POLITECNICO DI MILANO

Faculty of Biomedical Engineering

Computer-aided detection and classification of pigment network in pigmented skin lesions

Thesis supervisor: Enrico Caiani

Thesis Author: Ismael Huertas Fernández   ID number: 752257

Academic year 2010/2011

Index

List of figures

List of Equations

List of Tables

1. <u>Abstract</u>

Teledermoscopy is a recent medical practice where a primary care physician takes a picture of the Pigmented Skin Lesion and sends it telematically to the Dermatologist for a remote diagnosis. Its creation has boosted the development of computer-aided diagnosis with means of image processing techniques. Early Melanoma detection is the main target of this new tendency. Melanoma can be characterized by some structures and features. Association studies with histopathological diagnosis have demonstrated Atypical Pigment Network to be an indicative for the malign tumor. A pigment network can be either typical or atypical. Atypical Pigment Network consists of a net of hypopigmented holes non-uniformly distributed throughout the lesion, broad darker lines and hyperpigmented areas covering the meshes. A tool in Matlab has been created in order to detect the presence of a pigment network and if so, distinguish between typical and atypical, giving an estimation of risk. Given a dermatoscopic image, several stages have been followed, such artifacts' removal and contrast enhancement by filtering, segmentation of the lesion from the background by a clustering algorithm and feature extraction by morphological operators and some calculations. To classify the network certain created coefficients have been assessed such Spatial Uniformity of the meshes and thickness of the lines. Thresholds have been set for the differentiation. 35 images have been tested and as a result, 33 have been successfully classified, giving sensitivity, specificity and accuracy above 90%.

2. <u>Prefacio</u>

La teledermoscopia è una recente pratica medica nella quale il medico di medicina generale acquisisce una foto della lesione pigmentata della pelle del paziente e la invia telematicamente al Dermatologo per una diagnosi remota. La creazione di tale servizio a distanza ha spinto lo sviluppo di algoritmi e tecniche di diagnosi assistita dal computer tramite elaborazione delle immagini, atte a favorire il rilevamento precoce del Melanoma. Il Melanoma può essere caratterizzato da alcune strutture e caratteristiche. Una rete pigmentata può essere di due tipologie: tipica o atipica. Studi di associazione con diagnosi istologiche hanno dimostrato come la rete pigmentata atipica costituisca un fattore prognostico indicativo di tumore maligno. La rete pigmentata atipica presenta dei buchi ipopigmentati non uniformemente distribuiti lungo la lesione, contorni scuri e spessi, e regioni iperpigmentate. Scopo di questa tesi é stato lo sviluppo di un tool in Matlab per rilevare automaticamente la presenza di queste caratteristiche rete, e distinguere quindi tra rete pigmentata tipica e atipica. Data un'immagine dermatoscopica, alcuni step sono stati seguiti, come per esempio la rimozione di artifatti, il miglioramento del contrasto mediante operazioni su istogramma e filtraggio, segmentazione della lesione dallo sfondo mediante un algoritmo di clustering, e l'estrazione della caratteristica cercata mediante operatori morfologici. Per classificare correttamente la rete, sono stati valutate alcune caratteristiche quali l'uniformità spaziale dei buchi e lo spessore delle linee, definendo soglie opportune per la differenziazione. Gli algoritmi sviluppati sono stati testate su 35 immagini, risultando in 33 immagini classificate correttamente, con sensitività e specificità oltre il 90%.

3. <u>Introduction</u>

3.1. <u>Clinical Situation</u>

It is well-known the large waiting list existing in the last decades for every specialist, provoked by the increasing older population and if we talk about dermatology also by the large amount of cases of skin cancer. In order to avoid this expensive and bothersome trouble for the healthcare system that renders it ineffective, the tendency of these last years has been to give as greater as possible power of decision to the family doctor.

Furthermore, increasing subspecialization in medicine results in a concentration of experts in few medical centers and a lack of expert knowledge in other, predominantly rural areas. Physicians in these areas may have problems with the diagnosis and management of equivocal pigmented skin lesions. The patient has to be sent to an expert in a medical center for further diagnosis and, depending on the results, for adequate treatment.

This transportation means inconvenience for the patient and costs for the medical care system due to re-examination by the expert. Finally, the whole procedure of patient transportation and re-examination that is cost- and time-consuming might have been unnecessary if the expert could have been consulted beforehand.

In the last few years the continuous progress in computer technology has lead to the introduction of a revolutionary diagnostic tool, known as telemedicine, that is improving communication between physicians and medical specialists and will help decrease costs for the citizen and the health care system.

In several medical specialties, where digital images are crucial in diagnosis and management decisions, such as radiology or internal medicine (endoscopy; ultrasound) to name but a few, telemedicine already represents a well-integrated part of daily medical life.

In addition, it has recently been shown that teledermatology is a useful and suitable diagnostic tool, especially in communities where dermatologists are not available. An even newer diagnostic approach is **teledermoscopy that permits to send clinical as well as dermoscopic images of pigmented skin lesions using telematic networks[19]**.

Recently, Piccolo et al. performed a teledermoscopic study on 66 pigmented skin lesions by sending clinical and dermoscopic images from L'Aquila (Italy) to Graz (Austria) via internet. The results of this study, achieved by the store-and-forward approach, demonstrated that the degree of concordance between "face-to-face" diagnosis and telediagnosis was 91% when performed by experts in dermoscopy on both sides.

Definitely, teledermoscopy may contribute significantly to reduce waiting time and workload in dermatologic clinics, spare the patients unnecessary distress, and ultimately save public money whilst providing a faster, reliable and more efficient service.

This idea of Telediagnosis has led to the release of a new practice in dermatology, the **Computer-aided diagnosis**, which is basically a new way to diagnose implementing image processing techniques on image analyzers.

This means that, if we provide to a physician a dermatoscope and software that performs diagnosis, we are changing the Paradigm of diagnosis. A family doctor with no expertise in dermatology would be able to give an estimation of a skin lesion to be potentially a Melanoma. Thereby, according to the performed diagnosis, a priority sending the images can be set, putting the risky ones at the top of the queue. Or even, if the diagnosis doesn't hesitate at all, avoid the remote diagnosis reducing the workload.

Melanoma is a malignant tumor of melanocytes, which are melanin-producing cells located in the bottom layer of the skin's epidermis among others.

For one decade, experts have been working on it, trying to figure out which is the right criteria and the way to handle it computationally when dealing with Melanoma. However, it is just needed a quick search on literature to become aware there is remarkable approach to rely on and leads the state-of-art of the topic, the **Two-steps algorithm.**

This algorithm is later described in detail, but mainly it first classifies the lesion into Melanocytic and Non-melanocytic. Since Melanoma is a malignant tumor of melanocytes, non-melanocytic are discarded and melanoytic are further studied through a **Pattern Analysis** searching for the features highly related to the presence of Melanoma. Several were found: Atypical pigment network, Blue-whitish veil, Regression structures, Irregular Blotches, Irregular Streaks and some vascular structures. The description of each feature, including both Benign and Malignant patterns, is attached below in Table I [9].

	Benign	Malignant
Dots	Centrally located or situated right on the network	Unevenly distributed and scattered focally at the periphery
Globules	Uniform in size, shape, and color, symmetrically located at the periphery, centrally located, or uniform throughout the lesion as in a cobblestone pattern	Globules that are unevenly distributed and when reddish are highly suggestive of melanoma
Streaks	Radial streaming or pseudopods tend to be symmetrical and uniform at the periphery	Radial streaming or pseudopods tend to be focal and irregular at periphery
Blue-white veil	Tends to be centrally located	Tends to be asymmetrically located or diffuse almost over entire lesion
Blotch	Centrally located or may be diffuse hyperpigmented area that extends almost to periphery of the lesion	Asymmetrically located or there are often multiple asymmetrical blotches
Network	Typical network that consists of light to dark uniform pigmented lines and hypopigmented holes	Atypical network that may be nonuniform with black/brown or gray thickened lines and holes of different sizes and shapes
Network borders	Either fades into the periphery or is symmetrically sharp	Focally sharp

Table 1: Features related to Melanoma

Researchers and developers are currently working on the automated identification of these features. This thesis tries to face the extraction of an **Atypical Pigment Network**, which has been faithfully demonstrated being a high risk factor. It was proved through Association studies between hystopathologic diagnosis and Pattern Analysis diagnosis.

The rates of the survey are attached below in Table I [8], please notice the high values of Odds ratios in the mentioned features. Thus, we can truly believe that, applying Pattern Analysis and finding these features, we are probably dealing with a Melanoma.

	Distribution of assignments*			
Variable	Melanoma	Nonmelanoma	Odds ratio	95% CI
Global pattern				
Reticular	0.33	0.32	1.1	0.5-2.1
Globular	0.03	0.17	0.1	0.05-0.4|
Cobblestone	0.001	0.05	0.02	0.002-0.1†
Homogeneous	0.01	0.07	0.1	0.0-0.5†
Starburst	0.001	0.07	0.01	0.002-0.1†
Parallel	0.04	0.06	0.6	0.1-5.7
Multicomponent	0.51	0.19	4.3	2.5-7.4†
Nonspecific	0.09	0.06	1.4	0.7-3.0
Pigment network				
Absent	0.14	0.42	0.2	0.1 0.5†
Typical	0.09	0.32	0.2	0.1-0.4†
Atypical	0.77	0.27	9.0	5.1-16.0†
Dots/globules				
Absent	0.21	0.22	0.9	0.5 1.8
Regular	0.06	0.42	0.1	0.1-0.1†
Irregular	0.73	0.36	4.8	2.8-8.0†
Streaks				
Absent	0.45	0.70	0.4	0.2-0.6†
Regular	0.03	0.14	0.2	0.1-0.3|
Irregular	0.52	0.16	5.8	3.5-9.6†
Blue-whitish veil				
Present	0.49	0.25	2.9	1.6-5.0†
Blotches				
Absent	0.53	0.66	0.6	0.4-0.9†
Regular	0.04	0.19	0.2	0.1-0.3†
Irregular	0.42	0.15	4.1	2.7-6.4†
Hypopigmentation				
Present	0.44	0.28	2.0	1.3-3.0†
Regression structures				
Present	0.64	0.25	5.4	2.9-10.0†
Vascular structures				
Present	0.18	0.13	1.5	0.7-3.3

Table 2: Association between individual diagnostic criteria and melanoma diagnosis obtained by pattern analysis. Evaluated in 85 melanocytic lesions classified by at least 30 of 40 observers: odds ratios and 95% confidence intervals obtained by generalized estimating equation regression models.

Note a great value of percent, 0.77, is associated with the presence of Melanoma. This renders Atypical pigment network such a significant characteristic in the Pattern Analysis Diagnosis, and therefore must be accurately detected.

3.2. Pigment network

We just said Pigment Network is such a significant feature and it's what we are going to investigate, but, what exactly a pigment network is?

The pigment network represents the dermoscopic hallmark of melanocytic lesions, independent of their biologic behavior. Remarkably, the assessment of pigment network alterations is helpful for differentiating between benign and malignant melanocytic proliferations, especially when they are confined to the epidermis and superficial dermis.

The pigment network is a grid-like (honeycomb-like) network consisting of pigmented lines and hypopigmented holes. 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 pigment network can be either typical or atypical.

A typical network consist of light to dark brown network with small, relatively uniform-spaced network holes and homogeneous in color, and thin network lines distributed more or less regularly throughout the lesion, and usually thinning out at the periphery [9].
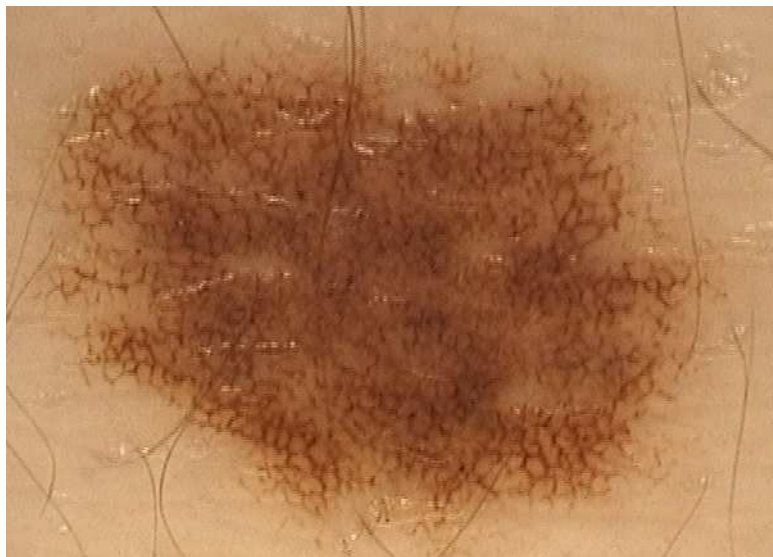
Figure 1: Typical Pigment Network

An atypical network is non-uniform, black, brown, or gray network with darker and/or broadened lines and irregularly distributed holes that are heterogeneous in area and shape. The lines are thick and often hyperpigmented and may end abruptly at the periphery [9]. You can notice in Figure 2 how "spots" of hyperpigmentation cover some areas rendering the distribution non-uniform.

Figure 2: Atypical Pigment Network

Slowing growing benign pigmented lesions such as a lentigo or junction nevus produce uniform network patterns. The lines are uniformly populated with benign pigment cells that grow at a slow rate. The network is homogeneous in color and usually thins out at the periphery.

Melanomas consist of malignant melanocytes that vary in size and degree of pigmentation. Malignant cells move through the epidermis in all directions at varying rates to produce structures with bizarre patterns. Cells are numerous in some locations and sparse in others.

A very explanatory video of what a pigment network is, could be found at this site:

http://www.youtube.com/watch?v=xeLdCrPkwdk

3.3. State-of-art in computer-aided diagnosis

Here we have to consider two kinds of state-of-art. The one related to the sort of criteria for facing the diagnosis, and then the one related to the techniques in image processing for facing the feature extraction.

3.3.1. State-of-art in Image Processing Techniques

Regarding the imaging techniques, we have to distinguish between the two main stages at the processing: Segmentation and Feature extraction.

With Segmentation we mean the process of taking apart any Region of Interest from a background, here we could even apply techniques not designed for Dermatologic images. Even though, since Dermoscopy has a long road in research, several have been described in literature.

We mainly highlight five methods evaluated in [2]:

- PCT/median cut algorithm [3]: This algorithm is based on the principal components transform (PCT). In this transform, the eigenvectors of the covariance matrix are used as a linear transform matrix on the original [RGB] vectors, so that the resulting vectors have components that are uncorrelated. The median cut component of the segmentation method is based on an algorithm that was developed for color compression, specifically, to map 24-bit-per-pixel color images into image requiring an average of 2 bits per pixel. The motivation behind this algorithm is to be able to display high-quality reproductions of color images with small frame buffers. In this experiment it was determined that segmentation into three colors was generally enough to allow for automatic determination of the tumor border. The color quantization task includes four phases:
    - Sampling the original image for color statistics
    - Choosing a colormap based on the color statistics
    - Mapping original colors to their nearest neighbors in the colormap
    - Quantizing and redrawing the original image

- Adaptive thresholding in the blue image plane [3]: The blue plane is used to generate a multiscale description of a histogram by convolving it with a series of

Gaussian kernels of gradually increasing width (standard deviation), and marking the location and direction of the sign change of zero-crossings in the second derivative. This process is known as scale-space filtering. The convolution is given by:

$$F(\xi, \tau) = f(x) * g(\xi, \tau) \quad = \int_{-\infty}^{\infty} f(u) \frac{1}{\sqrt{(2\pi)^\tau}} \exp\left[\frac{-(\xi - u)^2}{2\tau^2}\right] du$$

Equation 1: Scale-space filter convolution

The scale-space depends both on the independent variable, ξ, and the Gaussian deviation or scale constant, τ. The scale constant is inversely proportional to the number of peaks and valleys that can be extracted from the histograms. The importance of this fact is that if *a **priori*** knowledge is available about a particular type of image, τ can be set to be a constant to obtain the desired number of peaks and valleys, and the process can be made unsupervised. The resulting description, or fingerprint, is interpreted by relating pairs of zero-crossings to modes in the histogram, where each mode is modeled after a normal distribution. After histogram analysis, valid classes are determined from the modes of the histogram and mapped accordingly. Each class represents a different object within the image.

- Split-and-Merge algorithm [3]: Split and merge segmentation techniques are based upon a quad tree data representation, whereby a square image segment is broken *(split)* into four quadrants if the original image segment is nonuniform in attribute, as determined by a predefined homogeneity criteria based predicate test. If four neighboring squares are found to be uniform, then they are replaced *(merge)* by a single square composed of the four regions. Before performing the split and merge segmentation, the color information is first mapped onto a single image plane using the PCT. Once mapped, the average brightness level is determined for the plane and used in the predicate test of the split and merge procedure. The predicate test utilized here determines the standard deviation of each region to determine uniformity, and compares uniform regions with the average brightness level to determine to which class the region belongs, the

14

background or the object. The algorithm continues to split regions until all regions satisfy the homogeneity criterion.

- Stabilized Inverse Diffusion Equations (SIDE) non-linear diffusion technique [4] : SIDE allows the intensity value within a region to change in response to forces exerted by adjoining regions; these forces are a function of the difference in region intensity values:

$$u_i = \frac{1}{m_i} \sum_{j \in A_i} F(u_j - u_i) p_{ij}$$

Equation 2: SIDE algorithm equation

where u_i is the intensity value within the ith region, mi is the area of this region (i.e. the number of pixels in it), the sum is over all the neighbors of the ith region, and p_{ij} refers to the length of the boundary between region i and its neighbor. The force function is monotonically decreasing and discontinuous at the origin, for example, F(v) . 1 2 (v/L) for v . 0, F(v) . 2 1 2 (v/L) for v, 0, where L is a constant related to the dynamic range of the image. Two adjoining regions are merged when the difference in their intensities reaches zero (in practice, when the difference changes sign or falls below a small threshold). The algorithm is initialized by taking each pixel to be a separate region. Like all diffusions, the SIDE algorithm will ultimately abolish all region structure, merging all regions into a single region whose intensity is equal to the mean value for the image. Therefore, the number of regions to be produced by the segmentation must be specified as a parameter. This is supposed to be the most effective and accurate algorithm, but, to be honest, it has been proved unsuccessfully. Since the challenge of this project does not depend on a very accurate process of segmentation, and Fuzzy C-means method showed results good enough, this last has been the one used.

- C-means Clustering: This is the segmentation algorithm selected for this work. It is discussed in detail in the section Methods.

15

In addition, there are some other effective algorithms reported in literature that we won't describe. Anyway it is worthy to mention them:

- o A segmentation technique based on a Markov random Field (MRF)
- o Robust Snakes [20].
- o Level Sets [21].

With Feature Extraction we mean the capability of detecting some kind of feature from an image given. In our case, the feature to extract is the pigment network. Few papers have been reported in the matter.

We mainly highlight:

- Fleming et al. [2] report the detailed techniques for extracting and measuring the pigment network characteristics. 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 near the network periphery. They use morphological techniques in their method.

- Anantha et al. [13] 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. [14] 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 graphically, but appear to achieve a sensitivity of 50% with a specificity of 100%.

- Maryam Sadeghi et al. [11] describe a novel approach based on the fact that the edges of pigment network structures form cyclic graphs which can be automatically detected and analyzed. They first perform image enhancement and edge detection. The resulting binary edge image is converted to a graph and the defined feature patterns are extracted by finding cyclic subgraphs. Another high-level graph is created from each correctly extracted subgraph, with a node corresponding to a hole in the pigment network. Nodes are connected by edges according to their distances. Finally the image is classified according to the density ratio of the graph. Their results over a set of 500 images from the Atlas of dermoscopy show an accuracy of 94.3% on classification of the images as pigment network Present or Absent.

This last is the newest and the most effective according to the rates of accuracy. We base our processing almost entirely on this paper. Anyway some tips have been taken from the other papers.

3.3.2. <u>State-of-art in Diagnosis Criteria</u>

Regarding the criteria we basically found 4 options [9]: Pattern analysis, ABCD rule of dermoscopy, Menzies' method, and the 7-point checklist. These are basically, let's say, different ways to look at Melanoma, that is, different features to base on. They are all referred within a standard procedure in Dermatoscopy called **Two-steps algorithm**.

Let's describe briefly their fundamentals although we already introduce Pattern Analysis is the most successful and the one we are going to base on. We present below the features which every criteria searches for.

- ABCD rule of Dermoscopy: [1]
  - Asymmetry
  - Border
  - Color
  - Dermoscopic structures
- Menzies' method [15]:
  - Symmetry of pattern
    - Symmetrical pattern
    - Asymmetrical pattern
  - Color of the lesion
    - Single color
    - More than 1 color

- 7-point checklist [16]:
  - Atypical pigment network
  - Blue-whitish veil Atypical vascular pattern
  - Irregular streaks
  - Irregular pigmentation
  - Irregular dots/globules
  - Regression structures

- Pattern Analysis [8,9]: It's our reference. It firstly looks for a termed Global Pattern, which is the pattern observed viewing the lesion as a congregate. Then it looks for the called Local features, which are the more specific details, as it could be dots, veils or streaks. We list them below, their meaning can be checked out in Table 4.

- Global pattern
    - Reticular
    - Globula
    - Cobblestone
    - Homogeneous
    - Starburst
    - Parallel
    - Multicomponent
    - Nonspecific
- Pigment network
    - Absent
    - Typical
    - Atypical
- Dots/globules
    - Absent
        - Regular
        - Irregular
- Streaks
    - Absent
    - Regular
    - Irregular
- Blue-whitish veil
- Blotches
    - Absent
    - Regular
    - Irregular
- Hypopigmentation
- Regression structures
- Vascular structures

Two Steps algorithm, A criteria for the computer-aided diagnosis

In 2003 the International Dermoscopy Society embarked on creating a consensus document for the standardization and recommended criteria necessary to be able to effectively convey dermoscopic findings to consulting physicians and colleagues, the Consensus Net Meeting on Dermoscopy [8]. Dermoscopy Report Steering Committee created an extensive list of dermoscopic criteria obtained from an exhaustive search of the literature. This criteria was summarized into the 10-point dermoscopy report, which was culminated in 2006. One of the base-of-the-pillar points was the defined two-step method which is a method for the initial evaluation of pigmented Lesions. It helps separate lesions into one of two categories, melanocytic or non-melanocytic, based on structures seen under dermoscopy. As you can easily find out, it consists of two steps: The First discriminates between melanocytic and nonmelanocytic lesions. The melanocytic lesions are further evaluated with the Second step, which discriminates between benign melanocytic lesions and melanomas.
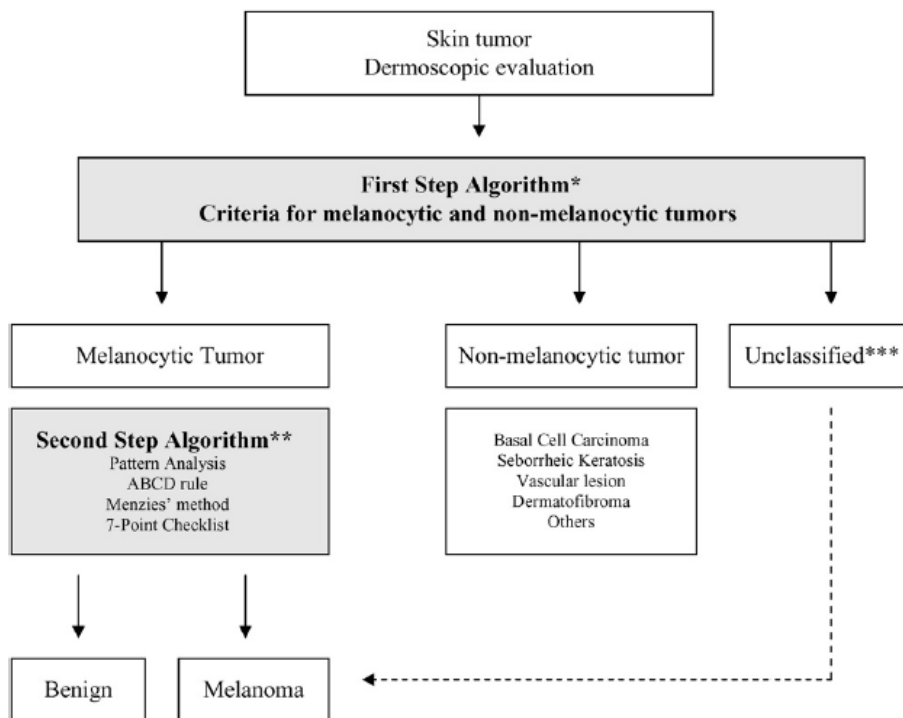

Figure 3: The two-step method as defined by the Consensus Net Meeting on dermoscopy. *, First step: discrimination between melanocytic and nonmelanocytic lesions. **, Second step: discrimination between benign melanocytic lesions and melanomas. ***, In the case of nonspecific criteria or indeterminate pattern, possibility of melanoma exists and should be in differential diagnosis

First Step

It is briefly shown and in an illustrative way how the first step works. In Figure 3 [9] it is noticeable how, according the kind of Dermoscopic criterion seen through Dermoscopy, the lesion is firstly classified into either Melanocytic or Non-melanocytic (Diagnostic significance). We don't go much into the physiological details since it is not the goal of this report. In Table III [9] are shown the features together with the Diagnostic significance. Seborrheic keratosis, Basal cell carcinoma and Vascular lesion are understood as Non-melanocytic.

Dermoscopic criterion	Definition	Diagnostic significance
Pigment network—pseudonetwork*	Network of brownish interconnected lines overlying background of tan diffuse pigmentation. In facial skin a peculiar pigment network, also called pseudonetwork, is typified by round, equally sized network holes corresponding to preexisting follicular ostia	Melanocytic lesion
Aggregated globules	Numerous, variously sized, often clustered, round to oval structures with various shades of brown and gray-black. They should be differentiated from multiple blue-gray globules.	Melanocytic lesion
Streaks	These have been previously described separately as pseudopods and radial streaming, but are now combined into one term. They are bulbous and often kinked or finger-like projections seen at the edge of a lesion. They may arise from network structures but more commonly do not. They range in color from tan to black.	Melanocytic lesion
Homogeneous blue pigmentation[†]	Structureless blue pigmentation in absence of pigment network or other discernable structures	Melanocytic lesion
Parallel pattern	Seen in melanocytic lesions of palms/soles and mucosal areas. On palms/soles pigmentation may follow sulci or cristae (ie, furrows or ridges) of the dermatoglyphics. Rarely arranged at right angles to these structures.	Melanocytic lesion
Multiple milia-like cysts	Numerous, variously sized, white or yellowish, roundish structures	Seborrheic keratosis
Comedo-like openings	Brown-yellowish to brown-black, round to oval, sharply circumscribed keratotic plugs in the ostia of hair follicles. Irregularly shaped comedo-like openings are also called irregular crypts.	Seborrheic keratosis
Light brown fingerprint-like structures	Light brown, delicate, network-like structures with fingerprint pattern	Seborrheic keratosis
Cerebriform pattern	Dark brown furrows between ridges producing brain-like appearance	Seborrheic keratosis
Arborizing vessels	Tree-like branching telangiectases	Basal cell carcinoma[‡]
Leaf-like structures	Brown to gray/blue discrete bulbous structures forming leaf-like patterns. They are discrete pigmented nests (islands) never arising from pigment network and usually not arising from adjacent confluent pigmented areas.	Basal cell carcinoma[‡]
Large blue-gray ovoid nests	Well-circumscribed, confluent or near confluent pigmented ovoid or elongated areas, larger than globules, and not intimately connected to pigmented tumor body	Basal cell carcinoma[‡]
Multiple blue-gray globules	Multiple globules (not dots) that should be differentiated from multiple blue-gray dots (melanophages).	Basal cell carcinoma[‡]
Spoke-wheel areas	Well-circumscribed radial projections, usually tan but sometimes blue or gray, meeting at often darker (dark brown, black, or blue) central axis	Basal cell carcinoma[‡]
Ulceration[§]	Absence of epidermis often associated with congealed blood, not due to well-described recent history of trauma.	Basal cell carcinoma[‡]
Red-blue lacunae	More or less sharply demarcated, roundish or oval areas with reddish, red-bluish, or dark-red to black	Vascular lesion
Red-bluish to reddish-black homogeneous areas	Structureless homogeneous red-bluish to red-black areas	Vascular lesion
None of listed criteria	Absence of above-mentioned criteria	Melanocytic lesion

Table 3: Features for the first step classification in Two-Steps algorithm

Second Step

As said, Melanocytic lesions are further assessed into the search of Melanoma. The Consensus has evaluated four algorithms for the second step: Pattern analysis, ABCD rule of dermoscopy, Menzies' method, and the 7-point checklist. According to the results, the sensitivity for melanoma of these 4 methods was similar, but **pattern analysis demonstrated a better specificity** as compared with the others and hence the best diagnostic performance. The algorithm consists of the evaluation of a sort of features and patterns contained in the lesion. These are listed below.

Dermoscopic criterion	Definition	Diagnostic significance
Global features		
Reticular pattern	Pigment network covering most parts of the lesion	Melanocytic nevus
Globular pattern	Numerous, variously sized, round to oval structures with various shades of brown and gray-black	Melanocytic nevus
Cobblestone pattern	Large, closely aggregated, somehow angulated globule-like structures resembling a cobblestone	Dermal nevus
Homogeneous pattern	Diffuse, brown, gray-blue to gray-black pigmentation in the absence of other distinctive local features	Melanocytic (blue) nevus
Starburst pattern	Pigmented streaks in a radial arrangement at edge of lesion	Spitz/Reed nevus
Parallel pattern	Pigmentation on palms/soles that follows sulci or cristae (furrows or ridges), occasionally arranged at right angles to these structures	Acral nevus/melanoma
Multicomponent pattern	Combination of ≥3 above-listed patterns	Melanoma
Nonspecific pattern	Pigmented lesion lacking above patterns	Possible melanoma
Local features		
Pigment network	Typical pigment network: light to dark brown network with small, uniformly spaced network holes and thin network lines distributed more or less regularly throughout lesion and usually thinning out at periphery.	Benign melanocytic lesion
	Atypical pigment network: black, brown, or gray network with irregular holes and thick lines	Melanoma
Dots/globules	Black, brown, round to oval, variously sized structures regularly or irregularly distributed within lesion	If regular, benign melanocytic lesion. If irregular, melanoma
Streaks (pseudopods and radial streaming)	These have been previously described separately as pseudopods and radial streaming. Streaks are bulbous and often kinked or finger-like projections seen at the edge of a lesion. They may arise from network structures but more commonly do not. They range in color from tan to black.	If regular, benign melanocytic lesion (Spitz/Reed nevus). If irregular, melanoma
Blue-whitish veil	Irregular, structureless area of confluent blue pigmentation with an overlying white "ground-glass" film. Pigmentation cannot occupy entire lesion and usually corresponds to a clinically elevated part of the lesion	Melanoma
Regression structures	White scar-like depigmentation and/or blue pepper-like granules usually corresponding to a clinically flat part of the lesion	Melanoma
Hypopigmented areas (structureless/ homogeneous)	Focal areas devoid of structures with less pigmentation than overall pigmentation of lesion and comprising at least 10% of total area	Nonspecific
Blotches	Black, dark brown, and/or gray structureless areas with symmetric or asymmetric distribution within lesion	If symmetric, benign melanocytic lesion. If asymmetric, melanoma
Vascular structures	Comma-like vessels	Dermal nevus
	Hairpin vessels	If uniformly distributed, seborrheic keratosis. If irregularly distributed, consider melanoma
	Dotted vessels	Melanoma
	Linear-irregular vessels	Melanoma
	Vessels and/or erythema within regression structures	Melanoma

Table 4: Patterns and Features of the Pattern Analysis criteria in the Two-steps algorithm. These are included together with its definition and diagnosis.

Note the Multicomponent pattern whereas the presence of local features such Atypical pigment network, Blue-whitish veil, Regression structures, Irregular Blotches, Irregular Streaks and some vascular structures are highly related to Melanoma. These statements have been faithfully proved through Association studies between hystopathologic diagnosis and Pattern Analysis diagnosis (carried out by the same Consensus).

The rates of the survey were shown above in Clinical Situation. We observed high values of Odds ratios in the mentioned features. Thus, we can truly believe that, applying Pattern Analysis and finding these features, we are probably dealing with a Melanoma.

We are going to focus on Pigment Network feature. Note a great value of percent, 0.77, is associated with the presence of Melanoma. This renders Atypical pigment network such a significant characteristic in the Pattern Analysis Diagnosis, and therefore must be accurately detected.

3.4. Goal of the Thesis

The clinical aim of this thesis is to provide, as far as it could be, a reliable and accurate tool to the family doctor that allows him a fast way of early diagnosis of malign lesions without a high expertise and deep knowledge of the dermatologic field, giving a good estimation of risk and reducing the high amount of derivations to the specialist doctor, that is the dermatologist. The tool to develop would contribute the computer-aided diagnosis by recognizing one of the main features characterizing Melanoma, the pigment network.

This tool should be as automatized and unsupervised as possible and will be implemented in Matlab©. The specific challenge is being capable to notice the presence of a pigment network, and if so, distinguish between typical or atypical. Many authors have tried to deal with the detection of this relevant feature, but no one yet has differentiated the both of types. This thesis then searches for a novel approach able to clear up the presence of an atypical pigment network from a given Pigmented Skin Lesion.

4. <u>Materials: Experimental Protocol.</u>

Pigmented Skin Lesions included in this thesis were obtained from:

- Dermatology section, Hospital Universitario Virgen del Rocío, Sevilla (Spain). All dermoscopic images were taken using **device for Dermoscopy**, which is composed by a **Contact oil immersion non-polarised Dermaphot Heine Delta 20**, and a digital camera with 5-megapixel Canon 400 SD, CCD sensor. The digitized image provided has a dimension of 576x767 pixels, RGB [red, green, blue] color space, and Joint Photographic Expert Group (JPEG) extension. All the images that this work is supposed to work on will come from this source. Ecography Gel was spread over the area before taking the picture.

Once the picture has been already acquired, it must be examined by means of any kind of image analyzer for the computer-aided diagnosis. In this thesis **Image Processing Toolbox from Matlab** has been used. This processing in any case must be based on some reliable criteria for the diagnosis of pigmented skin lesions. After a deep search in literature we did conclude the Two-steps algorithm developed by [8] is the most suitable for our purpose indeed. The recognition of patterns and local features, proposed as the second step (pattern analysis), has been actually the approach followed for the programming. That is, the code has been designed searching for the patterns and features described in Pattern analysis.

4.1. Dermoscopy

We have been referencing constantly Dermoscopy. The images obtained from this technique will be processed and the tool for the diagnosis, but, how do we take these pictures? Why Dermoscopy? What does that even mean? Let's go into details.

Dermoscopy, or also called epiluminescence microscopy or dermatoscopy [22], is a noninvasive diagnostic technique that uses optic magnification to permit the visualization of morphologic features that are not visible to the naked eye, allowing greater visualisation of structures in the skin up to the level of the papillary dermis. Structures in the reticular dermis cannot be seen. These features include the pigment network, brown globules, black dots, diffuse pigmentation, depigmentation, and blue-gray veil. These features correspond to histologic structures located in the upper dermis.

The rough surface scatters light reducing detail seen of the underlying structures. Dermoscopy overcomes this surface light scatter by either using contact with an interface medium or by using cross-polarised light.

Thus, it provides a more detailed inspection of the surface of pigmented skin lesions (PSL). These features include specific patterns, colors, and intensities of pigmentation, as well as the configuration, regularity, and other characteristics of both the margin and the surface of PSL.

The technique consists in placing mineral oil, alcohol or even water on the skin lesion that is subsequently inspected using a hand-held lens, a hand-held scope (also called dermatoscope), a stereomicroscope, a camera, or a digital imaging system. Both types of hand held dermatoscopes can usually be connected to digital cameras, either directly, or via adapter rings. Cameras and dermatoscopes vary and may not be compatible, therefore minding the manufacturer of the device is important for the camera selection.

Figure 4: Dermatoscope attached to a Digital Camera

The magnifications of these various instruments range from 6x to 40x and even up to 100x. The widely used dermatoscope has a 10-fold magnification permitting a sufficient assessment of pigmented skin lesions in daily routine. The fluid placed on the lesion eliminates surface reflection and renders the outer layer translucent, thus allowing a better visualization of pigmented structures within the epidermis, the dermoepidermal junction and the superficial dermis.

There are three main types of dermoscopy devices [17], contact oil immersion (non-polarised) devices; cross-polarised devices; and hybrid devices combining both polarised and non-polarised modes.

4.1.1. Contact oil immersion non-polarised devices

Oil immersion dermatoscopes (non-polarised) have a longer history of use in dermatology, however the inconvenience of applying an oil or interface fluid (alcohol gel or ultrasound gel) makes these instruments time consuming when multiple lesions are being examined. There is also the potential for cross contamination they involve contact with the patient. They have a bright image and may have a heavy handle although newer devices are more portable and compact.

These devices are the most established devices with a long history of use. An interface medium is required to overcome the surface light scatter, making these devices slow to use, messy and a potential for cross contamination. There are a variety of devices available, with variable illumination with either halogen or LED bulbs, optical lenses or power source, battery life / rechargability etc.

Advantages	Disdvantages
Established device	Interface medium required
Robust	Time consuming
White LED technology gives bright imaging	Naked-eye pre-selection of lesions
Excellent image quality	Cross contamination
	Messy

Table 5: Contact oil immersion non-polarised dermatoscopes Pros&Cons

Figure 5: Contact oil immersion non-polarised device

4.1.2. Polarised devices

Cross-polarised light dermatoscopes are becoming very popular with dermoscopy enthusiasts as they are now of a suitable quality that the device brightness is comparable with the oil immersion instruments. They are more expensive, however they are more versatile, smaller and lighter, and multiple lesions can be examined quickly without the need for interface fluid application. Newer devices have an extendable faceplate for use with interface fluid if needed; this can improve the image quality.

The development of polarised devices for dermoscopy at the start of the 21st century regenerated interest in dermoscopy. At last devices were able to overcome the surface light scatter without the need to apply copious amounts of interface fluid on the individual. This allowed many lesions to be examined quickly without the need for naked eye pre-selection. Since the earliest polarised devices were launched advances in imaging technology have now overcome the initial drawbacks making these devices now the standard for many dermatologists.

Advantages	Disdvantages
Portable	Pure polarised device
Quick to examine lesions	Less robust
No interface medium required	Expensive
No naked eye pre-selection	
Larger field of view	
Bright illumination	

Table 6: Polarised dermatoscopes Pros&Cons

Figure 6: Polarised device

4.1.3. Hybrid devices

Recently hybrid devices have been developed to combine the clarity of the contact oil immersion device with the practicality of the polarised devices.

Hybrid devices can combine both polarised and non-polarised modes for imaging. The first hybrid devices combined LED technology for a contact non-polarised mode as well as LEDs for polarised imaging. The launch of the Hybrid devices has produced a truly versatile dermoscopy device with excellent imaging quality in both modes, comparible to the best non-polarised device and the best polarised device.

Advantages	Disdvantages
Portable	Larger than earlier polarised devices
Quick to examine lesions	Less robust
No interface medium required	Expensive
No naked eye pre-selection	
Largest field of view	
Brightest illumination	

Table 7: Hybrid Dermatoscopes Pros&Cons

Figure 7: Hybrid device

The use of dermoscopy has uncovered a new and fascinating morphological dimension of pigmented skin lesions thus increasing the effectiveness of clinical diagnostic tools to differentiate melanoma from other pigmented skin lesions.

5. <u>Methods</u>

So far, it has been described what are we investigating and why, and we said we want to detect automatically the presence of a pigment network because it is associated with the presence of Melanoma. Automatically because we are working on the line of computer-aided diagnosis. Then we presented the materials, and we said the diagnosis will be based on images acquired with a dermatological practice called Dermoscopy and, these images will be processed in matlab for the feature extraction. What we haven't introduced yet is how we can detect these features with an image analyzer, and that's explained in the present chapter Methods.

The troubleshooting for this kind of surveys is highly extended. Any kind of image processing, whatever the field is, it is always faced in the same way. First you have to improve some image's properties from the acquisition, then you have to segment the image selecting what you are interested in, and finally you have to apply some mathematical techniques in order to get whatever you are looking for. And that's how this work is divided. We mainly differentiated 4 stages:

Pre-processing, Segmentation, and Feature Detection preceded by tape of Image Enhancement.

5.1. <u>Pre-processing</u>

Segmentation is a crucial stage since it is when the lesion is extracted from the health skin for a later study. We therefore must carry this stage out more or less accurately, it is important to have the lesion well delimited so that the study is done exclusively in the Region Of Interest (ROI). Segmentation will be done by some algorithm that will receive the image to segment as input. This image should be optimized according to the algorithm's working regarding some issues as it could be Artifacts Removal and Contrast Enhancing, and that is what we call Pre-processing stage, which is basically in charge of removing any interference and improving image's conditions.

These previous steps make the algorithm works much better. To demonstrate its relevance, Figures 8-9 show the outcomes of the segmentation algorithm with and without pre-

processing. It is noticeable how important is removing the artifacts and smoothing out the image.
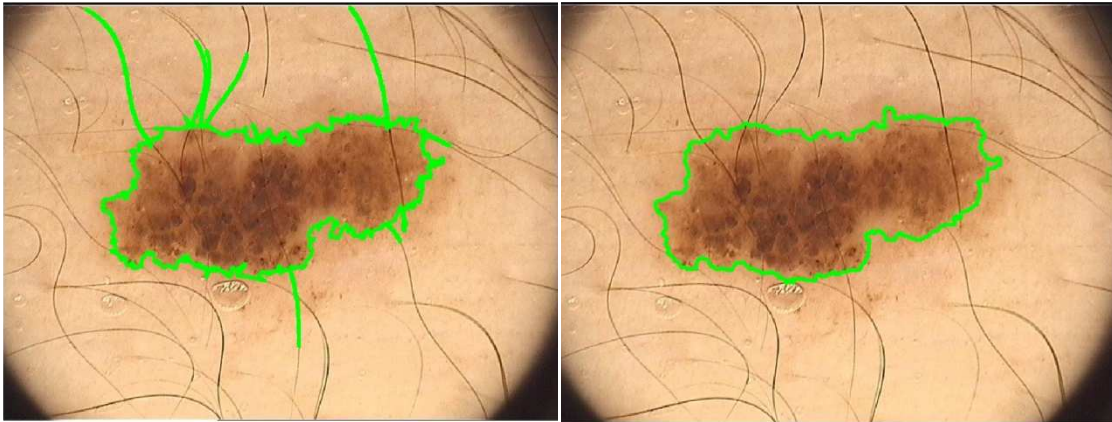
<div style="display:flex; justify-content:space-between;">
Figure 8: Segmentation without hair removal   Figure 9: Segmentation with hair removal
</div>

In Figure 8 hairs have been encircled as they belonged to the lesion, this is particularly undesired when dealing with pigment network detection since in a later stage we will detect lines to measure pigment network lines' thickness. The hairs will be then detected as lines and will be considered by mistake as network lines. Then, in Figure 9 you can notice the border is smoother, this is a proof contrast has been raised at the boundaries and therefore there is a much abrupt distinction between lesion and background that avoids the slithering throughout the contour when segmenting. The steps are described in detail below.

5.1.1. Artifacts Removal

In order to avoid the influence of artifacts in the segmentation, a clean-up capable to remove these artifacts must be done. [6] proposes a two-steps pre-processing, consisting of a filtering tape following by a morphological operator's tape.

Filtering is an important pre-processing step for the image segmentation when based on color information. Small structures should be removed in order to reduce the over-segmentation resulting from the color space clustering. A Median Filter has been used in this work. It consists of a non-linear filter which assigns to each pixel the median pixel-value over a neighborhood of a given size. This filter reduces indeed the influence of small

structures, like hair, on the segmentation result. The size of the filter depends on the image size and resolution.

Hair then can be removed using morphological closing (hair is dark, skin is light), the structuring element has to be a disk with diameter larger than the hair width.

The drawback is that these operations remove small texture structures at the lesion boundary that are part of the lesion and therefore must not be removed. In addition, it produces an effect of blurring throughout the entire image, making it fuzzy and consequently having a loss of accuracy especially at the edges when segmenting. To avoid this undesired effect as much as possible, the optimal parameters must be chosen. Optimal means in this context being big enough to produce a good effect of cleaning, at the same time of being little enough to produce a negligible blurring, that is a sharp image.

Figure 10: Blurred image due to an overfiltering

Improper sized filter spreads the contour out, causing the mentioned loss of precision. Sharp lines become smoothly spread areas.

According to the images provided, we noticed empirically that setting a 5x5 median filter and disk with diameter equal to 3 produce the best results.

Figure 11: Image with optimal artifacts removal. It is depicted next to the original image.

5.1.2. Contrast Enhancing

The second step consists of improving the color ratio between lesion and health skin. Since our goal is to extract the lesion apart, the further videointensity levels they have the better segmentation is performed. It has been used a histogram stretching.

Histogram stretching is a simple image enhancement technique that attempts to improve the contrast in an image by `stretching' the range of intensity values it contains to span a desired range of values, e.g. the the full range of pixel values that the image type concerned allows. It differs from the more sophisticated histogram equalization in that it can only apply a linear scaling function to the image pixel values. As a result the `enhancement' is less harsh [23,24].
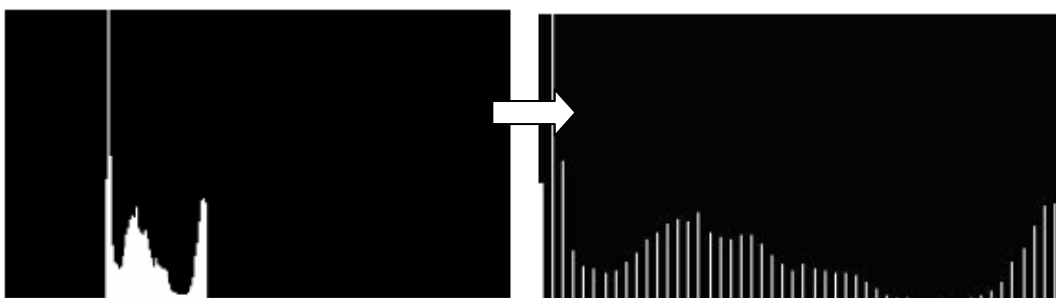
How It Works

Our RGB image is composed of 3 histograms, one per color plane. To obtain them we split out the RGB image to its 3 fundamental planes. Thus we have 3 gray images which histogram stretching can be performed to.

Before the stretching can be performed it is necessary to specify the upper and lower pixel value limits over which the image is to be normalized. Often these limits will just be the minimum and maximum pixel values that the image type concerned allows. For example for 8-bit graylevel images the lower and upper limits might be 0 and 255. Call the lower and the upper limits a and b respectively.

The simplest sort of normalization then scans the image to find the lowest and highest pixel values currently present in the image. Call these c and d. Then each pixel P is scaled using the following function:

$$P_{out} = (P_{in} - c)\left(\frac{b - a}{d - c}\right) + a$$

Values below 0 are set to 0 and values about 255 are set to 255.

The problem with this is that a single outlying pixel with either a very high or very low value can severely affect the value of c or d and this could lead to very unrepresentative scaling. Therefore a more robust approach is to first take a histogram of the image, and then select c and d at, say, the 2th and 98th percentile in the histogram (that is, 2% of the pixel in the histogram will have values lower than c, and 2% of the pixels will have values higher than d). This prevents outliers affecting the scaling so much.

This operation has been implemented in matlab through the command imadjust. Once we have the 3 stretched images, we merge them to achieve the RGB image again. The resulting output is shown below in Figure 12.

Figure 12: Image with enhanced contrast. Depicted next to the image before the enhancement.

5.2. <u>Segmentation</u>

The separation of lesion from background is a critical early step in any analytic procedure. In our case, the skin lesion to be assessed must be separated from the healthy skin. Several techniques have been reported for the segmentation of pigmented lesions, in dermatoscopic or conventional images.

Hance et al. [3] compared six procedures for the segmentation of macro images, including three thresholding algorithms (adaptive thresholding, spherical coordinate transform/center 2D split, and principal component transform [PCT]/median cut); a clustering algorithm (fuzzy c-means); and two split and merge procedures.

Thresholding and clustering approaches suffer from an important theoretical disadvantage: they make no use of spatial information at all. Pixels are classified entirely on the basis of their color values, without incorporating the constraint that lesional pixels should comprise a continuous and reasonably compact single region. It seems likely that an approach incorporating both chromatic and spatial information would produce better segmentations, and would follow more closely the human visual model.

As far as the sort of images in our possession let us test, we believe that Fuzzy C-Means clustering segmentation method is highly suitable for our further goal and covers all the needs released throughout the development of the project.

Anyway there are some negative points to keep in mind:
- As mentioned before, no spatial information is considered, it would be better having the main blotch and the rest of satellite ones circled together, and thus, the pigment network would be studied as a single aggregate.
- Adjacent clusters frequently overlap in color space, causing incorrect pixel classification.
- Clustering is more difficult when the number of clusters is unknown, as is typical for segmentation application. It requires prior knowledge of the number of regions existing in the image. Otherwise false results are returned.
- Exhaustive computational burden.

These weaknesses force us to say we don't discard there might be pictures in that our algorithm does not work 100% properly.

5.2.1. <u>Fuzzy C-Means Technique (FCM)</u>

FCM is an algorithm that searches for, given a number n of regions, n points that represent the centroids of the regions. Then, all the data will be clusterized according these centroids. In the case you are working with images, the centroids are pixel values of intensity and the clustering consists of matching every pixel to a centroid according to a minimum variance criterion. N groups will be formed. It is then required to know the number of classes (regions) in advance.

Furthermore it would be helpful to provide the algorithm data that somehow is optimal for the clustering. This is achieved by selecting the most proper color space.

Both discussions, the one related to the number of regions and the one related to the color space are presented after the technique's math description.

Next the algorithm has been successfully performed, there is a brief tape in which we have to decide what clusterized regions belong exclusively to the lesion and then the segmentation and after the mask will be correctly applied.

Fuzzy C-Means Math fundamentals

The FCM attempts to cluster the measurement vectors by finding local minima of the generalized within group sum of squared error objective function shown as follows [7, 18]:

$$J_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (U_{ik})^m \|x_k - v_i\|_A^2, \quad 1 \leqq m < \infty$$

Equation 4: FCM equation

where

- $x = \{x_1, x_2 \ldots x_n\}$ is a finite data set in Rd, if we apply this to image segmentation, our vector x_k are the values of the pixels of the image.

37

- $v=\{v_1, v_2 \ldots \ldots v_c\}$ is the set of class centers. $i=1,\ldots,c$ determines the indexes of class to match with each pixel once it has been already clustered, therefore c is the number of classes.

- $U=[U_{ik}]$ is a matrix cxn, where u_{ik} is the i-th value of membership of the k-th pixel x_k; these values satisfy the following conditions:

- $0<u_{ik}<1$ i=1,2,…..,c; k=1,2,…,n;

- $\sum U_{ik}$

- $0<\sum U_{ik}<n$, i=1,2,…, c.

- $m \in [1,\infty)$ is the weighting exponent, and defines the fuzziness of the membership values.

The Equation 4 is given by the sum of the Euclidean distances between the pixels x_k and the centers v_i of the clusters weighted by the corresponding values of fuzzy membership u_{ik}. The algorithm is iterative and uses these two recurrent equations, obtained from the minimization of J(U,v) and are given by:

$$U_{ik} = \left(1 + \sum_{\substack{j=1 \\ j \neq i}}^{c} \left(\frac{\| \mathbf{x}_k - \mathbf{v}_i \|}{\| \mathbf{x}_k - \mathbf{v}_j \|}\right)^{2/(m-1)}\right)^{-1}$$

Equation 5: Clusters' equation

$$v_i = \frac{\sum_{k=1}^{n} (U_{ik})^m \cdot \mathbf{x}_k}{\sum_{k=1}^{n} (U_{ik})^m},$$

Equation 6: Clusters' and Centers' equations

Where,

- v_i represents the center of the i-th cluster.

- U_{ik} represents the degree of belongingness of x_k to cluster i.

The centers v_i of every cluster are calculated considering all pixels x_k and weighting them by the corresponding value of membership u_{ik}. The degree of belongingness depends on the distance between the value of the pixel and the center of that cluster. The m factor reduces the influence of data with low value of membership: the greater m, the lower influence of data with low value of membership.

The algorithm FCM is composed by the following steps then:

- Initialization: it is stabilized the number c of clusters and the value of the m factor. The centers of the clusters are initialized $v(0)$. The counter a is set to 1.
- Membership matrix updating: Given the matrix of centers $v(a)$, the membership matrix U is calculated by the equation 5.
- End of iterations when $\max| u_{ik}(a) – u_{ik}(a-1)|<e$. Where e is the maximum variation accepted in U, otherwise:
  - The centers of the clusters are updated (Given $U(a)$, it is calculated the matrix of centers $v(a+1)$ by the equation 6)
  - The counter is incremented ($a=a+1$)
  - The loop is retaken by step 2.

Color Space

The color space of the image input to the Fuzzy C-means algorithm plays a crucial role in the efficiency of the method. Selecting a wrong color space would make the algorithm choose centroids that not represent the main regions containing the image. The algorithm must receive then an image with regions well-distinguishable for human perception regarding the contrast and texture. [6] states that the most appropriate color space transformation when using the proposed segmentation scheme is, as far as they could experience, the Karhunen-Loeve (KL) transform [25].

Karhunen-Loeve Transform (KLT):

A new color space is computed according to the direction of maximum variance. The KLT, also called Hotelling Transform or Principal Component Analysis (PCA), is the projection of the three components on the eigenvectors of their covariance matrix, computed using the following equations:

$$
\mathbf{m}_x = \frac{1}{M} \cdot \sum_{k=1}^{M} \mathbf{x}_k,
$$

$$
\mathbf{C}_x = \frac{1}{M} \cdot \sum_{k=1}^{M} \mathbf{x}_k \cdot \mathbf{x}_k^T - \mathbf{m}_x \cdot \mathbf{m}_x^T
$$

Equation 7: Mean and Covariance matrix

Where M is the number of RGB color vectors xk (or pixels in this case), and mx is the mean sample vector. Thus, defining a matrix A whose lines are composed of the eigenvectors placed in increasing order, the KLT is defined by:

$$
\mathbf{y} = \mathbf{A} \cdot (\mathbf{x} - \mathbf{m}_x)
$$

Equation 8: Karhunen-Loeve Transform equation

Figures 13 shows the two principal components obtained in decreasing order of variance. The third component is not included because of its small variance and therefore it is considered negligible and it is not performed anymore. The two principal components contain most of the structure information and may be used as a two dimensional

approximation of the RGB input image. Thus, combining this transform and the proposed segmentation scheme leads to the best results.
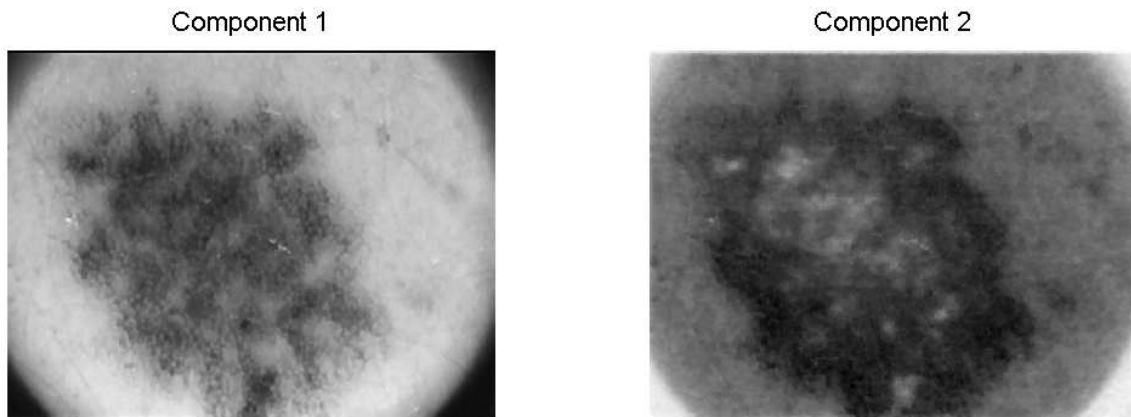
Figure 13: Principal Components Planes

Number of clusters

Once the two principal components of color spaces have been obtained, we are ready to perform the Fuzzy C-means Clustering algorithm. The goal of this method is, according to the color information, being able to classify the image into a number n of regions approximately equal in texture. Thus, a standard image would be divided into the main existing regions, which are skin, the one that represents the four dark corners surrounding the image, and up to three regions belonging the lesion.

Hence, firstly the number n of regions must be estimated. [7] proposes a histogram analysis so that the number of peaks would correspond to the number of regions. This number corresponds to the assessment of the zero-crossing of the second derivative (after a previous histogram smoothing and an obvious first derivative). This procedure has been tested unsuccessfully, so it drove us to try the second alternative, described below.

[18] proposes a fixed classification in five regions instead, although it is a crude approach and does not take into account the wide variety of possible images, that is, images more or less uniform and low number of regions(3, 4), or more complex images (regarding the texture) containing a higher number of regions(5, 6). Anyway, after proving it on an enough amount of lesions it has been demonstrated that n equal to five regions works properly for all of them.

Implementation

Fortunately, this algorithm has been already implemented in Matlab. It is done by the function fcm belonging to the Fuzzy Logic Toolbox. Its description is the following [26]:

[center, U, obj_fcn] = fcm(data, cluster_n, options)

applies the fuzzy c-means clustering method to a given data set. The input arguments are:

- data: data set to be clustered; each row is a sample data point
- cluster_n: number of clusters (greater than one)
- options, to control clustering parameters, introduce a stopping criteria:
  - options(1): exponent for the partition matrix U (default: 2.0)
  - options(2): maximum number of iterations (default: 100)
  - options(3): minimum amount of improvement (default: 1e-5)
  - options(4): info display during iteration (default: 1)

The output arguments of this function are

- center: matrix of final cluster centers where each row provides the center coordinates
- U: final fuzzy partition matrix (or membership function matrix)
- obj_fcn: values of the objective function during iterations

We used it with the following parameters:

[center,U,obj_fcn] = fcm(data,n,[2 40]);
Where,

Data is the matrix containing the two principal components obtained by the KL transform. Data=[pc1() pc2()];

N is the number of regions set to 4. Although [] set n to five, since our goal is just to segment the lesion, four seemed to be enough. Moreover, sometimes choosing five regions happens to be too discriminative, appearing small regions not desired to be distinguished. Do not forget our challenge is just to have the big main regions separated.

Yet it is not strictly right, the four regions may correspond as follows: one region for the corners, another for the skin, and other two for the lesion.

The m factor is set to 2 (as default) and the maximum number of iterations set to 40. This number has been set empirically, since a lower number does not cluster well enough and a higher number takes very long time.

As a result, four lesions clusterized are shown. It is observable how the algorithm is
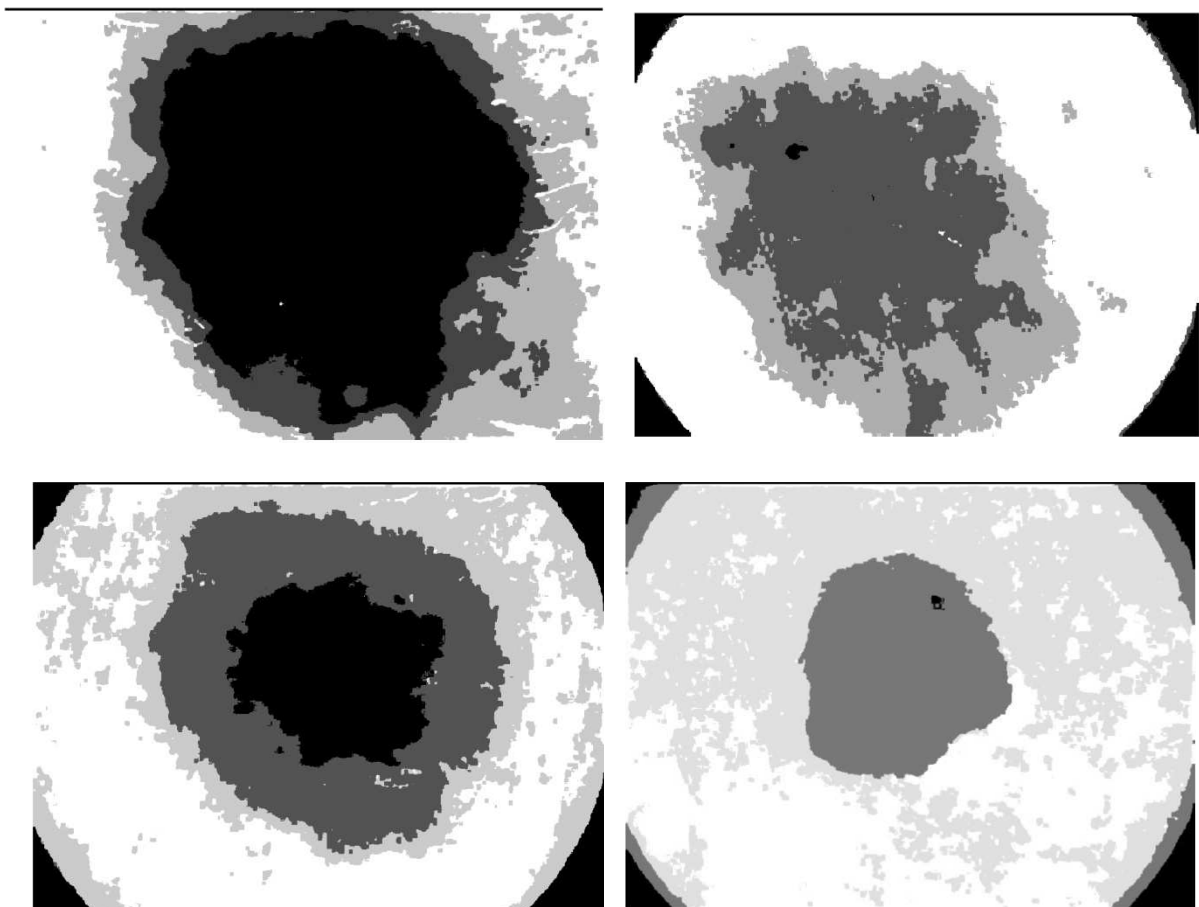
Figure 14: Clusterized images with Fuzzy C-means algorithm

capable to extract the lesion from the background pretty much accurately.

5.2.2. Region Assignment

The next logical step is to extract the binary mask. This is a binary image with only the pixels belonging to the lesion ON. This mask will multiply the original image so that only the target pixels will be kept, the rest of them will be set to zero. In order to do that, we first must to identify what clusters correspond to the lesion and what not. This means we have to assign the different levels of gray to the right regions.

For example, in the first clusterized image, we can truly say the darkest gray belongs mainly to the corners, the lightest gray belongs to the skin, and the two grays in the middle belong mainly to the lesion. So composing a binary mask with these two grays would segment the lesion correctly.
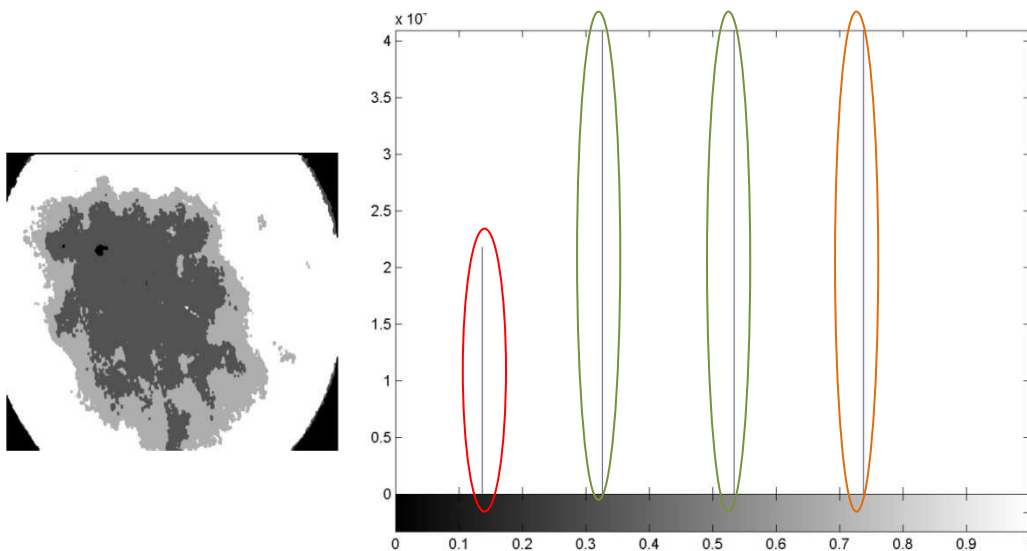
Figure 15: Histogram of example a. Red: Corners, Green: Lesion, Orange: Skin.

But it is not always this way. For example in the third image (Figure 14), we can see the darkest gray corresponds to the lesion and to the corners. So we cannot assign by color, because having 4 regions, it is likely to have the same color depicting both lesion and corners. Further spatial information is needed then.
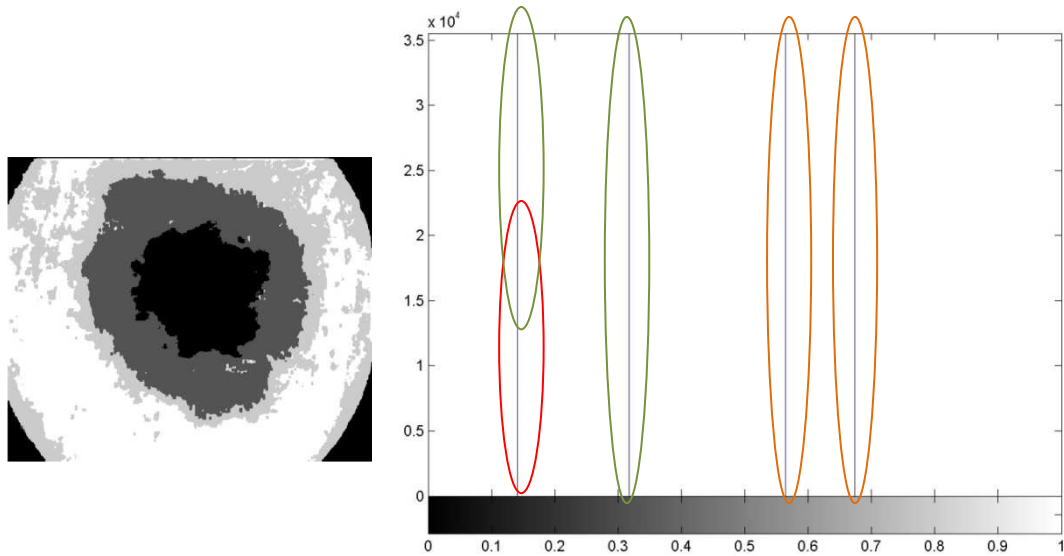
Figure 16: Histogram of example b. Red: Corners, Green: Lesion, Orange: Skin.

Then, in [18] is proposed to assign the regions according to the population of pixels. It is stated that a region containing more than the 15% of the pixels in the image, it is clearly assigned the health skin. Nevertheless we have images, as the fisrt shown above (figure 14), with a high percent of lesion covering the picture. So it is not fully correct assigning the regions by the number of pixels.

Summing up, we cannot assign the regions neither by the level of gray nor by the amount of pixels. The best solution found was turning the method supervised so that the user must to click on the different gray levels composing the lesion.
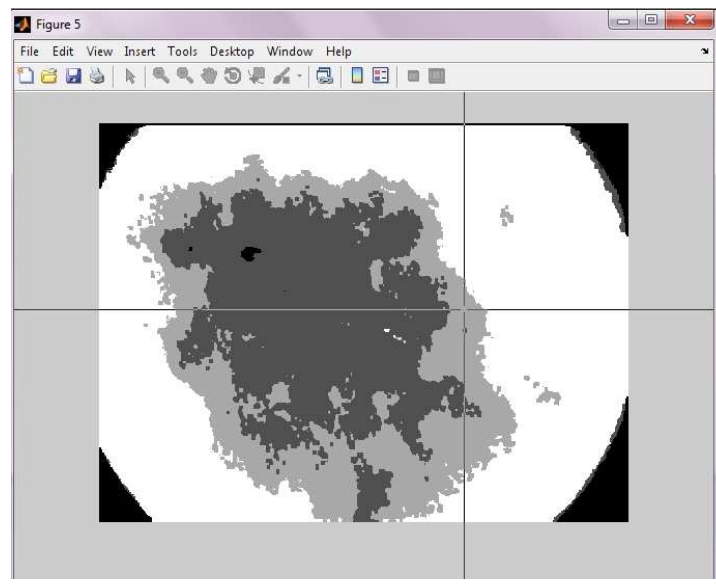
Figure 17: Selection of grays belonging to the lesion

Due to the fact there are images whose darkest gray is shared between the corners and some part of the lesion, we should introduce a kind of spatial filtering. The easiest way is to remove the corners so that they do not count for the mask. To do that, the centroid from each region is calculated and taken apart only if it is located out an imaginary frame bounding the picture.

This frame squares the center of the image 100 pixels far from the boundaries of the picture. It is depicted in Figure 18 in green.
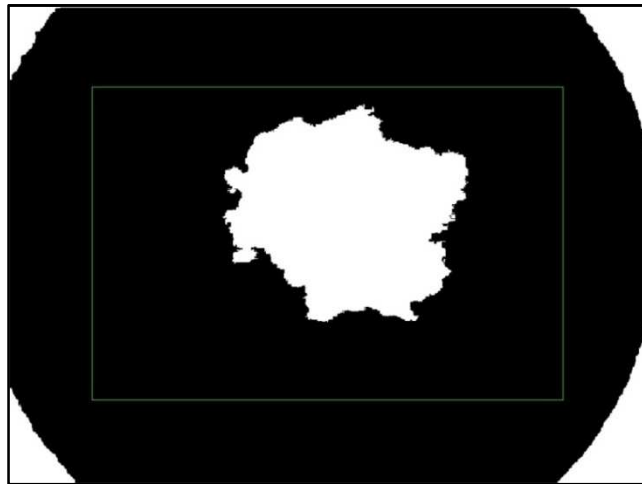
Figure 18: Lesion BW with corners

In the image above (Figure 18) it is noticeable the four regions belonging the corners will be removed since their centroids are outside the frame marked in green (Figure 19).
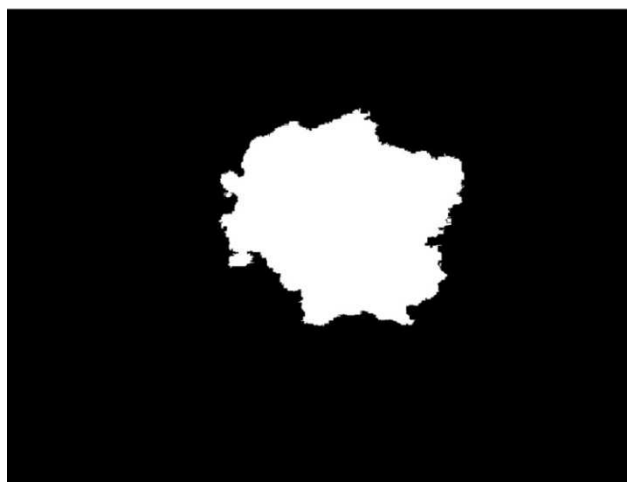
Figure 19: Lesion BW without corners

Once we have all the regions composing the lesion, we proceed cleaning up the binary images by morphological operators. More specifically we used the "fill" operator to remove the holes and "bwareaopen" which let us to remove objects smaller than 1000 pixels. Then, we can extract their contours through "bwboundaries" and plot them over the original image. The outcome will be the original image with the regions' boundaries highlighted in green as Figure 20.
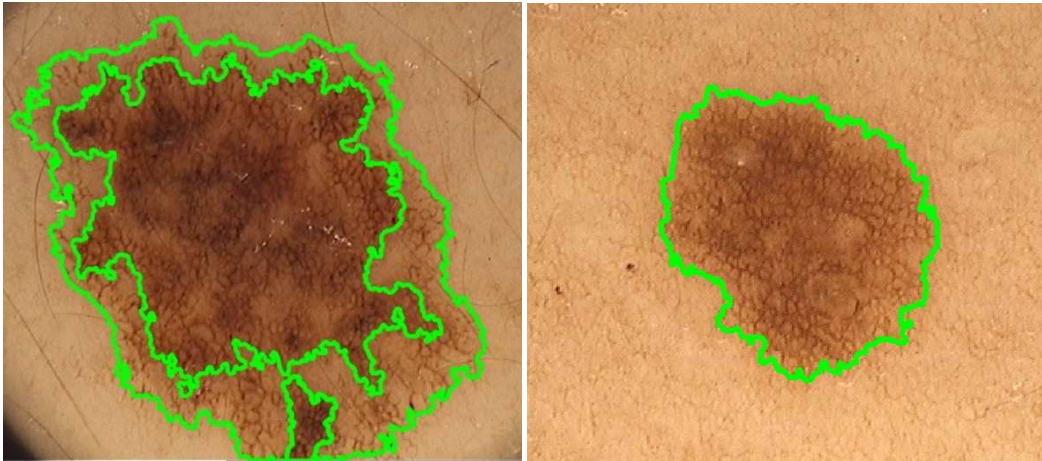
Figure 20: Segmented Images

Now we can state the lesion has been already segmented. Above it is observable how we have been able to separate the different regions of the lesion from the health skin. This means now we can apply the desired mask and therefore proceed with the study of the presence of the pigment network.

5.3. Image Enhancement

We have called Image Enhancement the stage prior to the Feature extraction. It consists of a couple of steps which improve some image's properties for facing the feature detection, the color plane selection and the pigment network's contrast enhancement.

5.3.1. Color Plane Selection

Once we have segmented the lesion, a mask will be formed from the most outer cluster surrounding the lesion. This mask is finally applied to a luminance plane of the original image achieving the searched portion of picture to be investigated.
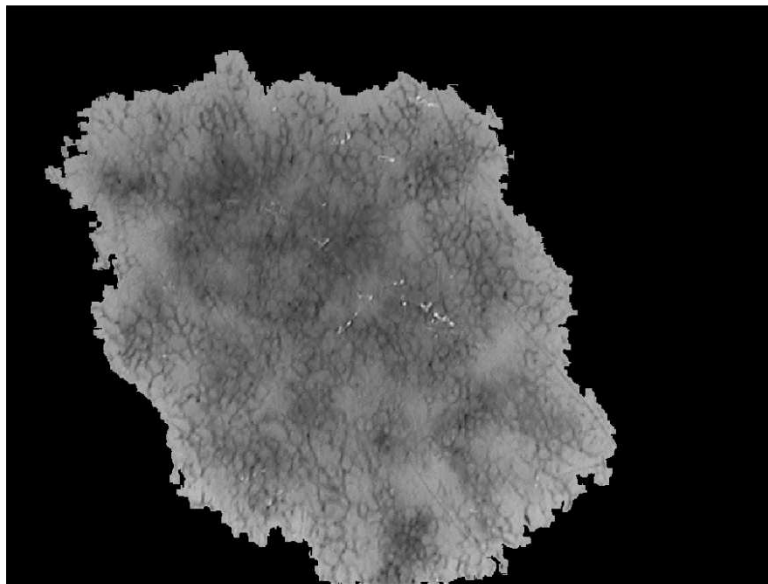
Figure 21: Image in color plane Luminance (Lab) with mask applied

The color plane to work on is a crucial decision. Every component of the color spaces Lab and HSV have been plotted through the function 'color_plane.m' [23]. The image selected contained a well-distinguished pigment network so that it was feasible to select the most proper color plane empirically. Both of them are shown below as illustration:
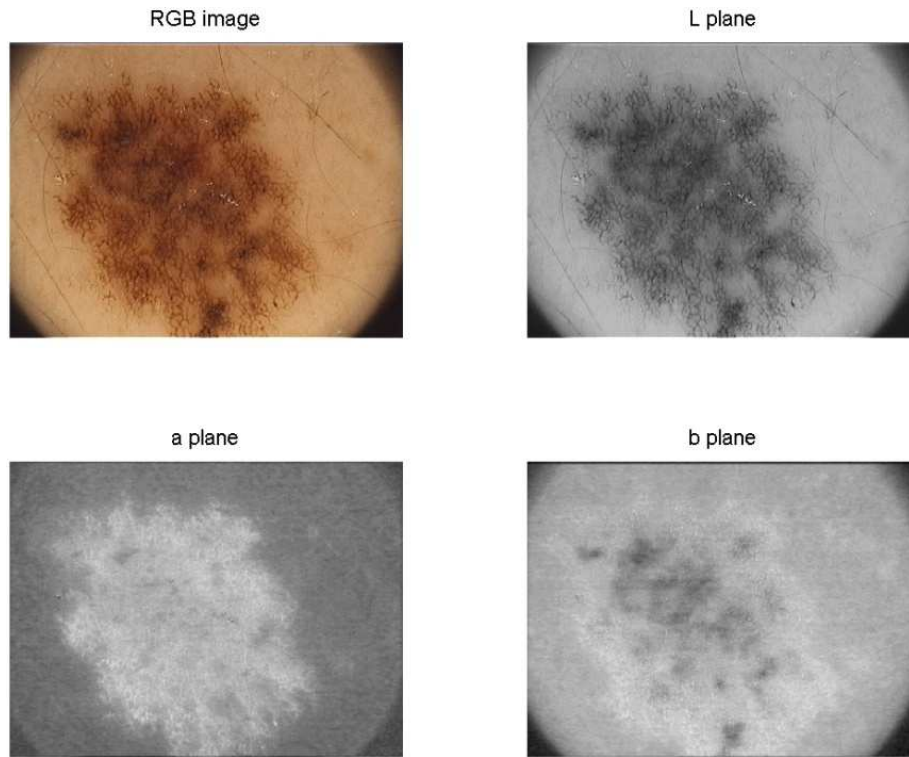
Figure 22: Transformation to Lab color space
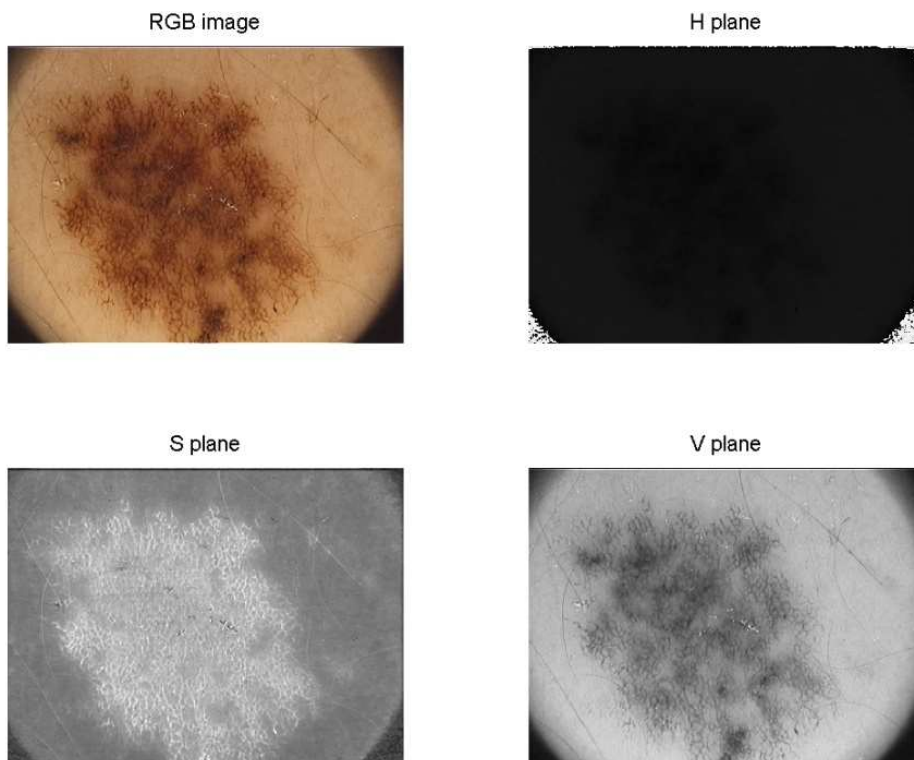
Figure 23: Transformation to HSV color space

Luminance plane from Lab color space seemed to be the most suitable due to the pigment network appears highly contrasted from the background.

5.3.2. Pigment Network's Contrast Enhancement

Once we have selected the most contrasted color plane, we do need to enhance this contrast even more yet. Keep in mind our goal is to evidence small hypopigmented holes within a hyperpigmented background. Sometimes these holes won't be as clear as we want, that is, will suffer from smooth color transitions or will be connected with other holes, etc…These not ideal situations must to be handled anyway. This means, until applying filters and algorithms for the holes identification, our previous concern must be to provide as input to these filters and algorithms an image with very high contrasted holes.

To do that, it has been used a grayscale morphological operator called 'imhmin'.

In an image, every small fluctuation in intensity represents a regional minimum or maximum. Regional minima are connected components of pixels with a constant intensity value, and whose external boundary pixels all have a higher value.

To remove the less significant minima but retain the significant minima we use imhmin(I,H,CONN) function. With this function, you can specify a contrast criteria or threshold level, H, that suppresses all minima in I whose depth is less than H [23,26].

```
a =

   10   10   10   10   10   10   10   10   10   10
   10    7    7    7   10   10   10   10   10   10
   10    7    7    7   10   10   10   10   10   10
   10    7    7    7   10   10   10   10   10   10
   10   10   10   10   10   10   10   10   10   10
   10   10   10   10   10    2    2    2   10   10
   10   10   10   10   10    2    2    2   10   10
   10   10   10   10   10    2    2    2   10   10
   10   10   10   10   10   10   10   10   10   10
   10   10   10   10   10   10   10   10   10   10
```

Suppress all minima below a specified value. Note how the region with pixel valued 7 disappears in the transformed image.

```
b = imhmin(a,4)

b =
```

10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	6	6	6	10	10
10	10	10	10	10	6	6	6	10	10
10	10	10	10	10	6	6	6	10	10
10	10	10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10

Note how the relative minimum 7 (under the threshold 4, since 10-7=3) has been removed replacing it for piece of background. With this tool, we can render more uniform the intensity of holes and lines. Isolated or darker pixels, responsible for the transitions and fuzzy holes, within a uniform hole (lighter pixels) will become part of the hole since the lines will become thicker and darker. This way we achieve a clearest and cleanest image prepared for the network identification.

A threshold of 24 has been set proving to be the most effective. The filter's effect is shown below:
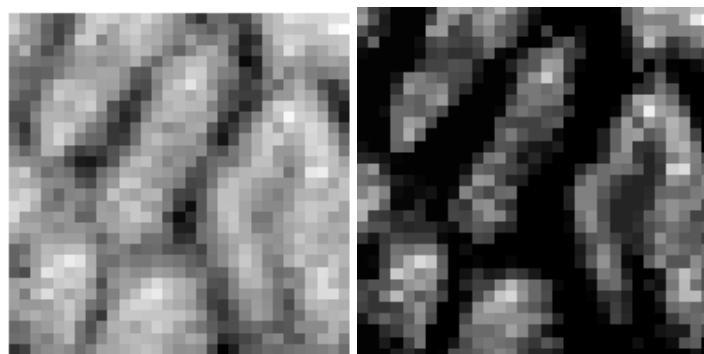
Figure 24: imhmin effect, holes' contrast enhancement

Note the effect of thickening and darkening of the lines, this emphasizes the holes as said.

5.4. Feature Extraction

This is the last main point of the chapter "Methods". Here in the first part, it is finally detected the presence of the searched pigment network. Next in the second part, we propose a novel method to distinguish the two types of pigment network, typical and atypical.

5.4.1. Detection of the presence of pigment network

A pigment network, either typical or atypical, is composed by a populated network of meshes. Both types share this feature, although have different shape and spatial arrangement. What we have to do to identify its presence is checking how populated the network is. This requires firstly identify the holes and then calculate a rate of density.

Holes identification

Now it is time to detect the hundreds-of-times mentioned holes. Once we ensure we put great efforts in emphasizing the holes and enhancing the contrast, we are ready to apply edge identification. [11] uses a Laplacian of Gaussian filter (LoG) for this purpose. Anyway it is not a surprising choice since this is widely used when dealing with contour's extraction.

Laplacian of Gaussian Filter

The filter's design departs from a Gaussian Lowpass filter with variable cut-off frequency. This variability is controlled by the adjustment of the standard deviation σ (inverse proportionality).

$$h(x, y) = \exp\left[-\frac{x^2 + y^2}{2\pi\sigma^2}\right]$$

$$H(\omega_x, \omega_y) = 2\pi^2\sigma^2 \exp\left[-\frac{\pi\sigma^2(\omega_x^2 + \omega_y^2)}{2}\right]$$

Equation 9: LoG equations in space and spatial frequency
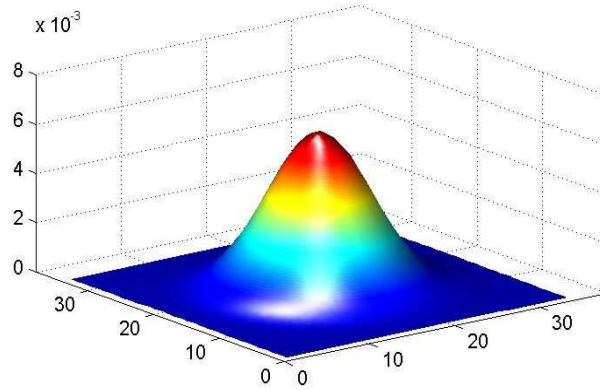
Figure 25: Gaussian Lowpass filter in 3D

The LoG filter [27] corresponds to the derivative of the Gaussian with respect to $2\sigma^2$

$$g(x,y) = \nabla^2\left(f(x,y) * h(x,y)\right)$$

$$g(x,y) = f(x,y) * \left[\nabla^2 h(x,y)\right]$$

$$\nabla^2 h(x,y) = \frac{x^2 + y^2 - 2\pi\sigma^2}{\pi^2\sigma^4} \exp\left[-\frac{x^2 + y^2}{2\sigma^2\pi}\right]$$

$$\Im\left[\nabla^2 h(x,y)\right] = 2\pi\sigma^2 \exp\left[-\pi\sigma 2 \frac{\omega_x^2 + \omega_y^2}{2}\right]\left(\omega_x^2 + \omega_y^2\right)$$

Equation 10: Laplacian of Gaussian Equation
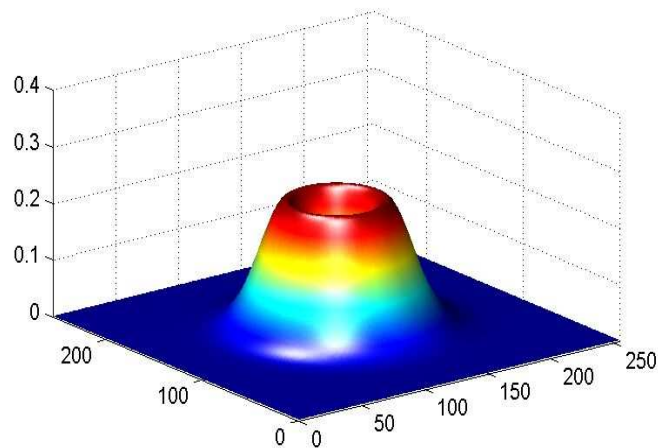
Figure 26: LoG filter in 3D

53

The application of this above transfer function detects the "lightdark-light" changes of the intensity. The detection criterion is the presence of a zero crossing in the second derivative with the corresponding large peak in the first derivative. LoG looks for zero crossings and their transposes. All zeros are kept and edges lie on the zero points. If there is no zero, it arbitrarily chooses the edge to be the negative point. Therefore when all zero responses of the filtered image are selected, the output image includes all closed contours of the zero crossing locations.

The LOG filtering step is restricted to the area within the lesion boundary, which was pre-determined thanks to the segmentation step.

We followed the MATLAB implementation of the LOG edge detection [26].

BW = edge(I,'log',thresh,sigma)

Edge ignores all edges that are not stronger than thresh (not include all the zero-crossing) and sigma is the standard deviation of the Gaussian filter (that is, the width). Both parameters have been optimized for the hole's encirclement empirically, resulting thresh=0.01 and sigma=1.8.

As a result, the LOG operation generates a closed contour surrounding the segmented lesion.
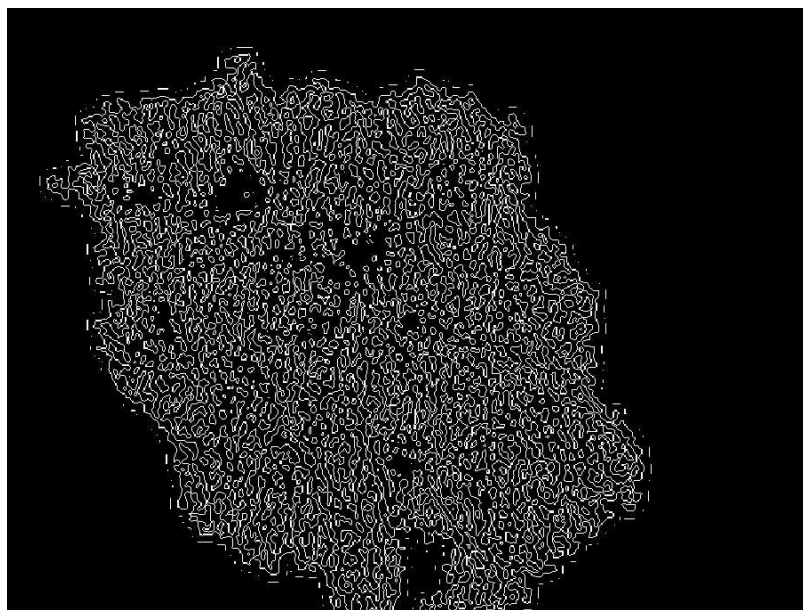
Figure 27: Lesion after LoG filtering

In order to clean a little bit of the binary image, morphological operators have been used. More specifically 'spur' operator to remove end points of lines and 'hbreak' operator to remove H-connected pixels, which will try to split some of the connected holes.

We following show a very explanatory image, where the original image and the holes' contour are plotted together, this last in RGB colors. It is highly noticeable how we successfully individuate most of the present holes composing the network.
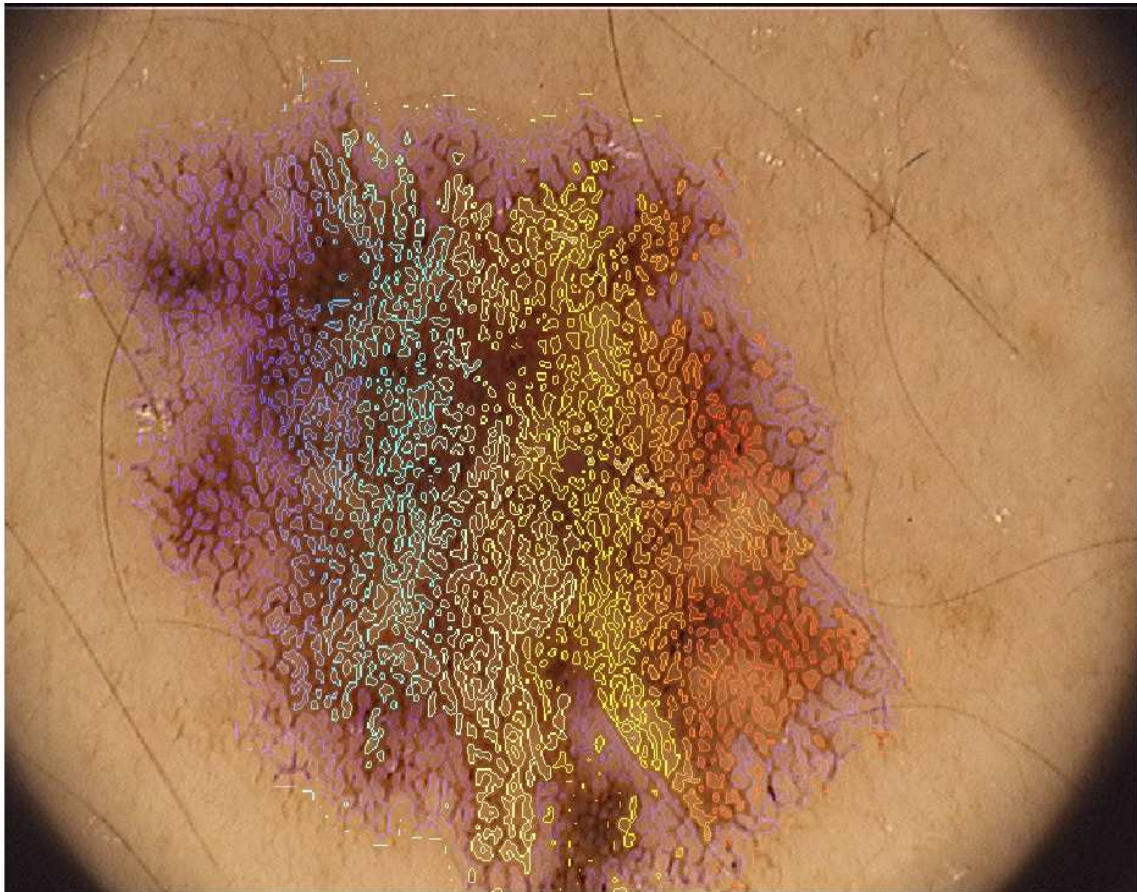
Figure 28: Contours from LoG filtering observed in the original RGB image.

Holes selection

The next logical step is selecting which are the correct encircled holes. Our filtering took them all even they do not correspond to the definition of hole. Besides, there are also contemplated lines and regions that do not form loop. It is obvious since the filter identifies all kinds of intensity variation. So we are in charge of deciding, to our knowledge, what a hole is and providing a touch of smartness to the code.

We based the selection on several statements:

- By definition a hole is small, although big enough to be a hole and not a dot. So we set a size threshold from 12 to 160 pixels. Every region smaller or bigger than the thresholds won't be taken into consideration.
- We do need to identify oil bubbles and white cysts and dots. This is an artifact due to the method of acquisition. Since we know these bubbles appear with very high levels of intensity, every region containing pixels with a greater value than 80% the maximum of intensity will be removed.
- In globules and brown dots, the mean intensity of the area inside the structure is lower than the intensity of border pixels. We need to discriminate these globules from holes of the pigment network. Therefore, we set a threshold for the difference between the average intensity of inner pixels and the average intensity of the border. For belonging to a pigment network, a detected region should have a higher intensity in the area contained by the network structures (holes) than in the network itself (lines).

The algorithm designed to carry out these statements has been the following:

We label the binary image resulting from the edge detection, this way we have numerated every encircled region. The size filtering just requires measuring the length of each region and deleting it if it is out of the bounds of the thresholds.

Since we have all the borders labeled, if we fill these regions we will obtain all the holes fulfilled. If we then extract to it the borders again, we will keep only with the inner pixels of the holes. After this simple procedure done using morphological operators, we have two labeled images. One image with the borders of the holes, and another one with the inner

pixels of the holes. Thus, we will able to detect oil bubbles and white cysts and dots (statement two) and to compare both averaged videointensities (statement three). Hence, every region containing either high intensity pixels or a mean intensity at the borders higher than at the interior will be discarded.
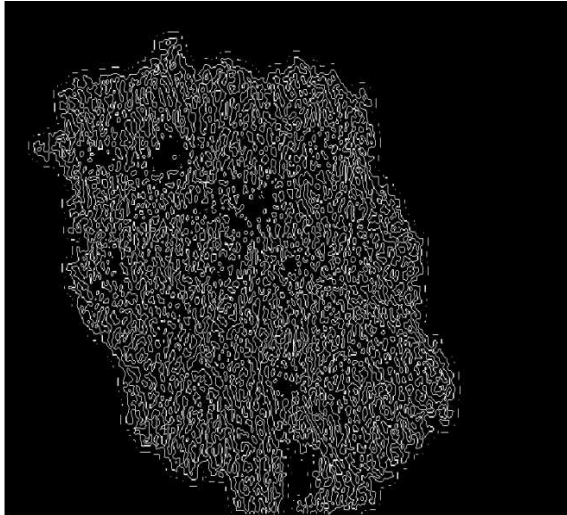
Figure 30: Labeled image of the holes' borders
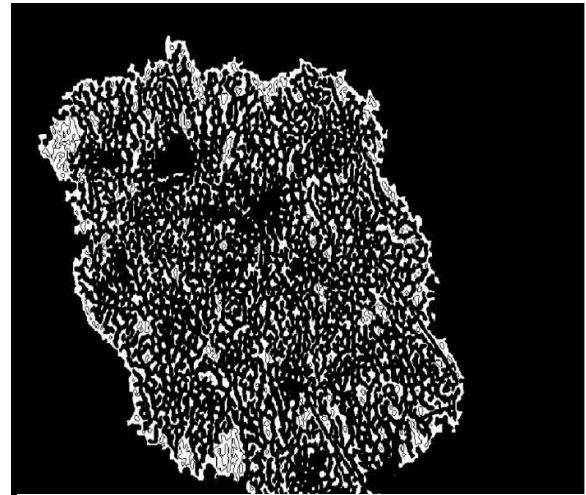
Figure 29: Labeled image of the holes' inner pixels

The resulting image after the holes' filtering remains only with the true candidates to holes in the pigment network. Note oil bubbles and white cysts are not encircled.

oil bubbles/white cysts excluded!!
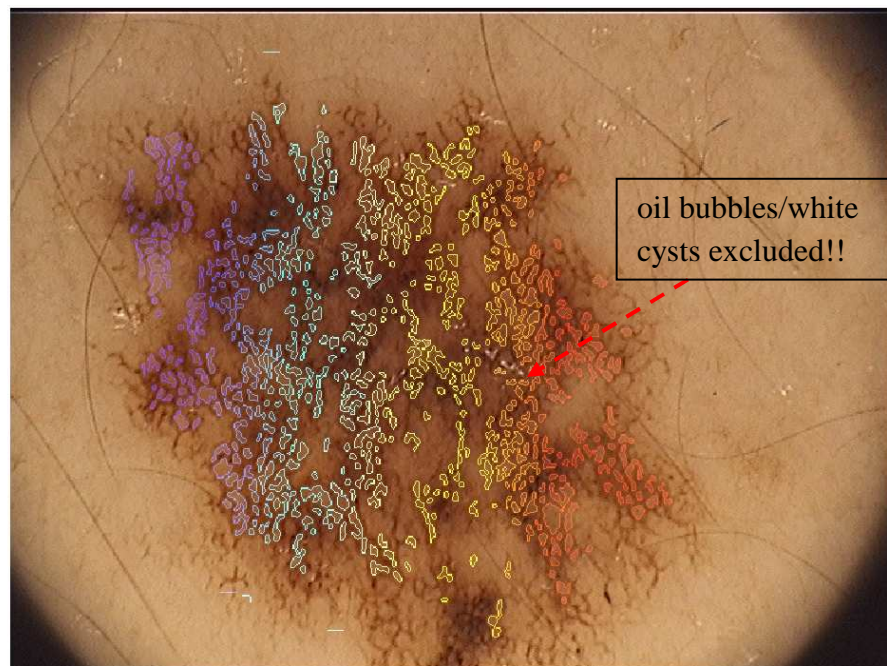
Figure 31: Lesion in RGB with target holes selected

Presence of the network

Once we have already identified the holes, we are prepared to detect the presence of a pigment network pattern. [11] proposes the creation of a graph from the image above (Figure 31). This is a graph whose nodes are the centers of the meshes belonging to the pigment network (colored holes).

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 meshes in the image. Based on the pigment network definition, meshes of the network are uniformly spaced. To consider this spatial arrangement, the MDT should be proportional to the size of holes and it is defined as alpha (set to 3) times the average diameter of meshes.

In order to calculate the average diameter of the meshes, it is used the option 'EquivDiameter' from the command 'regionprops'. It provides a scalar that specifies the diameter of a circle with the same area as the region. Computed as sqrt(4*Area/pi). The input is our more than once already used image with the holes labeled. Averaging the resulting vector we will obtain the searched diameter.

Then, the nodes of the graph will be the centroids of the meshes. We help ourselves out once again using the command 'regionprops'. This time by using the option 'Centroid'. Again the input is the labeled image. Next, the centers whose distance between them is lower than 3 times will be connected. This distance is estimated as the Euclidean Distance, both centers' coordinates are obtained, (x_1, y_1) and (x_2, y_2) respectively, and the Euclidean formula is computed:

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Equation 11: Euclidean distance

The resulting graph is the following below.
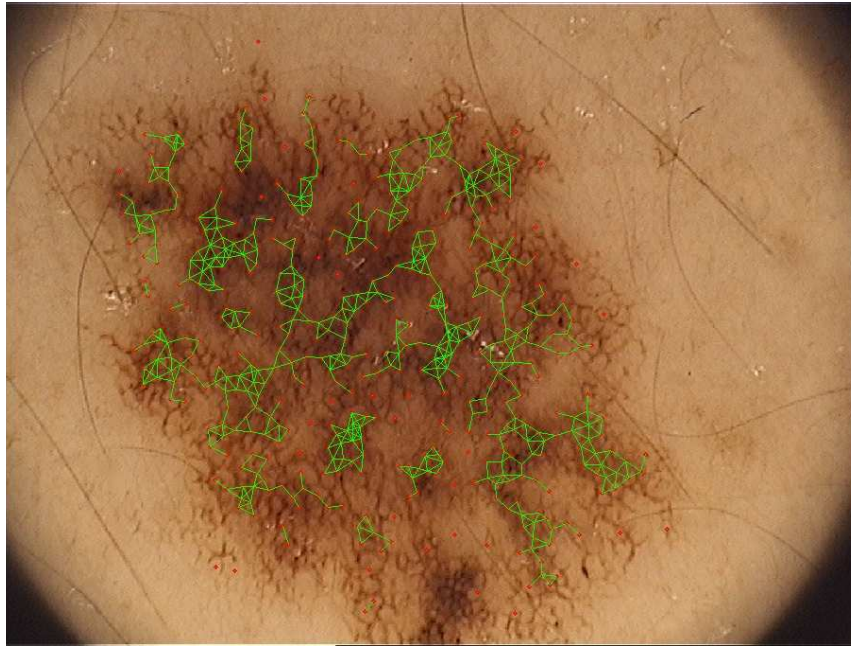
Figure 32: Graph from the network

Depending on the density of the graph, we may consider the presence of a pigment network, whatever the type it is. It is then assessed a density coefficient.

$$Density = \frac{|E|}{|V| \times \log(LesionSize)}$$

Equation 12: Density coefficient of the network

where E is the number of edges in the graph, 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. Being present is proportional to the number of edges since a bigger $|E|$ means that more holes are closer that MDT. Also, having a smaller $|V|$ for a fixed $|E|$ means that nodes or holes are uniformly spaced close to each other and graph of the pigment network is dense. Therefore, based on the pigment network definition, having a non null 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, [11] found experimentally that the logarithm of

Lesion-Size is more appropriate. Images containing a density ratio higher than a threshold (set to 0.2) are classified as Present and the rest as Absent.

5.4.2. <u>Pigment Network type identification</u>

In order to distinguish between a typical and an atypical pigment network, we do need to understand properly which the differences are. Let's remind them briefly:

- A typical network consist of light to dark brown network with small, relatively uniform-spaced network holes and homogeneous in color, and thin network lines distributed more or less regularly throughout the lesion, and usually thinning out at the periphery.

- An atypical network is non-uniform, black, brown, or gray network with darker and/or broadened lines and irregularly distributed holes that are heterogeneous in area and shape. The lines are thick and often hyperpigmented and may end abruptly at the periphery.

It is noticeable there are few features we must to investigate carefully:

- Spatial Uniformity of the meshes
- Color Uniformity of the meshes
- Thickness of the lines composing the network

Every feature will be characterized by a quantitative parameter. Thresholds will be set in each parameter so that the overcoming of it would mean the belongingness to a type of pigment network. If most of the features' characterization show a tendency to be a certain type then it will considered it belongs to the mentioned type.

Now we proceed to explain every feature's characterization for separate.

Spatial Uniformity of the meshes

The main difference at first sight is the spatial uniformity. The typical pigment network seems to show a uniform pattern regarding the meshes. It is necessary then to establish a condition of uniformity that allows us to carry this differentiation out. This condition will be a coefficient. Uniformity means the regular repetition of a determinate pattern. Therefore this coefficient must be calculated somehow by studying different parts of the lesion. We propose to divide the binary image containing the lesion into squared blocks. Every block will be composed (if any) by some meshes. If the lesion is spatially uniform regarding the holes, all the blocks will contain more or less the same amount of holes. We then create a variable that corresponds to the ratio between black pixels (holes) and white pixels (background). Taking all of them into consideration, we can average this quantity so that it sets the main pattern.

We propose the following method composed by several steps:

- The image to work on will be the binary one where the holes are in black among a white background.
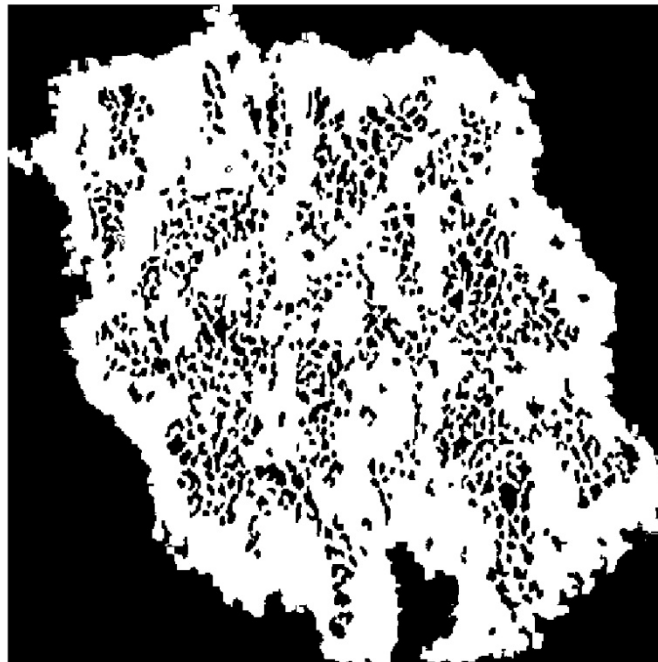
Figure 33: Holes in BW

- Next, this image is divided into blocks through a 'Split and Merge' algorithm: the quadtree decomposition. This method splits the whole image to squared blocks of different size.

Quadtree decomposition [23]

The process is the following:

1. Define a homogeneity measure, that is, determine a criterion the region must meet to pass the test. In this case, since we are dealing with a binary image, the criterion will be being black or white. Here we have to say that the image to split out CANNOT contain the holes, otherwise it would split the lesion out assuming holes as single regions, and our regions must contain groups of holes. However, then we will examine the blocks containing the holes.
2. Split the image into equal sized regions.
3. Calculate the homogeneity measure for each region.
4. If the homogeneity test is passed for a region, then a merge is attempted with its neighbor(s). If criterion is not met, the region is split.
5. Continue this process until all regions pass the test.

Summing up, it basically divides the image initially into 4 subimages (blocks): A condition will be applied to every block, if it is not verified; the block is further divided into 4 blocks, and so on.

Split and merge methods are referred to as state space techniques and use graph structures to represent the regions and their boundaries [23].
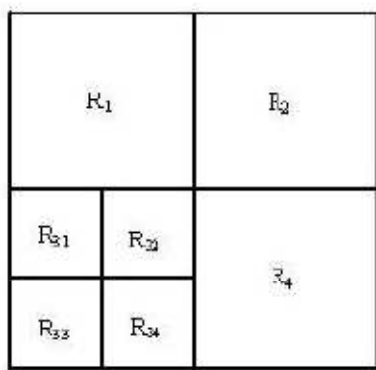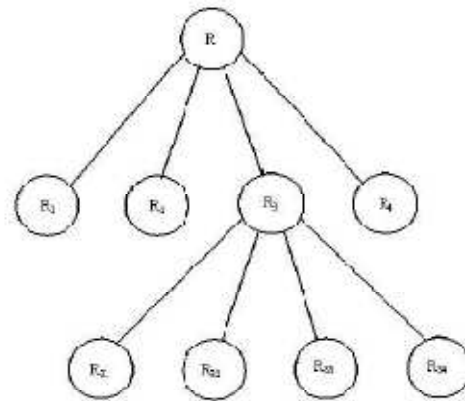
Figure 34: Quadtree decomposition schema

This operation is carried out by the command 'qtdecomp' [26]

QTDECOMP divides a square image into four equal-sized square blocks, and then tests each block to see if meets some criterion of homogeneity. If a block meets the criterion, it is not divided any further. If it does not meet the criterion, it is subdivided again into four blocks, and the test criterion is applied to those blocks. This process is repeated iteratively until each block meets the criterion. The result may have blocks of several different sizes. qtdecomp is appropriate primarily for square images whose dimensions are a power of 2, such as 128-by-128 or 512-by-512.

The image obtained after the acquisition process does not fit the power of 2 size. So it is required an adjustment to one of these sizes. The way to achieve it has been the following:

- It is first got an image which strictly fits the lesion, and that's achieved through the BoundingBox option from regionprops command. BoundingBox provides you the coordinates of the smallest rectangle containing the region (the lesion which is labeled to 1). These coordinates let us crop our target (the lesion again), and then resize the image to the closest power of 2 size to the original image.
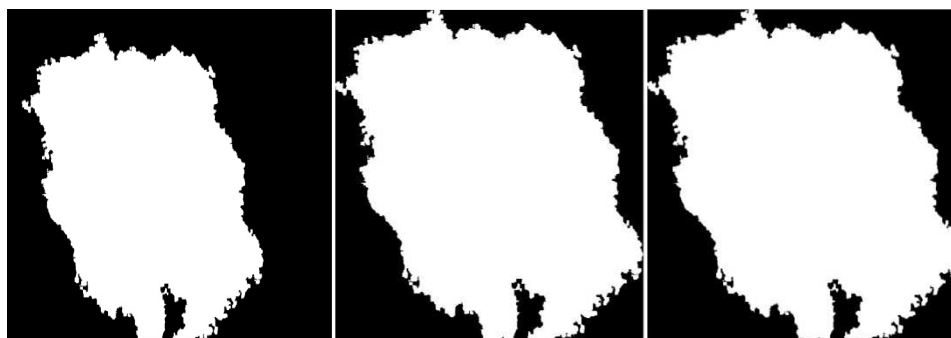
Figure 35: Resizing Process. A) original BW image 576x767. B) Bounding box fitting the lesion 531x568. C) Resized Lesion 512x512.

Once our image fits the required size for the quadtree decomposition, we are prepared to perform it over the image without holes. Then, the blocks found are translated into the image with the holes.
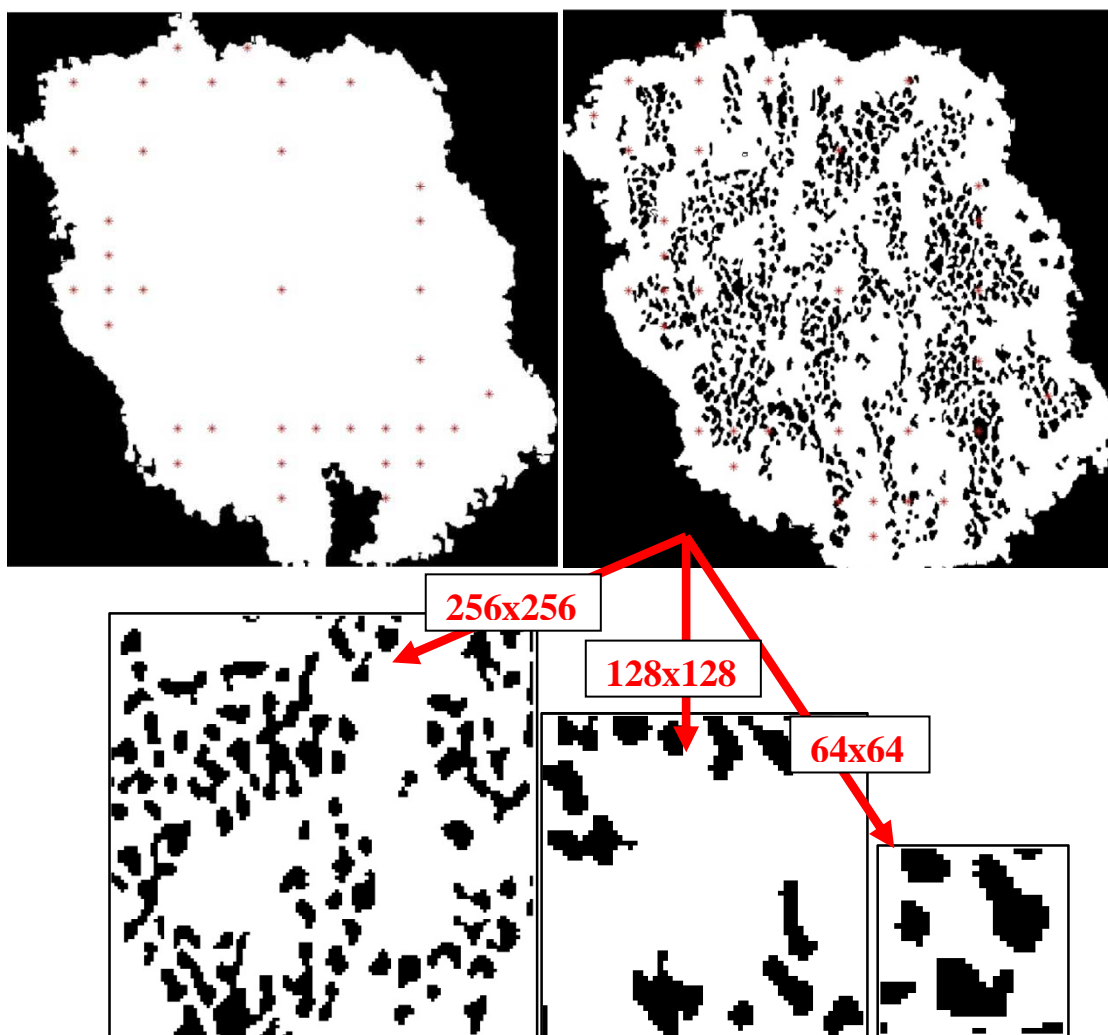
Figure 36: Quadtree decomposition of the lesion.

As mentioned before, it is assessed the black and white ratio for each block. This image for instance, has been split to 4 blocks of 256x256, 10 blocks of 128x128 and 22 blocks of 64x64 pixels. That means we have 36 ratios that somehow we have to average in order to obtain a mean pattern. Since they have different sizes, the fair way to do it it's by weighting each ratio according to the block's size. This way we guarantee a normalized comparison.

The best representation for a common pattern is not given by the mean value but by the median. We illustrate the concept with the following simple example of a vector.

[1 2 1 1 1 100]; a regular pattern should be represented by a number near 1.The median in this case is actually 1, giving an optimal approach. Instead the mean value is 17.67 because of the 100. Hence we can truly state the median is better characterizing patterns than the mean.

In this example, the median ratio is 0.7770, which means 77,7% of the pixels are white in average. Since the typical pigment network shows a uniform spatial arrangement, the mean deviation of the block's ratios with respect to this value will keep under a threshold mostly. We again weight the blocks for the mean deviation. This example deviation is 0.0626. The threshold will be set according to the values observed in the set of images with pigment networks.

Color Uniformity of the meshes

Typical pigment networks have meshes homogeneous in color. This could be another parameter to take into account. Since we have labeled all the meshes, calculating the color homogeneity of them is quite simple. We just arrange the standard deviation in R, G, and B color planes for every hole and take the deviation as the mean value of three. The final deviation is the mean value of the vector hole_std.

hole_std(k)=mean([std(red(ind2)) std(green(ind2)) std(blue(ind2))]);

Thickness of the lines composing the network

Typical pigment networks use to present thin lines, instead those atypical present thick lines. We then need to achieve an image with all the lines composing the network, and just them. After that we will be able to analyze the thickness of these lines.

After examining all color planes (including HSV, RGB, XYZ, HSI, YIQ, Lab, YCbCr), we observed that S plane from HSV individuated the network's lines remarkably.
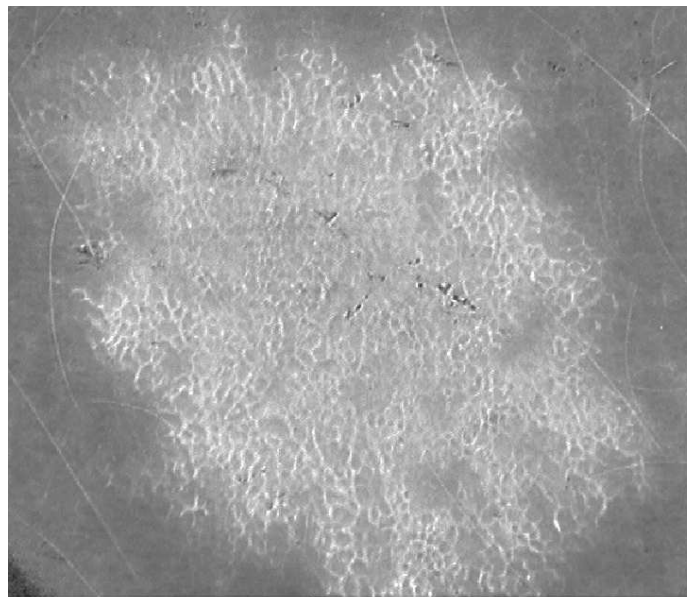
Figure 37: S plane of HSV color space.

Then contrast enhancement and some morphological operators are applied. The result after a Binary Thresholding with an appropriate value is a binary image with the main tracks of lines quite well noticeable. Value when Thresholding to black and white should be adapted according to the averaged value of the network's pixels.
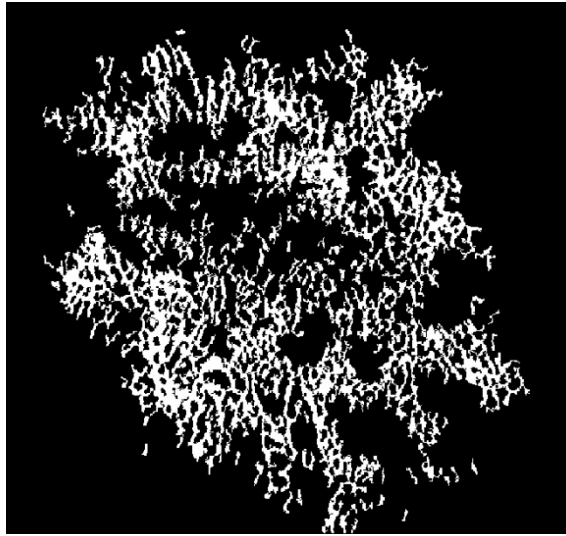
Figure 38: Lines thresholding

In order to measure the lines' width we have used morphological operators. The above image has been eroded twice consecutively, first with a thin vertical line and then with a thin horizontal one. Thus, lesions with typical pigment network are supposed to thin out white lines considerably. On the other hand, lesions with atypical pigment network are supposed to remain quite populated from white lines or regions, since they are supposed to have broader lines. Below in Figure 39 is plotted the erosion of the typical pigment network used so far. Notice the reduction of white pixel's percent.

imerode(imerode(lines,strel('line',2,90)),strel('line',2,0))
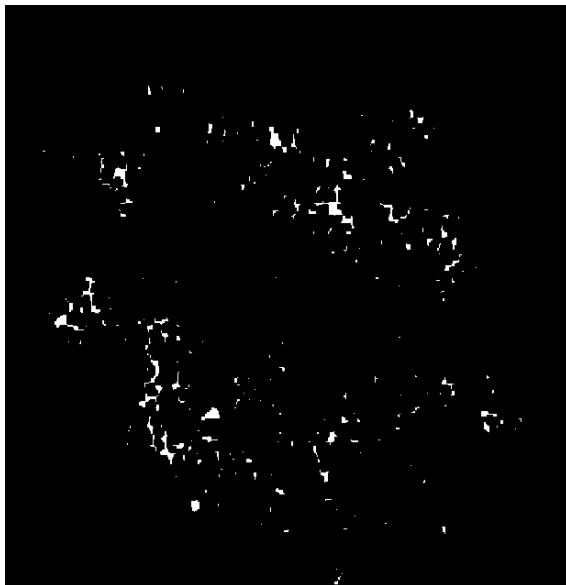
Figure 39: Typical Network Lines' erosion

However, if we do the same with a lesion containing an atypical pigment network, the reduction of white pixels is not that big.
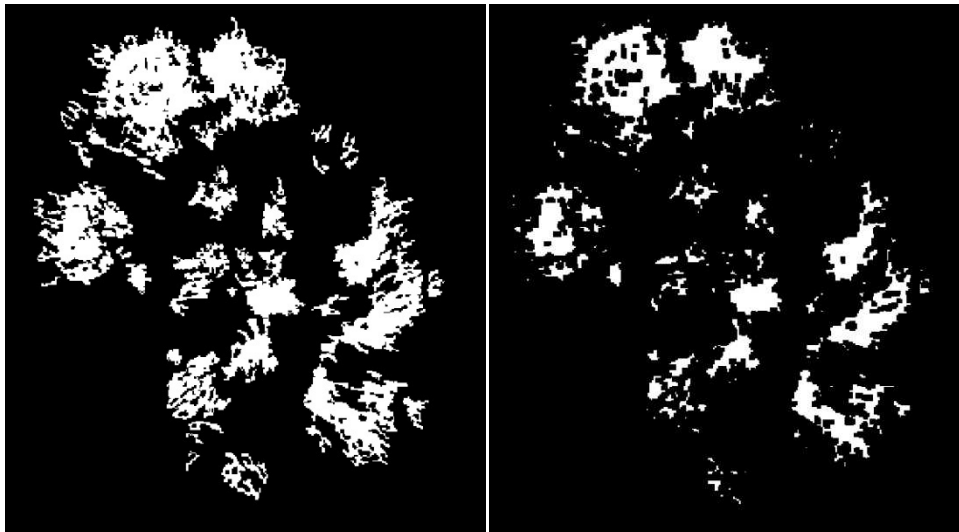
Figure 40: Atypical Network Lines' erosion

We can then set a coefficient which measures the reduction of white pixels after the erosion with thin lines. This coefficient will be an indicative of the lines' width.

6. <u>Results</u>

A set of 35 images has been tested. 11 of them contained a typical pigment network, 4 contained an atypical pigment network, and 20 were pigmented skin lesions but without reticular pattern. These last 20 contain every kind of possible patterns, which are Globular Pattern, Homogeneous Patten, Let's say the two first groups work as "cases", and the third group works as "controls".

Before the test, a small sort of images containing typical pigment network, atypical pigment network, and neither of them has been used for setting the thresholds between the different types regarding Spatial Uniformity, Color Uniformity and Thickness of lines (9 in total, 3 from each group). This sort is exclusively composed by images with a very clear feature, that is, model typical pigment network and model atypical pigment network.

Below, in the next point, it is presented the outcome of this brief survey.

6.1. <u>Small Test: Parameter's setting</u>

We have tested images from the three groups (Typical pigment networks, Atypical pigment networks and No Reticular). Density, Spatial Uniformity and Thickness of lines have been assessed for the thresholds' setting.

6.1.1. <u>Typical pigment networks</u>

- <u>Density and Spatial Uniformity</u>: It has been observed lesions containing typical pigment networks show a **high level of Density** in the graph, commonly greater than 30%, this means lots of red points (nodes) and mostly interconnected (green lines). Therefore, the median ratio which represents the pattern uses to be lower than 80%.
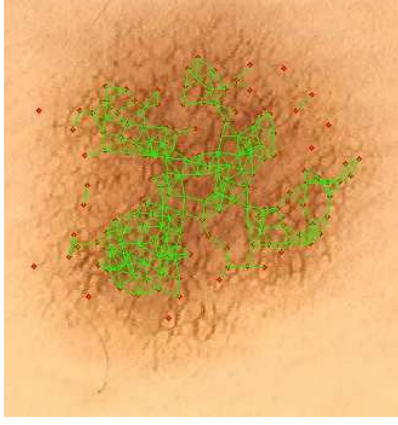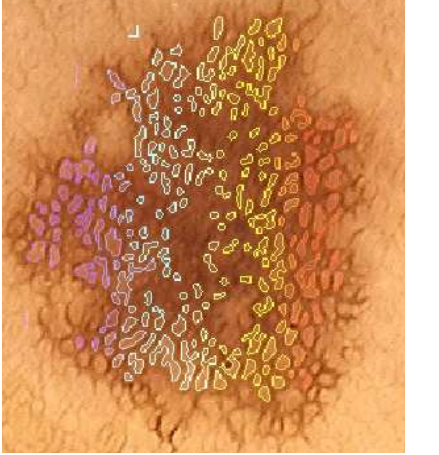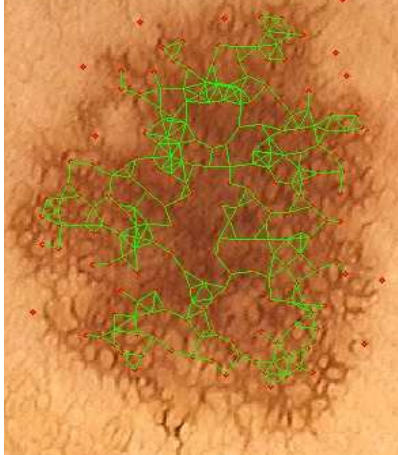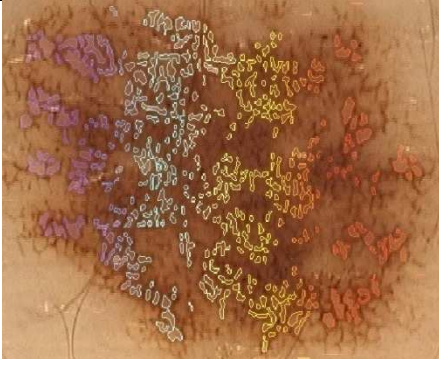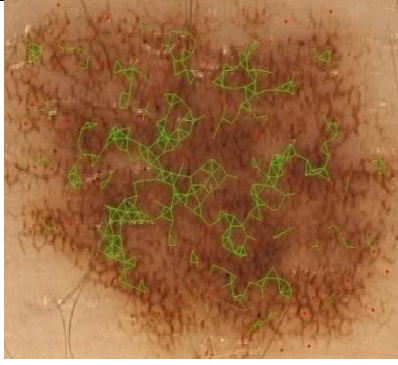
Identified Holes	Network	Parameters
		Density: 48,77% Spatial uniformity: 69,38%
		Density :39,82% Spatial uniformity: 74,46%
		Density: 31,64% Spatial uniformity: 77,73%

Table 8: Typical Pigment Network Density and Uniformity assessment

- Thickness of lines: It has been observed lesions containing typical pigment networks show a **high reduction of white lines** due to the erosion with thin lines, commonly greater than 60%.
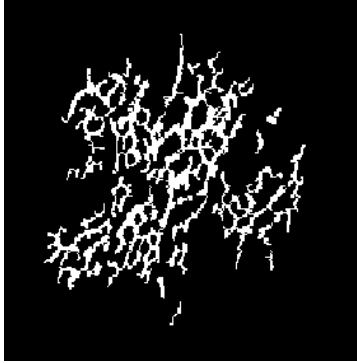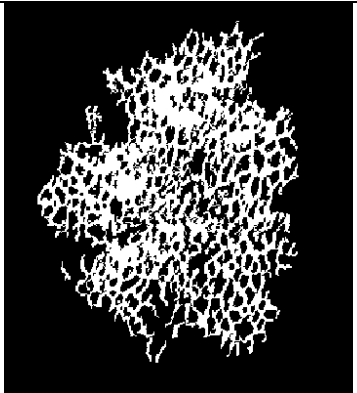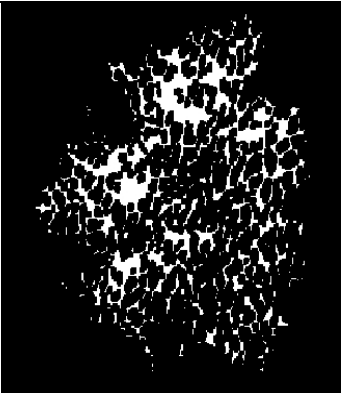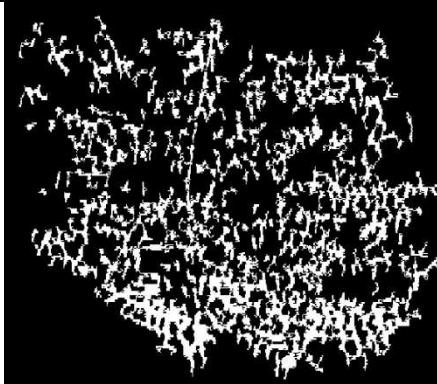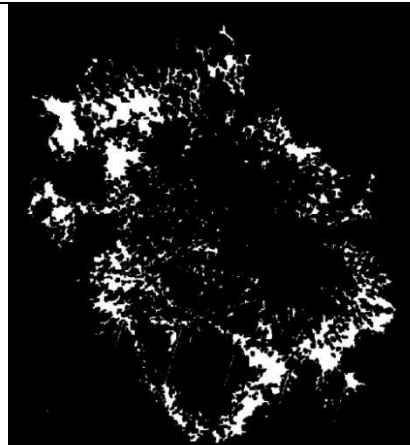
Network lines	Lines' erosion	Parameters
		Remaining white pixels: 7,06%
		Remaining white pixels: 36,90%
		Remaining white pixels: 20,24%

Table 9: Typical Pigment Network Lines' thickness assessment

6.1.2. Atypical pigment network

- Density: It has been observed lesions containing atypical pigment network show a **medium level of Density** in the graph, commonly between 15% and 30%. Therefore, the median ratio which represents the pattern uses to be compressed between 80% and 90%.
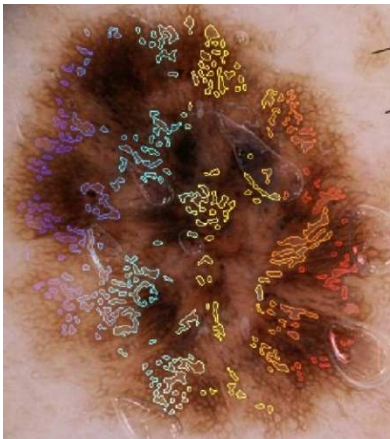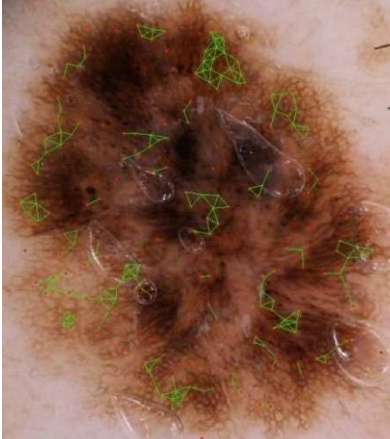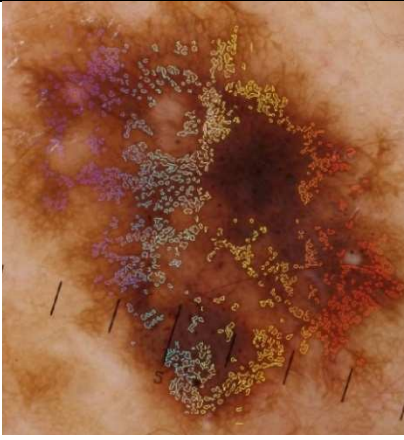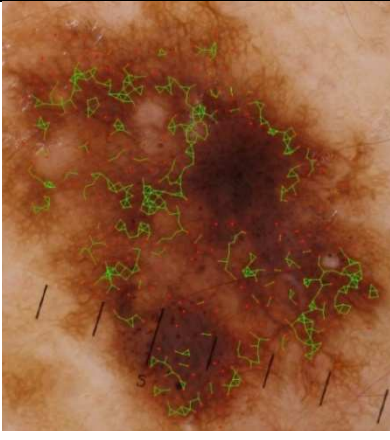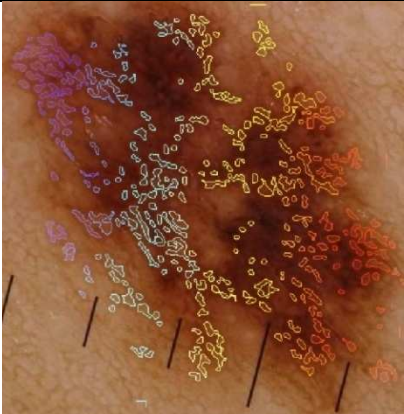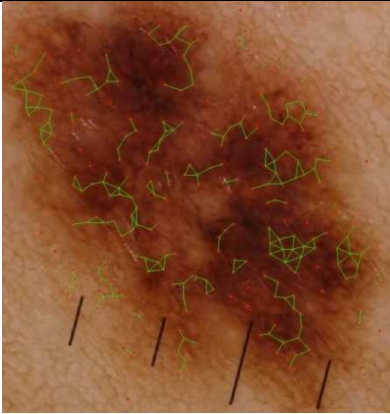
Identified Holes	Network	Parameters
		Density: 0.2835 Spatial uniformity: 0.8556
		Density : 0.2498 Spatial uniformity: 0.8655
		Density :0.2440 Spatial Uniformity: 0.8203

Table 10: ATypical Pigment Network Density and Uniformity assessment

- Thickness of lines: It has been observed lesions containing typical pigment networks show a **low reduction of white lines** due to the erosion with thin lines, commonly upon 60%.
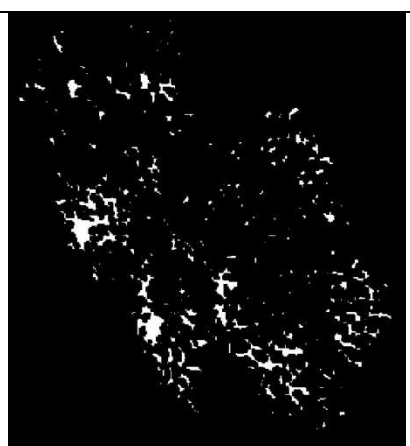
Network lines	Lines' erosion	Parameters
		Remaining white pixels: 51,80%
		Remaining white pixels: 59,43%
		Remaining white pixels: 45,08%

Table 11: ATypical Pigment Network Lines' thickness assessment

6.1.3. Non Reticular (None of networks)

- Density and Spatial Uniformity: It has been observed lesions non containing network show a **very low level of Density** in the graph, commonly upon 20%. Therefore, the median ratio which represents the pattern uses to be greater than 90%.
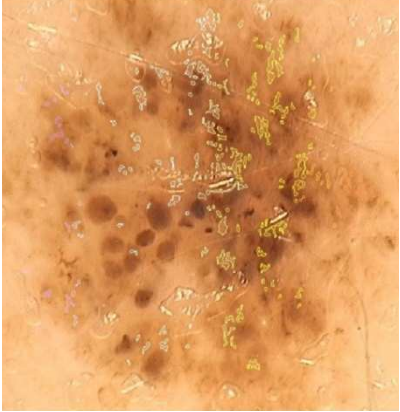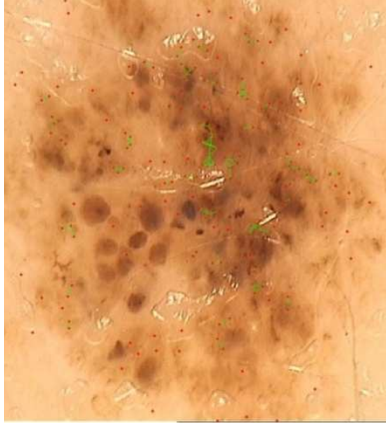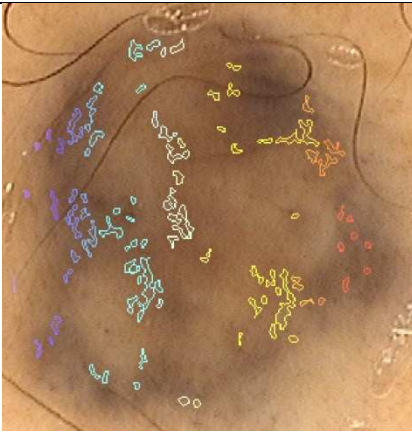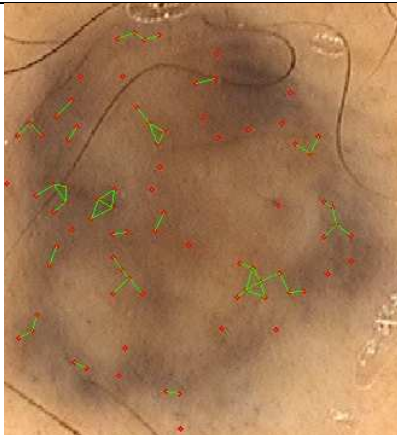
Identified Holes	Network	Parameters
		Density : 0.1613 Spatial Uniformity: 0.9584
		Density :0.0979 Spatial Uniformity: 0.9849
		Density :0.1788 Spatial Uniformity: 0.9287

Table 12: No reticular lesions Density and Uniformity assessment

- Thickness of lines: It has been observed lesions containing typical pigment networks show a **low reduction of white lines** due to the erosion with thin lines, commonly upper 60%.
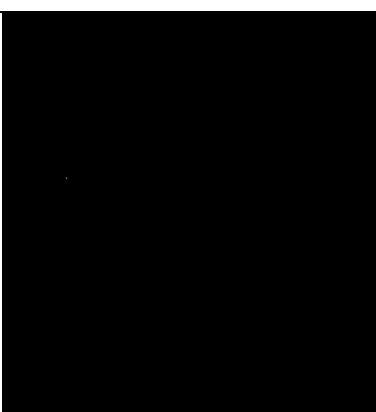
Network lines	Lines' erosion	Parameters
		Remaining white pixels: 32,44%
		Remaining white pixels: 22,69%
		Remaining white pixels: 0%

Table 13: No Reticular lesions Lines' thickness assessment

According to the previous results and taking into account a small safety interval, thresholds have been set to the following values:

	Typical PN	Atypical PN	No Reticular
Density	>25%	20%< D <40%	<30%
Spatial Uniformity	<80%	<90%	>85%
Thickness of lines (White pix. Reduction)	>60%	<60%	>70%

Table 14: Parameters' Thresholds setting

With these thresholds a tree of conditions has been created so that **there is no overlapping** in the selection. The total amount of pictures has been classified according to this criterion.

A 3x3 matrix is built for the calculation of Sensitivity and Specificity. The 2 rows indicate the classification of our test in Typical PN, Atypical PN and No Reticular. The 3 columns indicate Typical Pigment Network, Atypical Pigment Network and No Reticular lesions according to the histopathological assortment, that is, our ***Gold Standard***. This is the commonly created table which illustrates the True and False Positives, and the True and False Negatives. Most of cases have been classified properly.

		Gold Standard			
		Typical PN	Atypical PN	No Reticular	Specificity
Test outcome	Typical PN	10	0	0	100%
	Atypical PN	0	4	1	96,77%
	No Reticular	1	0	19	93,33%
	Sensitivity	90,90%	100%	95%	Accuracy=94,28%

Table 15: Final Test outcome: Calculation of Sensitivity and Specificity

As said, 33 of 35 images have been correctly classified, giving a test accuracy of 94,28%. The results are satisfactory although the amount of images is not as large as desired.

7. <u>Discussion</u>

In this work, most of the common effective guidelines described in literature to date for computer-aided diagnosis have been followed, since the most used techniques for segmentation until novel approaches for feature extraction. Anyway always the most straightforward method has been the one used in the end.

For example, we discussed SIDE algorithm was the finest segmentation technique, and could have avoided the step that rendered our code supervised, however its implementation was not trivial and following advice from other authors [6] and mates [18] Fuzzy C-means sorted the issue out, being more than sufficient to our needs.

In the same way, hairs were removed using a simple median filter although in literature were found finer algorithms as Steger's line detection or Hough Transform. Again the outcomes were satisfactory enough and more complicated strategies were not adopted.

Another important review is the one related to the holes' identification. Despite some effective techniques have been carried out in order to enhance the contrast between holes and lines for a better individualization, sometimes holes' overlapping has been observed, fact that could affect harmfully to the algorithm's decision. Watershed transform was used for holes' splitting but better approach to reality was achieved applying only contrast enhancement. The calculation of Density [11] and the novel coefficient Spatial Uniformity demonstrated to be great classifiers of lesions when dealing with pigment networks.

Regarding lines' thickness identification, S plane from HSV color space showed to be the one in which the network is best highlighted and Thresholding it to black and white showed to be a good method to estimate the width of lines.

The outcome of the test is good enough, although we believe that having a larger Gold Standard could have raised our statistics, since we worked with very few cases (11 typical networks and 4 atypical networks). Unfortunately, this practice has been running for little time in this Dermatology section and the lack of pictures is actually a bottleneck.

As conclusion, we hope this thesis contributes to the process of automation of computer-aided diagnosis, and together with other approaches that search for other features, build a

reliable tool for the physician, improving the Paradigm and effectiveness of nowadays Teledermoscopy.

8. <u>Bibliography</u>

[1] Stolz, Cognetta,et al. ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. J Am Acad Dermatol 1994;30(4):551-559.

[2] Fleming MG, Steger C, et al. Techniques for a structural analysis of dermatoscopic imagery. Computerized Medical Imaging and Graphics 22 (1998) 375-389.

[3] Hance GA, Umbaugh SE, Moss RH, Stoecker WV. Unsupervised color image segmentation with application to skin tumor borders. IEEE Engng Med Biol 1996;15(1):104-111.

[4] Pollak I, Willsky AS, Krim H. Image segementation and edge enhancement with stabilized inverse diffusion equations. Tech Rept LIDS-P-2368. Laboratory for Information and Decision Systems, MIT, 1996.

[5] Mendonça T, André R. S. et al. Comparison of Segmentation Methods for Automatic Diagnosis of Dermoscopy Images. Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France August 23-26, 2007.

[6] Schmid Ph. And Fischer S. Colour segmentation for the analysis of pigmented skin lesions. Signal Processing Laboratory5 Swiss Federal Institute of Technology.

[7] Won lim Y. and Uk Lee S. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. Pattern recognition, vol. 23, no. 9, pp. 935-952. 1990.

[8] Argenziano G, Soyer HP, Chimenti S, Talamini R, Corona R, Sera F, et al. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. J Am Acad Dermatol 2003;48:679-93.

[9] Malvehy J, Puig S, Argenziano G, et al. Dermoscopy report: Proposal for standardization Results of a consensus meeting of the International Dermoscopy Society.

[10] J AM ACAD DERMATOL Book reviews 807 Volume 50, number 5 Atlas.

[11] Maryam Sadeghia et al. A novel method for detection of pigment network in dermoscopic images using graphs. Computerized Medical Imaging and Graphics 35 (2011) 137–143.

[12] Argenziano G, Soyer HP, Giorgio VD, Piccolo D, Carli P, Delfino M, Ferrari A, Hofmann-Wellenhof R, Massi D, Mazzocchetti G, Scalvenzi M, Wolf IH. Interactive Atlas of Dermoscopy (Book and CD-ROM). Edra Medical Publishing and New Media; 2000.

[13] Anantha M, Moss RH, Stoecker WV. Detection of pigment network in dermatoscopy images using texture analysis. Computerized Medical Imaging and Graphics 2004;28(5):225–34.

[14] Betta G, Di Leo G, Fabbrocini G, Paolillo A, Sommella P. Dermoscopic imageanalysis system: estimation of atypical pigment network and atypical vascular pattern. In: Proceedings of the IEEE International Workshop on Medical Measurement and Applications. 2006. p. 63–7.

[15] Menzies S, Crotty K, Ingvar C, McCarthy W. An atlas of surface microscopy of pigmented skin lesions. Sydney: McGraw-Hill, 1996.

[16]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 1998;134:1563-70.

[17] http://www.dermoscopy.co.uk/index.html

[18] Albertini Claudio. Master Thesis. Politecnico di Milano.

[19] http://www.dermoscopy.org/atlas/base.htm

[20] J. Nascimento, J. S. Marques, Adaptive Snakes Using the EM Algorithm IEEE Trans. Image Processing, 2005, pp 1678-1686.

[21] T. E. Chan, B.Y. Sandberg,L.A. Vese, Active contours without edges for vector-valued images, Journal of Visual Communications amd Image Representation, vol.11, 2000, pp 130-141.

[22] Pehamberger H, Steiner A, Wolff K. In vivo epiluminescence microscopy of pigmented skin lesions. Pattern analysis of pigmented skin lesions. 1987.

[23] Caiani E. Course Material: Bioimages Processing Laboratory. Politecnico di Milano. Class notes 2010.

[24] http://homepages.inf.ed.ac.uk/rbf/HIPR2/stretch.htm

[25] Gonzalez R.C., Woods R. E., Eddins S.L., Digital Image Processing Using Matlab. ISBN 0-13-008519-7. 2004 by Pearson Education, Inc.

[26] Matlab Help.

[27] Marcon M. Course Material: Techniche Avanzate per il Trattamento delle Immagini (TATI). Politecnico di Milano. Class notes 2010.

About the author: ISMAEL HUERTAS FERNANDEZ is receiving his MD in Biomedical Engineering as Double Degree at Politecnico di Milano (Italy) by this work. He also receives BD and MD in Telecommunication Engineering at Escuela Superior de Ingenieros at University of Seville (Spain). His research interests concern bioimage and biosignal processing, as well as bioinformatics related to genetic studies. He is currently researcher at Instituto de Biomedicina de Sevilla (IBiS), laboratory of movement disorders, section of Neuroscience. His work is mainly focus on biodata analysis for genetic findings in diseases as Parkinson, Dystonia, Tourette or Huntington. He is kindly available for contact at: ihuertas-ibis@us.es