoods around s_k and s_0 are identical. The elements in the neighbourhoods are assumed to be independent, so the acceptance probability is the product of the probabilities from each site j.

$$\alpha(s_k|s_0) = \prod_j \frac{\frac{1}{2\pi\sigma}\exp\left[-\frac{(h_k[j]-h_0[j])^2}{2\sigma^2}\right]}{\lambda} \tag{3}$$

The candidate sample is accepted with a probability of α into the set Ω for estimating $p(i_{log}|v_{log})$. The selection and acceptance process is repeated until a desired number of samples were found in the search space. The posterior distribution is estimated as a weighted histogram, using α as the weights associated with each element. A sample histogram is shown in Fig. 2. The estimate of the log-transformed illumination map $\hat{i}_{log}$ is calculated using (2), as outlined in [26]. The initial illumination estimate $\hat{i}$ is acquired by taking the exponential of $\hat{i}_{log}$. An example of an image with visible illumination variance is shown in Fig. 3a and the associated non-parametric illumination model is shown in Fig. 3b.

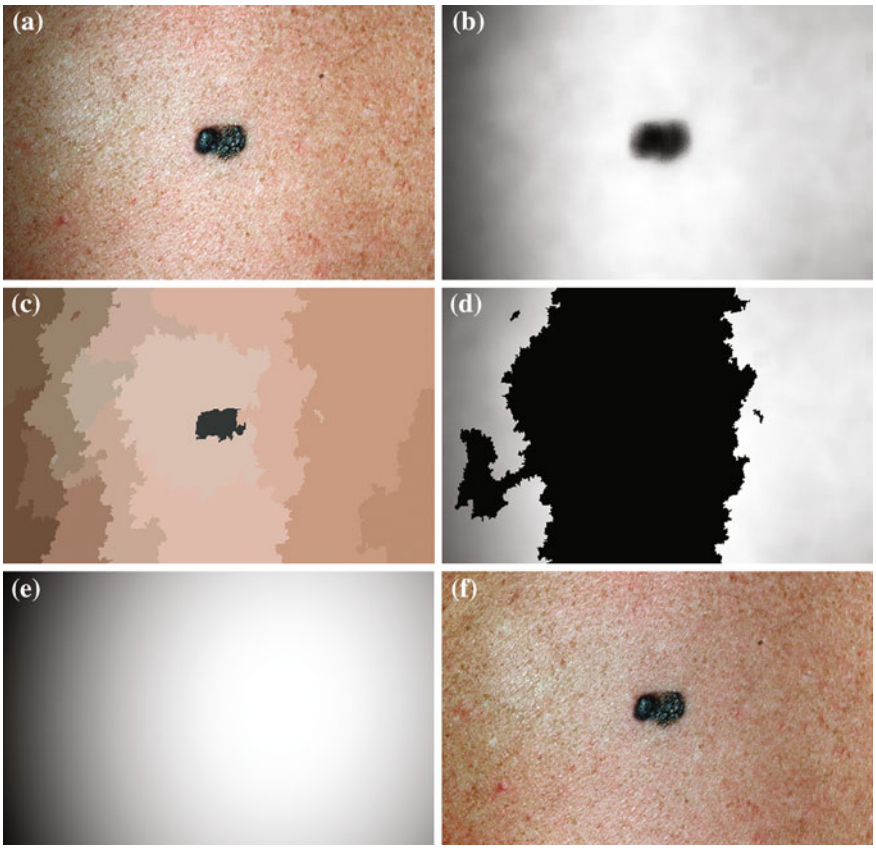


Fig. 3 Methodology to estimate illumination map: **a** original image of a skin lesion, where the *top* edge is noticeably darker than the *bottom* edge; **b** illumination map determined via non-parametric modeling using Monte Carlo sampling; **c** segmentation map found using Statistical Region Merging; **d** regions included in the subset of skin pixels, where pixels *black* in colour are not classified as normal skin; **e** new illumination map determined by using (**d**) as a prior to the quadratic surface model; **f** resulting corrected image using the multi-stage illumination correction algorithm

Final Parametric Illumination Modeling

The initial non-parametric illumination model results in an estimate of the illumination variation in healthy skin, but does not properly model illumination near the lesion. Instead, the initial model identifies the lesion as a shadow. Using the initial model to correct the image would result in a significant bright spot around the lesion, which is obviously undesirable. To better model the illumination, a second stage is added, which results in a parametric model of illumination that uses the initial illumination pixel values. The parametric model can adequately estimate the illumination variation because illumination is assumed to change slowly throughout the image.

The subset of pixels that are used to fit the parametric surface correspond to healthy skin in the original image.

To find the subset of healthy skin pixels, the original image is segmented into many regions. The segmentation algorithm used in this implementation is Statistical Region Merging [43]. The resulting segmented image is shown in Fig. 3c, where each region is represented as a single colour. Any regions that touched 20×20 pixel regions in the four corners of the image are considered part of the "healthy skin" class. While this method does not yield a perfect segmentation of the healthy skin and lesion classes, only an estimate of healthy skin pixels is required for fitting the parametric model. The regions that are considered 'healthy skin" are shown in Fig. 3d.

The final illumination model $\hat{i}'$ is estimated as a parametric surface (4) with coefficients c_1 to c_6, which is fit to the initial illumination values $\hat{i}$ corresponding to pixels in the "healthy skin" subset S using maximum likelihood estimation (5). The final parametric illumination model is shown in Fig. 3e.

$$i'(x, y) = c_1 x^2 + c_2 xy + c_3 y^2 + c_4 x + c_5 y + c_6 \tag{4}$$

$$\hat{i}' = \underset{\hat{i}'}{\operatorname{argmax}} \prod_{(x,y)\in S} P(\hat{i}(x, y)|\hat{i}'(x, y)) \tag{5}$$

$$\text{where } P(\hat{i}(x, y)|\hat{i}'(x, y)) \text{ i.i.d. } \mathcal{N}(\hat{i}'(x, y), \sigma^2)$$

Reflectance Map Estimation

The reflectance map is calculated by dividing the V channel v from the original image in the HSV colour space by $\hat{i}'$. The reflectance map $\hat{r}$ replaces the original V channel and is combined with the original hue (H) and saturation (S) channels. The resulting image is corrected for illumination. An example of a corrected image is shown in Fig. 3f.

Feature Extraction

Once the image has been preprocessed, descriptive features are extracted to describe the lesion as a vector of real numbers. One of the most prominent clinical methods for diagnosing a skin lesion is using the ABCD rubric [41, 50], where the dermatologist looks for signs of asymmetry, border irregularity, colour variations, and diameter. However, this is done in a very subjective manner, and results in discrete categorical values. For example, the score assigned to a lesion's asymmetry is determined by identifying whether the lesion is asymmetric across two orthogonal axes chosen by

the dermatologist, resulting in a score $A \in \{0, 1, 2\}$ [50]. This type of subjective visual analysis leads to large inter-observer bias as well as some intra-observer bias [5]. We aim to create continuous high-level intuitive features (HLIFs) that represent objective calculations modeled on a human's interpretation of the characteristic.

High-Level Intuitive Features

A "High-Level Intuitive Feature" (HLIF) is defined as a feature calculation that has been designed to model a human-observable phenomenon (e.g., amount of asymmetry about a shape), and whose score can be qualitatively intuited. As discussed in "Background", most skin lesion features are low-level features. That is, they are recycled mathematical calculations that were not designed for the specific purpose of analysing a characteristic of the lesion shape.

Although designing HLIFs is more time-consuming than amalgamating a set of low-level features, we show in "Results" that the discriminative ability of a small set of HLIFs is comparable to a large set of low-level features. Since the HLIF set is small, the amount of required computation for classification decreases, and the risk of overfitting a classifier in a highly dimensional space is reduced, especially with small data sets.

In the next two sections, we describe nine HLIFs to evaluate the asymmetry and border irregularity of a segmented skin lesion. These HLIFs are designed to model characteristics that dermatologists identify. This work is described in detail in [3, 4], so we limit our analysis here to a general overview of the features.

Asymmetry Features

Dermatologists try to identify asymmetry with respect to shape or colour to indicate malignancy. These visual cues result due to the unconstrained metastasis of melanocytes in the skin.

HLIF for Colour Asymmetry

Asymmetry with respect to colour can be quantified by separating the lesion along an axis passing through the centre of mass (centroid) such that it represents the maximal amount of colour asymmetry. This "maximal" axis was found iteratively. First, we calculate the major axis of the lesion. The major axis is that which passes through the centroid and describes the maximal variance of the shape. The lesion image was then converted to the Hue-Saturation-Value (HSV) space so we can use the illumination-invariant hue measure to analyse the colours. The normalized hue histograms of both sides of the axis were smoothed using a Gaussian filter for robustness and were then

compared to generate the following HLIF value:

$$f_1^A = \max_{\theta} \left\{ \frac{1}{2} \sum_{i=1}^{nbins} |H_1^{\theta}(i) - H_2^{\theta}(i)| \right\} \tag{6}$$

where H_1^{θ} and H_2^{θ} are the normalized smoothed hue histograms according to the separation axis defined by rotating the major axis by θ, and $nbins$ is the number of discretized histogram bins used for binning hue values (we used 256 bins). Noticing that $f_1^A \in [0, 1]$, f_1^A can be intuited as an asymmetry score ranging from 0 (completely symmetric) to 1 (completely asymmetric).

Figure 4 depicts an example of this calculation. The lesion has a dark blotch on one side of it, rendering it asymmetric with respect to colour, which is reflected in the calculated value of $f_1^A = 0.3866$.

HLIF for Structural Asymmetry

Making the observation that symmetry can usually not be found in highly irregular shapes, we express the segmented lesion as a structure and analyse its simplicity. Fourier descriptors apply the Fourier series decomposition theory to decomposing some arbitrary shape into low-frequency and high-frequency components. In particular, the points on the shape border are mapped to 1D complex number via $F : (x, y) \mapsto x + \mathrm{i}y$, where i is the complex number. The Fourier transform is performed on this set of complex numbers. We can then compare a low-frequency ("simple") reconstruction with the original shape to determine the amount of asymmetry due to shape irregularity.

First, since we must use the discretized Fourier transform, the lesion border was uniformly sampled at a fixed rate to ensure the same decomposition in the frequency domain across all images. The Fast Fourier Transform (FFT) was then applied, and only the two lowest frequency components were preserved. These two frequencies represent the zero-frequency mean and the minimum amount of information needed to reconstruct a representative border. The inverse FFT was applied on these two frequencies to reconstruct a low-frequency representation of the lesion structure. Comparing this shape with the original is ill-advised, as this would yield a metric that is more suitable for border irregularity. Instead, we applied the same procedure to reconstruct a structure using k frequencies, were k is some small number that reconstructs the general shape of the lesion. This shape is compared to the original shape to generate the following HLIF calculation:

$$f_2^A = \frac{area(S_k \oplus S_2)}{area(S_k \bigcup S_2)} \tag{7}$$

where S_k and S_2 are the k-frequency and 2-frequency reconstructions of the original lesion shape. This feature value can be intuited as a score representing the general

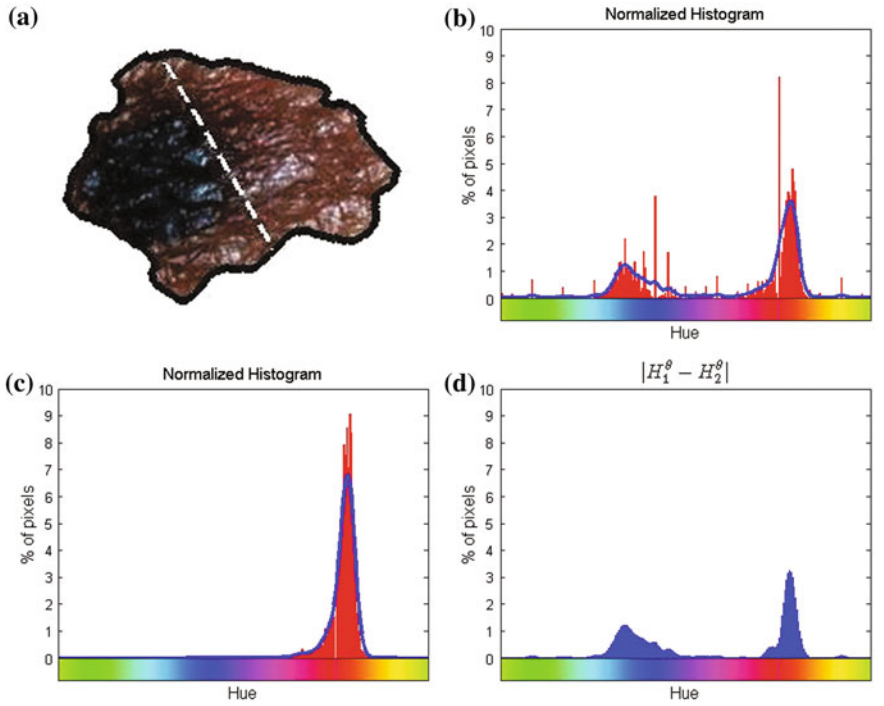


Fig. 4 Example of the design of f_1^A by comparing the normalized hue histograms of both sides of the separation axis. The *red bars* represent the original binned histogram of hue values, and the *blue line* represents these histograms smoothed by a Gaussian function ($\sigma = 2$ bins), which allows us to compare hue histograms robustly. In this example, $f_1^A = 0.3866$, representing a lesion with asymmetric colour distributions. **a** Image separated by the axis that produces maximal hue difference. **b** Normalized hue histogram of the *left* side of the lesion. Note the prominence of *blue* pixels along with the *red* ones in the histogram due do the *dark* blotch in the image. **c** Normalized hue histogram of the *right* side of the lesion. Note the prominence of *red* pixels in the histogram in correspondence with the image. **d** Absolute difference of the two hue histograms. The amount of *blue* and lack of *red* pixels in the first histogram are reflected by the two "humps" in the difference histogram

structural variations. We found empirically that linearly sampling the border with 1000 points and setting $k = 5$ yielded good results.

Figure 5 depicts an example of this calculation. Notice how the lesion has a very abnormal shape that does not seem to contain any symmetry. The logical XOR (Fig. 5b) captures this structural variation, and the calculation represents the dark area with respect to the union in Fig. 5c.

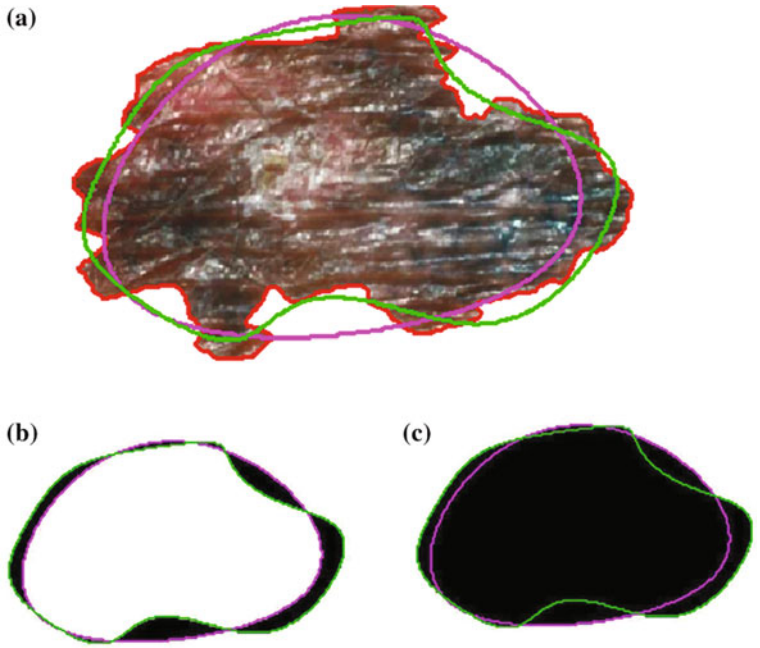


Fig. 5 Example of the design of f_2^A by comparing a baseline reconstruction of the lesion (*magenta*) with a low-frequency reconstruction (*green*) that incorporates the structural variability, it any exists. In this example, $f_2^A = 0.1609$, representing a structurally asymmetric lesion. **a** Original lesion, and reconstruction using 2 (*pink*) and 5 (*green*) frequency components. **b** Logical XOR of the two reconstructions (*dark area*). **c** Union of the two reconstructions (*dark area*)

HLIF for Asymmetric Areas

Cavalcanti and Scharcanski propose the following four HLIFs as a measure of area asymmetry [9]:

- $f_8 : (B_1 - B_2)/A$ with respect to L_1
- $f_9 : (B_1 - B_2)/A$ with respect to L_2
- $f_{10} : B_1/B_2$ with respect to L_1
- $f_{11} : B_1/B_2$ with respect to L_2

where L_1, L_2 are the major and minor axes of the lesion, and B_1, B_2 are the areas of each side of L_1 or L_2. No explicit constraint exists on the relative sizes of B_1 vs B_2. Thus, for very asymmetric shapes, $\{f_8, f_9\}$ can either before positive or negative, and $\{f_{10}, f_{11}\}$ can be large or small. For clarity, we add the following constraint to the features:

$$f_3^A = (A_1 - A_2)/A \text{ with respect to } L_1, \tag{8}$$

$$f_4^A = (A_1 - A_2)/A \text{ with respect to } L_2, \tag{9}$$

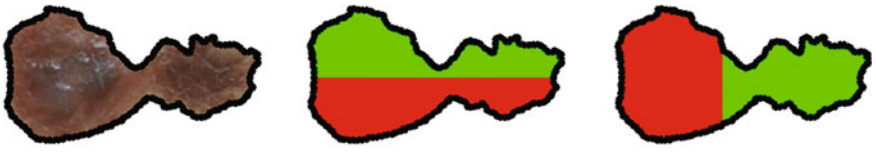


Fig. 6 Example of the design of $\{f_3^A, f_4^A, f_5^A, f_6^A\}$ by comparing the areas to each side of the lesion's major and minor axes. Notice that the red area appears much larger than the *green* area when separated by the minor axis. In this example, $f_3^A = 0.0052$, $f_4^A = 0.1560$, $f_5^A = 0.0105$, $f_6^A = 0.3698$, representing a structurally asymmetric lesion about the minor axis. **a** Original lesion. **b** Major axis (L_1). **c** Minor axis (L_2)

$$f_5^A = (A_1 - A_2)/A_2 \text{ with respect to } L_1, \tag{10}$$

$$f_6^A = (A_1 - A_2)/A_2 \text{ with respect to } L_2 \tag{11}$$

such that:

$$A_1 = \max\{B_1, B_2\},$$
$$A_2 = \min\{B_1, B_2\}$$

where B_1, B_2 are as before. This way, $\{f_3^A, f_4^A\}$ represent the positive difference in areas with respect to the total area, and $\{f_5^A, f_6^A\}$ represent the positive difference in areas with respect to the smaller area.

Figure 6 depicts an example of this calculation. When this lesion is separated by the minor axis which passes through the centroid, the red area is much larger than the green area, indicating a large amount of structural asymmetry, which is reflected in the scores $f_4^A = 0.1560$ and $f_6^A = 0.3698$.

Border Irregularity Features

Clinically, border irregularities are usually defined by spiky non-uniform pigmentation. Benign lesions usually have relatively smooth borders that form an oval-like shape. We therefore want to design features to capture this "spiky" nature.

HLIF for Fine Irregularities

One indicator of the malignancy of a lesion is the amount of small spiky deviations from a theoretical smooth border. We therefore wish to quantify the degree to which a lesion border contains these fine irregularities. To do so, it seems natural to compare the border to a "smoothed" version of itself, thus preserving the overall structure and

capturing the spikes. This can be accomplished using morphological operations. In particular, given an arbitrary shape and a structuring element, morphological closing tends to smooth over abrupt *exterior* peaks. Conversely, morphological opening tends to smooth over abrupt *internal* peaks. If the morphological operation yields any changes to the original shape due to abrupt structural elements, the modified shapes will have a different area than the original. These areas are compared to generate the following HLIF calculation:

$$f_1^B = \frac{A_{closed} - A_{lesion}}{A_{lesion}} + \frac{A_{lesion} - A_{opened}}{A_{lesion}} \tag{12}$$

where A_{lesion} is the area of the original lesion shape, and A_{closed} and A_{opened} are the areas of the modified shape under the specific morphological operations. The sum of the two normalized areas indicates the level of fine irregularities in a lesion's border.

Figure 7 depicts an example of this calculation. The red border denotes the resulting area from the morphological operation. notice how morphological closing produces a larger area that fills in the gaps from extreme irregularities of the border, and morphological opening produces a smaller area that crops these extreme irregularities.

HLIF for Coarse Irregularities

Another indicator of the malignancy of the lesion is the amount of structural deviation from a standard circular shape. This shape is influenced by the non-linear spatial reproduction of melanocytes in the skin. We can use Fourier descriptors again to characterize these coarse irregularities. In particular, we can compare a low-frequency reconstruction of the lesion shape to the original lesion shape. These two shapes will differ significantly if the lesion has a varying border. We capture this information by comparing the perimeters of the two shapes to generate the following HLIF

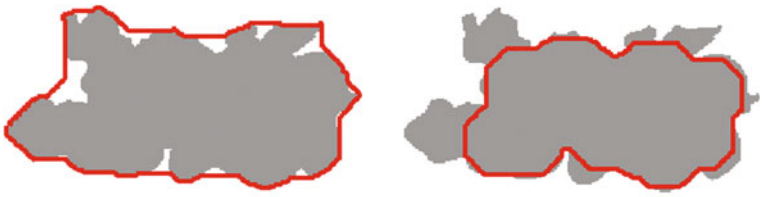


Fig. 7 Example of the design of f_1^B by comparing the areas resulting from morphological closing and opening. The *red* borders denote the resulting area from the respective morphological operation. Notice how the morphological closing fills in the *white* areas that are present due to the external spikes of the border, and morphological opening crops those spikes. In this example, $f_1^B = 0.3063$ representing a border with abrupt spikes. **a** Morphological closing. **b** Morphological opening

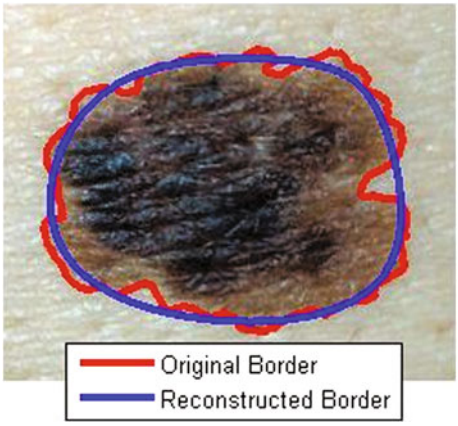


Fig. 8 Example of the design of f_2^B by comparing the perimeters of the original (*red*) and low-frequency reconstructed (*blue*) borders. Notice how the reconstructed border follows the general shape of the lesion, but does not take into account the large amounts of irregularity. In this example, $f_2^B = 0.24951$ representing a border with coarse irregularities

calculation:

$$f_2^B = \frac{|P_{lesion} - P_{low}|}{P_{lesion}} \tag{13}$$

where P_{lesion} and P_{low} are the perimeters of the original and low-frequency reconstruction of the lesion.

Figure 8 depicts an example of this calculation. The reconstructed (blue) border follows the general shape of the original lesion border (red), however it does not account for the coarse irregularities present in the actual border.

HLIF for Comparing Against Average Malignant Lesion

Over time a doctor will start to recognize malignant lesions based on the visual similarity to previously diagnosed cases. We can model this "learning" procedure by comparing a new case against the average malignant lesion border found across the training data. To perform this comparison, we must be able to compare the Fourier descriptors in a translation-, scale-, rotation-invariant manner. Given a set of Fourier coefficients $C = \{C_0, C_1, \ldots, C_{N-1}\}$, Fourier descriptor normalization is performed using the following three steps [52]:

1. *Translation Invariance*: set the first Fourier component (i.e., the DC component) to 0 ($C_0^* = 0$).
2. *Scale Invariance*: divide each $k-1$ Fourier coefficient by the complex magnitude of the second Fourier coefficient ($C_k^* = \frac{C_k}{|C_1|}$).
3. *Rotation/Point-Order Invariance*: consider only the real-valued complex magnitude of each Fourier component ($C_k^* = |C_k^*|$).

For each image in the training set, we sampled the border at a fixed rate (1,000 points on the border using linear interpolation produced good results), applied this normalization process, and computed the average normalized Fourier coefficients

across these images to obtain $\bar{C}^* = \{\bar{C}_0^*, \bar{C}_1^*, \ldots, \bar{C}_{N-1}^*\}$, where $C_k^* = |C_k/|C_1||$. This "average malignant" frequency representation is compared to a new case using the sum of squared differences to generate the following HLIF calculation:

$$f_3^B = \sum_{u=0}^{N-1} (|C_u^*| - |\bar{C}_u^*|)^2 \tag{14}$$

where $C^* = \{C_0^*, C_1^*, \ldots, C_{N-1}^*\}$ is the normalized set of Fourier coefficients for a new lesion image. Intuitively, the feature value corresponds to the amount of deviation from the "typical" malignant lesion.

Figure 9 depicts an example of this calculation. Note that the frequency components of the average malignant lesion (computed from the training data) and the extracted frequency components of the new lesion's border are quite similar, even in log space. This represents a "typical" malignant lesion in terms of structure.

Results

We validated the proposed framework in two phases using publicly accessible skin lesion images. First, we validated the resulting photographs after applying the initial illumination correction preprocessing algorithm. The photographs before and after correction were compared visually. We then tested the complete framework by extracting the HLIFs and low-level features from the corrected photographs and

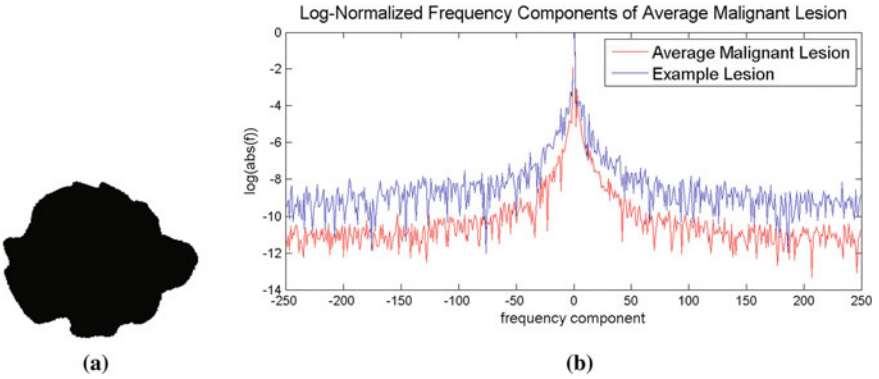


Fig. 9 Example of the design of f_3^B by comparing the frequency components of a lesion to the computed average malignant frequency components via sum of squared differences. This will give a low value for cases that look like a standard malignant lesion, based on training data. In this example, $f_3^B = 0.0087$, representing a typical malignant case. **a** Example melanoma lesion shape. **b** Log-frequency components of the average malignant lesion in the training data (*red*) and the example lesion on the *left* (*blue*). Log-space was used for visualization only

performing classification using a linear SVM model. This allowed us to validate the separability of the proposed framework's feature space.

Data and Setup

Our data set comprises 206 standard camera images which were taken in completely unconstrained and varying environments. These images were obtained from the Dermatology Information System [22] (43 malignant melanomas, 26 nevi) and DermQuest [23] (76 malignant melanomas, 61 nevi). Each image contains a single lesion of interest, which was manually segmented by the authors. In order to ensure rotation- and scale-invariance, each image was automatically rotated and scaled such that the major axis of the lesion lies parallel to the horizontal axis, and the lesion bounding box fit within a 200×200 rectangle prior to feature extraction. The test infrastructure was implemented in MATLAB.

Illumination Correction Results

After applying the illumination correction algorithm, the images were compared visually to the original images. The images were also compared visually to the Cavalcanti et al. illumination correction results. Figure 10 shows a set of images for comparison. To allow for a fair comparison between corrected and uncorrected images, the dynamic range of the V channel intensities were normalized.

In Fig. 10a–e, there is a visual improvement between the corrected and uncorrected images. Furthermore, in Fig. 10a, the framework's correction algorithm performs better than the Cavalcanti et al. algorithm for correcting illumination variation. This is because the correction algorithm uses a much larger subset of pixels in the parametric illumination model. The Cavalcanti et al. algorithm only uses pixels in the four corners, whereas the framework's correction algorithm uses any regions that touch the corners.

Figure 10f is an example of poor correction of illumination variation by both algorithms. This occurs when the illumination is complicated and cannot be modelled using a quadratic surface. For example, in Fig. 10f, the quadratic surface model is a false assumption due to the large patch of dark hair in the top of the photograph. As a result, the top left corner of the photograph becomes too bright.

Feature Extraction and Framework Classification Results

Upon applying the illumination correction algorithm, we extracted the asymmetry and border irregularity HLIFs (see "Feature Extraction") as well as Cavalcanti and

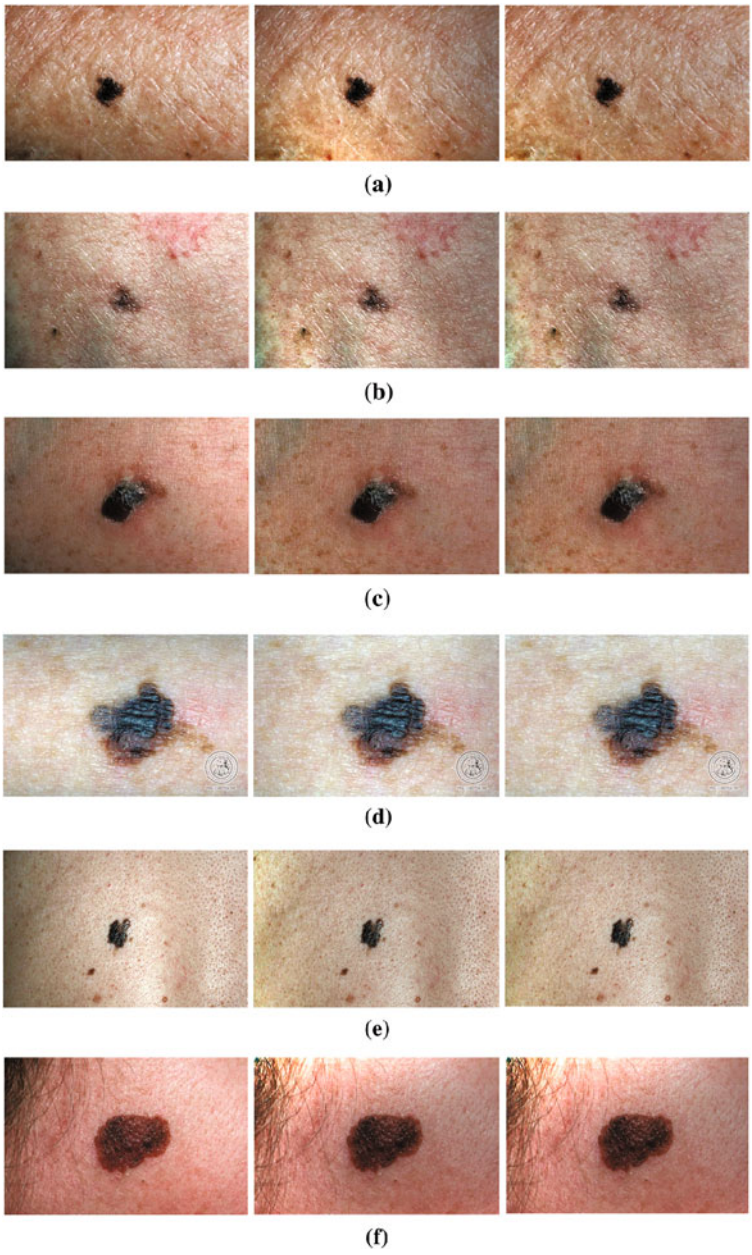


Fig. 10 Examples of dermatological photographs corrected for illumination variation. First column is the original photograph. Second column is the illumination correction results using the algorithm outlined in [10]. Third column is the illumination correction result using the proposed framework's multi-stage algorithm outlined in "Illumination Correction Preprocessing". In (**a**–**e**), the illumination variation is removed or reduced in the corrected image, while in (**f**), the illumination variation is not reduced. Examples of dermatological photographs corrected for illumination variation, continued

Scharcanski's low-level standard camera feature set for comparison [9]. For convenience, we use the following naming convention in the discussion:

- F_C: Cavalcanti and Scharcanski's low-level feature set.
- F_{HLIF}: set of HLIFs from "Feature Extraction".
- F_T: Total combined feature set containing both of the above feature sets ($F_T = F_C \bigcup F_{HLIF}$). Note that four features from F_C are replaced by the HLIFs in "Feature Extraction".

Due to the small data set, leave-one-out cross-validation (LOO CV) trials were used to calculate the success metrics.

Classification

We used a linear soft-margin SVM model [18] to calculate sensitivity, specificity, and accuracy. We used the LIBSVM implementation of SVM [14]. Linear SVM was chosen to emphasize the separability of the proposed framework's feature space due to illumination correction and descriptive features, rather than emphasizing the performance of an advanced classifier.

Parameter Optimization

For each feature set we found the optimal soft-margin constraint (c) and benign class weight (w) using a grid-search algorithm. In particular, for each set of parameters (c, w), we calculated the F-score [45] over 100 CV trials using randomized 80 %/20 % splits of the data for training and testing, respectively. Recall the standard F-score formula:

$$F_\beta = (1 + \beta^2)\frac{precision \times recall}{\beta^2 \times precision + recall} \tag{15}$$

where

$$precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{16}$$

$$recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{17}$$

$$accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{18}$$

where TP, TN, FP, FN are the number of true positive (i.e., malignant), true negative (i.e., benign), false positive, and false negative cases from classification. The F-score measure is the weighted harmonic mean of the precision and recall. For stability of the results due to the relatively restricted data set, we use $\beta = 1$ to calculate the harmonic mean of precision and recall.

The training and testing data points were chosen at random from the entire data set for each iteration. In accordance with the authors' suggestions [32], we varied the value of $c \in \{2^i | i = -5, \ldots, 11\}$ and $w \in \{2^i | i = -3, \ldots, 3\}$. Let (c^*, w^*) be the parameter values that exhibit the highest F_1 score. Upon determining (c^*, w^*), we sub-sampled the parameter space $\{(\frac{1}{2}c^*, \frac{1}{2}w^*), (2c^*, 2w^*)\}$ at $c_i = 2^{0.15}c_{i-1}$ and $w_i = 2^{0.5}w_{i-1}$. The optimal parameter values found according to F_1 in this sub-sampling over 100 independent cross-validation trials with 80 %/20 % train/test split were used in the classification.

Discussion

Table 1 shows that classification using the nine HLIFs exhibits higher test-retest reliability as compared to classification using the 52 low-level features. In particular, although mean sensitivity (95 % confidence interval (CI), 73.1–73.5 %) and specificity (95 % CI, 72.0–72.4 %) of the small HLIF set are slightly lower than the sensitivity (95 % CI, 78.1–79.7 %) and specificity (95 % CI, 75.3–76.3 %) of the large low-level set, classification using the HLIFs is much more reproducible as shown through the more narrow confidence interval. This is a powerful observation, since the HLIF set is only one-fifth the size of the low-level feature set. Since the features are in a lower dimensional space, the cost of computation and curse of dimensionality are not as pervasive as with the large low-level feature set.

Moreover, combining the HLIFs with the low-level features yields by far the best results (95 % CI, 83.3–84.8 % sensitivity, 79.7–80.1 % specificity). Adding only the nine HLIFs to the low-level feature set increases the number of features to 59, but yields non-trivial sensitivity and specificity improvements while also increasing reproducibility (i.e., decreasing standard deviation). This can be attributed to the HLIFs' ability to replicate human-observable phenomena in the data, whereas using many low-level features to model a high-level characteristic introduces a lot of variability in the measure, since the features were not designed specifically for the intended purpose of diagnosing skin cancer.

Statistical Significance

We wish to investigate each feature's ability to uniquely separate a malignant lesion from a benign lesion. The *t*-test indicates whether there is enough statistical signifi-

Table 1 Comparing classification results of different feature sets over 100 cross-validation trials see "Results" for feature set descriptions

Feature set	# features	Sensitivity (%)		Specificity (%)		Accuracy (%)	
		μ	σ	μ	σ	μ	σ
F_C	52	78.89	4.21	75.80	2.40	76.51	1.08
F_{HLIF}	9	73.32	**0.92**	72.21	0.99	72.52	**0.49**
F_T	59	**84.04**	3.67	**79.91**	**0.98**	**81.26**	1.31

μ mean, σ standard deviation

cance to reject a null hypothesis about a population using data from a sample of that population. In a two-sample t-test, the null hypothesis states that the means of two normally-distributed populations are equal ($H_0 : \Delta\mu = \mu_1 - \mu_2 = 0$). A p-value can be calculated using $\Delta\mu$, which indicates probability that we can observe a test statistic at least as extreme as the one observed assuming H_0 is indeed valid.

Representing the populations as a particular feature's scores across malignant (population 1) or benign (population 2) cases, we seek to reject this null hypothesis with a low p-value, thus showing that a particular feature's scores are significantly different between malignant and benign classes (i.e., the feature separates malignant and benign cases well). We therefore assume that the population responses of the malignant and benign cases follow a normal distribution. We use Welch's t-test, which assumes populations with different variances.

Table 2 summarizes the p-value scores for each HLIF using Welch's t-test. That is, for each feature, we set the null hypothesis $H_0 : \Delta\mu = \mu_m - \mu_b = 0$, where μ_m and μ_b are the mean response values for malignant and benign feature scores. The t-statistic is calculated as follows:

$$t = \frac{\mu_m - \mu_b}{\sqrt{\frac{s_m^2}{N_m} + \frac{s_b^2}{N_b}}}, \tag{19}$$

where s_m, s_b are the sample standard deviations of malignant and benign feature scores, and N_m, N_b are the number of malignant and benign cases in the data set. The associated p-value is the area under the normal curve to the right of the calculated t-score.

Most of the p-values are relatively low, indicating a high ability to separate malignant from benign. In particular, $\{f_1^A, f_2^A, f_1^B, f_2^B\}$ seem to be very good predictors of malignancy as indicated by their very low p-values. Recall that f_3^A and f_5^A are the relative area differences with respect to the major axis. So, although these p-values are high, it is probably worth keeping them in the feature set because they complement the minor axis features f_4^A and f_6^A.

Sources of Error

Figure 11 provides some examples of false negative cases (i.e., misidentified malignant cases) using the framework classification results. Using the F_{HLIF} feature set (Fig. 11a), we see lesions that have a fairly smooth and regular border, and most are fairly symmetric as well. It appears as though the colour patterns would be the

Table 2 Performing Welch's two-sample t-test on the set of HLIFs

HLIF	f_1^A	f_2^A	f_3^A	f_4^A	f_5^A	f_6^A	f_1^B	f_2^B	f_3^B
***p*-value**	< 0.0001	0.0340	0.7208	0.1215	0.6822	0.1372	0.0006	< 0.0001	0.3308

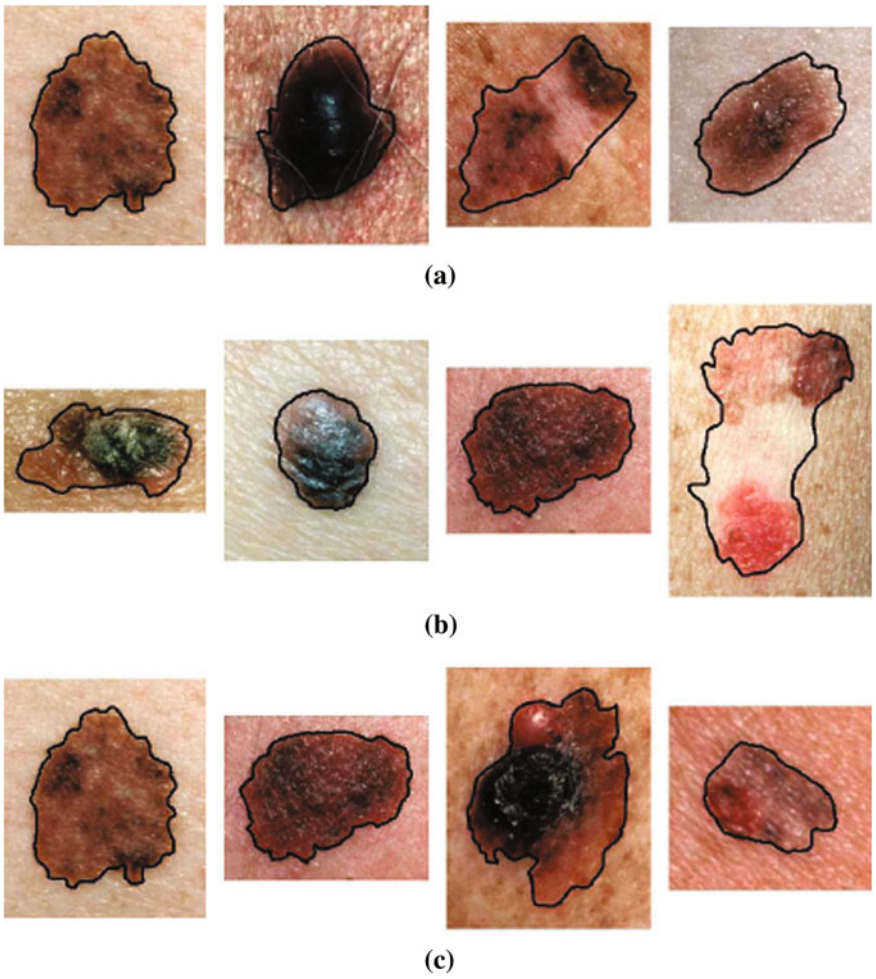


Fig. 11 Classification false negative examples using each feature set. **a** False negatives using F_{HLIF}. **b** False negatives using F_C. **c** False negatives using F_T

primary characteristic in determining the malignancy, of which there are no features in F_{HLIF}. Using the F_C feature set (Fig. 11b), we see lesions with varying colour distributions and symmetries. It is therefore no surprise that using the F_T feature set (Fig. 11c) also misclassifies lesions with prominent colour patterns.

Figure 12 provides some examples of false positive cases (i.e., misidentified benign cases) in the above classification results. Using the F_{HLIF} feature set (Fig. 12a), we see lesions that have asymmetric colour distributions and irregular borders, thus being classified as malignant. However, similar to the false negative cases above, their colour is fairly uniform, and could be an indicator of its benign nature. Using the F_C feature set (Fig. 12b), we see lesions with varying borders and some colour

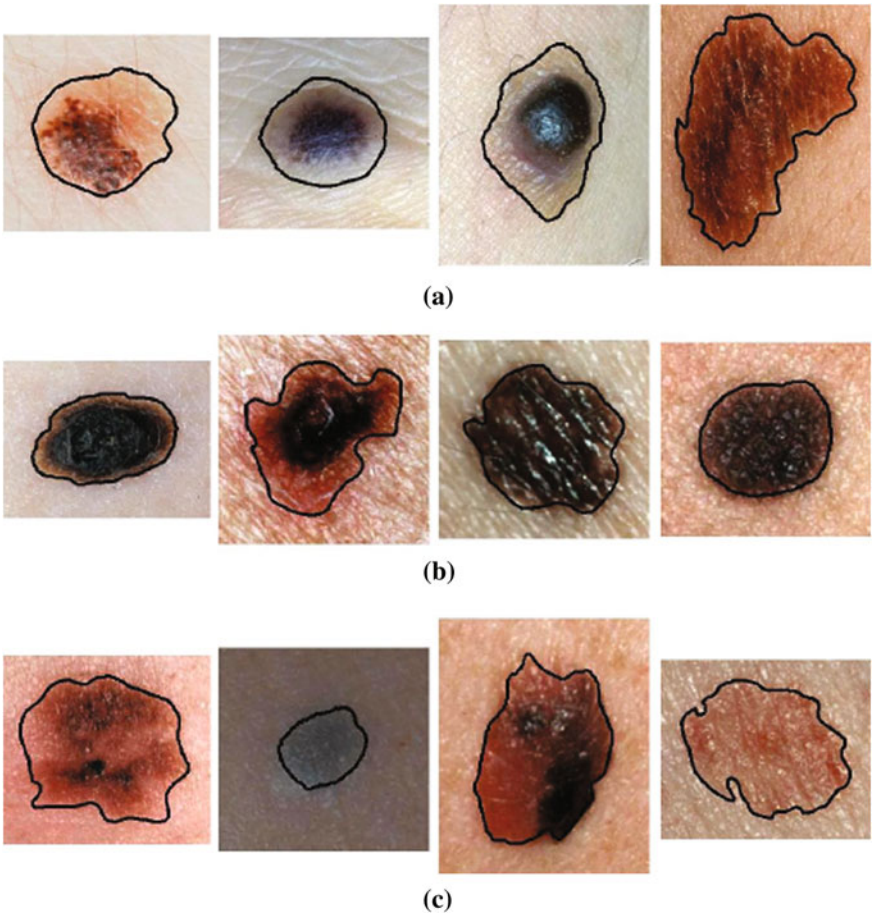


Fig. 12 Classification false positive examples using each feature set. **a** False positives using F_{HLIF}. **b** False positives using F_C. **c** False positives using F_T

asymmetry, although some appear symmetric, smooth border, and uniform colour, making it peculiar that some were classified as malignant. However, by adding F_{HLIF} to this set, we see that the false positive cases of F_T (Fig. 12c) are very suspicious lesions, with irregular borders and strange symmetries.

Conclusion

In this chapter we have proposed a novel framework for aiding in the diagnosis of skin lesions that uses lighting-corrected intuitive feature models. The multi-stage preprocessing step correctly adjusted the illumination across the standard camera image to ensure consistent analysis of that lesion. High-level intuitive features (HLIF) that characterize asymmetry and border irregularity were extracted and combined

with state-of-the-art low-level features. Designing HLIFs ensures that the feature scores can be conveyed to the doctor with intuitive rationale, as they are modeled in an intuitive manner.

Experimental results indicate that the illumination correction algorithm produces photographs taken with standard cameras that have been better corrected for illumination variation compared to a state-of-the-art algorithm [10]. Furthermore, linear classification using the small set of extracted HLIFs produces accurate and reliable results compared to the large state-of-the-art low-level feature set, and when combined with this large set we obtained improvements in sensitivity, specificity, and test-retest reliability. This unified framework can be used along with segmentation and advanced classification methods to provide a robust automatic diagnostic system for analysing skin lesions.

Acknowledgments This research was sponsored by Agfa Healthcare Inc., Ontario Centres of Excellence (OCE), and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

1. Alcon, J.F., Ciuhu, C., Ten Kate, W., Heinrich, A., Uzunbajakava, N., Krekels, G., Siem, D., de Haan, G.: Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J. Sel. Top. Signal Proces. **3**(1), 14–25 (2009)
2. Aldridge, R.B., Glodzik, D., Ballerini, L., Fisher, R.B.: Utility of non-rule-based visual matching as a strategy to allow novices to achieve skin lesion diagnosis. Acta Derm. Venereol. **91**(3), 279–283 (2011)
3. Amelard, R., Wong, A., Clausi, D.A.: Extracting high-level intuitive features (HLIF) for classifying skin lesions using standard camera images. In: CRV'12: Ninth Conference on Computer and Robot Vision, Toronto, pp. 396–403 (2012a)
4. Amelard, R., Wong, A., Clausi, D.A.: Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification. In: EMBC'12: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, pp. 4458–4461 (2012b)
5. Argenziano, G., Soyer, H.P., Chimenti, S., Talamini, R., Corona, R., Sera, F., Binder, M., Cerroni, L., De Rosa, G., Ferrara, G.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. J. Am. Acad. Dermatol. **48**(5), 679–693 (2003)
6. Aribisala, B.S., Claridge, E.: A border irregularity measure using a modified conditional entropy method as a malignant melanoma predictor. In: Kamel, M., Campilho, A. (eds.) Image Analysis and Recognition, Lecture Notes in Computer Science, vol. 3656. Springer, Heidelberg, pp. 914–921 (2005)
7. Ballerini, L., Li, X., Fisher, R.B., Rees, J.: A query-by-example content-based image retrieval system of non-melanoma skin lesions. In: Caputo, B., Müller, H., Syeda-Mahmood, T., Duncan, J., Wang, F., Kalpathy-Cramer, J. (eds.) Medical Content-Based Retrieval for Clinical Decision Support, Lecture Notes in Computer Science, vol. 5853, pp. 31–38. Springer, Heidelberg (2010)
8. Ballerini, L., Fisher, R.B., Aldridge, B., Rees, J.: A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions. In: Celebi, M.E., Schaefer, G. (eds.) Color Medical Image Analysis, Lecture Notes in Computational Vision and Biomechanics, vol. 6, pp 63–86. Springer, Netherlands (2013)
9. Cavalcanti, P.G., Scharcanski, J.: Automated prescreening of pigmented skin lesions using standard cameras. Comput. Med. Imaging Graph. **35**(6), 481–491 (2011)

10. Cavalcanti, P.G.., Scharcanski, J., Lopes, C.B.O.: Shading attenuation in human skin color images. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Chung, R., Hammoud, R., Hussain, M., Kar-Han, T., Crawfis, R., Thalmann, D., Kao, D., Avila, L. (eds.) Advances in Visual Computing, Lecture Notes in Computer Science, vol. 6453, pp. 190–198. Springer, Heidelberg (2010)
11. Celebi, M.E., Aslandogan, Y.A.: Content-based image retrieval incorporating models of human perception. In: ITCC'04: International Conference on Information Technology: Coding and Computing, vol. 2, pp. 241–245 Las Vegas (2004)
12. Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imaging Graph. **31**(6), 362–373 (2007)
13. Celebi, M.E., Iyatomi, H., Schaefer, G.: Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure. In: ICIP'09: 16th IEEE International Conference on Image Processing, Cairo, pp 2601–2604 (2009)
14. Chang, C.C., Lin, C.J.: Libsvm: A library for support vector machines. ACM Trans. Intel. Syst. Technol.**2**(3), 27:1–27:27 (2011). http://www.csie.ntu.edu.tw/ cjlin/libsvm
15. Chen, M.H.: Importance-weighted marginal bayesian posterior density estimation. J. Am. Stat. Assoc. **89**(427), 818–824 (1994)
16. Chen, T., Yin, W., Zhou, X.S., Comaniciu, D., Huang, T.S.: Pattern Anal. Mach. Intel. IEEE Trans. **28**(9), 1519–1524 (2006)
17. Claridge, E., Cotton, S., Hall, P., Moncrieff, M.: From colour to tissue histology: physics-based interpretation of images of pigmented skin lesions. Med. Image Anal. **7**(4), 489–502 (2003)
18. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**, 273–297 (1995)
19. Cotton, S.D.: A non-invasive imaging system for assisting in the diagnosis of malignant melanoma. PhD thesis, University of Birmingham, UK (1998)
20. D'Alessandro, B., Dhawan, A.P.: 3-d volume reconstruction of skin lesions for melanin and blood volume estimation and lesion severity analysis. IEEE Trans. Med. Imaging **31**(11), 2083–2092 (2012)
21. Day, G.R., Barbour, R.H.: Automated melanoma diagnosis: where are we at? Skin Res. Technol. **6**(1), 1–5 (2000)
22. Dermatology Information System: (2012). http://www.dermis.net. Accessed 08 Nov 2012
23. DermQuest: (2012). http://www.dermquest.com. Accessed 08 Nov 2012
24. Elad, M.: Retinex by two bilateral filters. In: Kimmel, R., Sochen, N.A., Weickert, J. (eds.) Scale Space and PDE Methods in Computer Vision, Lecture Notes in Computer Science, vol. 3459, pp. 217–229. Springer, Heidelberg (2005)
25. Engasser, H.C., Warshaw, E.M.: Dermatoscopy use by US dermatologists: a cross-sectional survey. J. Am. Acad. Dermatol. **63**(3), 412–419 (2010)
26. Fieguth, P.: Statistical image processing and multidimensional modeling, vol. 155, p. 65. Springer, New York (2010)
27. Frankle, J.A., McCann, J.J.: Method and apparatus for lightness imaging. US Patent 4,384,336 (1983)
28. Glaister, J., Wong, A., Clausi, D.A.: Illumination correction in dermatological photographs using multi-stage illumination modeling for skin lesion analysis. In: EMBC'12: 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, pp 102–105 (2012)
29. Haeghen, Y.V., Naeyaert, J.M.A.D., Lemahieu, I., Philips, W.: An imaging system with calibrated color image acquisition for use in dermatology. IEEE Trans. Med. Imaging **19**(7), 722–730 (2000)
30. Herzog, C., Pappo, A., Bondy, M., Bleyer, A., Kirkwood, J.: Cancer Epidemiology in Older Adolescents and Young Adults 15 to 29 Years of Age, National Cancer Institute, Bethesda, MD, chap Malignant Melanoma, pp 53–64. NIH Pub. No. 06–5767 (2006)
31. Howlader, N., Noone, A.M., Krapcho, M., Neyman, N., Aminou, R., Altekruse, S.F., Kosary, C.L., Ruhl, J., Tatalovich, Z., Cho, H., Mariotto, A., Eisner, M.P., Lewish, D.R., Chen, H.S., Feuer, E.J.: Seer cancer statistics review, 1975–2009 (vintage 2009 populations). Technical report, Bethesda, MD (2012)

32. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification (2010). http://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf. Accessed 22 Nov 2012
33. Iyatomi, H., Celebi, M.E., Schaefer, G., Tanaka, M.: Automated color calibration method for dermoscopy images. Comput. Med. Imaging Graph. **35**(2), 89–98 (2011)
34. Jemal, A., Siegel, R., Xu, J., Ward, E.: Cancer statistics, 2010. CA Cancer J. Clin. **60**(5), 277–300 (2010)
35. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: A review. Artif. Intell. Med. **56**(2), 69–90 (2012)
36. Land, E.H., McCann, J.J.: Lightness and retinex theory. J. Opt. Soc. Am. **61**(1), 1–11 (1971)
37. Lee, T.K., Atkins, M.S., Gallagher, R.P., MacAulay, C.E., Coldman, A., McLean, D.I.: Describing the structural shape of melanocytic lesions. In: SPIE Medical Imaging, pp. 1170–1179 (1999)
38. Madooei, A., Drew, M.S., Sadeghi, M., Atkins, M.S.: Intrinsic melanin and hemoglobin colour components for skin lesion malignancy detection. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) Medical Image Computing and Computer-Assisted Intervention MICCAI 2012, Lecture Notes in Computer Science, vol. 7510, pp. 315–322. Springer, Heidelberg (2012)
39. Maglogiannis, I., Doukas, C.N.: Overview of advanced computer vision systems for skin lesions characterization. IEEE Trans. Inf. Technol. Biomed. **13**(5), 721–733 (2009)
40. Moncrieff, M., Cotton, S., Claridge, E., Hall, P.: Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. Br. J. Dermatol. **146**(3), 448–457 (2002)
41. Nachbar, F., Stolz, W., Merkle, T., Cognetta, A.B., Vogt, T., Landthaler, M., Bilek, P., Braun-Falco, O., Plewig, G.: The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. J. Am. Acad. Dermatol. **30**(4), 551–559 (1994)
42. National Center for Biotechnology Information: Melanoma—PubMed Health (2012). http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001853. Accessed 08 Nov 2012
43. Nock, R., Nielsen, F.: Statistical region merging. IEEE Trans. Pattern Anal. Mach. Intel. **26**(11), 1452–1458 (2004)
44. Piatkowska, W., Martyna, J., Nowak, L., Przystalski, K.: A decision support system based on the semantic analysis of melanoma images using multi-elitist PSO and SVM. In: Perner, P. (ed.) Machine Learning and Data Mining in Pattern Recognition, Lecture Notes in Computer Science, vol. 6871, pp. 362–374. Springer, Heidelberg (2011)
45. van Rijsbergen, C.: Information Retrieval, 2nd edn. Butterworth-Heinemann, Newton (1979)
46. Schaefer, G., Rajab, M.I., Iyatomi, H.: Colour and contrast enhancement for improved skin lesion segmentation. Comput. Med. Imaging Graph. **35**(2), 99–104 (2011)
47. Shan, S., Gao, W., Cao, B., Zhao, D.: Illumination normalization for robust face recognition against varying lighting conditions. In: AMFG'03: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, Nice, pp. 157–164 (2003)
48. Smith, A.R.: Color gamut transform pairs. SIGGRAPH Comput. Graph. **12**(3), 12–19 (1978)
49. Soille, P.: Morphological operators. Handb. Comput. Vis. Appl. **2**, 627–682 (1999)
50. Stolz, W., Riemann, A., Cognetta, A., Pillet, L., Abmayr, W., Holzel, D., Bilek, P., Nachbar, F., Landthaler, M., Braun-Falco, O.: ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur. J. Dermatol. **4**(7), 521–527 (1994)
51. Tsumura, N., Ojima, N., Sato, K., Shiraishi, M., Shimizu, H., Nabeshima, H., Akazaki, S., Hori, K., Miyake, Y.: Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. ACM Trans. Graph. **22**(3), 770–779 (2003)
52. Wallace, T.P., Wintz, P.A.: An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors. Comput. Graph. Image Proces. **13**(2), 99–126 (1980)
53. Wong, A., Clausi, D.A., Fieguth, P.: Adaptive monte carlo retinex method for illumination and reflectance separation and color image enhancement. In: CRV'09: Canadian Conference on Computer and Robot Vision, Kelowna, pp. 108–115 (2009)
54. World Health Organization: WHO |Skin cancers (2012). http://www.who.int/uv/faq/skincancer/en/index1.html. Accessed 08 Nov 2012

Texture Information in Melanocytic Skin Lesion Analysis Based on Standard Camera Images

Pablo G. Cavalcanti and Jacob Scharcanski

Abstract The classification of melanocytic skin lesions is a very difficult task, and usually computer-aided diagnosis systems or screening systems focus on reproducing medical criteria as the ABCD rule. However, the texture information can also contribute significantly for the lesion classification, since malignant cases tends to present texture patterns different from benign cases. In this chapter, we detail five representative sets of features that have been proposed in the literature for the representation of melanocytic lesions texture information, and then we analyze how these features distinguish between malignant and benign classes using two well known classifiers.

Keywords Melanoma · Melanocytic skin lesion · Standard camera imaging · Texture analysis · Computer vision

Introduction

Discriminating benign from malignant cases of melanocytic skin lesions can be a very challenging task. Malignant melanomas and benign atypical nevi share at least some, and sometimes all, clinical characteristics detected visually [18, 35].

Different methods can be used to help diagnosing melanocytic lesions. The ABCD rule is probably the most frequently used methodology, trying to differentiate malignant and benign cases using a set of lesion characteristics, namely **A**symmetry, **B**order irregularity, **C**olor variation and lesion **D**iameter. However, there is no default protocol for acquiring the image, and the diameter measurement despite being a

P. G. Cavalcanti · J. Scharcanski (✉)
Instituto de Informática, UFRGS, Porto Alegre, Brazil
e-mail: pgcavalcanti@inf.ufrgs.br

J. Scharcanski
e-mail: jacobs@inf.ufrgs.br

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_8,

very important feature, is difficult to measure in practice. There is also a similar methodology called ABCD rule of dermoscopy [29], which is applicable of dermoscopy images and changes the D letter to **D**ifferential structures, i.e. submacroscopic morphologic and vascular structures (e.g., pigment network, dots, globules) that occur more frequently in malignant cases. These characteristics also can be an important part of other dermoscopic methodologies for the lesion diagnosis, such as the Menzies Scoring Method or the 7-point Checklist [23]. Nevertheless, these characteristics usually are not visible on standard camera images, and texture analysis can used as an attempt to overcome this difficulty. For example, the Pattern Analysis methodology [33] suggests using a type of global texture pattern (reticular, globular, homogeneous, etc) as an important feature for the lesion diagnosis.

Consequently, several texture analysis techniques have been tested in the recent years trying to quantify textural patterns in skin lesions, using both dermoscopy and standard camera images. The texture descriptors based on the gray-level co-occurrence matrices are probably the most frequently used technique [1, 10, 22], but other texture analysis approaches, such as those based on the intensity variability [6], based on independent component analysis (ICA) [39], fractal measurements [26] or local binary patterns (LBPs) [43] have been studied.

Specialists state that in the early evolution stages of malignant lesions, dermoscopy may not be helpful since it often does not improve the diagnosis accuracy [37]. Moreover, a recent study indicate that only 48 % of dermatologists in the United States use dermoscopy [17] when performing the visual inspection of the lesion (however, a definitive diagnosis is only given after the biopsy). So, in this chapter we focus on standard camera images only, which are more accessible to dermatologists and also provides the possible utilization on telemedicine [28]. Our objective is to present different techniques for the quantification of the texture patterns of melanocytic skin lesions and evaluate on a image dataset which features can better differentiate malignant and benign cases.

Texture Features Often Used for Melanocytic Skin Lesions Image Analysis

In many computer vision and image processing applications, the texture analysis is an important procedure for the success of the technique. However, many characteristics can define a texture, such as uniformity, density, roughness, regularity, direction, etc [40]. Consequently, several approaches based on different representations have been proposed in the literature aiming to represent the texture patterns of different kinds of images.

As already mentioned in section “Introduction”, texture descriptors approaches for melanocytic skin lesion images have also been proposed in the literature. In the following subsections we describe representative state-of-art approaches that are often used to differentiate benign and malignant cases of melanocytic skin

lesions. These methods are organized inspired on the classification of texture analysis methods proposed by Tuceryan and Jain [40].

Statistical Methods

The spatial distribution of gray values is a very important quality of texture, and the use of statistical features to represent the texture information is one of the early and most used methods in the computer vision literature [40]. We present next the use of co-occurrence matrices, a statistical tool frequently used for the analysis of melanocytic skin lesions.

Co-occurrence Matrices Features

A co-occurrence matrix is a matrix representing the distribution of co-occurring values at a given offset. Mathematically, a co-occurrence matrix C is defined over an $n \times m$ image I, parameterized by an offset $(\Delta x, \Delta y)$, as:

$$C_{\Delta x,\Delta y}(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1, \; if \;\; I(p,q) = i \;\; and \;\; I(p+\Delta x, q+\Delta y) = j \\ 0, \; otherwise \end{cases}. \quad (1)$$

The indexes i and j, and consequently the size of C, are defined by the range of possible values in I. For instance, if I is binary, the generated co-occurrence matrix C has size of 2×2, while a 24-bit color image generate a $2^{24} \times 2^{24}$ co-occurrence matrix. Usually, texture analysis methods compute these matrices for 8-bit grayscale images, frequently quantized for even less than 256 values, and are referred as gray-level co-occurrence matrices (GLCM).

Moreover, it should be observed that the $(\Delta x, \Delta y)$ parametrization makes the co-occurrence matrix sensitive to rotation. Unless the image is rotated in 180 degrees, any image rotation will result in a different co-occurrence distribution. So, to achieve a degree of rotation invariance, usually texture analysis procedures compute co-occurrence matrices considering rotations of 0, 45, 90, and 135 degrees. For instance, if we are considering one single pixel offsets (a reference pixel and its immediate neighbour), four co-occurrence matrices are computed using $(\Delta x, \Delta y) = \{(1,0), \; (1,-1), \; (0,-1), \; (-1,-1)\}$. This process is exemplified in Fig. 1 using an example image with four possible pixel values.

Because co-occurrence matrices are typically large and sparse, these matrices are not directly used for image analysis. In 1973, Haralick et al. [19] proposed a set of 14 metrics (frequently referred as the Haralick features) computed through the co-occurrence matrices to represent the image textural patterns, and this set of metrics and other similar metrics proposed in the literature [12, 38] have been used in the last decades for many different applications.

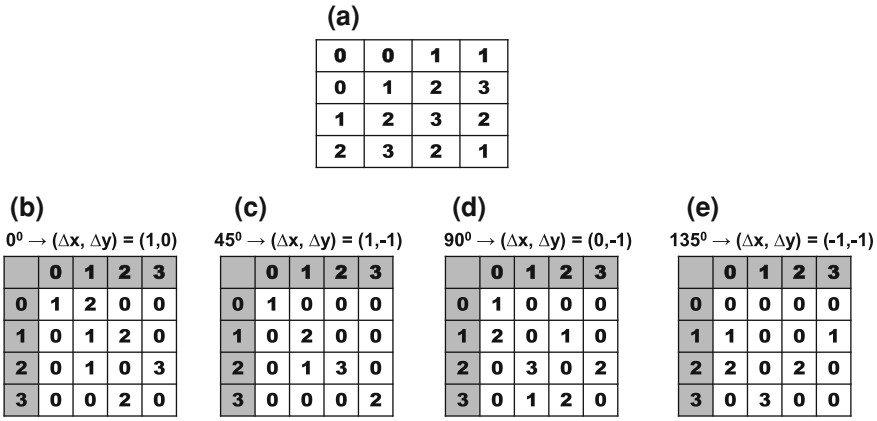


Fig. 1 Co-occurrence matrices computation example. **a** Reference example image; **b–e** Co-occurrence matrices for the rotations of 0°, 45°, 90° and 135°, respectively

As mentioned in section "Introduction", textural features based on the co-occurrence matrices are also very frequently used in melanocytic lesions analysis, though not existing a default scheme for that. Celebi et al. [10], for instance, suggest to uniformly quantize the images to 64 gray levels, and then extract 8 Haralick features for each one of the four orientations using single pixel offsets. Iyatomi et al. [22] proposed using only 4 features, but 11 different offsets. While Celebi et al. and Iyatomi et al. applied these techniques on dermoscopy images, Alcón et al. [1] used the Haralick features for the quantification of the textural structures on standard camera images. However, Alcón et al. is probably the simplest algorithm, computing only 4 Haralick features using single pixel offsets. In our experiments (see section "Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis") we will follow the Celebi et al. algorithm, as described next, which in our opinion is the most complete procedure.

Celebi et al. [10] initially divide the image in three regions. Based on a predefined segmentation, it is defined the lesion region, an outer periphery and an inner periphery. These peripheral regions are defined as adjacent regions with areas equal to 20 % of the lesion area, respectively outside and inside the lesion region. To reduce the effects of segmentation inaccuracies, these peripheral regions omit areas equal to 10 % of the lesion area next to the lesion border. Then, the images are uniformly quantized to 64 gray levels, and co-occurrence matrices C considering the four orientations are defined for each one of three regions. Finally, these matrices are normalized (i.e., divided by the total of co-occurrences), and eight Haralick features [12, 19, 38] are computed:

$$Maximum\ probability = \max_{i,j} c(i,j)\ , \tag{2}$$

$$Energy = \sum_i \sum_j c(i,j)^2\ , \tag{3}$$

$$Entropy = -\sum_i \sum_j p(i,j)\log(c(i,j))\ , \tag{4}$$

$$Dissimilarity = |i-j|\,c(i,j)\ , \tag{5}$$

$$Contrast = \sum_{n=0}^{N-1} \left\{ \sum_i \sum_j c(i,j) \,\middle|\, |i-j| = n \right\}\ , \tag{6}$$

$$Inverse\ difference = \sum \frac{c(i,j)}{1+|i-j|}\ , \tag{7}$$

$$Inverse\ difference\ moment = \sum \frac{c(i,j)}{1+(i-j)^2}\ , \tag{8}$$

$$Correlation = \frac{(ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}\ , \tag{9}$$

where $c(i,j)$ is the value of the normalized co-occurrence matrix at indexes (i,j), N is the number of gray levels, μ_x and μ_y are, respectively, the mean values of the rows and columns of c, σ_x and σ_y are the respective standard deviations, and in Eq. 6 the symbol $'|'$ indicates a condition that must be valid.

To obtain rotation invariant features, the 8 statistics are averaged over the four orientations, obtaining 24 features to represent the textural information in the three image regions. Celebi et al. also add the ratios and differences of the 8 statistics in each one of these regions, amounting 72 generated features for each single image.

In Celebi et al. experiments [10], these texture features were combined with 11 features based on the lesion shape and 354 color features to classify dermoscopy images. In section "Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis" we present some experiments and discuss its performance to classify standard camera images using these texture features standalone.

Model and Pattern Oriented Methods

The objective of model and pattern oriented texture analysis methods is to capture the essence of the texture, not only describing this information but also synthesizing it [40]. We present next typical approaches based on those assumptions often used in melanocytic skin lesion analysis.

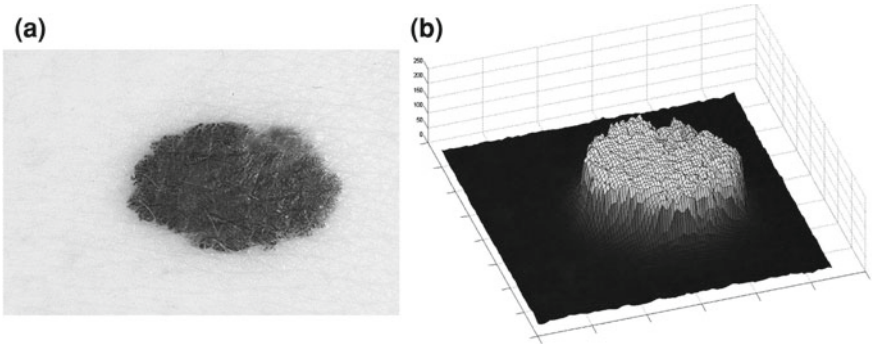


Fig. 2 Pseudoelevation surface example. **a** Melanocytic skin lesion image; **b** 3D fractal surface generated from the lesion image

Fractal Features

Given a melanocytic skin lesion image, Manousaki et al. [26, 27] proposed to create a three-dimensional pseudo-elevation surface by using the image intensities. Attributing the value of 255 to black and 0 to white (i.e., the complement of the grayscale intensities) to the z-coordinate component, the two-dimensional image ($I(x, y)$) is converted to 3D ($I(x, y, z)$). Consequently, these pseudoelevations reveals delicate differences of texture within the lesions, and appear as spatial isotropic surfaces [27]. An example of the generated 3D fractal surface can be seen in Fig. 2.

As we know, points, curves, surfaces and cubes are described in Euclidean geometry using integer dimensions of 0, 1, 2, and 3, respectively. A measure of an object such as the length of a line, the area of a surface and the volume of a cube are associated with each dimensions, and these measurements are invariant with respect to the used unit. However, almost any object found in nature appear disordered and irregular for which the measures of length, area and volume are scale-dependent. This suggests that the dimensions of such objects cannot be integers, i.e. these objects should be represented using fractal dimensions. For instance, a melanocytic lesion must have a dimension between 2 and 3, i.e. the dimension of an object that is not actually regular in shape and not a cube.

Many methods were proposed to compute the fractal dimension of an object [36]. The most used technique is probably the Minkowski–Bouligand dimension, also known as Minkowski dimension or box-counting dimension. Suposing the fractal on an evenly-space grid, we must count the number of boxes required to cover the object. The box-counting dimension is calculated by seeing how this number changes as we make the grid finer. Assuming $N(\varepsilon)$ as the number of boxes of side ε required to cover the object, the fractal dimension D is defined as:

$$D = \lim_{\varepsilon \to 0} \frac{\log\ N(\varepsilon)}{\log\,(1/\varepsilon)}. \tag{10}$$

In [27], Manousaki et al. used a modified algorithm of the Minkowski-Bouligand dimension proposed by Dubuc et al [16], using disks instead of boxes in its implementation. They performed experiments with 132 melanocytic skin lesion images, 23 of them being melanomas, 44 atypical nevi and 65 common nevi. While the melanomas achieved fractal dimension of 2.49 ± 0.10 (average $\pm$ standard deviation), the atypical nevi resulted 2.44 ± 0.11, and the common nevi resulted 2.46 ± 0.07.

In combination with fractal dimension, Manousaki et al. [26, 27] proposed the use of lacunarity parameter to classify the different textures. Lacunarity is a quantitative measure to describe differences in appearance of fractal surfaces with the same fractal dimension [25], measuring the distribution and size of "holes" within a textured surface.

Again, many different methods were proposed to compute the surface lacunarity [2]. Manousaki et al. suggested the gliding box method proposed by Allain and Cloitre [2]. A gliding box with size $r \times r$ is placed directly over the image and moves as a simple window, and the box mass M is determined as the sum of the box content. The lacunarity at scale r is defined as the variation of the computed M values over the image divided by its square mean. Using the same image database with 132 lesion images, Manousaki et al. computed for melanomas lacunarity of 0.42 ± 0.35 (average $\pm$ standard deviation), while aytpical nevi resulted 0.25 ± 0.22, and common nevi resulted 0.19 ± 0.14.

Manousaki et al. did not experimented these computed textural descriptors in a classification step. However, they observed through statistical tests (Kruskal-Wallis and Mann-Whitney tests) that lacunarity can be potentially used to differentiate malignant and benign skin lesions, while fractal dimension do not provided significant statistical differentiation.

LBP Features

Local binnary pattern (LBP) is a texture operator that was proposed by Ojala et al. [30]. and has already been used for many different applications. This technique labels the pixels of an image by thresholding the neighborhood of each pixel based on its value, and generates a binary number to quantify the local texture. Let us consider the following 3×3 image to illustrate its functionality:

As can be observed in Fig. 3, all neighbor pixels with value higher or equal to the center pixel value assume '1' in the LBP, while the neighbor pixels with lower value assume '0'. Arranging these 8 binary values in sequence we obtain a 8-digit binary number, i.e. the local texture of the 3×3 neighborhood can be represented by $2^8 = 256$ different labels. After computing the LBP for each image position, usually a histogram is used to describe the texture information of the whole image. This methodology also has been extended for larger neighborhoods, considering circles of larger radius and a large number of pixels [31].

However, if we observe again Fig. 3, it is obvious that the binary number generated from the neighborhood is dependent of the initial position to arrange the val-

(a)

10	15	60
16	55	70
22	55	50

(b)

0	0	1
0		1
0	1	0

Fig. 3 Local Binnary Pattern (LBP) computation example. **a** Reference example image; **b** Computed LBP (the 8 neighbor pixels are thresholded by the center pixel value)

ues. Although usually the clockwise direction is adopted, we can obtain 8 different numbers (or more, if larger neighborhoods are used). For instance, if we start at the top-left pixel we obtain 00110100, while we obtain 01000011 if we start arranging from the bottom-right pixel. So, to create a rotation invariant descriptor, Ojala et al. proposed to generated all the possible binary numbers, and the resultant LBP is the minimum possible value. In this case, the minimum possible binary value is 00001101 (i.e., starting the arrangement from the bottom-left pixel), and that would be the resultant rotation variant LBP to the example image in Fig. 3a.

Also, Ojala et al. proposed a step to reduce the feature vector. For instance, the 256 bins histogram can be a too high dimensional vector to be used in a posterior classification step. Considering that many applications use 16 neighbor pixels, the generated $2^{16} = 65,536$ bins histograms make unpractical the use of LBPs. So, they also proposed the concept of uniform pattern, which was inspired by the fact that some binary patterns occur more commonly in texture images than others. A LBP is called uniform if its binary number contains at most two bitwise transitions from 0 to 1 or vice versa. For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010010 (6 transitions) are not. Since LBPs with lower number of transitions are much more frequent, the final histogram is computed in such way that all the non-uniform patterns are counted in a single bin. For instance, in the 256 possible patterns of a 3×3 neighborhood, 58 of which are uniform, and consequently the generated feature vector will be 59-dimensional. Moreover, if we consider the rotation invariant descriptors, this dimensionality is reduced to only 10 bins.

Zortea et al. [43] proposed to use LBPs for the texture analysis of melanocytic skin lesions. Although their experiments have been performed with dermoscopy images, we present their proposed algorithm here understanding that it is a valuable contribution to the literature. Moreover, our experiments (see section "Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis") indicate that this technique is a potentially useful also for standard camera images. Their algorithm starts by dividing the whole image in windows of $w \times w$ pixels, with 50 % overlapping of adjacent windows. Then, uniform rotation invariant LBPs of 16 neighbor pixels are computed, and a 18-dimensional feature vector is generated to each one of these windows. Finally, the whole set of computed features vectors is clustered in K sub-classes (using the K-means algorithm) and we obtain a dictionary of possible LBPs presented in melanocytic lesion images.

We may observe that there is no segmentation before the texture feature extraction, so the "skin" texture pattern is also presented in the generated dictionary of K patterns. Zortea et al. understand that the "skin" pattern will be present in both malignant and benign classes, and consequently will not influence the posterior classification. Furthermore, that would make this technique not-dependent of segmentation inaccuracies.

So, after obtained the K-dimensional texture dictionary, LBPs are computed in the testing image set, and the frequency of occurrence of each of the K textural patterns is used as a textural representative vector. Additionally, they included a feature accounting the number of different patterns in each single image, providing a $K + 1$ feature vector. The algorithm is based on the idea that malignant cases would generated higher frequencies for some textural patterns, while benign cases higher frequencies for different patterns of the dictionary.

In their experiments, Zortea et al. [43] used a Support vector machine (SVM) classifier with Radial Basis Function kernel and tested different parameters specification. The best classification results have been obtained using $w = 40$ pixels and $K = 22$ clusters, achieving sensitivity of 73.3 % and specificity of 73.9 %. In our experiments (see section "Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis"), we randomly selected 30 images from each class for the computation of the texture dictionary, and we also varied the K parameter to check the performance of these LBP features in a classification procedure.

Signal Oriented Methods

Many texture analysis techniques have been proposed relying on signal processing techniques, such as features computed from filtered images [40]. We present next two approaches often used for the melanocytic skin lesions analysis focused on that idea.

ICA Features

Tabatabaie et al. [39] proposed a filtering method based on independent component analysis (ICA) to describe the texture information of melanocytic skin lesion images. ICA is a method to construct independent components from an ensemble of data [34]. Let X be a linear mixture of mutually statistically independent source signals S:

$$X = AS, \tag{11}$$

where A is the mixing matrix. The objective of ICA is to identify the independent component of S, given the matrix X. However, since both A and S are unknown, this is a typical case of blind source separation, i.e. no information is provided about the set of signals that should be estimated. So, several methods have been proposed to

estimate a matrix W:

$$Y = WX, \tag{12}$$

where the inverse of W is a estimation of A, and consequently Y is a estimation of S. The most frequently used methods to compute W is the Non-Gaussianity family of ICA algorithms. According to the central limit theorem, the probability density function of a sum of n independent random variables tends to a Gaussian, as n tends to infinity. So, these ICA algorithms goals to compute W in a way that each recovered signal in Y is as non-Gaussianity as possible [34].

Tabatabaie et al. [39] proposed to use FastICA [20, 21], an ICA algorithm based on non-Gaussianity, to compute texture features of melanocytic skin lesions. The approach starts by linearly normalizing each image, i.e. the processed images have zero mean and unit variance. Then, thousands of small image patches of 16×16 pixels are extracted from inside the lesion areas. From each patch the local mean is subtracted to remove the DC component, and all the patches are concatenated to create a matrix of mixed signals X. Finally, the FastICA is applied. This algorithm initially perform data whitening and dimension reduction through principal component analysis (PCA), and then identify the independent components using a fixed-point iteration scheme. Tabatabaie et al. suggest the hyperbolic tangent as the nonlinear function to estimate the non-Gaussianity of the signals, and obtained 100 independent components from FastICA. This technique has been used both for the set of malignant images and for the set of benign images, resulting in 200 independent components.

After the ICA application, each obtained independent component is used as an image filter, and the energy of an image I in response to the filters is used to represent the texture features. This energy is defined as:

$$E_i = \| I \otimes filter_i \| , \tag{13}$$

where $\otimes$ represents the 2D convolution, and $\|.\|$ denotes the Euclidean norm. Considering that we obtain 200 components/filters (i.e., $i \in [1, 200]$), we obviously also obtain 200 energy values or texture features E_i for each image. Ideally, malignant cases may result stronger energies from the filters generated from the malignant patches, while benign cases may result stronger energies from the filters generated from the benign patches. Consequently, these features can be used in a classification scheme aiming to differentiate malignant and benign cases.

In their experiments, Tabatabaie et al. [39] ranked the 200 computed features by t-test separability criteria [42], and a Support Vector Machine (SVM) classifier to identify the images as malignant or benign using from 1 to 200 features. The best results were obtained when a radial basis function (RBF) kernel was used, and approximately 25 features,[1] achieving 73.7 % of accuracy. These experiments were performed in 160 images collected from online atlases [13, 14].

[1] The results are presented in a line chart, and the exact number is not provided.

Tabatabaie et al. also proposed the combination of these ICA-based textural features with some color descriptors to improve the classification results, achieving 88.7 % of accuracy. However, since color features are out of the scopus of this chapter, we refer the reader to their work [39] to obtain more details about that.

Section "Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis" presents some experimental results, comparing the classification results of melanocytic lesions using these features with other texture descriptors features. In these experiments, we randomly selected 30 images of each class and extracted 100 patches from each image, obtaining 3,000 patches to generate the 100 components/filters for each class.

Intensity Variability Features

Skin lesions often have more local textural variability than healthy skin areas, and that can be used to differentiate healthy and unhealthy skin regions. Cavalcanti and Scharcanski [6] proposed to quantify this texture information in a image pixel $I(x, y)$ through the intensity variability of its neighborhood. Their method starts by computing $\tau(x, y, \sigma)$:

$$\tau(x, y, \sigma) = \frac{I(x, y)}{S(x, y, \sigma)} - I(x, y), \tag{14}$$

where, $S(x, y, \sigma) = I(x, y) \otimes G(\sigma)$ (i.e., image I is smoothed by a Gaussian filter with standard deviation σ). Re-arranging terms in Eq. 14, $\tau(x, y, \sigma)$ can also be defined as:

$$\begin{aligned} \tau(x, y, \sigma) &= \frac{I(x, y) - I(x, y) \cdot S(x, y, \sigma)}{S(x, y, \sigma)} \\ &= \frac{I(x, y)\,(1 - S(x, y, \sigma)))}{S(x, y, \sigma)} \\ &= I(x, y)\frac{\widetilde{S}(x, y, \sigma)}{S(x, y, \sigma)}, \end{aligned} \tag{15}$$

where, $\widetilde{S}(x, y, \sigma)$ represents the complement of the $\bar{L}$ smoothed image [(i.e. $\widetilde{S}(x, y, \sigma) = 1 - S(x, y, \sigma)$]. If an image region is dark (i.e. has low pixel intensities, as often occurs in skin lesions), $\widetilde{S}(x, y, \sigma) > S(x, y, \sigma)$ and the ratio $\widetilde{S}(x, y, \sigma)/S(x, y, \sigma)$ tends to increase, and the local region and textural information are emphasized; if the region is bright (e.g. as in healthy skin regions), $\widetilde{S}(x, y, \sigma) < S(x, y, \sigma)$ and the ratio $\widetilde{S}(x, y, \sigma)/S(x, y, \sigma)$ tends to decrease, and the local region and its textural information are de-emphasized. This process is illustrated in Fig. 4. Usually, most pixels in I correspond to healthy skin (i.e., have higher intensity values). However, most skin pixels have lower values (closer to zero) in

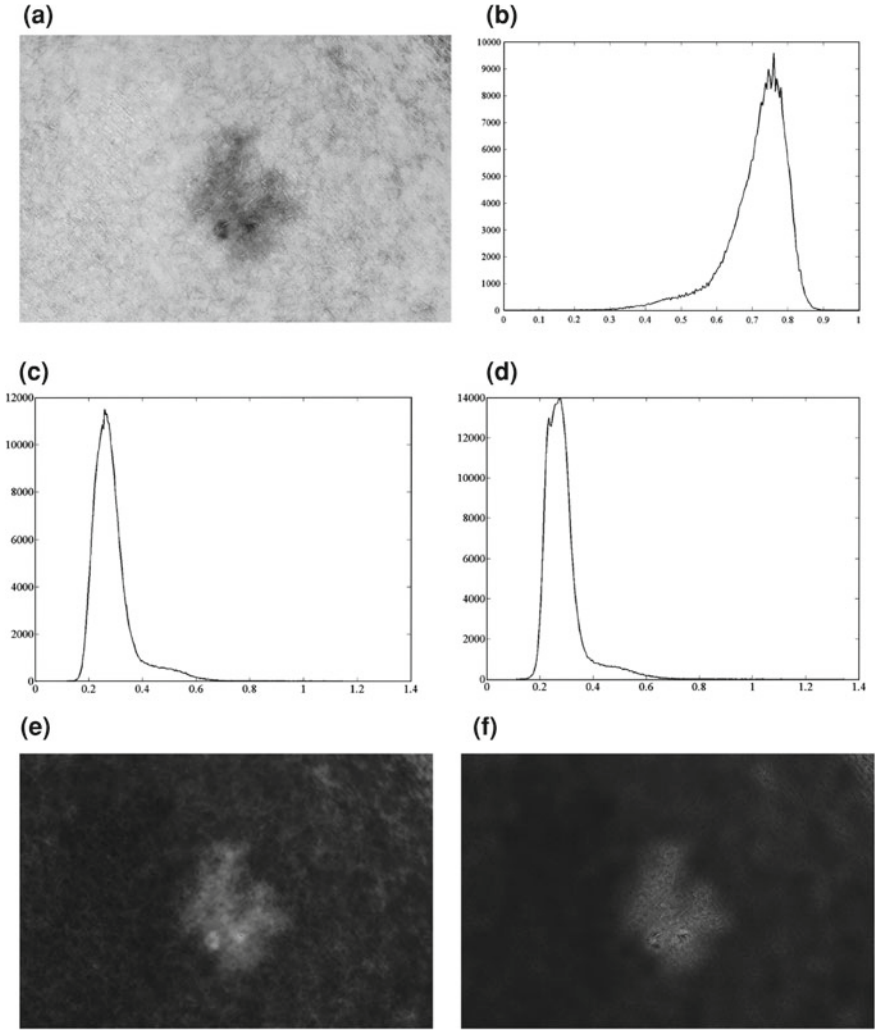


Fig. 4 Texture variability quantification $\tau(x, y, \sigma)$. **a** Melanocytic skin lesion image; **b** Histogram of the image; **c** and **d** show the histograms of $I(x, y)\widetilde{S}(x, y, \sigma)/S(x, y, \sigma)$ (see Eq. 15) using $\sigma = 1$ and $\sigma = \frac{43}{7}$, respectively; **e** and **f** show the image instances associated to the histograms shown in Figs. **c** and **d**

$\tau(x, y, \sigma)$, while skin lesion pixels tend to be brighter, as the histogram peaks show in Figs. 4c and d. [2]

A single Gaussian filter may not be sufficient to capture the intensity variability of different types of lesions in generic images. So, $\tau(x, y, \sigma)$ is computed for different

[2] The higher histogram peaks are associated to healthy skin regions in Figs. 4c and d.

σ values[3] $\{\sigma_1, \sigma_2, ..., \sigma_N\}$, and select at each pixel the maximum $\tau(x, y, \sigma)$ among all scales σ:

$$T(x, y) = \max_{\sigma}[\tau(x, y, \sigma)],\ \sigma \in \{\sigma_1, \sigma_2, ..., \sigma_N\}. \tag{16}$$

The procedure described in Eq. 16 was designed to capture the local intensity variability information at each pixel (x, y), in lesions of different sizes, shapes and texture patterns. And, finally, texture variability channel I_{text} is computed by normalizing T:

$$I_{text}(x, y) = (T(x, y) - \min(T)) \,/\, (\max(T) - \min(T)). \tag{17}$$

As mentioned in the beginning of this section, the main objective when proposing this intensity variation channel I_{text} was differentiate healthy and unhealthy skin texture patterns, aiming to help the segmentation task. However, Cavalcanti and Scharcanski [6] also included 4 features based on the channel I_{text} on their classification system to differentiate malignant and benign melanocytic lesions in standard camera images. These 4 features are the maximum, minimum, average and variance of the I_{text} values inside the lesion region, based on a precomputed segmentation. Combining these features with other 48 features based on the lesion shape, border and color, Cavalcanti and Scharcanski obtained high accuracy levels. In section "Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis" we analyze the performance of these features alone for classifying melanocytic lesions.

Experimental Comparison of Texture Features for Melanocytic Skin Lesion Image Analysis

In order to compare the performances of the five state-of-art sets of texture features for melanocytic skin lesions images presented before, we use the image dataset proposed in Alcón et al. [1], which contains 152 images that have been collected from the Dermnet atlas [14]. This dataset consists of 107 melanomas and 45 atypical nevi, a benign kind of lesion that present similar characteristics to melanomas. We extracted the texture features from these 152 images, and provided a label to each image using two different classifiers. However, before we present these results, we detail the preprocessing steps that were necessary before the feature extraction task.

[3] Cavalcanti and Scharcanski [6] suggest using $\sigma = 1, \frac{11}{7}, \frac{15}{7}, ..., \frac{43}{7}$, and filter window sizes of $7\sigma \times 7\sigma$.

Preprocessing and Skin Lesion Segmentation

As already mentioned, our experiments were performed with standard camera images, i.e. standard photographs of melanocytic skin lesions. However, it is not trivial to acquire reliable standard photographs of these lesions, since these skin lesions usually are only a few millimeters width, the camera should be placed near the the skin and the acquired image usually is affected by shading effects. So, before to realize any processing or analysis with these images, we perform a preprocessing step to attenuate the image shading effects using the Cavalcanti et al. [8] algorithm [5], as described next.

Cavalcanti et al. [9] proposed a shading attenuation method that is adaptive to the image data. Their method assumes that images are acquired in a way that the lesion appears in the image center, and it does not touch the image outer borders. The first step of the method is to convert the image from the original RGB color space to the HSV color space, and retain the Value channel V. This is justified by the fact that this channel presents the higher visibility of the shading effects. A region of 20×20 pixels is extracted from each one of the four V corners, and the union of these four sets define the pixel set S (with a total of 1600 image pixels). This pixel set is used to adjust the following quadric function $z(x, y)$:

$$z(x, y) = P_1 x^2 + P_2 y^2 + P_3 xy + P_4 x + P_5 y + P_6, \tag{18}$$

where the six quadric function parameters P_i $(i = 1, ..., 6)$ are chosen to minimize the error ε:

$$\varepsilon = \sum_{j=1}^{1600} \left[V(S_{j,x}, S_{j,y}) - z(S_{j,x}, S_{j,y}) \right]^2, \tag{19}$$

where, $S_{j,x}$ and $S_{j,y}$ are the x and y coordinates of the jth element of the set S, respectively.

Calculating the quadric function $z(x, y)$ for each image spatial location (x, y), we have an estimate $z(x, y)$ of the local illumination intensity in the image $V(x, y)$. Dividing the original $V(x, y)$ channel by $z(x, y)$, we obtain a new Value channel where the shading effects have been attenuated. The final step is to replace the original Value channel by this new Value channel, and convert the image from the HSV color space to the original RGB color space. In Fig. 5, an example of applying this method to a skin lesion image is presented. As can be seen, the result is a color image easier to be analyzed once the lesion region is more emphasized in relation to the healthy skin. Since the methods presented in section "Texture Features Often Used for Melanocytic Skin Lesions Image Analysis" were proposed for grayscale images, we convert the color images to grayscale using the weighted sum $(0.2989\ R + 0.5870\ G + 0.1140\ B)$ from the NTSC/YIQ standard [24].

Some of the texture methods presented in section "Texture Features Often Used for Melanocytic Skin Lesions Image Analysis" require a precomputed image segmentation, i.e. the texture features are extracted based on a previous identification

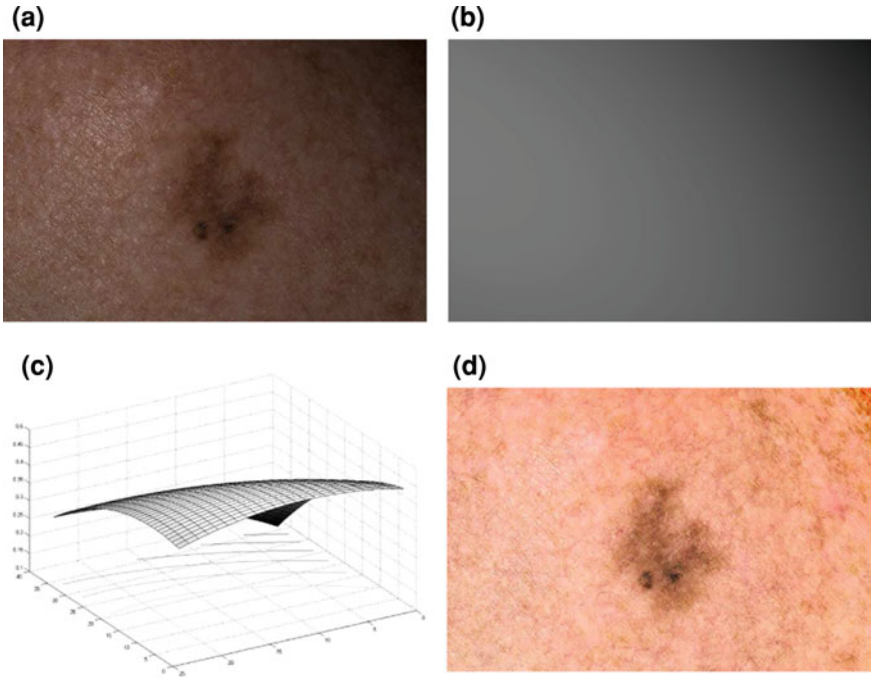


Fig. 5 Shading attenuation example. **a** Input image; **b** Obtained quadric model using the corners of the input Value channel; **c** Obtained quadric model in 3D; **d** Result obtained by the division of the Value channel by the obtained quadric model

of the lesion region. To perform this task we used the ICA-Based Active-Contours segmentation method as proposed by Cavalcanti et al. [9]. They observed that when applying ICA to the melanocytic skin lesions color images, the independent component with most non-Gaussian histogram correspond to a large part of the lesion region. So, this component is thresholded using the Otsu's method [32], and the thresholded image is used as initialization to the Chan-Vese Active-contours method for vector-valued images [11]. After some post-processing techniques, this segmentation method achieved the lower segmentation errors in previous experiments with this same dataset [7]. Some segmentation results are presented in Fig. 6, and we indicate the original paper [9] for more details about this method.

Comparative Results and Discussion

To compare how the presented sets of texture features are able to distinguish the textural patterns of malignant and benign cases of melanocytic skin lesions, we

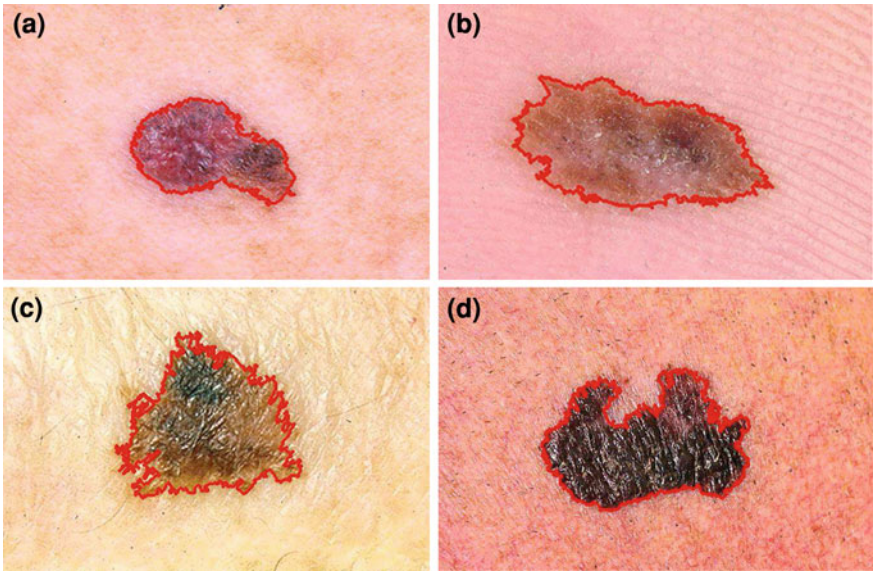


Fig. 6 Segmentation results using the ICA-Based Active-Contours method

applied two classifiers on the computed features by each one of these methods. These classifiers are:

- K-nearest neighbors (KNN): classifies samples based on K closest matches in the N-dimensional feature space [3]. We use $K = 1$ and each sample is assigned to its nearest neighbor class by using the Euclidean distance. This classifier has already been used in previous approaches for the classification of melanocytic skin lesions [4, 15], including very accurate results for this same dataset [6].
- support vector machine (SVM): based on a training set, this classifier [41] constructs a hyperplane representing the largest separation between the two classes in the N-dimensional feature space, i.e. a hyperplane that maximizes the distance from it to the nearest sample on each side. To facilitate the hyperplane computation, usually SVMs are used with the advantage of the "kernel trick", projecting the data to a higher dimensional space. Then, testing samples are classified based on its localization in relation to the hyperplane. This method has also been used in previous approaches for the classification of melanocytic skin lesions [10, 15, 39]. In our experiments, we used SVM with linear kernel.

During these experiments, we used 5-fold cross-validation. It means that 80 % of the samples are used for training, while 20 % of the samples are tested. This process is repeated 5 times, alternating the samples in the test set and assuring that all samples receive a label. We also varied the number of features used. Following the approach proposed by Tabatabaie et al. [39] (see section "ICA Features"), we ranked the features using the t-test separability criteria [42], and classified each sample using

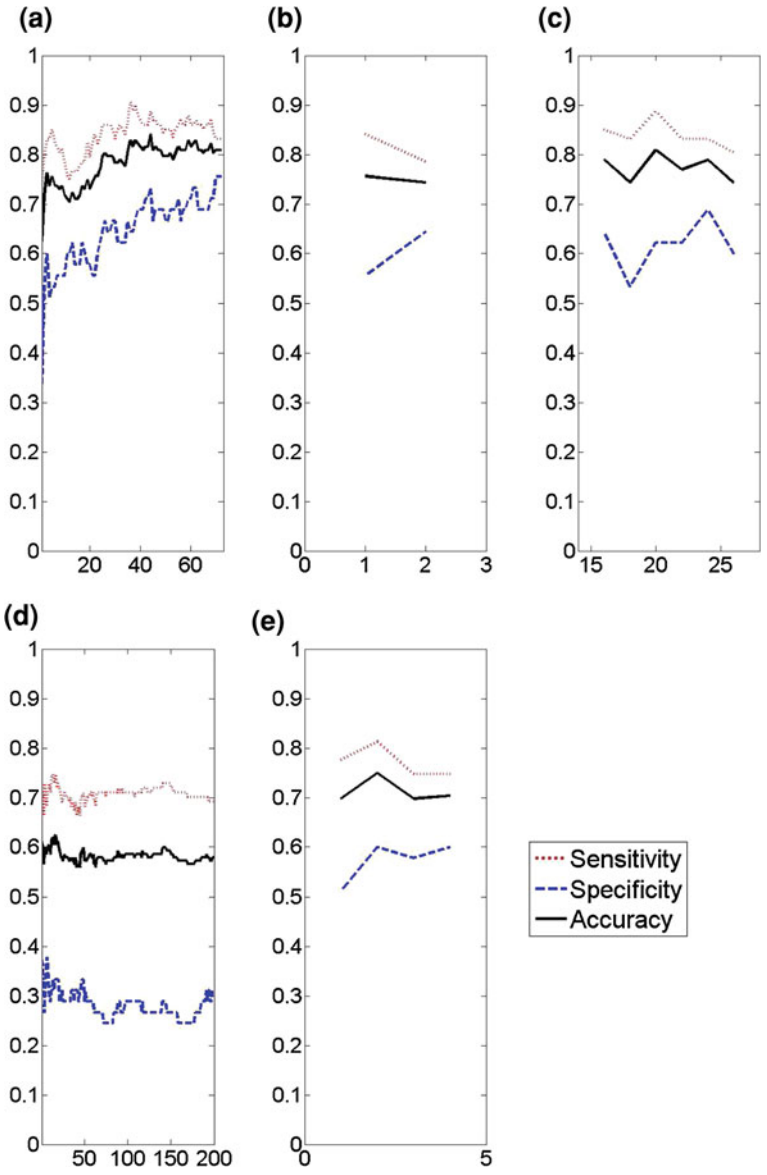


Fig. 7 Classification results by the KNN classifier using: **a** Co-occurrence matrices features; **b** Fractal features; **c** LBP features; **d** ICA features; **e** Intensity variability features. The experiments were performed varying the number of features, as indicate in the *x*-axis

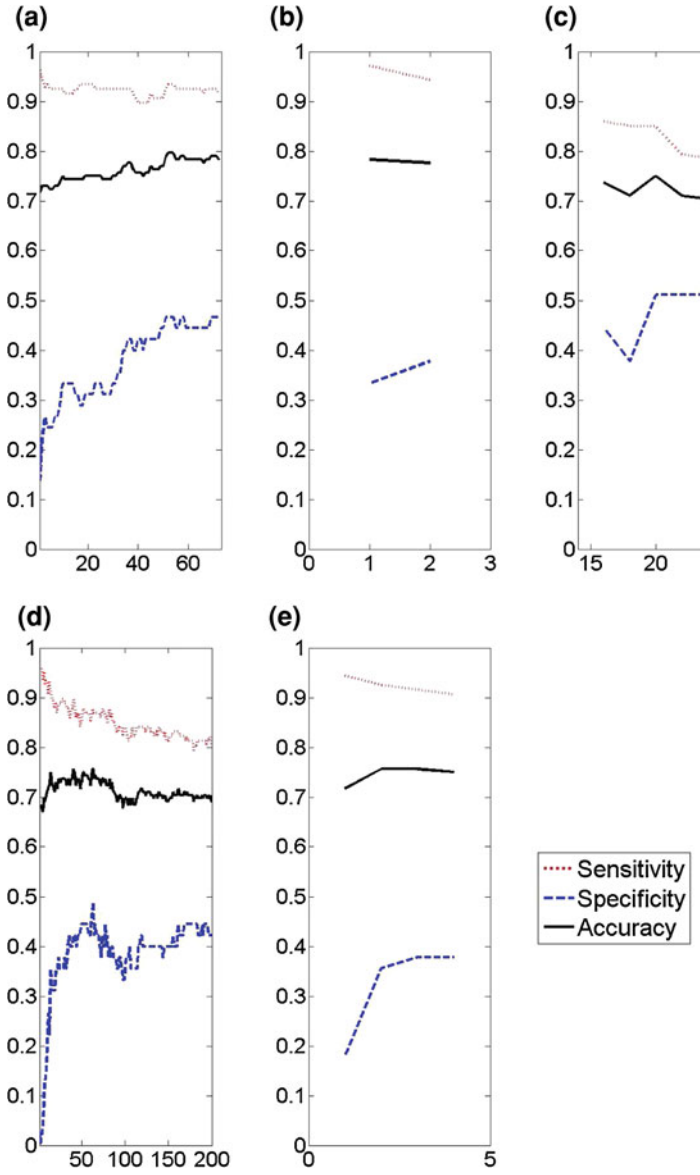


Fig. 8 Classification results by the SVM classifier using: **a** Co-occurrence matrices features; **b** Fractal features; **c** LBP features; **d** ICA features; **e** Intensity variability features. The experiments were performed varying the number of features, as indicate in the *x*-axis

from 1 to N features (where N is the maximum number of features generated by the texture descriptor). The only exception was the LBP features (see section "LBP Features"), which we varied the number of features by changing the K parameter,

Table 1 Best classification results by the KNN classifier

Approaches	Sensitivity (%)	Specificity (%)	Accuracy (%)	Number of features
Co-occurrence matrices	88.78	73.33	84.21	44
Fractal	84.11	55.55	75.65	1
LBP	88.78	62.22	80.92	20
ICA	74.76	33.33	62.50	17
Intensity variability	81.30	60.00	75.00	2

Table 2 Best classification results by the SVM classifier

Approaches	Sensitivity (%)	Specificity (%)	Accuracy (%)	Number of features
Co-occurrence matrices	93.45	46.66	79.60	52
Fractal	97.19	33.33	78.28	1
LBP	85.04	51.11	75.00	20
ICA	89.71	42.22	75.65	40
Intensity variability	92.52	35.55	75.65	2

i.e. the number of clusters identified by the K-means algorithm and consequently the size of the texture dictionary.

We analyze the classification results in terms of three metrics ordinarily used for this purpose: (a) Sensitivity (i.e, the percentage of images correctly classified in the malignant class); (b) Specificity (i.e., the percentage of images correctly classified in the benign class); (c) Accuracy (i.e., the percentage of images correctly classified overall, considering all images). The obtained results for the texture features obtained by the implementation of the five presented sets of texture features are presented in Figs. 7 and 8. Figure 7 shows the classification results using the KNN classifier, and the best results (in terms of accuracy) for each one of the five methods are presented in Table 1. Figure 8 and Table 2 present the respective results using the SVM classifier.

Analyzing the classification results we can observe that the best results are not usually obtained using all the computed texture features, and the dimensionality reduction is important to enhance the classifier performance. Furthermore, the higher number of malignant cases than benign cases in the image dataset contributed for better sensitivity than specificity percentages. Finally, the results indicate that the use of texture feature standalone hardly generate accurate results as other literature approaches that combine those features with shape, color and other kind of features, and obtain accuracies higher than 90 % [6, 10].

Not coincidentally, the texture descriptor that obtained the best classification results in our experiments is the Co-occurrence matrices features, which is the method that received more attention in the literature [1, 10, 22]. However, the implemented algorithm proposed by Celebi et al. [10] takes advantage of other image information and not only the lesion region, such as the outer and inner periphery of the lesion. In opposite way, the fractal features do not require preprocessing steps and with only one feature (the lacunarity measure, see section “Fractal Features”) obtained promising classification results.

Summary and Future Trends

In this chapter, we reviewed five approaches proposed in the literature for the numerical representation of the texture information of melanocytic skin lesions. Then, we extracted these five sets of features from a dataset of standard camera images containing malignant and benign cases, and verified how these features help classifying these lesions.

According to our experiments, the set of texture features based on Co-occurrence matrices is the best option to represent the texture information distinction between malignant and benign lesions. However, we also observe that simpler approaches also obtained promising results, and may receive attention in future works. Moreover, the combination of texture features from different algorithms is a possibility that has not been tested, as far as we know.

In particular, we believe that computational texture analysis methods is an open research topic, and future works will contribute to better representing the texture information of melanocytic lesions and earlier identifying malignant melanomas cases.

References

1. Alcon, J.F., Ciuhu, C., ten Kate, W., Heinrich, A., Uzunbajakava, N., Krekels, G., Siem, D., de Haan, G.: Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J. Sel. Top. Sig. Process. **3**(1), 14–25 (2009)
2. Allain, C., Cloitre, M.: Characterizing the lacunarity of random and deterministic fractal sets. Phys. Rev. A **44**, 3552–3558 (1991)
3. Alpaydin, E.: Introduction to Machine Learning (Adaptive Computation and Machine Learning). The MIT Press, Cambridge (2004)
4. Burroni, M., Corona, R., DellÉva, G., Sera, F., Bono, R., Puddu, P., Perotti, R., Nobile, F., Andreassi, L., Rubegni, P.: Melanoma computer-aided diagnosis: Reliability and feasibility study. Clin. Cancer Res. **10**, 1881–1886 (2004)
5. Cavalcanti, P.G., Scharcanski, J.: Shading Correction in Human Skin Color Images. In: Mishra, A., Nawaz, Z., Shahid, Z. (eds.) An Introductory Guide to the Emerging Áreas in Digital Image Procesing. 1ed. Sunnybank Hills, Australia: iConcept Press, 1–16 (2013a)
6. Cavalcanti, P.G., Scharcanski, J.: Automated prescreening of pigmented skin lesions using standard cameras. Comput. Med. Imaging Graph. **35**(6), 481–491 (2011)
7. Cavalcanti, P.G., Scharcanski, J.: Macroscopic pigmented skin lesion segmentation and its influence on lesion classification and diagnosis. In: Celebi, M.E. and Schaefer G. (eds.) Color Medical Image Analysis, Lecture Notes in Computational Vision and Biomechanics, vol. 6, pp. 15–39. Springer Netherlands (2013b)
8. Cavalcanti, P.G., Scharcanski, J., Lopes, C.B.O.: Shading attenuation in human skin color images. In: Proceedings 6th International Symposium on Visual Computing (ISVC 2010) (2010)
9. Cavalcanti, P.G., Scharcanski, J., Persia, L.E.D., Milone, D.H.: An ICA-based method for the segmentation of pigmented skin lesions in macroscopic images. In: Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology, Society., pp. 5993–5996 (2011)

10. Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imaging Graph. **31**(6), 362–373 (2007)
11. Chan, T.F., Sandberg, B.Y., Vese, L.A.: Active contours without edges for vector-valued images. J. Vis. Commun. Image Represent. **11**(2), 130–141 (2000)
12. Clausi, D.A.: An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. **28**(1), 45–62 (2002)
13. Dermis: Dermatology information system. http://www.dermis.net (2011)
14. Dermnet skin disease image atlas. http://www.dermnet.com (2009)
15. Dreiseitl, S., Ohno-Machado, L., Kittler, H., Vinterbo, S., Billhardt, H., Binder, M.: A comparison of machine learning methods for the diagnosis of pigmented skin lesions. J. Biomed. Inform. **34**(1), 28–36 (2001)
16. Dubuc, B., Quiniou, J.F., Roques-Carmes, C., Tricot, C., Zucker, S.W.: Evaluating the fractal dimension of profiles. Phys. Rev. A **39**, 1500–1512 (1989)
17. Engasser, H.C., Warshaw, E.M.: Dermatoscopy use by us dermatologists: A cross-sectional survey. J. Am. Acad. Dermatol. **63**(3), 412–419 (2010)
18. Fikrle, T., Pizinger, K.: Digital computer analysis of dermatoscopical images of 260 melanocytic skin lesions; perimeter/area ratio for the differentiation between malignant melanomas and melanocytic nevi. J. Eur. Acad. Dermatol. Venereol. **21**(1), 48–55 (2007)
19. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **6**, 610–621 (1973)
20. Hyvärinen, A.: Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Netw. **10**(3), 626–634 (1999)
21. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley, New York. (2001)
22. Iyatomi, H., Oka, H., Celebi, M., Hashimoto, M., Hagiwara, M., Tanaka, M., Ogawa, K.: An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. Comput. Med. Imaging Graph. **32**(7), 566–579 (2008)
23. Johr, R.H.: Dermoscopy: alternative melanocytic algorithms - the ABCD rule of dermatoscopy, Menzies scoring method, and 7-point checklist. Clin. Dermatol. **20**(3), 240–247 (2002)
24. Lee, H.C.: Introduction to Color Imaging Science. Cambridge University Press, New York (2005)
25. Mandelbrot, B.B.: The Fractal Geometry of Nature. W. H. Freedman and Co., New York (1983)
26. Manousaki, A.G., Manios, A.G., Tsompanaki, E.I., Panayiotides, J.G., Tsiftsis, D.D., Kostaki, A.K., Tosca, A.D.: A simple digital image processing system to aid in melanoma diagnosis in an everyday melanocytic skin lesion unit: a preliminary report. Int. J. Dermatol. **45**(4), 402–410 (2006)
27. Manousaki, A.G., Manios, A.G., Tsompanaki, E.I., Tosca, A.D.: Use of color texture in determining the nature of melanocytic skin lesionsa qualitative and quantitative approach. Comput. Biol. Med. **36**(4), 419–427 (2006)
28. Massone, C., Wurm, E.M.T., Hofmann-Wellenhof, R., Soyer, H.P.: Teledermatology: An update. Semin. Cutan. Med. Surg. **27**(1), 101–105 (2008)
29. Nachbar, F., Stolz, W., Merkle, T., Cognetta, A.B., Vogt, T., Landthaler, M., Bilek, P., Braun-Falco, O., Plewig, G.: The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. J. Am. Acad. Dermatol. **30**(4), 551–559 (1994)
30. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recogn. **29**(1), 51–59 (1996)
31. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machin. Intell. **24**(7), 971–987 (2002)
32. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
33. Pehamberger, H., Steiner, A., Wolff, K.: In vivo epiluminescence microscopy of pigmented skin lesions. i. pattern analysis of pigmented skin lesions. J. Am. Acad. Dermatol. **17**, 571–583 (1987)

34. Petrou, M., Petrou, C.: Image Processing: The Fundamentals, 2nd. edn. Wiley, New York (2010)
35. Rao, B.K., Marghoob, A.A., Stolz, W., Kopf, A.W., Slade, J., Wasti, Q., Schoenbach, S.P., De-David, M., Bart, R.S.: Can early malignant melanoma be differentiated from atypical melanocytic nevi by in vivo techniques? Skin Res. Technol. **3**(1), 8–14 (1997)
36. Sahoo, P., Barman, T., Davim, J.: Fractal Analysis in Machining. Computational Mechanics. Springer, SpringerBriefs in Applied Sciences and Technology, Berlin (2011)
37. Skvara, H., Teban, L., Fiebiger, M., Binder, M., Kittler, H.: Limitations of dermoscopy in the recognition of melanoma. Arch. Dermatol. **141**, 155–160 (2005)
38. Soh, L., Tsatsoulis, C.: Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. IEEE Trans. Geosci. Remote Sens. **37**(2), 780–795 (1999)
39. Tabatabaie, K., Esteki, A., Toossi, P.: Extraction of skin lesion texture features based on independent component analysis. Skin Res. Technol. **15**(4), 433–439 (2009)
40. Tuceryan, M., Jain, A.K.: Texture analysis. In: Chen, C.H. (ed.) Handbook of Pattern Recognition and Computer Vision, 4th edn. World Scientific Publishing, River Edge (2010)
41. Vapnik, V.N.: Statistical Learning Theory, 1st edn. Wiley-Interscience, New York (1998)
42. Zhu, W., Wang, X., Ma, Y., Rao, M., Glimm, J., Kovach, J.: Detection of cancer-specific markers amid massive mass spectral data. Proc. Natl. Acad. Sci. **100**(25), 14666–14671 (2003)
43. Zortea, M., Skrovseth, S., Godtliebsen, F.: Automatic learning of spatial patterns for diagnosis of skin lesions. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp. 5601–5604 (2010)

Recovering Skin Reflectance and Geometry for Diagnosis of Melanoma

Jiuai Sun, Zhao Liu, Yi Ding and Melvyn Smith

Abstract In order to achieve an early detection of skin cancers, various state-of-the-art imaging modalities have been investigated from optical, impedance, biomechanical and physiological perspectives to find out the potential biomarkers. However multilayered skin microstructure and a wide spectrum of dynamical chromophores embedded underneath skin make it very challenging to quantify this mostly accessible, but very complex and heterogeneous largest organ of the human body. Rather than concentrating on characterizing those internal features in a microscopic level, both lesion reflectance (colour) and 3D geometry have been suggested to recover through a relative easy and cost effective way towards an improved diagnosis of melanoma. The reflectance recovered can be used as a good replacement for conventional photograph for the measurement of the ABCD criteria, while the geometry of lesion surface provides extra dimension for characterizing the topography disruption of lesion region. As both the reflectance and geometry of skin surface generally reflect the growth of chromophore cells under the skin, any external abnormalities indicating the change of skin conditions must accompany with some irregular evolvement and change with these cells. For example, a blurred and asymmetrical border possibly reveals an abnormal growth of melanocytes in the horizontal phase; whiles 3D surface indentations and protrusions accompanying variegated pigmentation may indicate an aggressive penetration of melanin into the dermal layer. We compared new features derived from reflectance and geometrical information with those traditional ones and demonstrated their significance as additional clues for melanoma diagnosis.

Keywords Melanoma · Skin surface · Reflectance · Profile · Chromophore index · Skin analyzer · Topographic disruption · Photometric stereo · Segmentation

J. Sun(✉) · Z. Liu · Y. Ding · M. Smith
Machine Vision Lab, University of the West of England, BristolBS16 1QY, UK
e-mail: jiuai2.sun@uwe.ac.uk

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_9,

Introduction

As people tend to expose themselves under the sun for longer periods of time, excessive ultraviolet rays may penetrate deeply into the skin to cause sunburn and other forms of skin disease. Most of these diseases are curable relatively easily; however, malignant melanoma (MM)—the most fatal form of skin cancer—is one important exception. Perhaps even more worrying, the incidence of malignant melanoma is increasing rapidly worldwide, with an annual increase second only to lung cancer in women. Currently Australia/New Zealand have highest incidence rates with almost 1 in 50 populations in the risk of developing melanoma. Apart from high degree of sun exposure, unhealthy lifestyle such as less exercise and high calorie diet may be also blamed for the high prevalence rate. The average lifetime risk of developing malignant melanoma in the UK has also doubled every 10 years, making it the fastest growing type of cancer among all major cancers over the past few decades [1]. Even so, early diagnosis is the most effective way to reduce the mortality rate of melanoma, unfortunately the successful diagnosis of this cancer at its early stage can be practically difficult due to limitations of the naked human eye in identifying subtle suspicious details and potential changes from the appearance of skin lesions. In order to improve the early detection of melanoma, various imaging techniques have been investigated to extend clinician's capability to pick up the fine structures of suspicious lesions for differentiating MM from benign lesions (BLs).

Dermoscopy is a valuable handy tool being widely used in dermatology to reveal the penetration and distribution of pigments for predicting the malignancy of skin tumours. In order to enhance the observation of morphological structures and pigmentation patterns, the polarized light or immersion oils are often employed to remove the specular reflection from the skin surface and render the epidermis translucent. This enhanced visibility gives experienced examiners more confidence and improves their diagnostic accuracy of melanoma, but inexperienced or less trained clinicians may find it difficult to identify and explain those complex morphological structures such as pigmented network, dots and globules within the lesion [2]. Another similar approach, spectrophotometric intracutaneous analysis (SIAScopy), uses several lights with different spectrum to characterize the chromophore and structure of skin subsurface. The penetration of melanin within the dermis is estimated through the Kubelka–Munk model by concerning both the optical scattering and absorption mechanisms. This additional capability of detecting the presence of melanin below the dermo-epidermal junction gives the SIAScopy potentiality to pick up the metastatic invasion signal of skin cancers [3]. Other approaches being investigated to enhance the melanoma diagnosis may include ultrasound imaging modality working either independently or with the optical approach to quantify general thickness and volume of skin cancers [4], optical coherence tomography and Terahertz pulsed imaging modalities to differentiate between normal and cancerous tissues [5].

Melanoma is a dark-pigmented, widely metastasizing tumour arising from melanocytes, which give our skin different colours through the melanogenesis process. Apart from melanin produced by the melanocytes human skin also con-

tains other forms of chromophores such as bilirubin, beta-carotene and haemoglobin whose optical absorption characteristics all present very different from each other. For those microstructures with size comparable to the wavelength of incident light, the light interaction will become dominated by the scattering mechanism. All these chromophores associated with the microstructure of skin sub-layers make up of a complex, dynamic, variegated and multilayered optical turbid medium. It is approved very challenging to identify rigorous formulas to accurately model all this nature complexity [6]. Fortunately any development and evolvement of these chromophores and microstructures will finally reflect on its external appearances. For example, melanoma may pathologically present with variegated colour or distorted skin line or net patterns [7]. This allows the diagnosis of melanoma from the observation of the skin external appearances directly. As a good example, the ABCD rules are two dimensional geometry and colour visual descriptors summarized by clinicians and being most frequently used for melanoma diagnosis in clinics.

Direct visual observations of skin lesion or its photograph have been used for diagnosis of melanoma for decades; however they tend to suffer from the problems of low accuracy and subjective judgment based on personal experience. Meanwhile observation under one specific set of conditions or one photograph normally has difficulty to represent the full characteristics of skin surface as the appearance of skin surface is jointly determined by several internal and external factors such as geometry and reflectance of the skin surface, the ambient lighting and the spectral sensitivity of cameras. A simple example in Fig. 1 shows two different impressions of a skin replica illuminated by a light from two different directions. Fig. 1a reveals more detail in a horizontal primary direction, whereas the details in both horizontal and vertical secondary direction can be found in Fig. 1b. To improve the diagnosis of skin diseases, an objective and comprehensive representation of skin surface must be obtained at first hand.

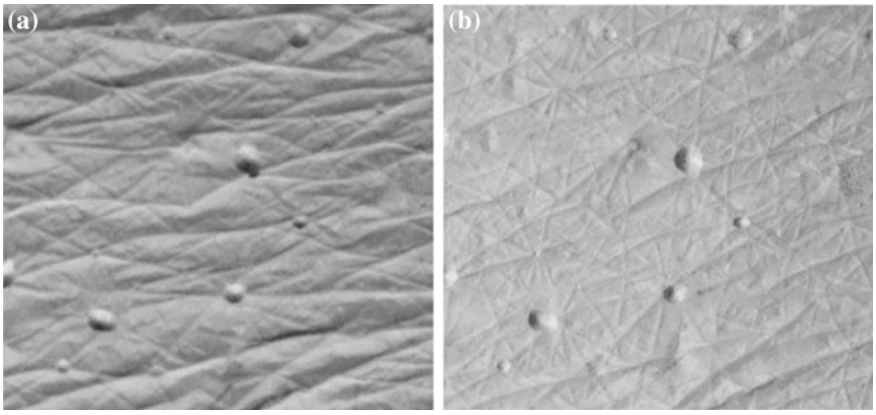


Fig. 1 One skin replica illuminated from *top* (**a**) and *right hand side* (**b**)

At a micro scale the skin surface is not simply flat but able to be characterized by a particular micro-relief representing the keratinisation progress of epidermis and 3D organization of dermis and subcutaneous tissue [8]. The periodical cellular updating and frequent mechanical interaction with external environment also give skin the macro level net-like topographic structures in regular geometrical forms like triangles, polygons and line patterns. The roughness and palpation parameters of these geometrical structures can be used to estimate an overall size and depth of a cutaneous tumour [9]. On the other hand, skin colour/reflectance, another important element to give skin general appearance, originates from the chromophores including melanin, one main and particular chromophore directly indicating the invasive depth of the melanoma. It usually gives the melanoma bluish colour as the pigments penetrating in the deep dermal layer absorb a large amount of light with long wavelengths [10].

Therefore both the geometric and colour/reflectance information of skin lesion are potentially important for differentiating the MM from BLs. A working device called skin analyzer (SA) has been purposely developed with the capability to recover the orientation and reflectance of skin surface simultaneously by using multiple images captured from same viewpoint but illuminated from different directions [11]. This provides an effective way to simulate the procedure adopted by dermatologists in making a clinical diagnosis, i.e. look those suspicious lesions from different angles, pick up suspicious physical features, and then possibly feed the features into a specialized intelligent classification system (doctor's brain) to conclude which class the lesion may belong to. Such an imaging system is expected to help objectively detect and quantify features which are not easily picked up by naked human eyes.

Characterizing Skin Reflectance and Geometry

When a beam of light projects onto the surface of skin, it will undergo hundreds of absorption and scattering before emission. This process makes the radiation distribution isotropic in all directions. The micro facet structure of the stratum corneum further contributes diffusive impression of the skin appearance [12]. Therefore a diffuse reflection model is used here to approximate the mechanism of the photons interacting with the chromophores and layered structure of skin. Based on this reasonably simplified skin optical reflection model the working SA device is able to recover both orientation and reflectance of skin surface simultaneously from the theory of photometric stereo.

Colour Photometric Stereo

The image intensity I of a Lambertian surface can be simply expressed as the product of the composite reflectance K and the cosine of the incidence angle θ, i.e. the angle between the direction of the incident light $\boldsymbol{l}$ and the surface normal $\boldsymbol{n}$:

$$I = K cos(\theta_i) = K(\boldsymbol{l} \cdot \boldsymbol{n}) \tag{1}$$

Without any prior geometry knowledge of an object's surface, more than one image is required to recover the surface normal and reflectance. In general several images are taken using the same imaging system from the same observation point, but for each the object is illuminated by a single light source from a different direction $\boldsymbol{l}^j$ (j = 1, 2,…, N, N $\geq$ 3). A six evenly distributed light source photometric stereo method has been proposed to ensure that all visible points on a convex surface can be illuminated by at least three light sources, the minimum required for traditional photometric stereo [13]. Shadows and highlight can be eliminated by directly using the intensity values through a hierarchical light-source selection stratagem. The redundant information from the images illuminated through the use of the extra light sources can improve the accuracy of the recovery as well as overcoming the effects of highlights and shadows. With the wide accessibility of colour photographs which provide three values at three channels, an improved solution can be achieved by averaging the results on three channels after implementing the traditional photometric stereo approach separately [14].

Suppose I_c^j (c = r, g, b) represents the intensity values of each pixel in three colour channels respectively within the jth image, 6×3 intensity values can form three groups of intensity vectors $\boldsymbol{I}_c = (I_c^1, I_c^2, \ldots, I_c^6)^T$. In the same way, an illumination matrix can be obtained through arranging the lighting directions as: $\boldsymbol{L} = (\boldsymbol{l}^1, \boldsymbol{l}^2, \ldots, \boldsymbol{l}^6)^T$. So total 3×6 equations from Eq. (1), corresponding to the same pixel in six images, can be written as the following linear system of equations:

$$\boldsymbol{I}_c = K_c(\boldsymbol{L} \cdot \boldsymbol{n}) \tag{2}$$

When the matrix $\boldsymbol{L}$ is know through calibration, and is of at least rank 3, three groups of albedo and surface normal can be uniquely calculated using a linear least-squares method:

$$K_c = |(\boldsymbol{L}^T\boldsymbol{L})^{-1} \cdot \boldsymbol{L}^T \cdot \boldsymbol{I}^c| \tag{3}$$

$$n = (\boldsymbol{L}^T\boldsymbol{L})^{-1} \cdot \boldsymbol{L}^T \cdot \boldsymbol{I}^c / K_c \tag{4}$$

After discarding those problematic pixels, the above procedure is carried out for each of the three channels of the captured colour images respectively. Theoretically the surface normal calculated at three channels should be same. However some slight variation always exists in practice as there is usually noise within the acquired images and uncertainty in the estimation of lighting location. To minimize these errors, the mean of the three surface normal values is calculated as a standard optimal value. After the surface normal is determined, the composite albedo can be recalculated from Eq. (3) and then converted into RGB colour space to form the colour reflectance image.

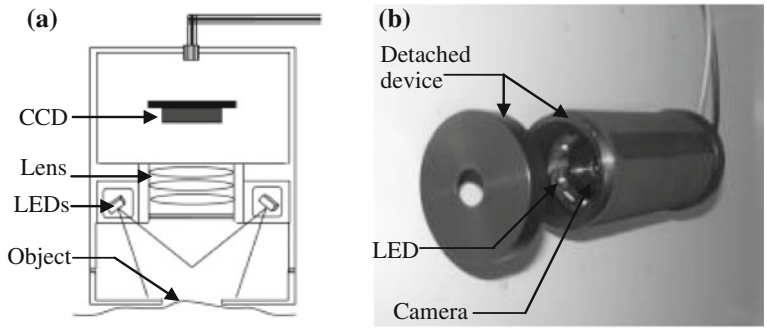


Fig. 2 Schematic (**a**) and developed handheld skin analyzer (**b**)

Different from the available techniques which only measure either skin reflectance or skin topography [15, 16], this simplified reflection model allows both the reflectance and shape information recovered simultaneously and then provided as an improved input representation for differentiating suspicious lesions.

Evaluation of the Skin Analyzer

The SA is an ergonomically and aesthetically pleasing hand-held device designed and manufactured for undertaking pilot clinical studies. Figure 2 shows the schematic and prototype of the device which is finished in a black colour with a compact digital CCD camera and six surface mounted LEDs whose relative positions are optimized to achieve better recovery accuracy [13]. It can be easily operated through customized software interfacing with a PC or laptop. Due to the light-weight design, it can be held in a single hand without touching the skin being measured. This ensures the measurements have good repeatability and stability because any external distortion on the skin due to the contact is largely excluded from the skin surface during this non-contact imaging procedure.

Eight images are totally collected within one second for each clinical case. The first image, taken without any lights on is considered as the image under illumination from ambient light only. Following the first image acquisition, six images are acquired with each spot light in a different position sequentially activated. The eighth image is acquired with three equally distributed lights switched on at the same time to simulate a diffuse lighting condition. The effect of ambient light is eliminated by subtracting the first image from each of the later seven images. The device is evaluated through its performance in 3D reconstruction and rendering synthetic images for diagnosis purposes. The experiments are carried out on both skin replica and in vivo skin.

3D Reconstruction

A ground truth of skin surface is required in order to evaluate the performance of 3D recovery of the SA; however it is normally difficult to obtain a good ground truth of skin surface because of its elastic, dynamic and multilayered microstructure. A negative skin replica collected by smearing a silicon rubber material mixed with a catalyst is used. This provides a simple and repeatable means to record the external detailed structures of the skin without affecting the skin function. Fine detail in the form of the furrows and peaks of the skin relief can be reproduced exactly proving this process is carefully undertaken. After the skin replica has been produced, the topography of skin can be obtained by scanning the surface of the replica through either a contacting mechanical stylus or noncontact optical systems. These measurement techniques have proved useful for recording skin topography through static silicone replicas with satisfactory accuracy in analyzing skin microstructure and anisotropy. Here a commercial optical scanning device, phase shifting rapid in-vivo measurement of skin device (PRIMOS) is used to obtain the profile of skin replica as a standard reference for comparing with the data acquired by the skin analyzer [15]. The PRIMOS recovers a height map of skin surface from the images illuminated by a series of phase shifting stripe lights onto the skin or skin replica.

To make the surface normal extracted from the photometric stereo technique comparable to that output from the PRIMOS device, the 3D data from the PRIMOS is transformed into a gradient representation format through partial differentiation, which will not introduce any global error due to the local calculating operation. On the other hand, the gradient data obtained from photometric stereo approach is also integrated into a format of skin profile which can be compared with the PRIMOS height map directly, though the integration and comparison are only applied on a relative small area to minimize the effect from the accumulated errors with an integration procedure.

Figure 3 shows skin replica taken from normal skin on the back of a hand and the 3D topography representation in the form of needle maps and 3D reconstruction results. The rectangle area in Fig. 3a formed by the black solid lines specifies the interested area where the integration of surface normal and comparison are carried out.

Figure 3b and c are needle maps extracted from the 3D data acquired by the PRIMOS and derived from the surface normal data captured by the SA. The SA demonstrates a higher sensitivity as more vectors are presented with long amplitudes. Meanwhile the reconstructed skin profile from the skin analyzer changes more smoothly, while sharp edges can be found with the results from the PRIMOS. This can be observed from the reconstructed results of three identified small holes in Fig. 3d and e. This failure of the PRIMOS in the recovery of sharp edges may be explained by the obstruction of light projected from only one direction. The skin analyzer is able to take advantage of the multiple light sources to remove the presence of shadow and specularities before recovery. This makes those edges and holes reconstructed from photometric stereo look more reliable and credible.

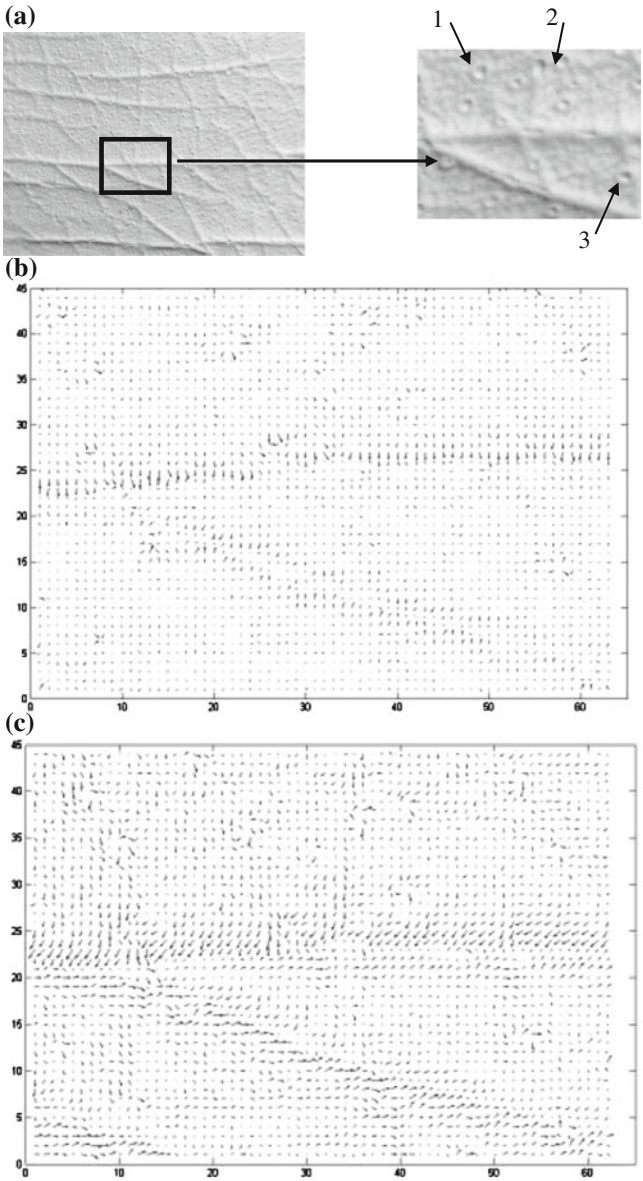


Fig. 3 Evaluation of geometry recovery on skin replica (**a**) through comparing the surface normal (**b**) and (**c**), 3D reconstruction (**d**) and (**e**) obtained from PRIMOS and SA

Skin replicas have been used as an acceptable means to investigate the topography of the skin. However it requires considerable operator skill to successfully copy the skin structure. Most importantly it cannot replicate skin colour information which is

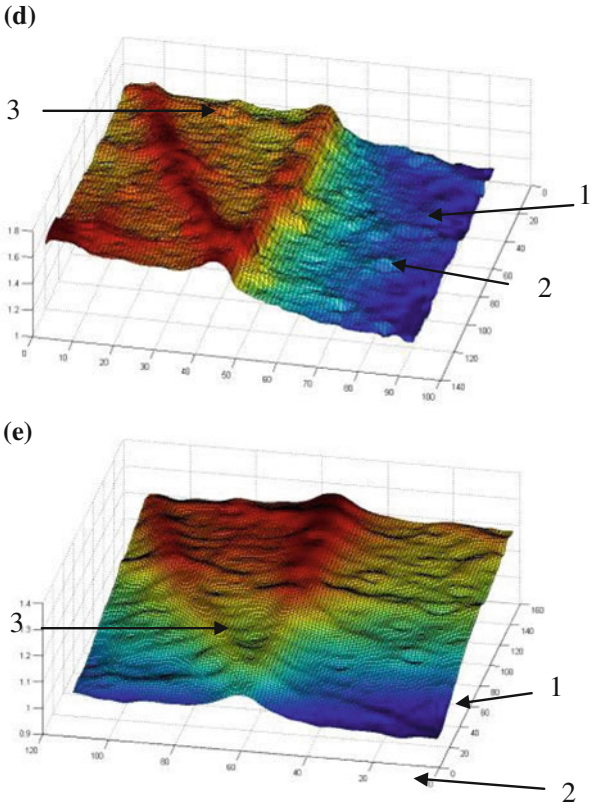


Fig. 3 continued

sometimes more important than skin microstructure in clinics. Meanwhile the serial of parallel black/white light stripes projected by the PRIMOS onto the skin surface may be concealed by the heavy pigmentation of skin lesions. It may deteriorate the accuracy and resolution of the reconstruction results [17]. The bulky weight also prevents the PRIMOS in vivo measurement of skin lesions distributing arbitrarily around the body. On the contrary the SA provides a new flexible mean to record the appearance of skin surface for either clinical or cosmetic purposes.

Synthetic Colour Images

Figure 4 shows four types of images used to assess the colour representation achieved by the SA. They are traditional photography captured with only one illumination (a), diffuse photograph illuminated by a ring-type of flash (b), reflectance reconstructed from multiple images with the elimination of specular and shadow effects (c), and

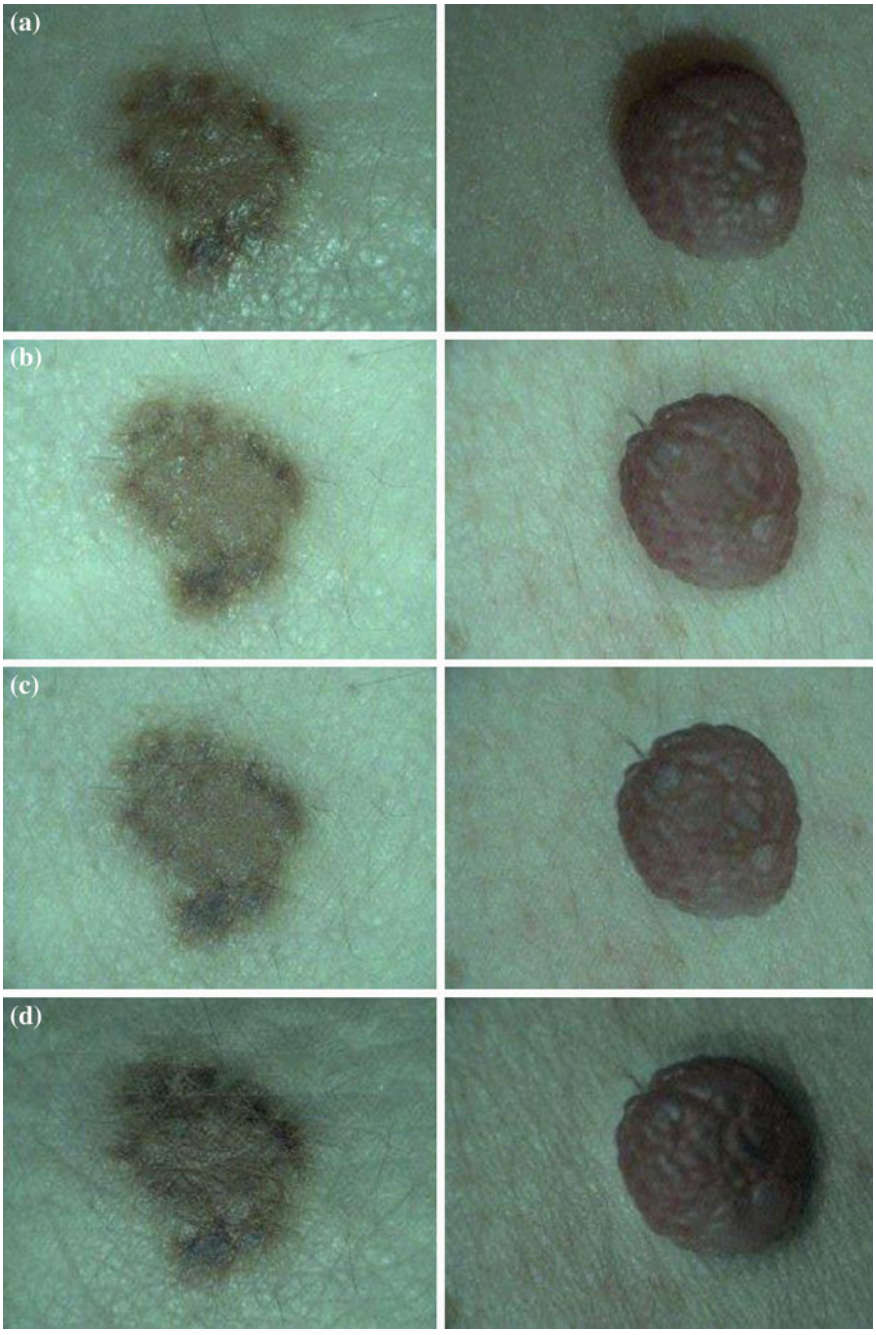


Fig. 4 Four type of images of two sample lesions (**a**) illuminated from one direction, (**b**) illuminated from all directions, (**c**) recovered with surface reflectance, and (**d**) rendered into virtual photographs

images rendered by using a virtual light source (d). The eighth image illuminated by three equally distributed lights is used to simulate the idea diffuse or ring-type flash photograph, one of the popular manners to reliably record visual information of the skin lesions [18].

The photographs illuminated by one light in Fig. 4a present heavy highlight on the left of the lesion and shadows on the right side of the lesion. The diffuse (multiple) lights can resolve such problems to some extent in Fig. 4b, although illumination by diffuse light often does not produce an image with perceived depth and structure details as may be seen by the human naked eye. The fine scaling of valleys and bumps can hardly be discerned due to the effects of the diffuse illumination. It may make this method unacceptable due to the reduced fidelity and 'unreal' nature of the image especially when the topographic information is interested.

The PS approach can average the specular or shadow effects to obtain an approximation of the surface reflectance in Fig. 4c. Similar to diffuse photograph, the three-dimensional topographic details have been eliminated in reflectance image too. However the disappearance of skin structure texture in Fig. 4b is caused by the illumination from multiple directions, whereas the reflectance is the reflective properties of the skin and independent from the topography of the skin. When the reflectance is combined with the recovered surface orientation information, the skin surface can be fully characterized. Figure 4d shows a synthetic skin image, illuminated by a virtual light source. It is noticeable that both the colour and profile information are rendered effectively.

Reflectance Mapping for Segmentation of Lesions

The reflectance image reflecting the type, quantity and distribution of the chromophores under the skin is largely independent from the position and intensity of the illumination. Such objective information potentially offers significant benefit for outlining the lesion area and then differentiating between malignant and benign tumors. Here we examine its applications for the first step of automated diagnosis, i.e. the segmentation of skin lesion. Due to the lack of a reference method which could provide the exact boundary of lesion area, experienced dermatologists are invited to outline the border of the lesion on the conventional photographs and indicate using black lines. An established automated segmentation algorithm is used to separate the lesion from normal skin for both photographs in red line and reflectance images in green line [19].

As the automated segmentation algorithm works on the intensity of the images, the results of border tracking are sensitive to changes in the intensity values. Figure 5a shows that the boundary of the lesion tends to be larger than those in Fig. 5b and c. This is caused by shadows occurring during the photographing of the bumped and raised lesion, even though multiple light sources have been used.

To compare quantitatively the automatic segmentation results obtained from conventional photographs with those obtained from reflectance images, a performance

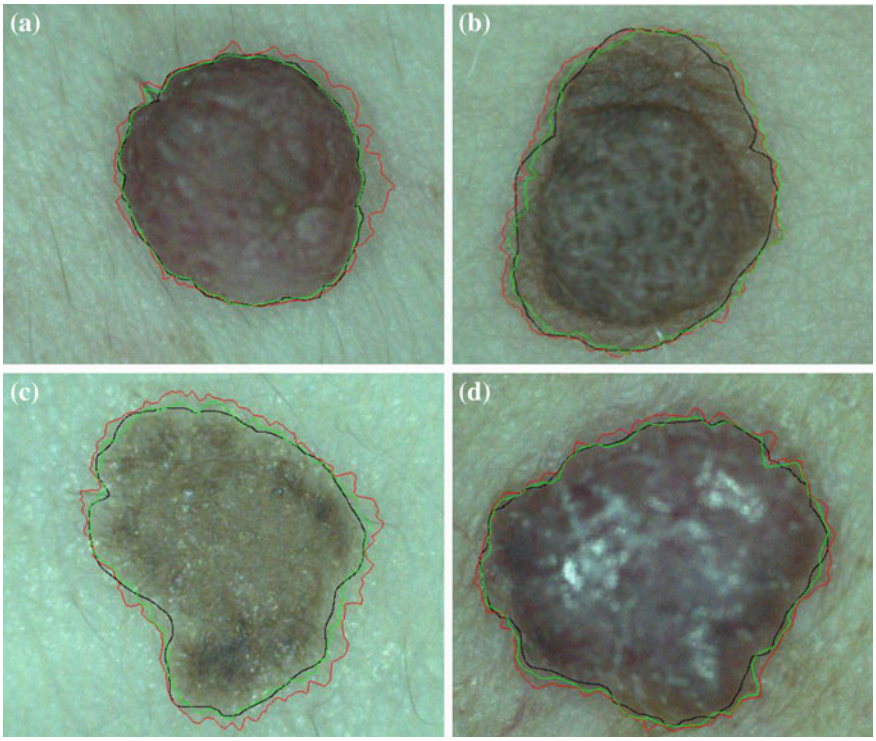


Fig. 5 Four pigmented skin lesions **a** intradermal naevus, **b** intradermal naevus, **c** basal cell carcinoma and **d** dermatofibroma, superimposed with manually segmentation result (*black line*), automatic segmentation result from the diffuse photograph (*red line*), and automatic segmentation result from the photometrically derived reflectance image (*green line*)

indicator was calculated by subtracting the area enclosed within the manually derived outline from that enclosed within the automatically derived outline, and then expressing this difference as a percentage of the area within the manually derived outline. Figure 6 provides the results in this form for all four lesions, as well as for a series of inverted binary difference images in which white areas indicate regions that lie within the automatically derived boundary but not within the corresponding manually derived boundary (the top row of results is from the reflectance images and the bottom row is from the conventional photographs).

It is readily apparent that the segmentation from the reflectance images and the photographs are both deviated from the results outlined manually by the dermatologists. However the segmentation results based on the reflectance images approach the ground truth closer than those based on the diffuse photographs. The results for sample 1 and 2 make it clear that shadows and specularities can heavily affect the segmented results when diffuse photographs are used as sources to be segmented. However, such errors are effectively eliminated when reflectance images are employed.

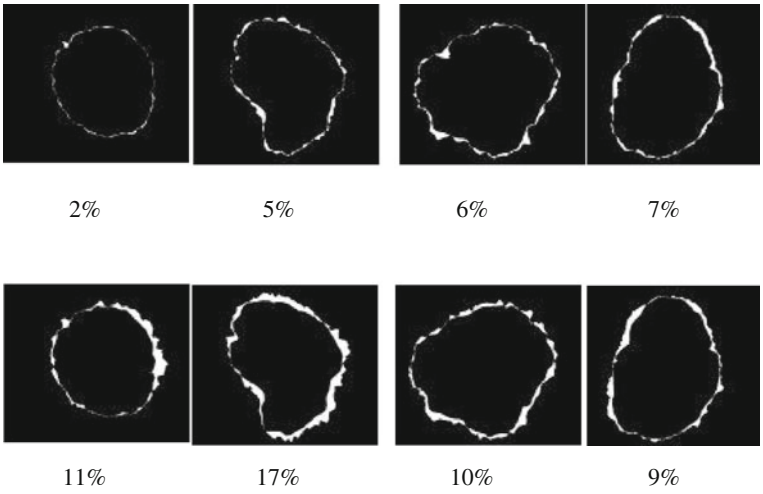


Fig. 6 Comparison of the performance of automated segmentation applied to photometrically derived colour reflectance images (*top row*), with that applied to conventional photographs taken with diffuse lighting condition (*bottom row*)

Therefore the employment of the surface reflectance photograph extracted from photometric stereo, rather than a conventional photograph, can obtain an improved segmentation result. It will then be able to provide a more objective and reliable extraction of the ABCD features as important input for subsequent automatic diagnosis of melanoma.

Diagnosis Enhancement Through Chromophore Indexes

Measurement of lesion reflectance provides a more objective way to characterize skin lesions as skin reflectance reflects the concentration and distribution of those major chromophores like melanin and haemoglobin, important diagnosis evidence linking to the invasiveness of cancers. These biological indexes can be extracted from the reflectance image as alternatives to describe the colour variegation of lesion for melanoma diagnosis.

Melanin Index and Erythema Index

When only absorption are concerned, human skin can be simplified as two uniform optical media layers over an ideal diffusive surface. The first and second layers represent epidermis and dermis, and the lowest surface is equivalent to the perfect diffusive

fat layer with no chromophores. The direct reflection on the epidermis surface and the interface between epidermis and dermis are negligible and the absorption is one main manner of light transportation within the first two layers [20]. Melanin, the major pigmented chromophore, effectively absorbs light of all wavelengths in the spectral range from 300 to 1000 nm, but the strongest absorption occurs at shorter wavelengths, especially near UV radiations [21].

Within long wavelength light spectrum the absorption caused by the haemoglobin is negligible so that light attenuation is dominated by the absorption of melanin, which can be considered as a linear function of the melanin content providing the absorption of bottom layer is constant. Therefore the melanin index can be approximated by the diffuse reflectance at a long wavelength,

$$\mathrm{MI} = 100 * log_{10} \frac{1}{R_r} \tag{5}$$

where R_r represents the reflectance (or albedo) at the red wavelength recovered from the above colour photometric stereo approach.

Similarly when melanin content is low, the difference of the absorption at the green and red wavelength can be approximated as a linear function of haemoglobin content. When melanin pigmentation is not negligible, a correction coefficient can be adapted to weight the contribution from melanin component. Therefore the erythema index, a parameter to reflect the blood component within the skin, can be expressed as:

$$EI = 100 * log_{10} \left(\frac{1}{R_g} - \xi \cdot \frac{1}{R_r} \right) \tag{6}$$

where is the reflectance (or albedo) at green wavelength, and ξ is the correction coefficient.

This calculation works with the reflectance image recovered from the PS approach. It may also apply to any other colour images providing an ideal white balanced reference point can be found within the images. Figure 7 demonstrates the calculated MI and EI images of one lesion sample captured from the SA. Clearly the contribution

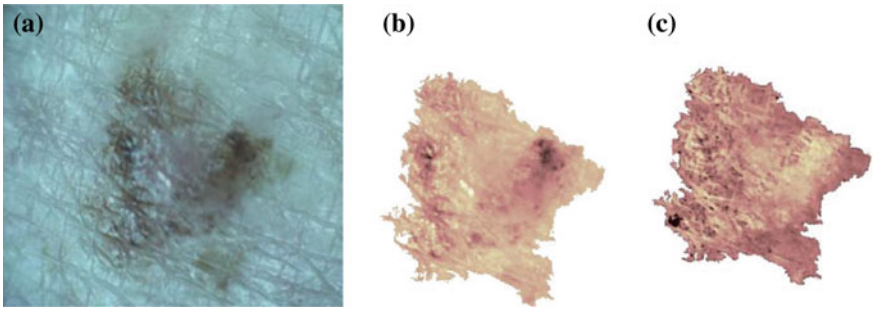


Fig. 7 Reflectance image (**a**) and extracted MI (**b**) and EI (**c**) images

of different pigmentation becomes more obvious by working with the MI and EI mappings separately rather than the reflectance image directly.

Colour Variegation of Chromophore Indexes

As the content of cutaneous melanin and haemoglobin are important clues characterizing the pigmentation of skin surface, the direct employment of the MI and EI should give a better description on colour variegation for early diagnosis of melanoma. Similar to conventional representation of colour variegation in cutaneous lesion images, the variegation of chromophore indexes is also expressed through the difference between the lesion and its own mean value or mean value of the surrounding normal skin.

Suppose there are a set of pixels $\mathrm{P_i}$ in the lesion areas Π_{lesion} with $\mathrm{P_i} \in \Pi_{\mathrm{lesion}}$. By quantizing the colour range (intensity value) into $\mathbf{N}(16)$ bins, the colour variegation CV can be defined as,

$$CV = \sqrt{\sum_{\mathrm{i}=0}^{15} (\mathrm{H_i} - \bar{\mathrm{H}}_\mathrm{i})^2 \mathrm{PD}(\mathrm{H_i})} \tag{7}$$

where H_i describes the quantised absolute colour value in an image (MI, EI, reflectance and colour image intensity), $\bar{\mathrm{H}}_i$ is the mean value within the overall lesion, i is the index of the quantised bin, and $PD(\mathrm{H}_i) = \frac{\Pi_{\mathrm{H}_i}}{\Pi_{lesion}}$ is the probability density function, with Π_{H_i} standing for the number of pixels in a quantized bin.

A four-dimensional feature vector is generated by the chromophore indexes of lesion area and all normal skin area surrounding the lesion, while a six-dimensional feature vector is constructed using the conventional colour spaces of RGB, as illustrated as follows:

$$\mathrm{CV_{chrom}} = \{\mathrm{CV}^{\mathrm{MI}}_{\mathrm{les}}, \mathrm{CV}^{\mathrm{EI}}_{\mathrm{les}}, \mathrm{CV}^{\mathrm{MI}}_{\mathrm{sur}}, \mathrm{CV}^{\mathrm{EI}}_{\mathrm{sur}}\} \tag{8}$$

$$\mathrm{CV_{RGB}} = \{\mathrm{CV}^{\mathrm{R}}_{\mathrm{les}}, \mathrm{CV}^{\mathrm{G}}_{\mathrm{les}}, \mathrm{CV}^{\mathrm{B}}_{\mathrm{les}}, \mathrm{CV}^{\mathrm{R}}_{\mathrm{sur}}, \mathrm{CV}^{\mathrm{G}}_{\mathrm{sur}}, \mathrm{CV}^{\mathrm{B}}_{\mathrm{sur}}\} \tag{9}$$

Chromophore Indexes for Melanoma Diagnosis

The effect of the colour variegation extracted from chromophore indexes (as well as the topography disruption to be addressed in the next session) is evaluated through the following clinical database: 253 cutaneous lesions including 69 MMs and 184 non-melanoma cutaneous lesions acquired by the Skin Analyser device. A linear support vector machine (SVM) is applied as the classifier, and a tenfold cross validation is employed as the training-testing strategy for evaluating the classification performance. Sensitivity (SE), specificity (SP) and overall accuracy (Acc.) are used

Table 1 Classification results using colour variegation descriptors derived from RGB colour space and the chromophore indexes

Pigmentation	RGB	MI+EI
SE (%)	67.85	76.81
SP (%)	76.27	80.97
Acc. (%)	73.97	79.84
AUC	0.71	0.80

as the quantitative measurements of correct discriminations. Specifically, sensitivity represents the proportion of melanomas correctly recognized as malignant lesions, and specificity denotes the proportion of non-melanoma naevi correctly classified as non-melanoma lesions. Accuracy is the value showing the correct classification within the whole database. The receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC) values are calculated with a confidence interval of 95 %.

Colour variegation as the only diagnostic descriptor is investigated in this section to demonstrate the effectiveness of the chromophore indexes for melanoma diagnosis and compare the classification results with those derived from reflectance images. To demonstrate the usefulness of the major chromophores for melanoma identification, colour variegation descriptors, computed from RGB colour space, and the major chromophore indices as shown in Eqs. (8) and (9), are respectively applied as the feature vectors, to automatically classify melanoma and non-melanoma lesions.

The classification results in Table 1 shows the RGB colour gives the lowest diagnostic accuracy, with 67.85 % SE and 76.27 % SP for the reflectance images obtained by the SA. The highest classification is achieved by the chromophore indices, which boost the sensitivity by 9 % for the reflectance data. As expected these results demonstrate that the melanin index and erythema index can characterize the major chromophores underneath the skin better than the colorimetric representations based on human vision system for melanoma diagnosis.

Obtaining 3D Malignancy Indicators

Early works have found that the loss of skin topographic markings may imply a late sign of the tumour invasion into the deep dermis [22–24]. The statistical parameters of skin topography such as co-occurrence matrices, Fourier power spectrum and fractal features obtained through the 3D laser scanning measurement on the negative skin replicas also show distinctive difference between the surface of benign lesions and that of malignant lesions [25]. The skin line patterns have also been developed to represent the local skin line direction (SLD) and the variation of skin primary line direction (SLV) between lesion area and surrounding normal skin area [26]. The skin line patterns have the advantages of being able to be extracted from only one white-light photograph without requiring any special 3D scanning devices, though they are

inherently sensitive to the selection of parameters for extracting high frequency of skin line patterns.

We investigate the potentiality of 3D lesion representation extracted from the PS approach for identifying the abnormality of skin lesions. Specifically the surface orientation rather than an integrated 3D height or depth maps are used directly for this purpose as the process of integrating surface normal into a height map is prone to accumulated errors.

Adaptive Modeling of Skin Disruption

Relative to the observation center the orientation of a skin patch can be expressed by tilt angle ϕ and slant angle θ which are able to be calculated from surface normal and therefore they are equivalent to each other. Figure 8 shows three different representation of a sample patch isolated from an image of a benign lesion. The regular skin line patterns are not immediately visible on the intensity image in Fig. 8b as some 3D information projected onto a 2D optical sensor may be concealed during photographing. While those fine topographic details become more visible with the

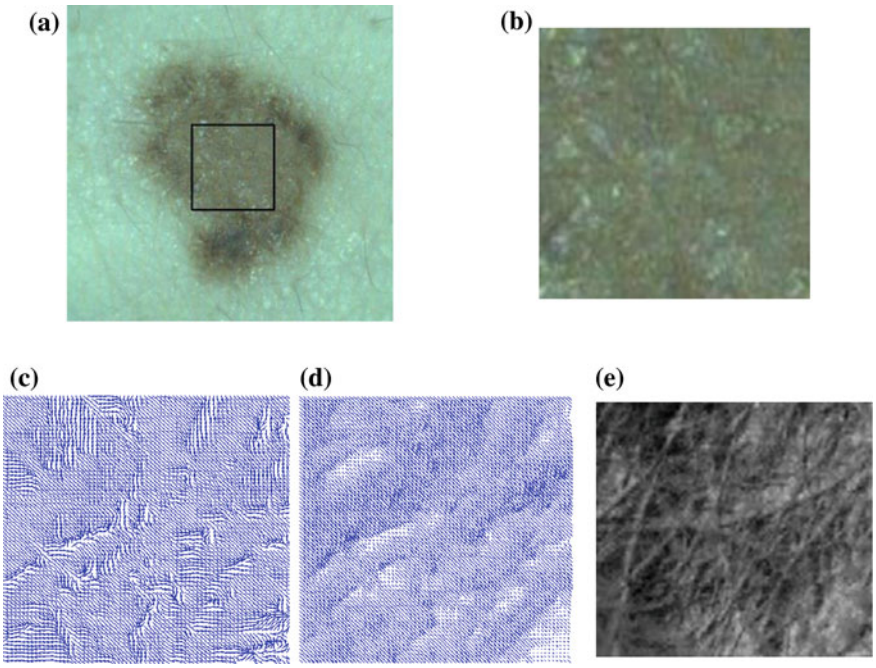


Fig. 8 A colour image of a skin lesion (**a**), zoom-in view of the sample patch (**b**) isolated by the rectangle, skin tilt pattern (**c**) and slant pattern (**d**) of the isolated region expressed as needle maps, and a rendered grayscale image (**e**)

tilt/slant representation. This shows the advantages of a 3D representation of skin surface over the conventional 2D image format. This enhanced power to visualize subtle disruptions on skin surface can help to identify those features less obvious to the naked eyes [27].

The skin tilt and slant representations characterize the topographic variations in x-y (horizontal) direction and in z (vertical) direction respectively. In order to quantify the topographic disruptions of skin lesion, two reference skin models are developed to simulate the skin tilt/slant patterns of the ideal normal skin lesions closest to the acquired skin tilt (slant) patterns captured by the Skin Analyser. The deviation between the acquired patterns and those ideal reference skin patterns are be used to indicate the malignancy of the lesions [28].

Skin Tilt Pattern

Figure 9 shows a local dominant skin tilt distribution of a malignant melanoma and a benign seborrheic keratosis. The benign lesion presents more regularly distribution with respect to its centre across all directions than that for the MM. Therefore a 2D isotropic function uniformly distributed with respect to its centre is chosen to generate an ideal skin tilt pattern.

Since the distribution centre of lesion's acquired skin tilt patterns could be anywhere within the lesion region, an arbitrary skin tilt pattern can be generated by searching every location (u, v) being the distribution centre. However there should be only one simulated skin tilt pattern best-fits (i.e., having the smallest difference to) lesion's acquired skin tilt patterns. The best-fit skin tilt pattern is estimated as the one having the minimum sum of difference to the acquired lesion's skin tilt patterns. This can be expressed by the following equation.

$$(u_c^*, v_c^*) = argmin \sum\nolimits_{\{(u,v),(u_c^*,v_c^*)\}\in\Pi_{\text{lesion}}} ||\phi(u, v) - \phi^*(u, v)|| \tag{10}$$

where (u_c^*, v_c^*) is the estimated centre of the skin tilt pattern model, $||\cdot||$ denotes the Euclidean distance, ϕ are the acquired skin tilt patterns, ϕ^* is the estimated skin tilt pattern. The star sign "*" indicates an estimated variable.

After all skin tilt models are searched across the whole lesion region, an estimated distribution centre and a skin tilt pattern is guaranteed to correspond to a global minimum. Having found the best-fit skin tilt pattern model, the resultant disruptions in skin tilt pattern can be described by the following two measures.

(A) Overall disruption in skin Tilt pattern (OT): it represents an average of the sum of differences between the best-fit skin tilt pattern and the acquired skin tilt patterns over the lesion region:

$$\bar{\phi}_{\Delta}(u, v) = \frac{\sum_{\{(u,v)\}\in\Pi_{\text{lesion}}} ||\phi(u, v) - \phi^*_{min}(u, v)||}{\Pi_{\text{lesion}}} \tag{11}$$

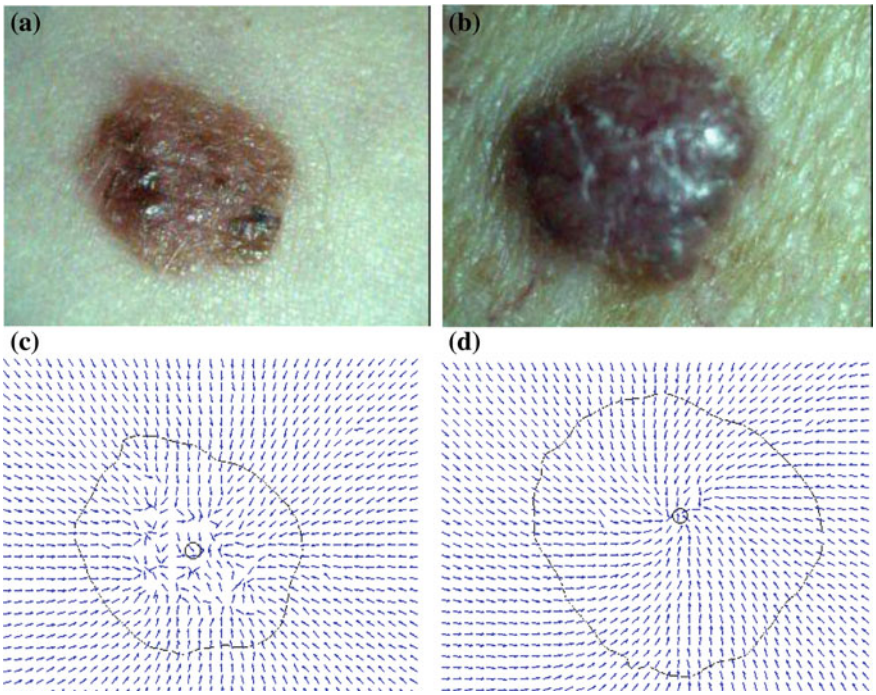


Fig. 9 Conventional photographs (**a**), (**b**) of a malignant melanoma and a benign seborrheic keratosis and their skin tilt pattern plots (**c**), (**d**)

(B) Most disruption in skin Tilt pattern (MT): a square region of size $T \times T\,(5 \times 5)$ centred at (u, v) is firstly chosen to estimate local disruptions in skin tilt pattern and reduce the noise effects from individual pixels. Its local skin tilt pattern disruptions are then defined as the average of the sum of differences between the best-fit skin tilt patterns and the acquired skin tilt patterns.

$$\phi_{local}(u, v) = \frac{\sum_{i=-T/2}^{T/2} \sum_{j=-T/2}^{T/2} ||\phi(u+i, v+j) - \phi^*_{min}(u+i, v+j)||}{\mathrm{T}^2} \quad (12)$$

The MT is finally defined as the maximum local skin tilt pattern disruptions within the whole lesion region:

$$\phi_{MT}(u, v) = max\{\phi_{local}(u, v) | (u, v) \in \Pi_{\text{lesion}}\} \quad (13)$$

Skin Slant Disruptions

Although the exact 3D topography of an acquired lesion is unknown without reconstruction, the skin slant pattern will be able to sense and adapt to the 3D topographic variation on the lesion surface by comparing with the lesion's acquired skin slant patterns. In order to generate the candidate skin slant patterns, a series of 3D Gaussian topographies are generated by varying the amplitude A, the variances in the x-axis σ_x and y-axis σ_y and the rotation angle τ of a 2D Gaussian function. This 3D skin model is adaptive to variable 3D topographies of lesions, i.e. able to match for both flat and non-flat lesions' topographies by adjusting its parameters, but also able to pick up those subtle topographic variations.

Among the generated skin slant pattern models, only the one that best-fits a lesion's acquired skin slant patterns is used to estimate the disruptions in skin slant pattern. The best-fit skin slant pattern θ is estimated as the one having the minimum sum of differences to the acquired skin slant patterns among all the skin slant models θ^* generated by the 2D Gaussian functions. This can be expressed as

$$(A^*, \sigma_x^*, \sigma_y^*, \tau^*) = argmin \sum\nolimits_{(m,n)\in\Pi_{\text{lesion}}} ||\theta(u,v) - \theta^*(u,v)|| \tag{14}$$

A 2D Gaussian function may have the following form,

$$Z^*(u,v) = A^* e^{-(a(u-u_c^*)^2 + 2b(u-u_c^*)(v-v_c^*) + c(u-u_c^*)^2)} \tag{15}$$

where $a = \frac{cos^2\tau^*}{2(\sigma_x^*)^2} + \frac{sin^2\tau^*}{2(\sigma_y^*)^2}$, $b = \frac{sin^2\tau^*}{4(\sigma_x^*)^2} + \frac{sin^2\tau^*}{4(\sigma_y^*)^2}$, and $c\frac{sin^2\tau^*}{2(\sigma_x^*)^2} + \frac{cos^2\tau^*}{2(\sigma_y^*)^2}$, the (u_c^*, v_c^*) is the same distribution centre as that used by the best-fit skin tilt model. The possible values for the variances and the amplitude are chosen empirically to cover a sufficient large number range while the data range for the rotation angles of the 2D Gaussian function is $[0\ \pi)$.

Having found the best-fit skin slant pattern, the disruptions in skin slant pattern at a pixel is defined as the difference between the slant pattern of the best-fit skin slant pattern and the acquired skin slant pattern:

$$\theta_\Delta(u,v) = ||\theta(u,v) - \theta^*(u,v)|| \tag{16}$$

Upon the estimation of the best-fit skin slant pattern model, the resultant disruptions in skin slant pattern can be described by two similar measures.

(A) Overall disruptions in skin Slant pattern (OS): the average of the sum of differences between the best-fit skin slant pattern and the acquired skin slant patterns over the whole lesion region:

$$\bar{\theta}_\Delta(u,v) = \frac{min \sum_{\{(u,v)\}\in\Pi_{\text{lesion}}} ||\theta(u,v) - \theta^*(u,v)||}{\Pi_{\text{lesion}}} \tag{17}$$

Table 2 Classification performances with different combination of skin topographic patterns

Lesion disruption patterns	SE % / SP %
OT + OS	79.3/80.5
MT + MS	75.1/83.3
OT + MT	79.5/78.2
OS + MS	75.3/80.8
SLD + SLV	77.5/63.4

(B) Most disruptions in skin Slant pattern (MS): a local skin slant pattern disruption is firstly defined as the average of the sum of differences between the best-fit skin slant patterns and the acquired skin slant patterns over a window region with the size $T \times T$ centred at (u, v):

$$\theta_{local}(u, v) = \frac{\sum_{i=-T/2}^{T/2} \sum_{j=-T/2}^{T/2} ||\theta(u+i, v+j) - \theta^*_{min}(u+i, v+j)||}{\mathrm{T}^2} \quad (18)$$

The MS is then defined as the maximum local skin slant pattern disruptions within the whole lesion region, i.e.

$$\theta_{MT}(u, v) = max\{\theta_{local}(u, v) | (u, v) \in \Pi_{\text{lesion}}\} \quad (19)$$

Skin Disruption Descriptors for Melanoma Diagnosis

The effectiveness of the proposed skin tilt pattern and skin slant pattern derived from recovered skin topography for melanoma diagnosis is assessed through the linear SVM classifier to differentiate the benign and malignant lesions and compare with the results from the 2D skin line patterns. Table 2 summarizes the classification performances of different combination of the proposed skin surface patterns on the same database as used in last section for evaluating the discrimination power of MI and EI. The two global features in the tilt and slant patterns OT and OS demonstrate complementary to each other and their combination offers a better classification between the MMs and benign lesions than those using only the tilt pattern (OT and MT) or slant pattern (OS and MS) features. Both the line patterns and the proposed 3D skin surface patterns demonstrate comparably good performance in sensitivity, however the skin line patterns shows a poorer performance in specificity. This could be caused by the line pattern (either SLD or SLV) which is extracted from one image, but inherently has difficulty to reveal the topographic variations in vertical direction.

Conclusion and Future Work

This chapter looks into the potentiality of the new recovered reflectance and 3D geometry of lesion surface for diagnosis of melanoma. This unique avenue produces an easy and cheap way to quantify the malignancy of suspicious skin lesions. The features derived from the proposed approach have been approved useful in providing complementary information for differentiating the melanoma from benign lesions. It is also understood that each individual feature has its own weaknesses and strengths for lesion classification. An optimal combination of these newly-found features with those traditional ABCD rules may offer new hope for improving the diagnosis accuracy. The simplicity and cost effectiveness of this approach may also lend itself an applicable mean to monitoring suspicious skin lesions over time especially for those with family history and under environmental risks.

References

1. Garbe, C., Leiter, U.: Melanoma epidemiology and trends. Clin. Dermatol. **27**, 3–9 (2009)
2. Binder, M., Schwarz, M., Winkler, A., Steiner, A., Kaider, A., Wolff, K., Pehamberger, H.: Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch. Dermatol. **31**, 286–291 (1995)
3. Cotton, S.D., Claridge, E.: Developing a predictive model of human skin colouring. Proc. SPIE Med. Imaging **2708**, 814–825 (1996)
4. Jasaitiene, D., Valiukeviciene, S., Linkeviciute, G., Raisutis, R., Jasiuniene, E., Kazys, R.: Principles of high-frequency ultrasonography for investigation of skin pathology. J. Eur. Acad. Dermatol. Venereol. **25**(4), 375–382 (2011)
5. Woodward, R.M., Wallace, V.P., Arnone, D.D., Linfield, E.H., Pepper, M.: Terahertz pulsed imaging of skin cancer in the time and frequency domain. J. Biol. Phys. **29**, 257–259 (2003)
6. Anderson, R.R., Parrish, J.A.: The optics of human skin. J. Invest. Dermatol. **77**, 13–19 (1981)
7. Lee, T.K.: Measuring border irregularity and shape of cutaneous melanocytic lesions. Ph.D. thesis, Simon Fraser University, Vancouver (2001)
8. Leveque, J.L.: EEMCO guidance for the assessment of skin topography. J. Eur. Acad. Dermatol. Venereol. **12**, 103–114 (1999)
9. Egawa, M., Oguri, M.H., Takahashi, M., Miyakawa, M.: The evaluation of skin friction using a frictional feel analyzer. Skin Res. Technol. **8**, 41–51 (2002)
10. Numahara, T.: From the standpoint of dermatology in digital color imaging in biomedicine. In: Tanaka,H., Miyake,Y., Nishibori M., Mukhophadhyay, D. (eds.) Digital Biocolor Society, 67–72 (2001)
11. Smith, L.N., Smith, M.L., Farooq, A.R., Sun, J., Warr, R.: Machine vision 3D skin texture analysis for detection of melanoma. Sens. Rev. J. **31**(2), 111–119 (2011)
12. Krishnaswamy, A., Baranoski, G.V.G.: A study on skin optics. Technical report CS-2004-01, School of Computer Science, University of Waterloo, Canada, (2004)
13. Sun, J., Smith, M., Smith, L., Midha, S., Bamber, J.: Object surface recovery using a multi-light photometric stereo technique. Image Vis. Comput. **25**(7), 1050–1057 (2007)
14. Sun, J., Smith, M., Smith, L., Coutts, L., Dabis, R.: Reflectance of human skin using colour photometric stereo: with particular application to pigmented lesion analysis. Skin Res. Technol. **14**, 173–179 (2008)
15. Jaspers, S., Hopermann, H., Sauermann, G., Hoppe, U., Lunderstadt, R., Ennen, J.: Rapid in vivo measurement of the topography of human skin by active image triangulation using a digital micro-mirror device. Skin Res. Technol. **5**(3), 195–207 (1999)

16. Takiwaki, H., Overgaard, L., Serup, J.: Comparison of narrowband reflectance spectrophotometric and tristimulus colorimetric measurement of skin color—23 anatomical sites evaluated by the DermaSpectrometer and the Chroma Meter CR-200. Skin Pharmacol. **7**, 217–225 (1994)
17. Rosén, B.G., Blunt, L., Thomas, T.R.: On in-vivo skin topography metrology and replication techniques. J. Phys. Conf. Ser. **13**, 325–329 (2005)
18. Ikeda, I., Urushihara, K., Ono, T.: A pitfall in clinical photography: the appearance of skin lesions depends upon the illumination device. Arch. Dermatol.l Res. **294**, 438–443 (2003)
19. Xu, L., Jackowski, M., Goshtasby, A., Yu, C., Dhawand, A., Huntleye, A.: Segmentation of skin cancer images. Image Vis. Comput. **17**(1), 65–74 (1999)
20. Dawson, J.B., Barker, D.J., Ellis, D.J., Grassam, E., Cotterill, J.A., Fisher, G.W., Feather, J.W.: A theoretical and experimental study of light absorption and scattering by in vivo skin. Phys. Med. Biol. **25**, 695–709 (1980)
21. Kollias, N., Bager, A.: Spectroscopic characteristics of human melanin in vivo. J. Invest. Dermatol. **85**, 593–601 (1985)
22. Connemann, B.J., Busche, H., Kreusch, J., Teichert, H.M., Wolff, H.: Quantitative surface topography as a tool in the differential diagnosis between melanoma and naevus. Skin Res. Technol. **1**, 180–186 (1995)
23. Mazzarello, V., Soggiu, D., Masia, D.R., Ena, P., Rubino, C.: Melanoma versus dysplastic naevi: microtopographic skin study with noninvasive method. J. Plast. Reconstr. Aesthetic Surg. **59**, 700–705 (2006)
24. Bondi, E.E., Elder, D.E., Dupont, G., Clark, W.H.: Skin markings in malignant melanoma. J. Am. Med. Assoc. **250**, 503 (1984)
25. Handels, H., Ross, T., Kreusch, J., Wolff, H.H., Poppl, S.J.: Computer-supported diagnosis of melanoma in profilometry. Methods Inf. Med. **38**, 43–49 (1999)
26. Round, A.J., Duller, A.W.G., Fish, P.J.: Lesion classification using skin patterning. Skin Res. Technol. **6**, 183–192 (2000)
27. Ding, Y., Smith, L., Smith, M., Warr, R., Sun, J. Enhancement of skin tilt pattern for lesion classification. IASTED Conference on Visualization, Imaging and Image Processing, pp. 1–6 (2008)
28. Ding, Y., Smith, L., Smith, M., Sun, J., Warr, R.: Obtaining malignant melanoma indicators through statistical analysis of 3D skin surface disruptions. Skin Res. Technol. **15**, 262–270 (2009)

Melanoma Diagnosis with Multiple Decision Trees

Yu Zhou and Zhuoyi Song

Abstract This chapter highlights the application of multiple binary decision trees in melanoma diagnosis. Since the clinical rules in diagnosing melanoma involve inhomogeneous/non-metric data and various 'if-then' statements, direct utilization of machine learning techniques such as neural networks can not perform satisfactorily in modelling the clinical diagnostic knowledge which, typically, is nonlinear and fuzzy. As a versatile and intuitive paradigm in pattern classification, the decision tree is perhaps the optimal mechanism in mimicking the clinical diagnostic rules. This chapter compares the performances of two different designs of the multiple decision trees via experiments. Digital image attributes, including both geometric and colorimetric ones, are all examined in detail. Receiver operating characteristic curves of varying ensemble sizes are presented, illustrating the effectiveness of decision trees in melanoma diagnosis.

Keywords Computer aided diagnosis · Decision tree · Melanoma

Introduction

Computer based early diagnosis of melanoma has been studied for more than 2 decades [13]. One of the key aims of this research field is to build an assistance system for clinical diagnosis applications. This task is important because the manual inspection, while common in clinical practice, has undesirable features such as repetitiveness and subjectivity. Computer based methods, however, have huge potential in

Y. Zhou (✉)
School of Computing, Engineering and Physical Sciences, University of Central Lancashire, Preston PR1 2HE, UK
e-mail: kylepub@gmail.com

Z. Song
Department of Biomedical Science, University of Sheffield, Sheffield S10 3TN, UK

J. Scharcanski and M. E. Celebi (eds.), *Computer Vision Techniques for the Diagnosis of Skin Cancer*, Series in BioEngineering, DOI: 10.1007/978-3-642-39608-3_10,

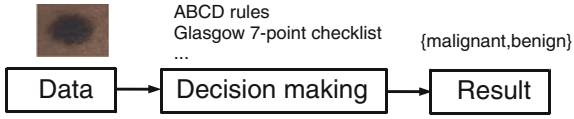


Fig. 1 General melanoma diagnosis process

alleviating these shortcomings and providing an important clinical alternative when second opinion is needed.

For a computerized melanoma diagnosis system, there are two key components: data and decision-making. As shown in the flowchart in Fig. 1, the data are mainly images while the decision-making process refers to applying certain machine learning techniques to label the sample as either benign or malignant. In the literature, a variety of imaging protocols, such as digital dermoscopy, infrared imaging, multispectral imaging, confocal microscopy, consumer level digital cameras, have been applied in collecting digital data of lesions [6, 15, 16].

Though data is the inalienable part of a computer aided diagnosis (CAD) system, the decision-making is of utmost importance, especially with the abundance of numerous imaging devices on the market. Here for a CAD system, decision-making refers to carrying a diagnosis based on certain algorithms/classifiers. In other words, a CAD system mimics the clinical diagnosis process, which typically involves computing techniques such as feature extraction and classification.

Although CAD based melanoma diagnosis has been an active research area, its clinical application is still stagnating currently. Before bringing CAD based melanoma inspection into real clinical practices, it seems there are two critical challenges which deserve highlights here:

(1) Data gap: The feature descriptors describing pigmented lesion properties may include both metric and non-metric values. Some attributes might involve physical meanings such as the diameter, some other features, such as Euler number of the lesion area, might not have any unit. If this is the case, the decision-making process has data gap. Typical learning paradigms like support vector machines [11] and neural networks [9] treat these data as purely numerical 'inputs' without paying attention to the physical/clinical meanings of these attributes.
(2) Decision-making gap: This refers to the decision making styles discriminating clinical diagnosis and computer based early diagnosis of melanoma. For instance, [1, 14, 21] only used border attributes in melanoma diagnosis. This decision-making style focuses on just one property of pigmented lesions. Though the one-feature based CAD is effective in justifying the usefulness of certain descriptors, there is a gap between these CAD systems and the clinical diagnosis.

There are a few publications covering the decision gap [2, 9, 18], which are of varying performances in dealing with the decision-making gap. However, to our knowledge, the data gap is still neglected by and large in CAD based melanoma diagnosis. To avoid data gap in CAD systems, it is important to choose attributes with equivalent metrics if numerical operations are going to be carried out between

these attributes. In addition, numerical operations for attributes of different metrics shall be avoided.

In the literature, there are a few protocols proposed for decision-making in melanoma diagnosis.

In [9], a neural network was proposed to carry diagnosis of melanoma. The feature descriptors cover asymmetry/border/colour properties and overall there are 14 entries in the feature vector. In experiments, one of the tested neural networks includes 14 input neurons, 7 hidden neurons and 1 output neuron. This neural network was trained with the well-known back-propagation method.

Sboner et al. [17] formulated a multiple classifier system, including linear discriminant analysis (LDA), decision trees and k-nearest neighbor. The feature descriptors cover geometrical and colorimetric domains, resembling the ABCD rules proposed in [10].

In [18], another ABCD rules based system for melanoma diagnosis was tested. In this system, there are 8 feature descriptors: skin line direction, skin line intensity, asymmetry, border irregularity, red component variegation, green component variegation, blue component variegation and diameter of lesion. After dimension reduction for the feature vectors, the first two principal components of these 8 components are selected for designing a linear classifier. The area under the receiver operating characteristic (ROC) curve obtained is 0.94.

She et al. [18] represents a typical approach in designing computer aided diagnosis of melanoma, which involves dimension reduction after feature extraction. In [18] and [9, 17], there is one hidden assumption behind this method: different features representing different properties are numerically computable. Here by computable, it means they are allowed to be mixed together numerically, including addition/deduction etc. Due to the non-homogeneity of the attributes, this operation, though commonly accepted, neglects the data gap between different attributes. The associated undesirable risks include: Firstly, the result depends on the chosen metrics. Secondly, the result only has numerical meaning but the physical/clinical meaning might be elusive. For instance, suppose the skin line direction, blue component variegation and diameter of lesion are f_1, f_7 and f_8 respectively, adding them together as $f_1+f_7+f_8$ gives a numerical number, which is hardly of any clinical significance.

This chapter utilizes multiple decision trees for melanoma diagnosis, aiming at solving the data gap and the decision gap simultaneously. Section "Method" discusses the main designs of two multiple decision tree systems. Section "Experiments and Results" presents experimental results on a test dataset. For an early version of this paper, see [20].

Method

Since a typical decision-making process involves feature extraction and classification, this section is divided into two subsections. In section "Feature descriptors", the shape, size and colour features for the pigmented lesions are examined, which are

Table 1 2D features for lesion images

Asymmetry	x1	a_1	Asymmetry w.r.t. horizontal axis
	x2	a_2	Asymmetry w.r.t. diagonal axis
	x3	a_3	Asymmetry w.r.t. off-diagonal axis
	x4	a_4	Asymmetry w.r.t. vertical axis
Border	x5	b_1	Indentation irregularity index
	x6	b_2	Protrusion irregularity index
	x7	b_3	Maximum indentation irregularity index
	x8	b_4	Maximum protrusion irregularity index
Colour	x9	c_1	Colour variation in red channel
	x10	c_2	Colour variation in green channel
	x11	c_3	Colour variation in blue channel
	x12	c_4	Colour variation in gray intensity
Diameter	x13	d_1	Equivalent diameter
	x14	d_2	Equivalent diameter of convex hull
	x15	d_3	First singular value of covariance matrix of coordinate vector
	x16	d_4	Second singular value of covariance matrix of coordinate vector

inspired by the ABCD rules [10, 16]. In section "Decision Trees", unlike [17] which used three heterogeneous classifiers the decision tree is the only classifier employed here to conduct diagnosis.

Feature Descriptors

This is a classical topic in CAD based melanoma diagnosis. Here a set of features are extracted following the ABCD rules. As shown in Table 1, there are 16 attributes altogether.

Asymmetry

In [10], asymmetry rule means the more asymmetrical the lesion is, the more likely it is malignant. Celebi et al. and She et al. [7, 18] use principal axis based methods to describe the asymmetry features. In this chapter, a four-axis method is utilized to describe the asymmetry features. Figure 2 shows the idea of how to construct the asymmetry features. For instance, in Fig. 2a, the solid line is the first axis selected and the original lesion area lies within the region enclosed by the solid boundary. The dashed line encloses an area which, given the axis, is symmetrical with respect to (w.r.t.) the original lesion area. Apart from the horizontal line in Fig. 2a serving as the symmetrical axis, the diagonal line, off-diagonal line and vertical line can also be used as symmetrical axes. Figure 2b–d show these symmetrical axes and their corresponding symmetrical images of the original lesion area respectively.

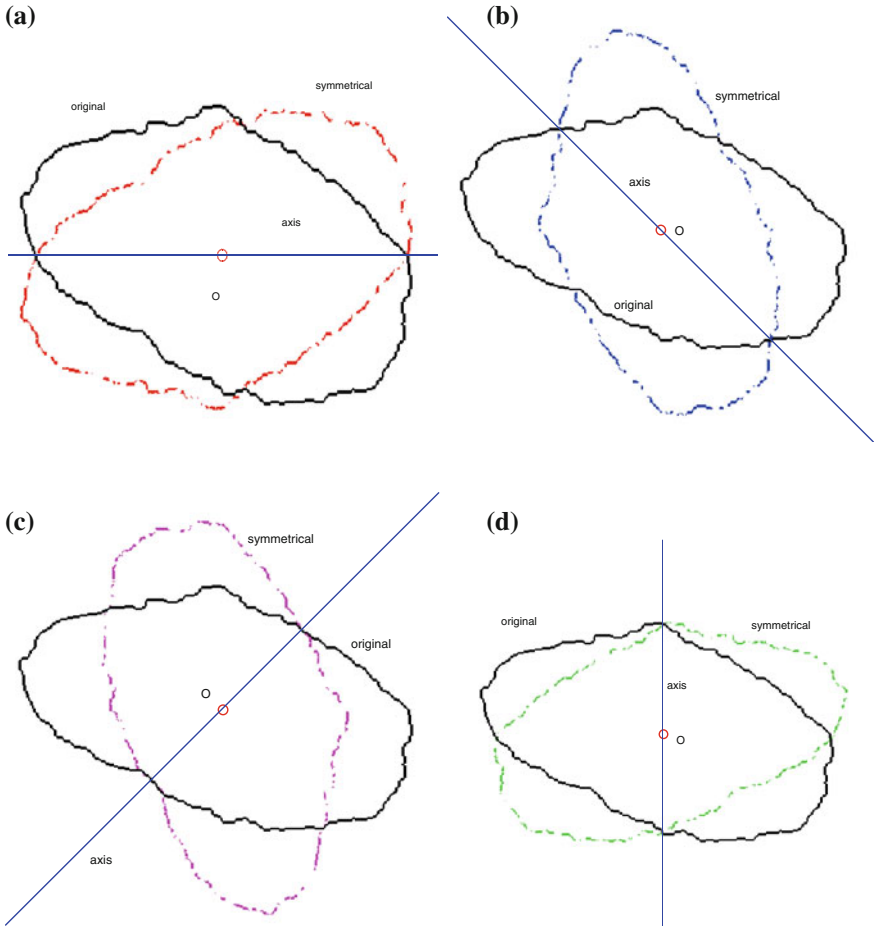


Fig. 2 Extract asymmetrical features for suspicious pigmented lesions

To construct one symmetrical area as shown in Fig. 2, there are two steps. Firstly, one detects the centroid of the lesion area. Then a reference axis, i.e., the symmetrical axis, is selected. In Fig. 2, these symmetrical axes are 0, 45, 90, 135 degrees w.r.t. the horizontal axis. Denote the original lesion area as A_0, and the symmetrical lesion area as A_i, the asymmetry indices are defined as:

$$a_i = \frac{\#\mathrm{XOR}(A_0, A_i)}{2A_0}, i = 1, 2, 3, 4 \tag{1}$$

where XOR means exclusive-or operation at pixel level; '#' is to take the number of all non-zero elements. Thus these four attributes are all non-metric variables.

Border Irregularity

To maintain the equivalence in terms of units for the border irregularity descriptors, indentation/protrusion index proposed in [14] is used here.

The method in [14] constructs an area based irregularity index in several steps: Firstly, a smoothed outline of the lesion is extracted via a series of multiscale Gaussian filters. Then the area enclosed within the smoothed outline is compared with the original lesion area. Denote the original lesion area as A_0 and the smoothed area as A_s, the indentation area lies within A_s while outside A_0. The protrusion area is the opposite: it lies within A_0while outside A_s. So the indentation and protrusion maps can be obtained as follows:

$$I_{ind} = A_0 - A_s, \tag{2}$$

$$I_{pro} = A_s - A_0, \tag{3}$$

where I_{ind} and I_{pro} represents the indentation and protrusion images respectively. Specifically, for pixels in I_{ind} and I_{pro}, the definitions are as follows:

$$I_{ind}(x, y) = \begin{cases} 1, & A_0(x, y) == 1 \& A_s(x, y) == 0 \\ 0, & otherwise \end{cases}, \tag{4}$$

$$I_{pro}(x, y) = \begin{cases} 1, & A_0(x, y) == 0 \& A_s(x, y) == 1 \\ 0, & otherwise \end{cases}. \tag{5}$$

With the above definitions, one can extract the border irregularity features as follows:

$$b_1 := \frac{\#(I_{ind})}{\#(A_0)}, \tag{6}$$

$$b_2 := \frac{\#(I_{pro})}{\#(A_0)}, \tag{7}$$

$$b_3 := \frac{\#(\widetilde{I}_{ind})}{\#(A_0)}, \tag{8}$$

$$b_4 := \frac{\#(\widetilde{I}_{pro})}{\#(A_0)}, \tag{9}$$

where #(.) represents the total area in terms of pixels; ($\sim$) refers to the region of the maximum indentation/protrusion area. Therefore, from the above definitions of the border irregularity features provided by [14], it can be seen that these features are non-metric and they lie within [0, 1].

Colour Variation

The colour variations are extracted via four different channels: red, green, blue and intensity channels Fig. 3. The intensity channel is obtained by fusing the RGB channels as follows:

$$I(x, y) = \sqrt{R^2(x, y) + G^2(x, y) + B^2(x, y)}, \tag{10}$$

where x and y denotes the coordinates of the pixels. By calculating the statistics of the images, the four features of colour variations in Table 1 are defined as follows:

$$c_1 := \log \frac{\sigma(R)}{\mu(R)}, \tag{11}$$

$$c_2 := \log \frac{\sigma(G)}{\mu(G)}, \tag{12}$$

$$c_3 := \log \frac{\sigma(B)}{\mu(B)}, \tag{13}$$

$$c_4 := \log \frac{\sigma(I)}{\mu(I)}, \tag{14}$$

where μ and σ denotes the mean value and standard deviation of the lesion area pixels of the given colour channel. Since the mean value and the standard deviation are of the same unit, i.e., the image intensity in different colour channels, the ratios between these two are free of metric unit. For computational conveniences, one can set the range for the above colour variation descriptors, e.g., $[-10, 10]$.

Diameter

Unlike the original diameter feature expressed in [10] which states that 6mm diameter is the critical threshold in judging a suspicious lesion as malignant, here a 4-element diameter feature vector is extracted, which includes the following attributes:

$$d_1 := 2\sqrt{\#(A_0)/\pi}, \tag{15}$$

$$d_2 := 2\sqrt{\#(A_c)/\pi}, \tag{16}$$

$$d_3 := svd(S, 1) \times 2, \tag{17}$$

$$d_4 := svd(S, 2) \times 2, \tag{18}$$

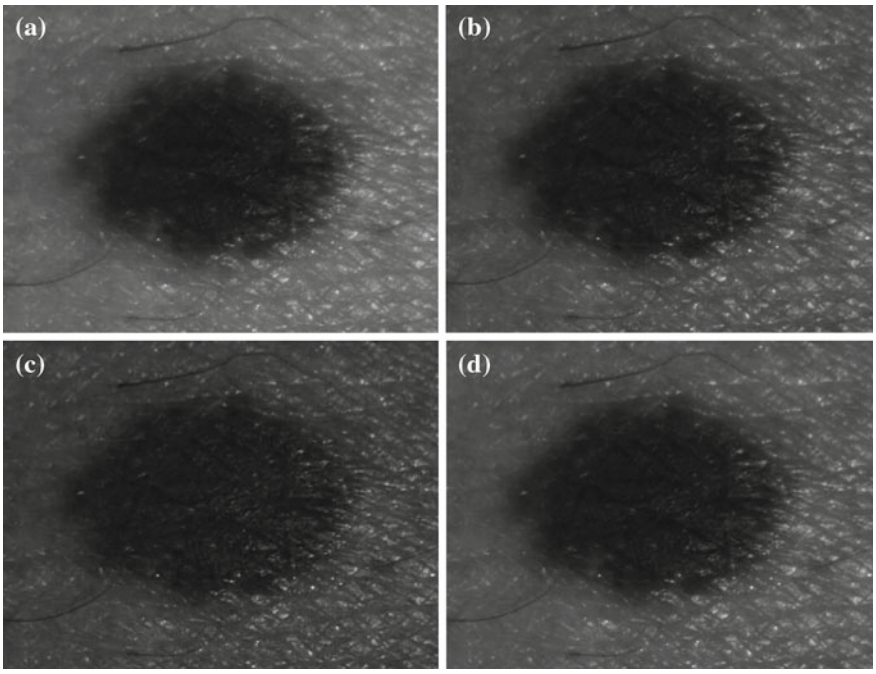


Fig. 3 Example of one raw image in red channel (**a**), green channel (**b**), blue channel (**c**) and gray intensity channel (**d**)

where A_c is the convex hull extracted with the given lesion area. Figure 4 gives an example of a lesion area and its convex hull. Therefore from (15), d_1 is the equivalent diameter of the original lesion area. Likewise, (16) means that d_2 is the equivalent diameter of the convex hull.

S in (17) and (18) is the covariance matrix constructed by the coordinates of the lesion area pixels. Since the image plane is 2D and S is a 2 by 2 positive definite matrix. Therefore, there are two singular values for S and $svd(S, 1)$ and $svd(S, 2)$ represent the first and the second singular values of S respectively.

Clearly the four diameter feature descriptors above are of length units expressed in image pixels. In addition, the numerical values of these four attributes are non-negative.

Therefore, in the above subsections, the ABCD rules have been implemented in a way aiming at reducing the data gap between attributes within each feature group. As there are 4 elements for every feature group, altogether there are 16 attributes extracted for one 2D image sample of pigmented skin lesions.

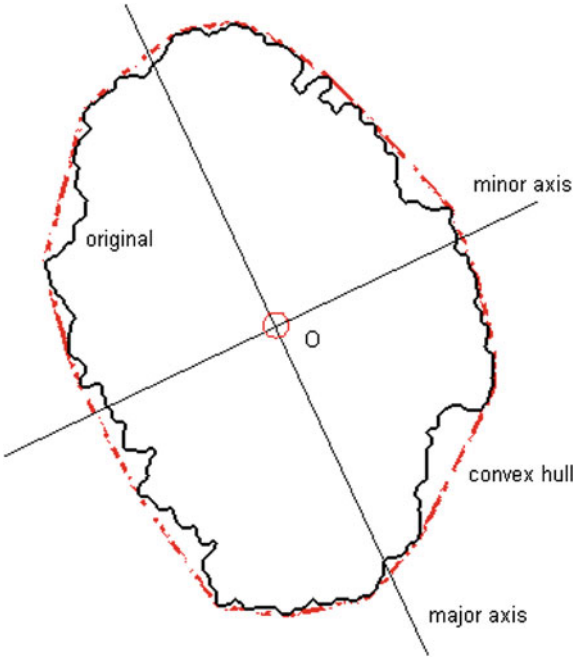


Fig. 4 Example of a lesion area and its convex hull

Decision Trees

Here decision trees [8] and only decision trees are used for decision making. They are selected not only because of the learning capability, but also because of the interpretability and intuitiveness [19].

Recently, [19] proposes a decision tree based classification system which allows the end-users to tune the trees manually. This technique, called Visually Tuned Decision Tree (VTDT), illustrates that decision tree not only can generate comprehensible rules to domain experts, it also has the potential to allow human experts to embed their domain knowledge in designing the decision making models.

Figure 5 shows one example of the binary decision tree generated during experiments. As can be seen from Fig. 5, every non-leaf node for this decision tree splits into 2 branches and the split is based on univariate nodes only.

In utilizing decision trees to design the CAD system for melanoma diagnosis, there exist two approaches. As shown in Fig. 6a, this method stacks the feature vectors together and feeds the overall feature vector into decision trees. The final output is generated by majority voting of the member decision trees. This approach has the data gap problem since the diameter attributes are in pixels while the other values are non-metric.

Figure 6b offers an alternative which eliminates the data gap. This is achieved by forwarding only homogeneous data into one given decision tree, i.e., no mixture or stacking of variables bearing different metric units. In addition, for the decision

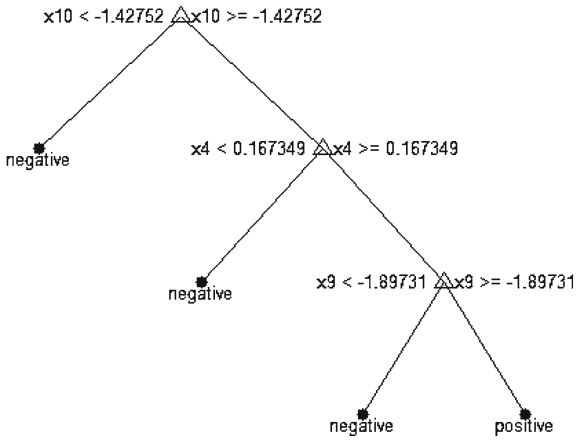


Fig. 5 Example of a binary decision tree based on univariate splits for melanoma diagnosis

making part, since Fig. 6b combines both the geometric and colorimetric properties of skin lesion in diagnosis, the decision-making gap is avoided as well, making it more similar to the clinical diagnosis than using a single diagnosis rule.

Experiments and Results

This section presents the CAD experimental results, which were obtained by testing the diagnosis systems in Fig. 6a, b via cross-validations.

Experimental Setup

In the collected 2D macroscopic image dataset, there are 110 malignant samples and 125 benign lesions. The lesion areas in these data were segmented by using the online graph-cut based algorithm [3–5, 12]. Feature descriptors were extracted with the above methods. Figure 7 shows scatter plots of the samples by applying principal component analysis. In addition, to run the experiments, Matlab 7.12 is used as the experimental platform.

In standard k-fold cross validation, the data are divided into k-subsets first. Then k-1 subsets are used to train the classifier and the remaining subset is left for testing. As multiple decision trees are used in both Fig. 6a and b, employing standard k-fold cross-validation will introduce significant training data overlaps among different decision trees. Thus in training the multiple decision trees, only 1 of the k subsets was selected to train the classifier. Hence for different decision trees, the probability for two of them holding the same training data is generally small ($1/k^2$). In the following experiments, k is chosen as 5 and for multiple decision trees, N is set as

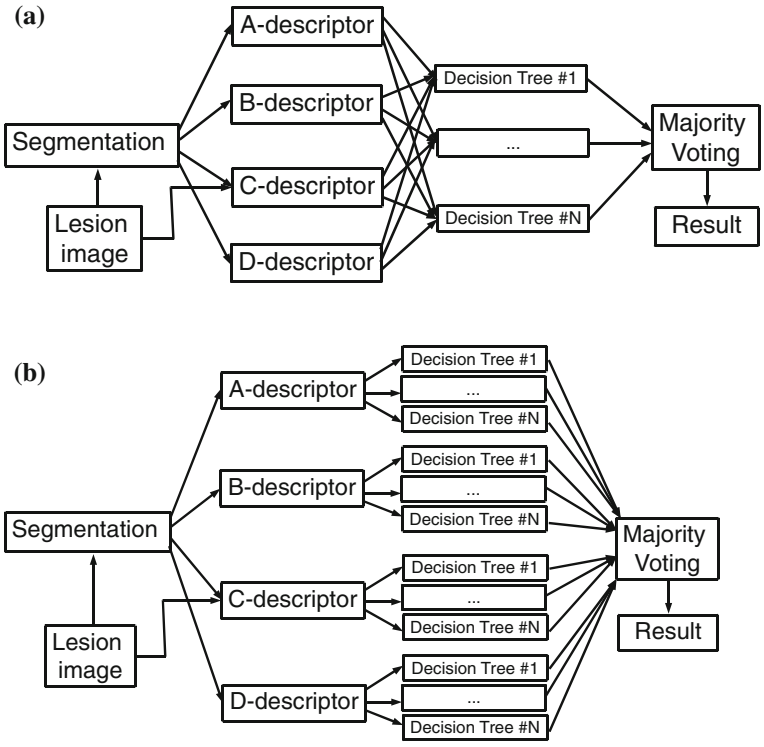


Fig. 6 Melanoma diagnosis using multiple decision trees. **a** with data gap; **b** data gap free

3. In this case, the Fig. 6b design has 12 trees. When making a majority voting for Fig. 6b, if there is a draw, the outputs of the decision trees using colour attributes are then chosen to make a final decision. This draw-breaking strategy is supported by Fig. 8.

For the design of nodes in the binary decision trees in Fig. 6, one can typically use either univariate nodes, or multivariate nodes with linear discriminant analysis to split a mother node into two daughter nodes. Specifically, for a univariate node, there is no data gap since no direct numerical operation between attributes arises. However, it also limits the flexibility in choosing the decision boundaries [8]. For multivariate nodes with linear classifiers, as shown in Table 2, receiver operating characteristic (ROC) curves can be generated which offers an effective evaluation for different methods.

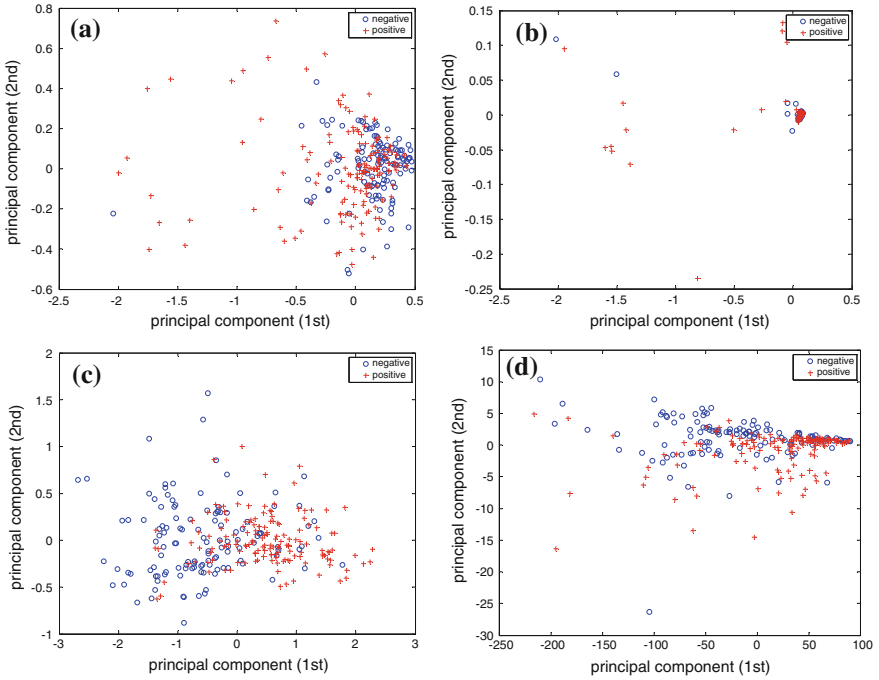


Fig. 7 Scatter plots of samples based on principal component analysis: **a** asymmetry; **b** border; **c** colour; **d** diameter

Table 2 Two setups for evaluating multiple decision trees

	Nodes with univariate splits		Nodes with linear classifiers	
	Data gap	ROC curve	Data gap	ROC curve
Fig. 6a	No	–	Yes	Available
Fig. 6b	No	–	No	Available

ROC Analysis

Figure 8 shows the ROC curves for the ABCD features. Each ROC curve corresponds to one feature group processed via linear discriminant analysis based multivariate nodes. The areas under curves (AUCs) in Fig. 8 are 0.69, 0.71, 0.74 and 0.62 respectively.

Figure 9a presents the ROC curve by stacking the descriptors directly as one vector and then feeding the vector into a decision tree with multivariate nodes. As can be seen from Fig. 9a, the AUC is 0.88. Figure 9b shows the ROC curve for decision system in Fig. 6b. This AUC is 0.93, slightly higher than Fig. 9a.

In addition, from Figs. 8 and 9, it can be seen that by applying multiple decision trees to mimic the ABCD rules, the diagnostic performance can be enhanced effec-

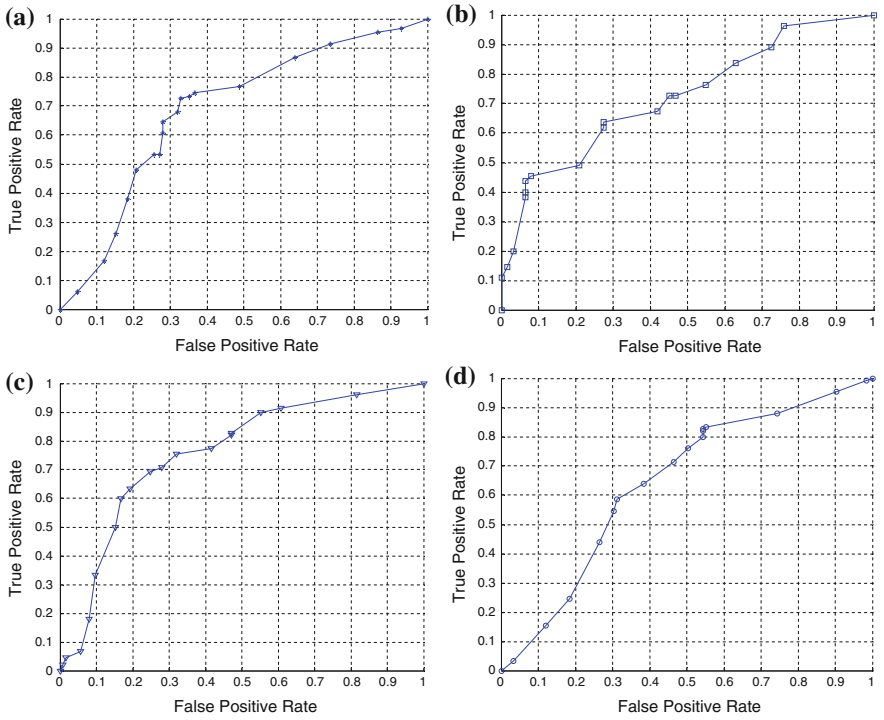


Fig. 8 ROC curves for ABCD descriptors respectively. **a** Asymmetry (AUC = 0.69); **b** border irregularity (AUC = 0.71); **c** colour variations (AUC = 0.74); **d** diameter (AUC = 0.62)

tively. Although there exists data gap in using Figs. 6a, 9a still suggests that the direct stacking of features deserves certain efforts. The multiple decision trees in Fig. 6b eliminate the data gap while performing even better, at least qualitatively as shown in Fig. 8's ROC analysis.

Classification Result

Apart from the ROC analysis, the two multiple decision tree systems in Fig. 6 are also evaluated in classification. Table 3 shows the classification results, in which the performance indicators, including accuracy/sensitivity/specificity, positive predictive value (NPV) and negative predictive value (NPV), are defined as follows:

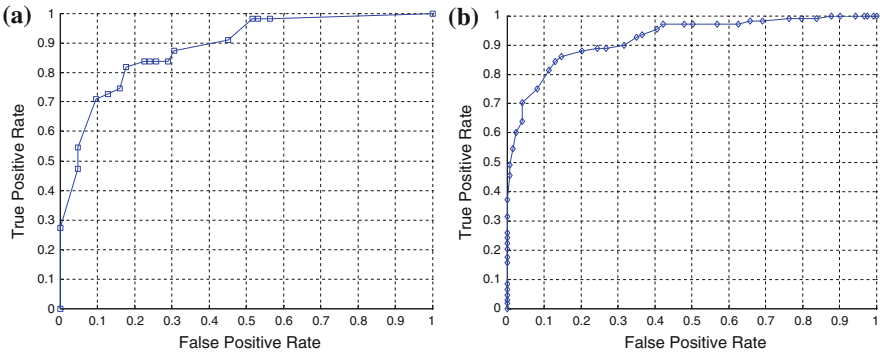


Fig. 9 ROC curve for multiple decision trees in Fig. 6a and b. **a** Fig. 6a: AUC = 0.88; **b** Fig. 6b: AUC = 0.93

Table 3 Simulation results with different multiple decision tree designs (%)

	Accuracy	Sensitivity	Specificity	PPV	NPV
Fig. 6a	79.6(±0.39)	73.8(±0.50)	84.7(±0.39)	81.1(±0.44)	78.4(±0.42)
Fig. 6b	82.9(±0.24)	77.2(±0.51)	87.8(±0.01)	84.8(±0.08)	81.5(±0.34)

(*PPV* positive predicative value; *NPV* negative predicative value)

$$f_{acc} := \frac{TN + TP}{TN + FN + TP + FP}, \tag{19}$$

$$f_{spe} := \frac{TN}{TN + FP}, \tag{20}$$

$$f_{sen} := \frac{TP}{TP + FN}, \tag{21}$$

$$f_{ppv} := \frac{TP}{TP + FP}, \tag{22}$$

$$f_{npv} = \frac{TN}{TN + FN}, \tag{23}$$

where TN/FN/TP/FP refer to the number of test results which are 'true negative', 'false negative', 'true positive' and 'false positive'.

From Table 3, it can be seen that Fig. 6b is better than Fig. 6a in terms of mean values of the above indicators, though with varying levels. The standard deviations of these indicators for Fig. 6b's design are generally less than results obtained from using Fig. 6a, indicating a slightly better consistency in performance.

Conclusion

By working on the data gap as well as the decision-making gap, this chapter examined the application of multiple decision trees in computer aided diagnosis of melanoma. Two structurally different designs are compared: the classic design (Fig. 6a) and the new design (Fig. 6b). In experimental studies on a dataset containing 235 samples (125 benign and 110 malignant), while the first design gives good results in ROC analysis and classification tests, the second design can perform even better. One disadvantage of the new design is that its structure appears less straightforward than the first design. However, if one needs a higher performance CAD system, as shown in the above experiments, choosing the system in Fig. 6b might be slightly better than Fig. 6a.

Though Fig. 6a and b are different, in fact Fig. 6b can be obtained from Fig. 6a in two steps. Firstly, one can set the number of decision trees in Fig. 6a as 12, i.e., $N = 12$. Then by putting feature selections prior to feeding the data into the decision trees, one can generate the design in Fig. 6b.

References

1. Abbas, Q., Celebi, M.E., Garcia I.F., Rashid M.: Lesion border detection in dermoscopy images using dynamic programming. Skin Res. Technol. **17**(1), 91–100 (2011)
2. Alcón, J., Ciuhu, C., Kate, W., Heinrich, A., Uzunbajakava, U., Krekels, G., Siem, D., Haan, G.: Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE J. Sel. Top. Signal Process. **3**(1), 14–25 (2009)
3. Bagon, S.: Matlab wrapper for graph cut. (Dec 2006)
4. Boykov, Y., Veksler, O., Zabih, R.: Efficient approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. **20**(12), 1222–1239 (2001)
5. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. Pattern Anal. Mach. Intell. **26**(9), 1124–1137 (2004)
6. Cavalcanti, P.G., Scharcanski, J.: Automated prescreening of pigmented skin lesions using standard cameras. Comput. Med. Imag. Graph. **35**, 481–491 (2011)
7. Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. Comput. Med. Imag. Graph. **31**(6), 362–373 (2007)
8. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd edn. John Wiley, New York (2001)
9. Ercal, F., Chawla, A., Stoecker, W., Lee, H., Moss, R.: Neural network diagnosis of malignant melanoma from color images. IEEE Trans. Biomed. Eng. **41**(9), 837–845 (1994)
10. Friedman R, Rigel D, Kopf A.: Early detection of malignant melanoma: the role of physician examination and self-examination of the skin. CA Cancer J. Clin. **35**(3), 130–151 (1985)
11. Gilmore, S., Hofmann-Wellenhof, R., Soyer, H.: A support vector machine for decision support in melanoma recognition. Exp. Dermatol. **19**(9), 830–835 (2010)
12. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Trans. Pattern Anal. Mach. Intell. **26**(2), 147–159 (2004)
13. Korotkov, K., Garcia, R.: Computerized analysis of pigmented skin lesions: a review. Artif. Intell. Med. **56**(2), 69–90 (2012)

14. Lee, T.K., Claridge, E.: Predictive power of irregular border shapes for malignant melanomas. Skin Res. Technol. **11**(1), 1–8 (2005)
15. Manousaki, A.G., Manios, A.G., Tsompanaki, E.I., Panayiotides, J.G., Tsiftsis, D.D., Kostaki, A.K., Tosca, A.D.: A simple digital image processing system to aid in melanoma diagnosis in an everyday melanocytic skin lesion unit: a preliminary report. Int. J. Dermatol. **45**(4), 402–410 (2006)
16. Rigel, D., Russak, J., Friedman, R.: The evolution of melanoma diagnosis: 25 years beyond the ABCDs. CA Cancer J. Clin. **60**(5), 301–316 (2010)
17. Sboner, A., Eccher, C., Blanzieri, E., Bauer, P., Cristofolini, M., Zumiani, G., et al.: A multiple classifier system for early melanoma diagnosis. Artif. Intell. Med. **27**(1), 29–44 (2003)
18. She, Z., Liu, Y., Damatoa, A.: Combination of features from skin pattern and ABCD analysis for lesion classification. Skin Res. Technol. **13**(1), 25–33 (2007)
19. Stiglic, G., Kocbek, S., Pernek, I., Kokol, P.: Comprehensive decision tree models in bioinformatics. PLoS ONE **7**(3) (2012)
20. Zhou, Y., Song, Z.: Binary decision trees for melanoma diagnosis. In: 11th International Workshop on Multiple Classifier Systems, Nanjing, 15–17 May 2013
21. Zhou, Y., Smith, M., Smith, L., Warr, R.: A new method describing border irregularity of pigmented lesions. Skin Res. Technol. **16**(1), 66–76 (2010)