

## ALGORITHM DEVELOPMENT FOR DERMOSCOPIC IMAGES ANALYSIS OF SKIN LESION

**Talita S. Coelho<sup>1</sup>, Felipe Massicano<sup>1</sup>, Rafael G. Possani<sup>1</sup>, Giuliana Fedorenko<sup>1</sup>, Rodrigo S. Viana<sup>1</sup>, Vinícius A. Castro<sup>1</sup>, Marco A. R. Fernandes<sup>2</sup>, Hélio A. Miot<sup>2</sup> and Hélio Yoriyaz<sup>1</sup>**

<sup>1</sup> Instituto de Pesquisas Energéticas e Nucleares (IPEN / CNEN - SP)  
Av. Professor Lineu Prestes 2242  
05508-000 São Paulo, SP  
[tasallesc@gmail.com](mailto:tasallesc@gmail.com)

<sup>2</sup> Departamento de Dermatologia e Radioterapia da Faculdade de Medicina de Botucatu -  
Universidade Estadual de São Paulo  
Campus de Rubião Júnior, S/N  
18600-000 Botucatu, SP  
[marfernandes@fmb.unesp.br](mailto:marfernandes@fmb.unesp.br)

### ABSTRACT

Skin cancer is the most incident malignant neoplasm, in Brazil representing approximately 25% of all diagnosed tumors. The malignant melanoma (MM) is less frequent among all others, corresponding to about 4% of all diagnoses. As most types of cancer, a better prognosis is associated with early diagnosis, i.e., detection at an early stage (in situ melanoma). The discovery of lesions in an advanced stage and after its infiltration in the lower layers of the skin is correlated to a high probability of metastasis and the cure is practically null. The most common diagnosis method of melanoma is based on the so-called ABCD rule, where physical factors are analyzed: asymmetry, edge detection, color and total diameter, all observed directly at the lesions. The dermoscopy is a noninvasive technique for recognition of skin surface structures, served both for magnification (up to 400 times) and evaluation of images of cutaneous tissue. The parameters of the ABCD rule can be observed through the dermatoscope, requiring a certain degree of the operator experience and knowledge. These visual signs (images) greatly facilitate the skin tumors. Aiming to help solving this potential problem inherent to dermoscopy, i.e., the subjectivity parameter imposed by dermatologist and clinical operator, the use of methods for Processing Digital Images (DIP) combined with digital dermoscopy was developed in this work, with the objective of computationally analyzing dermoscopic lesion images. An algorithm was implemented using JAVA language, which segments the image and analyzes the ABCD parameters rule. Therefore, by applying the formula known as TDS (Total Dermatoscopy Score) after the extraction of tissue characteristics, particular scores (based on their results) for each parameter (ABCD), i.e., the probability of the lesion being malignant, benign or suspicious is calculated by a weight factor multiplication. As a whole it can be said that the software reached its goal, being able to segments lesions and classify them using the TDS formula and the ABCD parameters rule.

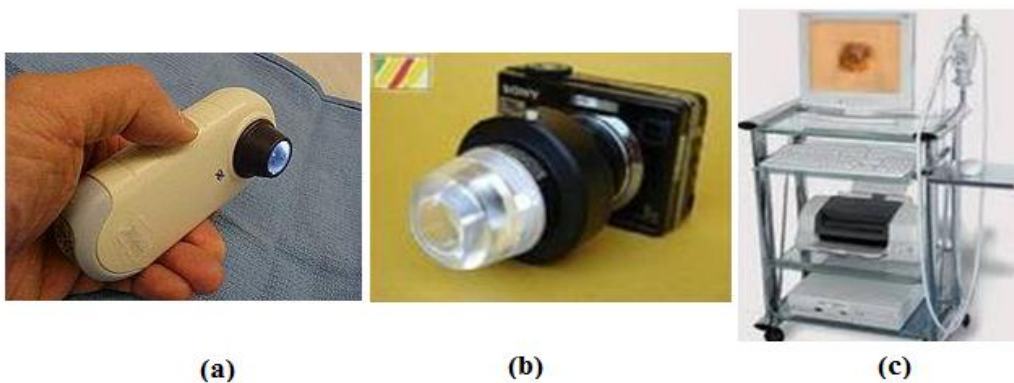
### 1. INTRODUCTION

Skin cancer is the most incident malignant neoplasm, in Brazil represents approximately 25% of all diagnosed tumor. The malignant melanoma (MM) is less frequent among all others, corresponding to about 4% of all diagnoses [1]. As most types of cancer, a better prognosis is associated with early diagnosis, i.e., detection at an early stage (in situ melanoma) [2]. The discovery of lesions in advanced stage and after its infiltration in the lower layers of the skin, is correlated to a high probability of metastasis and the cure is practically null [3].

The most common diagnosis method of melanoma is based on the so-called ABCD rule, where physical factors are analyzed: asymmetry, edge detection, color and total diameter, all observed directly at the lesions [4], [5], [6], [7], [8], [9], [10]. This methodology was

proposed in 1994 by Wilhelm Stolz [11], due to the difficulty of early diagnosis of melanoma, and its common confused with benign nevi and moles (in the initial stages).

The dermoscopy is a noninvasive technique for recognition of skin surface structures, served both for magnification (up to 400 times) and evaluation of images of cutaneous tissue. Figure 1a, 1b and 1c show the three existent types of dermatoscopes [12], [13], [14]. The dermatoscope in Figure 1a shows the simplest form, consisting of a convex lens with angulation of  $20^\circ$  and surrounded by various LED of polarized light, used by dermatologists in order to view the image manually with magnitude of about 400 times. Figure 1b shows a dermoscopic image which consists of a digital camera with a dermatoscope attached to it, used to take pictures at high magnitude. Figure 1c shows a digital dermoscopy system which consists of a dermoscopic image coupled to a computer, which allows one to scan images of skin lesions and also manipulate them by using image processing software.



**Figure 1. Types of dermatoscopes: a) standard dermatoscope, b) dermoscopic images, c) digital dermoscopy system.**

The parameters of the ABCD rule can be observed by dermatoscope, requiring some degree of experience and knowledge from the operator. This aspect is a differential factor in the task of skin tumor classification based on visual signs (pictures). As every presentation of skin cancer has its own characteristics regarding the shape and color. Each examination has as its first fundamental step the differentiation of non-melanocytic lesions of melanocytes, and then classifies them into benign, malignant or suspicious.

The correct use of the dermatoscope depends on the degree of sensitivity and knowledge of the dermatologist [15], [16], [17]. Thus, methods of digital image processing (DIP) can be used to minimize errors through software and hardware systems, where computational algorithms for preprocessing and segmentation are performed.

A Digital Image Processing system, or DIP, consists of several stages, with capacity to produce a response based on a problem domain and the expected result, where each step is performed by a series of computer algorithms in order to obtain an optimal result from digital image. The basic steps of a DIP system are: image acquisition, preprocessing, segmentation, feature extraction and recognition / interpretation [5], [6], [7].

In this work it was developed software that performs the pre-image processing, lesion segmentation, and extracts relevant features of the ABCD rule, which is finally applied to the Total Dermatoscopy Score (TDS) formula in order to classify the lesion.

## 2. METHODS

### 2.1. Descriptions of the stages of software development

The images used in this work were provided by the department of dermatology at the Medicine College of Botucatu - SP. These dermoscopic images were obtained using a digital camera in which a dermatoscope has attached to it, allowing a magnification up to 100 times the size of the lesion.

The software was developed in Java programming language which is an object-oriented, developed by Sun Microsystems which has a large advantage over others languages because its multi-platform characteristics. This portability is achieved because the language is interpreted, in order words; the compiler generates a code independent of the machine, called byte-code. At the time of running this byte-code is interpreted by a virtual machine (interpreter) installed on the machine [18].

The program was written through the development environment (IDE) NetBeans IDE 6.9.1, free and open code, aimed at developers of Java, C / C + +, PHP, among other.

All the steps applied in the software DIP is decribed bellow.

1) Sampling and quantization of images: the scan involves these two processes, where sampling is to generate an array of  $M \times N$  samples, corresponding to the  $x$  and  $y$  direction samples of the image, with possible values that the function  $f(x,y)$  can assign to their elements (pixels). In quantization, the image of each point is assigned to a positive integer  $L$ , corresponding to their color level. The processes of sampling and quantization are connected directly to the output display resolution.

Once the image is sampled in RGB was converted to a grayscale, using the method of luminance, which consists of applying a certain weight to each channel, the channel  $w_R=0.299$  to the R channel  $w_G=0.587$  to the channel G and  $w_B=0.114$  to the channel B [19]. Once the channel of a given pixel is balanced, the sum of them is performed in order to obtain a grayscale value for this pixel (Y).

$$Y = w_R \times R + w_G \times G + w_B \times B \quad (1)$$

This grayscale is applied to the RGB channel of the pixel in question, thus obtaining a corresponding grayscale of that pixel.

$$I(x,y) = (R,G,B) = (Y,Y,Y) \quad (2)$$

2) Image pre-processing: The pre-processing technique aims the improvement of certain characteristics of the images to making them more suitable for subsequent steps. This enhancement is necessary when the image undergoes a process of degradation or loss of

quality, a phenomenon characterized as noise, such as: loss of contrast, blurring, artifacts and brightness. The noise may be inherent in the acquisition technique chosen, such as improper lighting or other relevant aspects.

The technique used in images preprocessing was the equalization of the histogram in order to improve the image contrast. This technique seeks to redistribute the values of a grayscale image, so as to obtain a uniform histogram, where the number of pixels and the number of gray levels remain practically the same.

In addition to equalization, the background image was homogenized in order to remove noises such as air bubbles formed during the process of image acquisition. Through this technique is also possible to identify the edge of the lesion. The algorithm consists of obtaining an average value of four pixels from the four corners of the image, pixel (0, 0), pixel (0, width-1), pixel (height-1) and pixels (height-1, width-1). This is followed by scanning the original image, making sure that the original pixel value,  $I_o(x, y)$ , is within the range of  $\pm 10\%$  of the calculated average pixel, it is assigned the value of 255 (white) to this pixel; otherwise it keeps the original color. The result of this algorithm is the elimination of background image represented by white pixels and the injury image keeps their original grayscale.

3) Image segmentation: Segmenting an image is a process of partitioning a set of image data, in structures with relevant content to a desired application, aiming the improvement of the interpretation during the analysis stage. Often this step is based on the detection of lines, discontinuities (edges) or different regions. Edge detection is a technique that determines points of an image where the light intensity changes suddenly, that is the edge [19].

There are several types of image segmentation techniques. Among them the more conventional approaches are usually based on the basic properties of grayscale image, trying to detect discontinuities (filters) or similarities in the image.

The operation of these filters is based on templates, that is, a mask used to perform operations of a neighboring pixel. In practice, it is an array whose central element is positioned at the pixel of interest. The elements of the neighborhood, including the pixel in question, are multiplied by indicated values in the correspondent positions in the array. The sum of the results replaces the pixel value of interest in the resulting image [20].

The filters used in this work are high-pass filters very used to highlight features in the images, like edges, points and line segments. The purpose of this filter is to attenuate low frequencies, making sharp transitions between different regions (the lesion and background) [20].

The algorithms developed for edge detection were homogenization, thresholding and, filter application using templates: Sobel line, Sobel column, point detection and Laplacian. Below is a brief description of each one.

The thresholding, better known as binarization, is the classification of image pixels according to the specification of a threshold value. This algorithm uses the threshold value (T) for conversion of grayscale pixels in white or black, in such a way that separates the image into two regions, the lesion and the background according to the following rule:

$$I(x; y) = \begin{cases} 255 & \text{se } I_o(x; y) \geq T \\ 0 & \text{se } I_o(x; y) < T, \end{cases} \quad (3)$$

where  $I_o(x, y)$  is the original image and  $I(x, y)$  is the processed image. The pixels corresponding to 0 (black) represent the injury and the pixels corresponding to 255 (white) represent the background and  $T$  is the grayscale value, called threshold.

Sobel line and Sobel column: the masks of these operators are used to detect line segments. Figures 2a and 2b contain masks for detecting lines horizontally and vertically oriented. The detectors are based on derivatives that measure the rate of change of a function. These variations in the images are larger in the edges and smaller in constant regions of the image. Once the points where the derivatives have a maximum variation were identified, all the contours are marked [19].

-1	-2	-1
0	0	0
1	2	1

(a)

-1	0	1
-2	0	2
-1	0	1

(b)

**Figure 2. Line detection mask: (a) Sobel line mask and (b) Sobel column mask.**

Filter points detection: it is the simplest technique of detection. When all the pixels of the mask have the same intensity (homogeneous region), the response of the mask will be zero. Otherwise a point will be detected in the central position of the mask.

-1	-1	-1
-1	8	-1
-1	-1	-1

**Figure 3. Mask for the detection of isolated points.**

Laplacian Filter: is a local differential operator, by the application of the Laplacian operation the positive (internal injury) and negative (external injury) points are defined according to equation (4) which is a derivative of second order two-dimensional function, applied to the image through the mask shown in Figure 4.

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} i + \frac{\partial^2 f(x, y)}{\partial y^2} j \quad (4)$$

0	1	0
1	-4	1
0	1	0

**Figure 4. Mask to calculate the Laplacian filter.**

4) Images feature extraction: This step aims to extract characteristics of the resulting images segmentation through algorithms that allow extracting the information precisely. These algorithms must be represented by a data structure suitable for recognition of these features extraction. It is important to note that in this step the input is still an image, but the output is a set of data corresponding to the image.

After the segmentation four parameters were extracted which belongs to the ABCD rule, they are: asymmetry, border, color and lesion diameter. These parameters were calculated in four dermoscopic images of skin lesions.

Asymmetry: asymmetry of the calculation was based on the method proposed by Stolz [11], where the image of the lesion is divided into two segments perpendicular to each other (horizontal and vertical) and, positioned on the lesion to produce the lowest score of asymmetry, as shows the Figure 5a.

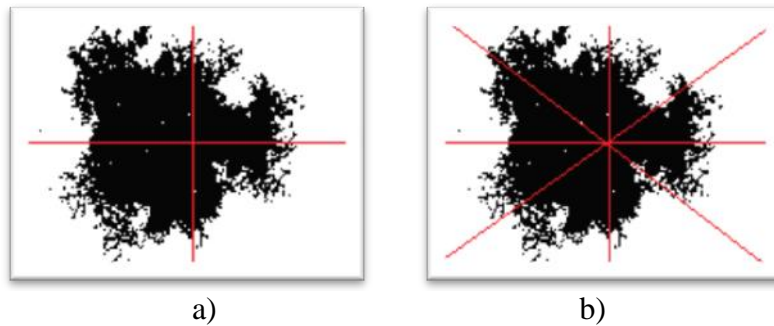
The calculation of asymmetry gives an analysis of the injury at a geometry level. So it was proposed an area test for analysis of the lesions, where the amount of black pixels (which corresponds to the region of the lesion) belonging to each axis (on both the positive and the negative), are counted. Then the difference in the amount of pixels on both sides is calculated. A variation of  $\pm 0.25\%$  is considered in the difference between the correspondent pixels, since in the segmentation process, there is a possibility to count false pixels in the border.

If both axes are asymmetrical, its score is 2. If the asymmetry is in only one of the axes its score is 1, and if there is no asymmetry its score is 0. Most melanomas have a score of 2, as opposed to benign nevi and melanocytes nevi.

Edge: the lesion is divided into eight equal parts, as shown in Figure 5b. The score is found by the analysis of eight parts, i.e., the amount of black pixels belonging to each slice is counted and this amount was compared to amount obtained in other slices. A variation of  $\pm 20\%$  was considered. For each slice presenting a difference greater than the established here 1 point is summed to the total score.

Colour: in the ABCD rule the observed colors indicating malignancy are black, white, blue-gray, light brown, dark brown and red. A range of possible colors for each of these six tones was used as a matter of evaluation in this test, as presented in Figure 6. The colors observed were only counted if it occupies an area of 0.1% of the total lesion area, since the variation is estimated as tolerance factor for each channel R, G and B. This variation was  $\pm 10$  color

levels for each channel. The presences of each shade of gray add more value in a score where the minimum is 1 and the maximum is 6.



**Figure 5. Lesion image divided into a) 4 equal parts to calculate asymmetry b) 8 equal parts to calculate edges.**

Black		White		Blue-Gray	
RGB Code	Color	RGB Code	Color	RGB Code	Color
(0,0,0)		(255,255,255)		(150,125,150)	
(10,10,10)		(245,245,245)		(125,125,150)	
(20,20,20)		(235,235,235)		(100,100,125)	
(30,30,30)		(225,225,225)		(100,125,150)	
(40,40,40)		(215,215,215)		(50,100,150)	
(50,50,50)		(205,205,205)		(0,100,150)	
Red		Light Brown		Dark Brown	
RGB Code	Color	RGB Code	Color	RGB Code	Color
(255,0,0)		(200,150,100)		(150,100,100)	
(255,50,50)		(200,100,0)		(125,75,75)	
(200,0,0)		(200,100,50)		(100,50,50)	
(200,50,50)		(150,100,50)		(100,50,0)	
(150,0,0)		(150,100,0)		(100,0,0)	
(150,50,50)		(150,50,0)		(50,0,0)	

**Figure 6. Six levels of common colors in MM.**

Diameter: values above 5-6 mm are alarms of malignancy, and the score is based on this length. To obtain this parameter, the amount of pixels in the width and height of the lesion were determined, therefore obtaining the average amount which is converted in millimeters. For each 1mm step 1 point is added to the score of the diameter.

Obtained after the scores of the four parameters belonging the ABCD rule, these were applied the formula for Total Dermoscopy Score (TDS), which interprets the analyzed lesion based on its total score, setting a diagnosis prior possible to the observed image. For scores <4.75 is the lesion is benign, for scores between 4.8 - 5.45 the lesion is suspicious and for > 5.45 a lesions with high probabilities of develop MM.

$$TDS = [ (A \times 1.3) + (B \times 0.1) + (C \times 0.5) + (D \times 0.5) ] \quad (4)$$

### 3. RESULTS AND DISCUSSION

Before starting the segmentation, the images were converted from RGB to grayscale, using the luminance method. Once the grayscale images are obtained, they are equalized in order to improve contrast and homogenized to obtain a homogeneous background reducing the noise and bubbles, which are typical characteristics of this type of images.

Segmentation methods were applied on several skin lesions images, whether benign nevi, melanocytes nevi or malignant melanoma, in order to verify the implementation of each of these methods in these types of lesions.

The methods were chosen based on an extensive literature review, where the most common of them are thresholding, homogenization, Sobel line, Sobel column, Laplacian filter and isolated points filter. Figures 7a, 7b, 7c, 7d, 7e e 7f show, respectively the segmented images produced by each one of this methods.

The homogenization was used in the image pre-processing that was subsequently segmented by the use of filters, and also as a segmentation method. It can be noticed that if the image has many artifacts or noise, the application of the threshold can provide better image segmentation in the background region. But when an image has few artifacts or noise, the homogenization method is a better alternative. Segmented images with filters showed a better edge enhancement over other methods, but not the most suitable for segmentation purposes.

For feature extraction step four binarized images were analyzed: a) melanocyte nevi with some aspects of malignancy, b) benign nevi and c) two malignant melanomas. Figures 8a, 8b, 8c and 8d show the lesions used in the software analysis.

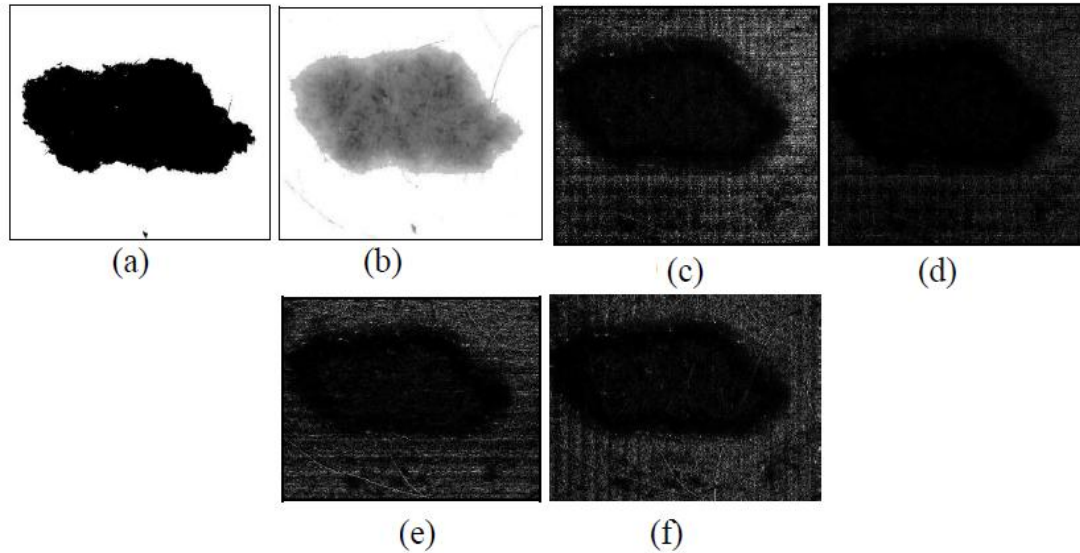
Figure 8a shows an asymmetry score equal to 0 indicating a symmetrical lesion, a score equal to 2 for irregular edges, 1 for color score and 5 for the diameter score. Applying these values to TDS formula a score of 3.2 is obtained, which indicates a benign lesion.

In Figure 8b shows an asymmetry and border scores equal to 0, coloration equal to 2 and diameter equal to 5 and a calculated TDS equal to 3.5, also indicating that it is a benign lesion.

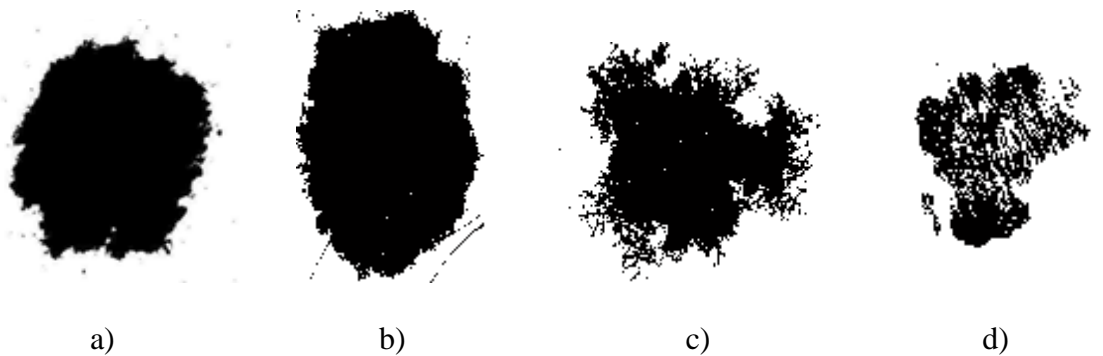
The Figure 8c shows an asymmetry score of 2, showing that the lesion is completely asymmetrical typically of malignant melanomas appearance. The edge score equal to 1, color score equal to 2 and diameter score equal to 5, obtaining a total score of 6.2 indicating high developing malignancy chances.



Figure 8d shows an asymmetry score of 2, an edge score equal to 3, color score equal to 2 and diameter score equal to 5. The calculated TDS was of 6.4 which show high probability of developing malignancy.



**Figure 7. Images of the same skin lesion segmented by different techniques: a) thresholding, b) homogenizing, c) isolated points filter, d) Laplacian filter, e) Sobel line filter and f) Sobel column filter.**



**Figure 8. Four analyzed lesions images.**

The obtained scores values for the four lesions were validated according to the results of biopsies, performed at the Department of Dermatology and Radiotherapy, Medicine College of Botucatu - SP.

## 4. CONCLUSIONS

Within of the DIP area, the work goal was covered and achieved namely: the pre-processing, the segmentation and the feature extraction (asymmetry, border, color and diameter) from skin lesions dermoscopic images.

The pre-processing was achieved by using the histogram equalization and the homogenization methods, to prepare the images and then apply to it the segmentation through template. The segmentation is one of the most important steps in DIP, once errors in segmented image, compromises the features extraction, especially when factors extracting is associated with border regions. Among various segmentation methods applied in this work, it is observed that the efficiency of each of these techniques is directly related to the input image quality.

Thus, it can be noted that there is no standard method of segmenting image, where one technique could be more efficient than others for a specific case. The feature extraction of skin lesions covered in these works which are: asymmetry, border, color and diameter were within the expected values. But it is very important to note that the quantification of these parameters needs to be as accurate as possible when one aims to apply it in a software designed to assist in early diagnosis of skin cancer, once erroneous analysis leads to a diagnose failure.

## ACKNOWLEDGMENTS

The authors thank the financial support of CAPES, CNEN, CNPq and FAPESP.

## REFERENCES

1. Faria S L, Salvajoli J V, Souhami L. *Radioterapia em Oncologia*, Medsi, Rio de Janeiro & Brazil (1999).
2. “Estimativa 2010: Incidência do câncer no Brasil,” <http://www.inca.com.br> (2009).
3. Fernandes MAR, Oliveira. Jr. B, Coelho TS, Yoriyaz H, Louzada MJQ, Fernandes MF, “Avaliação Densitométrica e Aspectos Clínicos da Betaterapia em Quelóides e Pterígio”, *UNIVERSITAS - Revista do Centro Universitário Católico Salesiano Auxilium*, Araçatuba, Vol.1, n1, pp.133-144 (2009).
4. Sobieranski AC, et al. Metodologia Computacional para Aplicação da Regra ABCD na Avaliação de Lesões Pigmentadas, VII Workshop de Informática Médica–WIM, Universidade Federal de Santa Catarina, pp 1-10 (2007).
5. Fikrle T, Pizinger K, “Digital computer analysis of dermoscopic images of 260 melanocytic skin lesions; perimeter/area ratio for the differentiation between malignant melanomas and melanocytic nevi”, *Academy of Dermatology and Venerology*, pp 48-55, (2006).
6. Johr RH, “Dermoscopy: Alternative Melanocytic Algorithms- The Menzies Socring Method, and 7-Point Checklist”, *Clinics in Dermatology*, pp 240-247, (2002).
7. Maglogiannis I, Pavlopoulos D, Koutsouris D, “An Integrated Computer Supported Acquisition, Handling and Characterization System for Pigmented Skin lesions in Dermatological Images”, *IEE Trans Inf Technol Biomed*, pp 86-98, (2005).

8. Ganster H, et al, "Automated Melanoma Recognition", IEE Trans Inf Technol Biomed, pp 233-239, (2001).
9. Bono A, et AL, "The ABCD System of Melanoma Detection. A Spectrophotometric Analysis of the Asymetry, Border, Color, and Dimension", American Cancer Society, pp 72-77, (1999).
10. Manousaki AG, et al, "A simple digital image processing system to aid in melanomandiagnosis in an everyday melanocytic skin lesion unit". A preliminary report, International Journal of Dermatology, pp 402-410, (2006).
11. Stolz W, et al, "ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma", Eur J Dermatol, pp 521-527, (1994).
12. Pivetti G, "Análise de Técnicas de Processamento Digital de Imagens para Diagnóstico de Tumores de Pele". [Trabalho de Conclusão de Curso]. Centro Universitário Católica Salesiano Auxilium-UniSalesiano, Araçatuba, (2010).
13. Binder M, et AL, "Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologist", Arch Dermatol, pp 286-291, (1995).
14. Soares HB, "Análise e Classificação de Imagens de Lesões da Pele por Atributos de Cor, Forma e Textura Utilizando Máquina de Vetor de Suporte". [Tese de Doutorado]. Universidade Federal do Rio Grande do Norte, (2008).
15. Frange VMN, Arruda LHF, Daldon PEC, "Dermatoscopia: importância para a prática clínica", Rev. Ciên. Méd, pp. 209-215, (2009).
16. Perrinaud A, et al, "Can automated dermoscopy image analysis instruments provide added benefit for the dermatologist. A study comparing the results of three systems", British Journal of Dermatology, pp 926-933, (2007).
17. Rigel DS, Russak J, Friedman R, "The Evolution of Melanoma Diagnosis: 25 Years Beyond the ABCDs", CA Cancer J Clin, pp 301-316, (2010).
18. Deitel HM, Deitel PJ, Java: Como Programar, Pearson Education, Brasil (2005).
19. Burger W, Burger MJ, Digital Image Processing – Na Algorithmic Introduction using Java, Springer Science, New York &NY (2008).
20. Gonzalez RC, Woods RE, Digital Image Processing, Prentice Hall, New Jersey (2002).