Original Article

# SOFTWARE FOR AUTOMATIC DIAGNOSTIC PREDICTION OF SKIN CLINICAL IMAGES BASED ON ABCD RULE

# SOFTWARE PARA PREDIÇÃO DIAGNÓSTICA AUTOMÁTICA DE IMAGENS CLÍNICAS DA PELE BASEADA NA REGRA ABCD

**Gleidson Brandão OSELAME[1]; Ionildo José SANCHES[1]; Alana KUNTZE[1]; Eduardo Borba NEVES[1]**

1. Postgraduate Program in Biomedical Engineering. Federal University of Technology - Paraná. Curitiba, Brazil.
gleidsonoselame@gmail.com

**ABSTRACT:** Cancer is responsible for about 7 million annual deaths worldwide. Among them, the melanoma type, responsible for 4% of the skin cancers, whose incidence has doubled in the last ten years. The processing of digital images has shown good potential for assistance in the early detection of melanomas. In this sense, the objective of the current study was to develop a software for clinical images processing and reach a score of accuracy higher than 95%. The ABCD rule was used as a guide for the development of computational analysis methods. MATLAB was used as programming environment for the development of the processing of digital images software. The images used were acquired from two banks of free images. They included images of melanomas (n=15) and nevi images (not cancer) (n=15). Images in RGB color channel were used, which were converted to grayscale, 8x8 median filter applications and 3x3 neighborhood approach technique. After, we proceeded to the binarization and inversion of black and white for later extraction of contour characteristics of the lesion. The classifier used was an artificial neural network of radial basis, getting accuracy for diagnosis of melanomas images of 100% and of 90.9% for not cancer images. Thus, global correction for diagnostic prediction was 95.5%. An area under the ROC graph 0.967 was achieved, suggesting a great diagnostic predictive ability. Besides, the software presents low cost use, since it can be run on most operating systems used nowadays.

**KEYWORDS:** Skin cancer. Processing of digital images. Computer vision. Automatic diagnosis.

## INTRODUCTION

Currently, cancer is seen as a result from a variety of causative factors, essentially those relating to genetic, environmental and lifestyle. It is responsible for more than 13% of all causes of death in the world; approximately seven million people die from the disease (INCA, 2008; POPIM et al., 2008; FACINA, 2011; 2014). In Brazil, like worldwide, the incidence of cancer is rising along with the pace of population aging, because of the transformations that have been changing the health situation of people through the process of urbanization, way of life and patterns of consumption (INCA, 2008; FACINA, 2011; BELLINI, DOS SANTOS; OSELAME, 2013, Curitiba; FACINA, 2014).

In Brazil, non-melanoma skin cancer is the most incident cancer, being considered a serious public health problem, corresponding to 25% of all cancers in all geographic regions. The estimated number of cases for 2014/2015 is of 182 thousand new cases (FACINA, 2014). Another estimate suggests the possibility that in 2015, in Brazil, there will be 595.92 cases of skin cancer for every 100 thousand inhabitants (REZENDE, DE PAULA CORRÊA; DUARTE, 2009). Special emphasis is given to the Southern region, as a result of multiple epidemiological factors such as high sun exposure, predominance of the white race and large amount of Caucasian immigrants (ROCHA et al., 2004; MANTESE et al., 2006; INCA, 2008).

Despite low lethality, usually cases of non-melanoma skin cancer lead to the occurrence of physical deformities and ulcerations. The greatest risk is associated with the melanoma skin cancer, which accounts for only 4% of skin cancers (ROCHA, 2005; INCA, 2008; OSELAME, NEVES; SANCHES, 2013).

Melanoma skin cancer is characterized mainly by the intensification, loss or change in color of a mole or nevi. Among the non-melanoma skin cancers, basal cell carcinoma accounts for about 70% of cases and can be cured when diagnosed early and operated correctly. Squamous cell carcinoma is approximately 20% of skin cancers and is curable, when operated early (INCA, 2008; ROSADO, 2009).

In clinical practice, melanoma can be recognized using the ABCD rule. This rule was proposed in 1994 by Whilhelm Stolz, the second algorithm developed after analysis of patterns, whose objective was to simplify and improve the process of diagnosing melanoma (SOBIERANSKI

*et al.*, 2007; ROSADO, 2009). It originated after analysis of thirty-one dermatological features, of which four were listed as crucial for the diagnosis of melanoma: asymmetry (A), irregularities of borders (B), color (C) and diameter (D). Feature A is important because recent lesions grow unevenly, while feature B refers to the jagged borders caused by uneven growth. Color of lesion is an important characteristic, since irregular growth causes new shades of black and light brown and dark (C). Besides, ABC lesions greater than 6 mm are considered suspicions (D) (GOLA ISASI, GARCÍA ZAPIRAIN; MÉNDEZ ZORRILLA, 2011; OSELAME, NEVES; SANCHES, 2013).

Pigmented lesions are sometimes difficult to diagnose just by their clinical characteristics, so it is necessary to use criteria for greatest accuracy, what can be achieved by a method named dermoscopy (SOYER et al., 2001; MIOT, PAIXÃO; WEN, 2005). Such method, also known as epiluminescence or surface microscopy, allows the visualization of structures located below the stratum corneum. Its main indication is establish the diagnosis of pigmented lesions of skin, aiming at the detection of melanoma in early stages (REZZE et al., 2006). It is estimated that the diagnosis performed by dermatologist without using dermatoscope accessories gives accuracy between 75 to 80%, while its use improves accuracy rate to about 90% (MENZIES et al., 2001; CHRISTENSEN et al., 2010).

Therefore, the knowledge of suspicious lesions or premalignant is essential to combat the disease. Especially the melanocytic nevi, which are pigmented, flat, brownish and of variable size lesions, which can be present since birth or appear during childhood or youth. In addition, there are the dysplastic nevi, which are melanocyte nevi, with special clinical and histological features, usually located in the trunk, maculopapular-shaped with ill-defined borders, irregular, with variable pigmentation, multiple, that arise during the period of adolescence until about 30 years-old (INCA, 2008; ROSADO, 2009). These lesions have high potential for the development of skin cancer (INCA, 2008).

In order to assist in the early detection of skin cancer, identifying injuries with potential malignancy, some studies have developed processing of digital image techniques to perform the automatic or semi-automatic diagnosis of these lesions (MANOUSAKI et al., 2006; CHRISTENSEN et al., 2010; PAROLIN, HERZER; JUNG, 2010; GOLA ISASI, GARCÍA ZAPIRAIN; MÉNDEZ ZORRILLA, 2011).

The study of Christensen *et al.*, (2010) presented 86% of sensitivity and specificity of 70.3% evaluating form, edges and color of the lesions. On the other hand, the study of Manousaki et al., (2006), has evaluated geometry, color and texture, reaching only 60.9% of sensitivity and specificity of 95.4%. Rubegni et al. (2002) have evaluated geometry, color, texture and color Islands, achieving sensitivity and specificity of 94%. The method proposed by Iyatomi et al., (2008) peaked sensitivity and specificity of 86%, based on ABCD rule. The study of Gangster et al., (2001) also took into account the ABCD rule and hit 73% of sensitivity and specificity of 89%. However, it is important to note that a score of accuracy greater than 95% (minimum probability considered in the area of health) on diagnostic prediction for this type of evaluation has not be reached yet. Thus, the objective of the present study was to develop a software for processing of dermatological digital images for automatic diagnosis based on ABCD rule to achieve a score of accuracy greater than 95% of cases.

**MATERIAL ANDA METHODS**

This was an analytical type study with quantitative data treatment. A computational analysis was used based on ABCD rule of dermatological images acquired through digital photography of hyperpigmented lesions, nevi and melanomas. The images used in the study came from a free access image data bank. Thirty-seven (37) images of melanoma-type skin cancer and 26 images of skin lesions not classified as cancer, considered nevi, were selected. The images of melanomas were taken from image data bank DermIS (*Dermatology Information System*) (n=25) and DermAtlas (*Interactive Dermatology Atlas*) (n=12). The not-cancer images were all taken from the image data bank DermIS (*Dermatology Information System*).

The criteria for the images inclusion were: positive diagnosis confirmed for skin cancer, noise-free images, shadows and angles in the procurement process. Among the images available in both banks of images, 22 images of melanomas and 11 images not cancer were no longer included in the study for presenting patterns that could interfere with the digital processing. Thus, the following exclusion criteria were adopted: shadows or objects in the background during the acquisition of images, angles and light reflexes. At last, the final sample for this study was composed of 15 melanoma images and 15 images that were not related cancer. We analyzed

the patterns of asymmetry (A), border irregularity (B), color (C) and diameter (D). Figure 1 illustrates

the stages of the process proposed in the software.
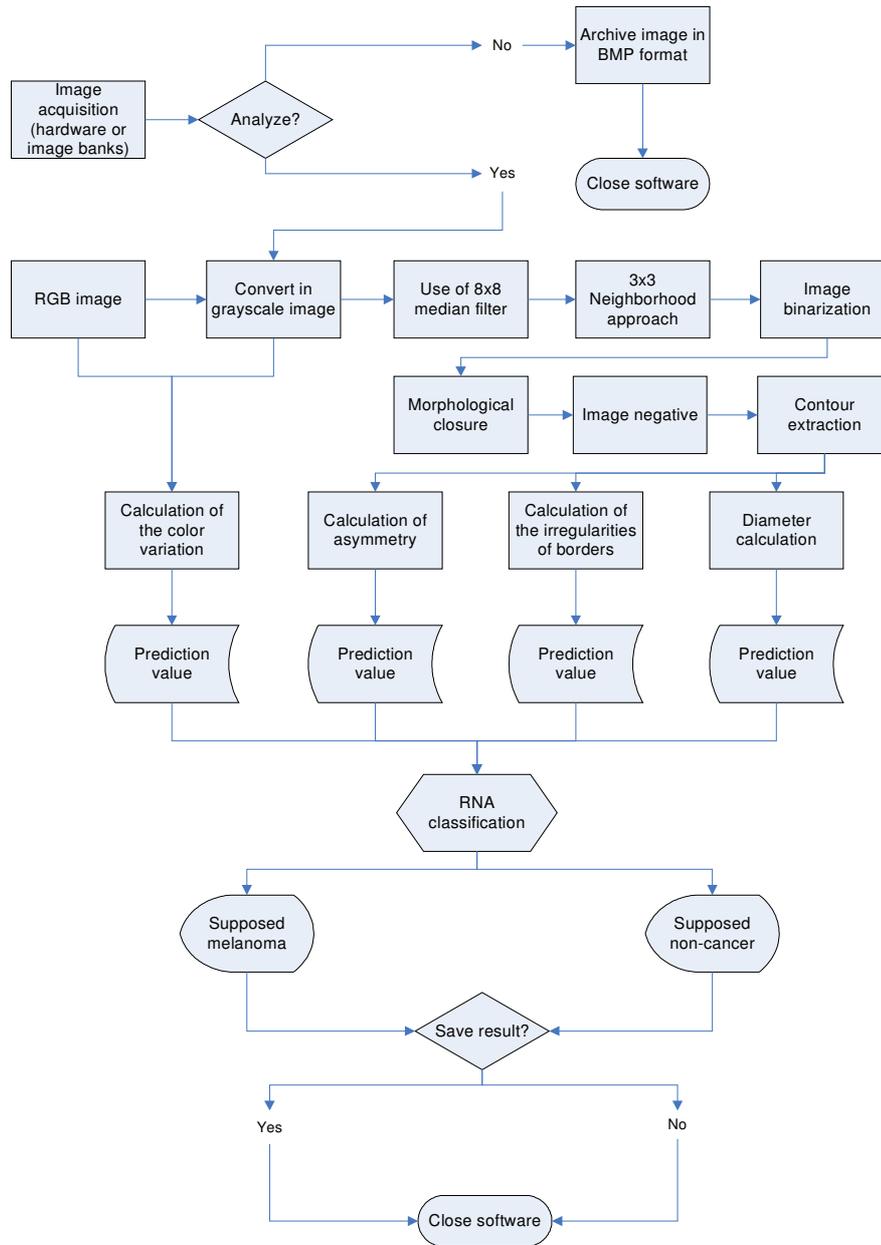


**Figure 1.** Stages of the method proposed for image processing.

**Asymmetry (A)**

The asymmetry was evaluated by the method proposed by Parolin et al. (2010). In this method, asymmetry was calculated for each axis that passes through the center of the lesion, being considered asymmetric by the smallest value considering all axes (PAROLIN, HERZER; JUNG, 2010). It is considered the centroid of the lesion (c) and its perimeter containing N pixels (Pi). For each point (Pi) the Euclidean distance (di) of the lesion's centroid is measured. This method is calculated as shown in Equation 1:

$$d_i = \|P_i - c\|, i = 1, \dots, N \tag{1}$$

In order to obtain the asymmetry judgment of lesion (Aj) around a point (Pi). Around an axis that passes through c and Pi it's necessary to calculate the average of the distances (di) between the left and right points. This model is expressed in Equation 2:

$$Aj = \frac{2}{N} \sum_{i=1}^{N/2} |d_{j-i} - d_{j+i}|, j = 1, \dots, N \tag{2}$$

The rate of asymmetry of the lesion is given by the smallest value of Aj (expressed in Equation 3):

$$A = \min_j Aj$$

(3)

Another method used from the ABCD rule, conducted by Gola Isasi *et al*., (2011) the asymmetry was expressed as shown in Equation 4:

$$\frac{\sum_{j=1}^{n/180}\left(Pj/\left(\frac{T}{2}\right)\right)}{\left(\frac{n}{180}\right)}$$

(4)

Where:
N represents the degrees of rotation;
P indicates the value of a pixel;
T indicates the number of pixels in the lesion.

The calculation of the asymmetry of the lesions by two straight lines drawn on the image, called long diagonal and its respective largest perpendicular was also used. These points were generated from the more distant points between the edges of the image. This method was based on the *Dermatoscopy Point Value* (DPV) that aims to assist the dermatologist during the diagnosis process of pigmented lesions. It takes into account the result of the ratio of the perpendicular segment and the long diagonal, where the score is quantified in (0 to 2 points). If the ratio is greater than 40 (0 point), it is a symmetric lesion. If the proportion is among 20 to 40 (1 point), it is a quite asymmetrical lesion. If the ratio is less than 20 (2 points), we have a very asymmetric lesion (SOBIERANSKI et al., 2007; ROSADO, 2009).

This score is applied in a score to classify the lesion in benign, malignant or suspicious, and to evaluate other characteristics of the ABCD rule. The method considered the two biggest segments of lines generated in the image. The distance in pixels was obtained between the points generated on the line. In order to get the value of the proportion, the value in pixels of the greatest perpendicular line was divided by the value in pixels of the longest diagonal line.

**Irregularity of edges (B)**

The edges of the lesion bring out aspects of both regularity and irregularity, meaning the possibility of benign or malignant lesions, respectively. The method proposed by Kusumoputro and Ariyanto (1998) and Parolin *et al*., (2010) was used, in which the irregularity of the edges was calculated according to Equation 5:

$$I = \frac{N^2}{4\pi A}$$

(5)

Where:
N is perimeter of the lesion in pixels;
A is the area of the lesion in pixels.

Another method for assessing the irregularity of borders was proposed in the study of Gola Isasi *et al*., (2011). Equation 6 illustrates this method developed with the selected images.

$$\frac{\sum_{j=1}^{n/360}(rj - rmin)/(rmax - rmin)}{\left(\frac{n}{360}\right)}$$

(6)

Where:
n refers to the degrees of rotation;
r is the distance of a central pixel to the border of the lesion.

In order to evaluate the irregularity of the borders, it was used a division of the image into eight segments of straight lines from the center. This method aimed to verify the existence of a significant difference between the diameters of these lines in its symmetric opposites, in melanomas lesions and non-cancer images.

**Color (C)**

The color variation of the lesion is characterized among all the pixels of the image, calculated for each color component of the RGB channels (Red, Green and Blue). It was used the method proposed by Parolin *et al*., (2010), in which the variation of these colors is expressed according to Equation 7:

$$\sigma_c^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^c - \mu c_{RGB})^2, c \in \{R, G, B\}$$

(7)

Where:
$x_i^c$ represents the value of the c component of pixels from i in the RGB color channels;
$\mu c_{RGB}$ relates to the average value of the same color component;
*n* represents the number of pixels in the lesion.

Another method proposed for the evaluation of the variation of color images in lesions used the intensity averages of the skin lesion in each color channel, for both the images of skin cancer and non-cancer. In order to designate the color variation the statistical method of amplitude calculation was employed.

The color variation was also examined using the images converted to grayscale. This choice is justified by the fact that the images have been extracted from image banks, and have not received the same treatment of lighting at the time of purchase. This way, in order to minimize the brightness interferences of the color channels, an option was to use grayscale processing with the calculation of amplitude.

**Diameter (D)**

The diameter was estimated in the study by calculating the area/perimeter ratio (Equation 4), whereas the lesions present circular shape. The perimeter of a closed flat figure is the length of the line that limits the figure. The Equation 8 expresses this calculation.

$$R\frac{A}{P} = \frac{r^2}{2r}$$

(8)

Where:
R represents ray
A represents area
P represents perimeter

The software has processed the images in order to identify the perimeter and its respective area in pixels. The method used considered that area grows exponentially in relationship to the radius. On the other hand, the perimeter grows linearly, thus, it is assumed that in lesions with circular format (non-cancer) the ratio area/perimeter assumes a greater value in lesions of larger radius than the lesions with shorter radius value.

**Statistical Analysis**

The statistical analysis between imaging groups was held in the software Statistical Package for the Social Sciences (SPSS) version 21. Descriptive statistics was used to compare the groups and the non-parametric test for two independent samples of Mann-Whitney U (*p value* 0.05). In order to choose the methods with better sensitivity and specificity it was used the *Receiver Operating Characteristic* (ROC). Thus, the method with better area results under the curve for each variable analyzed was selected to compose the classification system using neural network with radial basis activation (Radial Basis Function Neural Network), available in MATLAB.

**RESULTS AND DISCUSSION**

**Initial Processing of Images**

Initially the image was processed to extract features that could interfere in the analysis, such as shadows, hairs and other noises from the acquisition process. It should be noted that the image was converted to grayscale for further processing. Then we applied the 8x8 median filter, aiming at mitigating noise with preservation of the borders. This type of filter analyzes the values of neighboring pixels compared to the pixel being calculated. This means that the pixels are ordered in increasing values, with the objective of choosing the median of those values as a new value for the pixel that is being calculated.

Later It was used the technique of approach by neighborhood. This method used the 3x3 approach, aiming to improve the contrast of the image. Sequentially the image was binarized. This method transforms the grayscale image into black and white. To allow the processing only of the area of the skin lesion, the white background was reverted to black and the lesion from black to white (Figure 2).

The present study used the images in the RGB system, with conversion to grayscale. The RGB system was chosen for preserving original features of the image, with the aim of comparing human analysis of images in future works. In addition, it reduces the impact of light during the acquisition of the images, since the images used were from an images data bank.

A similar process was used in the study of Candeias and Silva (2008), in which the original RGB image has been converted to grayscale and was used a 8x8 median, for further segmentation (CANDEIAS; SILVA, 2008). The segmentation has been conducted by means of thresholds in the histogram for further binarization.

In this work, the approach by neighborhood technique was used. This method used the 3x3 approach, aiming to improve the contrast of the image and improve the contour of the lesion, as also expressed in the study of Gola Isasi et al., (2011). After this process, it happened the filtered image binarization and the subsequent extraction of contour characteristics of the lesion (GOLA ISASI, GARCÍA ZAPIRAIN; MÉNDEZ ZORRILLA, 2011). The method proposed in this study proved to be effective in the process of segmentation and extraction of contours of the selected images in the image data banks.
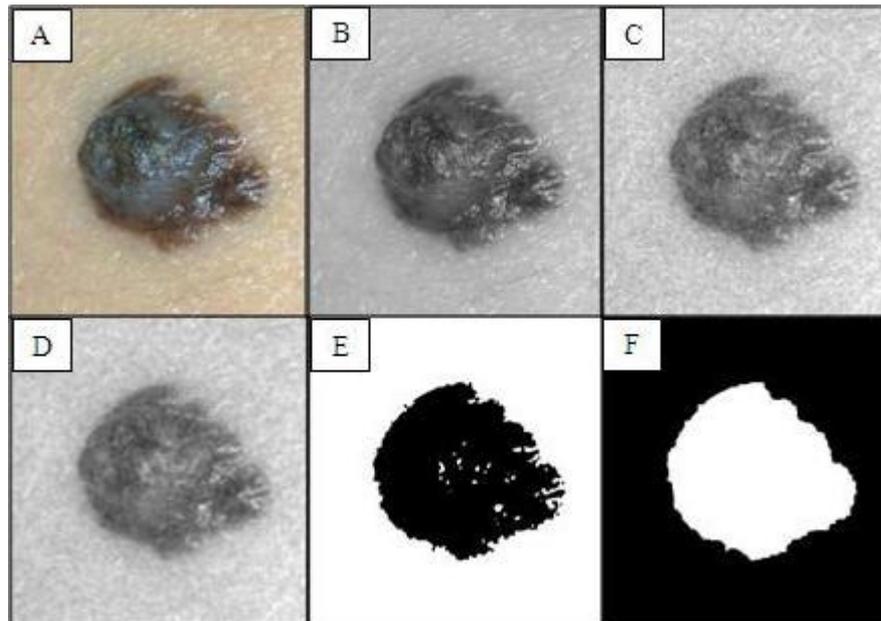
**Figure 2. Steps for processing the images. (A)** Original image; (**B**) Conversion to grayscale; **(C)** Use of the median filter; **(D)** Use of the 3x3 neighborhood approach filter; **(E)** Binary image; **(F)** Morphological Closing, and image negative. Original image source: http://dermis.net.

**Automatic classification using Artificial Neural Network**

To sort the images in melanomas and non-cancer we decided for the use of neural network with radial basis activation (Radial Basis Function Neural Network), available in MATLAB. Neural networks with radial basis activation functions (RBF) feature three main differences in relation to Perceptron-type multilayer networks: always feature a single intermediate layer; output neurons are always linear; intermediate layer neurons have only radial basis functions, as the activation function (VON ZUBEN; ATTUX, 2001; KOVÁCS, 2002).

The immediate consequence of using radial basis activation functions is in the way the entries are processed by intermediate layer neurons. Instead of internal activation of each middle layer neuron happening by the employment of the scalar product between the input vector and the vector of weights, as in the case of the *Perceptron*, it is obtained from a weighted norm of the difference between the two vectors (VON ZUBEN; ATTUX, 2001; KOVÁCS, 2002).

**Methods selection for composition of classifier**

The described methods have been assessed with the ROC graph. Table 1 expresses the area under the curve between the variables.

RBF networks are local learning networks, so that it is possible to get a good approximation since a sufficient number of data for training is provided in the area of interest. In contrast, multilayer Perceptrons are "global" learning networks (due to the nature of the activation functions) that make global effect approximations in compact regions of the space approach. Neural networks with ability of local approach are very efficient when the dimension of the input vector is reduced (KOVÁCS, 2002).

Therefore, we used the RBF classifier, where variables were associated with better results obtained with the ROC graph analysis, including: (A) proposal by Parolin *et al.*, (2010), (B) irregularities of borders by the distance from the center to the edges with eight straight segments; (C) amplitude of *Red, Green and Gray* and (D) by the estimate area/perimeter ratio. Table 2 presents the cases used for training and testing of the RBF. Subsequently, the RBF has classified the images according to the input variables in melanoma and non-cancer. Table 3 shows the score of accuracy in the training group and in the test group. To verify the sensitivity and specificity of the test, the ROC graph was generated. Figure 3 displays the result.

**Table 1.** Analysis of the ABCD variables with the ROC curve graph.

| Variable (s) of test result | Area Under the Curve |
|---|---|
| (A) Gola Isasi *et al*., (GOLA ISASI, GARCÍA ZAPIRAIN e MÉNDEZ ZORRILLA, 2011) | 0.414 |
| (A) Parolin *et al*., (PAROLIN, HERZER e JUNG, 2010) | 0.700 |
| (A) Dermatoscopy Point Value (DPV) | 0.443 |
| (B) Kusumoputro and Ariyanto (KUSUMOPUTRO e ARIYANTO, 1998); Parolin *et al.,* (PAROLIN, HERZER e JUNG, 2010) | 0.519 |
| (B) Irregularity of borders | 0.524 |
| (B) Gola Isasi *et al*., (2011) | 0.429 |
| (C) Amplitude *Red* | 0.807 |
| (C) Amplitude *Green* | 0.724 |
| (C) Amplitude *Blue* | 0.636 |
| (C) Parolin *et al*., (PAROLIN, HERZER e JUNG, 2010) *Red* | 0.686 |
| (C) Parolin *et al*., (PAROLIN, HERZER e JUNG, 2010) *Green* | 0.567 |
| (C) Parolin *et al*., (PAROLIN, HERZER e JUNG, 2010) *Blue* | 0.452 |
| (C) Amplitude *Gray* | 0.721 |
| (D) Area/Perimeter ratio | 0.586 |

**Table 2.** Cases targeted for training and testing on RBF.

| | N | Percentage |
|---|---|---|
| Training | 8 | 26.7% |
| Tests | 22 | 73.3% |
| Total | 30 | 100.0% |

**Table 3.** Result of classification of images using RBF.

| Sample | Observed | Melanoma | Non-cancer | Right Percentage |
|---|---|---|---|---|
| Training | Melanoma | 4 | 0 | 100.0% |
| | Non-cancer | 0 | 4 | 100.0% |
| | Global hit | 50.0% | 50.0% | 100.0% |
| Tests | Melanoma | 11 | 0 | 100.0% |
| | Non-cancer | 1 | 10 | 90.9% |
| | Global hit | 54.5% | 45.5% | 95.5% |

Dependent variable: Group

Receiver Operating Characteristic - ROC
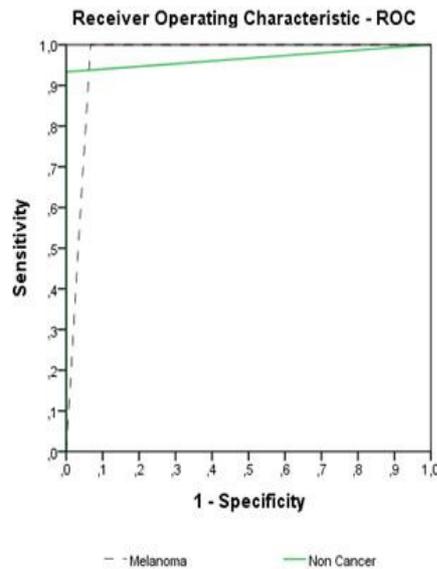


**Figure 3.** ROC curve graph result of classification by RBF for groups of images. The test sensitivity was 100.0% and specificity was 90.9%.

**Assessment of asymmetry (A)**

For assessment of asymmetry, the method described by Gola Isasi et al*., (2011) has been chosen. The equation applied to analyze the images in the present study was not able to differentiate melanomas from non-cancer images groups, thus there was no statistically significant difference between the groups (0.663; *p value* 0.05). The study used as a reference for calculating the asymmetry does not present the statistical description concerning the image group used to validate the equation. It should also be noted that only 30 images were used in this study, while Gola Isasi et al., utilized 160 images.

Another method described was the calculation of the Dermatoscopy Point Value (DPV). This method has been described in the study of Sobieranski et al., (2007), and takes into account the ratio between the longer diagonal and its respective perpendicular in the internal area of the lesion. In the 30 selected images, there was no statistically significant difference between groups (0.663; *p value* 0.05). Thus, the images were classified as symmetrical (above 40%) by the technique, both in melanomas (average = 43.181%) as in non-cancer images (average = 44.707%).

The method described by Parolin et al., (2010) was the one that showed better result, but with no real significant difference from the other methods (0.085; p value 0.05). In this study, statistical analyses for the validation of the equation were not presented.

Araújo (2010), calculated the diameter of the lesion been by searching for the pair of more distant points in its outline. Thus, a perpendicular

line was generated, aiming to identify the amount of spicules present in the lesion (ARAUJO, 2010).

The criteria used for selecting the method to be adopted as a component in the system of classification by artificial neural network has considered the best sensitivity (SE) and specificity (SP) described in the ROC graph. Thus, the method proposed by Parolin et al., (2010) got the best area under the curve (0,684).

**Evaluation of border irregularity (B)**

The equation proposed by Kusumoputo and Ariyanto (1998) and Parolin et al. (2010) assessed the irregularity of borders. Parolin *et al*. (2010) used an image bank with 290 images, while the present study used 30 images. It did not demonstrate a statistically significant difference between the groups of images (0.852; *p value* 0.05).

Another method was the one proposed by Gola Isasi et al., (2011), which did not show statistically significant difference between groups (0.513; *p value* 0.05).

In this study, a method to determine the irregularity of lesion borders was proposed. The method passes through the image and highlights eight segments of straight lines from the center. Supposedly, the more irregular is the lesion, the greater is the difference between the opposite segments. The distance of the segments between the center and the border was calculated in pixels.

It is important to note the difficulty encountered in this process because the images were originated from image banks, with distinct centralizations during the acquisition process. This could cause error in the determination of the

geometric center of the lesion. Thus, this method did not seem to be effective, with no statistically significant difference between analyzed groups (0.820; *p value* 0.05). It is expected that in future works, with standardized acquisition, the technique can be applied to a larger number of images.

Similar difficulty was described by Rosado (2009), who has used the calculation of centric for analysis of the asymmetry, seeking for the geometric center of the lesion. But in the same way as in this present study, various images were omitted from the initial set due to the fact that the lesion was not fully contained in the image. Thus, the author emphasizes the need for the development of routines that perform a step of selecting and acquiring appropriate images. In this sense, with a small number of images, the algorithm responsible for the extraction of the asymmetry presented error probability of 39% (ROSADO, 2009).

It should be noted that in this study a standardized method for acquisition of dermatological images was proposed and validated, what presented positive results regarding the reduction of interference at the time of the acquisition, especially those relating to focal length and brightness (NEVES; OSELAME, 2012; OSELAME, NEVES; SANCHES, 2013). An error margin was obtained for 1 (one) cm at the 0.03 mm image. The accuracy of the method described, with variation coefficient of 0.005 and margin of error of ca. 0.03 mm for each 1 cm was still lower than the one reported by Feliciano (FELICIANO, SOUZA; LETA, 2010), who reported the possibility of 0.25 mm of error in automated dimensioning for computer vision.

The proposed method establishes the use of photogrammetry, whose purpose is to obtain measurements of the shape and the dimensions of bodies or their parts by digital images. Various health sectors have been incorporating the technology of digital images. Studies with the use of photographic images as a tool are described in: diagnosis of ventilatory patterns in children (RIPKA et al., 2012); diagnosis of postural deviations (VACARI et al., 2012); determination of flexibility of the trunk (PERIN et al., 2012).

Thus, it should be noted that the choice of automatic image measurement techniques requires a detailed evaluation of the lighting conditions and of the object to be measured, reducing the so-called type B uncertainties, which are determined from ancillary information and also external to the measurement procedure (FELICIANO, SOUZA; LETA, 2010). It is expected, therefore, with the use of prototype developed to reduce the influence of

interfering variables during the acquisition process of dermatological images.

Therefore, among the three methods, the assessment of the irregularity of the border from the distance from the center of the lesion toward the edges was the method which showed improved ability for classification images, obtaining the best area under the ROC graph (0.524).

**Evaluation of color variation (C)**

The evaluation of color in RGB channels was calculated using the average, maximum and minimum intensity of each channel (Red, Green and Blue). We opted for this model because it is the most widely used for cameras and video monitors, not requiring conversion for software analysis.

Based on maximum and minimum intensity, we got the value of the amplitude, which was used in this study to identify the variation in both groups of analyzed images. Thus, the amplitude in the color channel Red was that one that presented the best result (0,003; *p value* 0.05), followed by the Green amplitude (0.24; *p value* 0.05). The method proposed by Parolin et al., (2010) showed no significant variation in RGB color channels (*Red* 0.59; *Green* 0.373; *Blue* 0.885; *p value* 0.05).

The color variation was expressed in the form of histograms. The histogram is a representation of the frequency distribution of a set of numbers within a selected context. Is represented by the number of pixels in the image with values assigned to its intensity (GONZALEZ; WOODS, 2002).

Initially the histogram provides global data of the image, represented by intervals between the gray levels in the image (ACHARYA; RAY, 2005a). Some indicators may appear in the histogram, as level of intensity, average, standard deviation and other useful information to the processing (ACHARYA; RAY, 2005b).

Another method employed was the analysis in grayscale. This method is justified by the RGB color channel being sensitive to the environmental interferences in the process of capturing images. The amplitude *Gray* also proved to be a good option for the analysis of color variation (0.41; *p value* 0.05).

In a similar study, the color extraction algorithm presented high probability of error in RBG channel, reaching ca. 40%.(ROSADO, 2009). Therefore, the option of associating the grayscale analysis should be considered.

Thus, to compose the classifier, the Red amplitude (area under the curve 0.820), the Green amplitude (area under the curve 0.742) and the Gray

amplitude (area under the curve 0.718) were considered by the better performance expressed in the ROC graph.

**Evaluation of the diameter (D)**

The method proposed in the present study, which relates the area and the perimeter, both in pixels, also showed no statistically significant differences between the groups of images (0,633; *p* value0,05). The method uses active contours from a threshold RGB color channels. It was used as the default for the images the threshold T = 120 to the channels *Red*, *Green* and *Blue*.

Thus, a routine was applied to determine the perimeter and the internal area of the lesion in pixels. A similar procedure was performed by Araújo (2010), where he used the method of growth of regions by Watershed, followed by the active contours model (ARAUJO, 2010).

Figure 3 shows the similarity between malignant and benign images, which makes clinical diagnosis difficult.
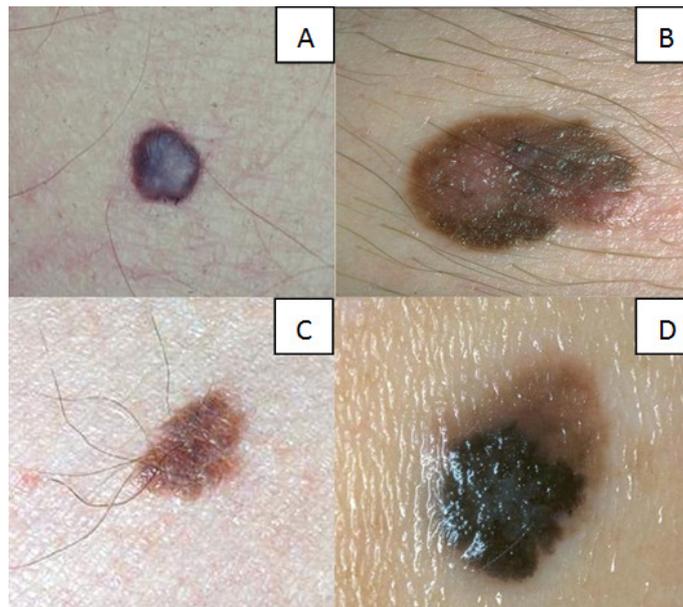


**Figure 3. (A-C)** Not melanoma; (**B-D**) Melanoma. Original image source: http://dermis.net.

**Automatic sorting using Artificial Neural Network**

Several studies have proposed the use of digital processing of images to aid diagnosis of dermatological lesions. Automated systems, with the use of the ABCD rule achieved good results when compared to human eye. The study by Gangster et al., (2001), pioneered by using the four components of the ABCD rule, involving 122 parameters. It was used the classifier called "K nearest neighbors (KNN)", which allows you to sort the image through a process based on closer training standards (GANSTER et al., 2001).

The KNN grading method considers the relationship of images with normal and abnormal patterns. It is a non-parametric method of determination that sorts an unknown image through the calculation of a metric, that is the measurement of distances. These metrics are vector distances which are associated to an unknown image to be classified and to every known image of the database previously classified under normal or abnormal

(GANSTER et al., 2001; CHRISTENSEN et al., 2010).

In 2002, Rubegni *et al*., (2002) reached sensitivity (SE) of 94.3% and specificity (SP) of 94.3%. In this study, we used an image data bank with 550 images, of which 350 were nevi (non-cancer) and 200 were melanomas. The adopted classifier was the artificial neural network (RUBEGNI et al., 2002).

Taouil and Romdhane (2006) aimed at an automatic system for segmentation and diagnosis of melanomas. 62 images were analyzed between benign and malignant, in which obtained a hit rate between 89.5% to 95.8% using artificial neural network (TAOUIL; ROMDHANE, 2006).

The neural networks method involves the process of mapping a large number of entries and a small number of outputs, and it is often applied in cases of classification of information. In these cases, the predictors form entries and a set of variables, which will generate the membership of a group that will represent the outputs. Thus, neural networks are

alternative instruments for the discriminate analysis of variables (RUBEGNI et al., 2002; IYATOMI et al., 2008; GOLA ISASI, GARCÍA ZAPIRAIN; MÉNDEZ ZORRILLA, 2011).

In the study conducted by Blum *et al.*, (2004), they used the regressive logistic model as classifier, obtaining SE of 82.3% and SP of 86.9%, in an image data bank with 837 images, being 753 nevi and 84 melanomas. (BLUM et al., 2004).

Melli et al., (2006) proposed segmentation by clustering (grouping by color) associated with an automatic system for border detection. In this study it was not declared the rate of diagnosis hit. However, certain values of sensitivity and specificity were determined to the proposed processing type, reaching between them the average of 0.96 (MELLI, GRANA; CUCCHIARA, 2006).

In the study conducted by Celebi et al., (2007), they used the Support Vector Machine, reaching 93.3% from SE and 92.3% of SP, in a set of 564 images, among them 476 nevi and 88 melanomas. In another study, conducted by Rosado (2009), the classifier method used was the Support Vector Machine, based on the principles of statistical learning, aiming at greater margin of separation between classes. The global correction for the melanoma images was of 92.7% and of 86% for benign and non-melanoma.

The Support Vector Machine consists of a technique based on statistical learning theory, which seeks to maximize the generalization ability through methods capable of extracting knowledge through data samples (CELEBI et al., 2008; SOARES, 2008).

In 2008, Iyatomi *et al.* (2008) proposed a classification system using the internet. In this project, any image uploaded on the page would be classified, and in addition, could form a public bank of images. The results showed a SE of 85.9% and a SP of 86.0%.

Parolin et al., (2010) scheduled a system to detect melanomas based on skin segmentation by WT (Transformed Wavelet) associated with the automatic detection of borders. Two hundred and ninety images were evaluated (139 benign images and 151 malignant images). They reached a correction diagnosis for melanomas of 88.41% (PAROLIN, HERZER; JUNG, 2010).

The transformed wavelet has been used in the processing of signals to improve its ranking. It is a tool to extract relevant characteristics of the sign for the improvement of classifiers to extract characteristics through the decomposition of images at different frequencies. It can be applied specifically to images with concentrated energy in low frequency regions (SOARES, 2008; PAROLIN, HERZER; JUNG, 2010).

The study proposed by Christensen et al., (2010) aimed to achieve a degree of accuracy of at least 64%, the same percentage estimated by dermatologists with analysis to the naked eye. Ninety-seven images were evaluated (43 were benign and 54 were malignant) from an image data bank. It has reached a degree of IF of 86% and SP of 70.37% on semi-automatic prediction in diagnosing melanomas. The rate of global hit for diagnosis was of 77% for melanomas. The analysis method used was the ROC graph.

Gola Isasi et al. (2011) performed the computational technique integrated to the ABCD rule with the analysis of 160 images. The present study developed an interface for the demonstration of the evaluated parameters, in which the risk predictive values for melanomas were presented. The method managed to hit 85% diagnosis for melanomas by pattern recognition in MATLAB.

Variables of asymmetry, border irregularity, color variation and diameter were associated. Among the methods proposed in the literature, the best individual results achieved by statistical analysis were developed in this study.

The calculation of the distance from the center of the lesion to its edges, generated by eight straight segments was used for the analysis of asymmetry. To assess irregularities of edges, the area/perimeter ratio was used. In addition, in the matter of color variation the Red, Green and Gray amplitudes. These variables were included in the grading method. The method got diagnostic correction for the melanoma images tested (n = 11) of 100%. On the other hand, for the non-cancer images (n = 11), the percentage of settlement was of 90.9%. Thus, the global correction for diagnostic prediction was of 95.5%.

In relation to the SE and the SP of the proposed method, it has obtained an area under the ROC curve of 0.967, resulting in a great ability of images classification.

## CONCLUSION

The software developed showed good capability of digital image processing, especially in the discrimination among images classified as melanoma and non-cancer with the use of artificial neural network of radial basis, reaching a global correction of 95.5% in diagnostic prediction. Therefore, joining the human analysis through sight with the prediction by computational means can improve the fight against skin cancer, mainly

providing the longitudinal follow-up of these          the medium and long terms.
lesions, ensuring the possibility of comparisons in

**RESUMO:** O câncer é responsável por cerca de 7 milhões de óbitos anuais em todo o mundo. Entre eles, o tipo melanoma, responsável por 4% dos cânceres de pele, cuja incidência dobrou mundialmente nos últimos dez anos. O processamento digital de imagens tem mostrado um bom potencial para auxiliar no diagnóstico precoce de melanomas. Neste sentido, objetivo do presente estudo foi desenvolver um *software* para processamento digital de imagens clínicas para diagnóstico automático baseado na regra ABCD que alcançasse um percentual de acerto maior do que 95% dos casos. Utilizou-se como norteador a regra ABCD para o desenvolvimento de métodos de análise computacional. Empregou-se o MATLAB como ambiente de programação para o desenvolvimento de um *software* para o processamento digital de imagens. As imagens utilizadas foram adquiridas de dois bancos de imagens de acesso livre. Foram inclusas imagens clínicas de melanomas (n=15) e imagens de nevos (lesão melanocítica benigna) (n=15). Utilizaram-se imagens no canal de cor RGB, as quais foram convertidas para escala de cinza, aplicação de filtro de mediana 8x8 e técnica de aproximação por vizinhança 3x3. Após, procedeu-se a binarização e inversão de preto e branco para posterior extração das características do contorno da lesão. O classificador utilizado foi uma rede neural artificial de base radial, obtendo acerto diagnóstico para as imagens melanomas de 100% e para imagens benignas de 90,9%. Desta forma, o acerto global para predição diagnóstica foi de 95,5%. Obteve-se uma área sob a curva ROC de 0,967, o que sugere uma excelente capacidade de predição diagnóstica, sobretudo, com baixo custo de utilização, visto que o *software* pode ser executado na grande maioria dos sistemas operacionais hoje utilizados.

**PALAVRAS-CHAVE:** Câncer de pele. Processamento digital de imagens. Visão computacional. Diagnóstico automático.

**REFERENCES**

ACHARYA, T.; RAY, A. K. **Image Enhancement and Restoration**. New Jersey: Wiley Online Library, 2005a. 105-129 ISBN 0471745790. https://doi.org/10.1002/0471745790.ch6

_____. **Image formation and representation. Image processing: principles and applications**. New Jersey: John Wiley & Sons, 2005b. ISBN 0471745782.

ARAUJO, A. F. D. **Método para extração e caracterização de lesões de pele usando difusão anisotrópica, crescimento de regiões, watersheds e contornos ativos**. 2010. Dissertação Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto. https://doi.org/10.5016/dt000607597

BELLINI, V. B. S.; DOS SANTOS, C.; OSELAME, G. B. Fatores de risco e proteção para câncer de mama na mulher. **Revista UNIANDRADE,** v. 14, n. 1, p. 45-64, 2013, Curitiba. ISSN 1519-5694.

BLUM, A. et al. Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. **British Journal of Dermatology,** v. 151, n. 5, p. 1029-1038, 2004. ISSN 1365-2133.

CANDEIAS, A. L. B.; SILVA, M. G. N. Um novo algoritmo para detecção de contorno de melanomas em imagens de pele. Brazilian Symposium on Computer Graphics and Image Processing, 2008, Campo Grande. Anais Eletrônicos. p.9-13.

CELEBI, M. E. et al. Automatic detection of blue-white veil and related structures in dermoscopy images. **Computerized Medical Imaging and Graphics,** v. 32, n. 8, p. 670-677, 2008. ISSN 0895-6111.

CHRISTENSEN, J. H. et al. Pre - diagnostic digital imaging prediction model to discriminate between malignant melanoma and benign pigmented skin lesion. **Skin Research and Technology,** v. 16, n. 1, p. 98-108, 2010. ISSN 1600-0846.

FACINA, T. Estimativa 2012–Incidência de Câncer no Brasil. **Revista Brasileira de Cancerologia,** v. 57, n. 4, p. 557,  2011.

_____. Estimativa 2014–Incidência de Câncer no Brasil. **Rev. Bras. Cancerol. ,** v. 60, n. 1, p. 63,  2014.

FELICIANO, F. F.; SOUZA, I. L.; LETA, F. R. Visão computacional aplicacada à metrologia dimensional automatizada: considerações sobre sua exatidão. **Engevista,** v. 7, n. 2,  2010. ISSN 1415-7314.

GANSTER, H.  et al. Automated melanoma recognition. **Medical Imaging, IEEE Transactions on,** v. 20, n. 3, p. 233-239,  2001. ISSN 0278-0062.

GOLA ISASI, A.; GARCÍA ZAPIRAIN, B.; MÉNDEZ ZORRILLA, A. Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. **Computers in Biology and Medicine,** v. 41, n. 9, p. 742-755,  2011. ISSN 0010-4825.

GONZALEZ, R. C.; WOODS, R. E. **Digital image fundamentals**. 2. New Jersey: Prentice, 2002. 52-54.

INCA. **Instituto Nacional do Câncer (INCA). Ações de enfermagem para o controle do câncer: uma proposta de integração ensino-serviço**. Rio de Janeiro: 3ª 2008.

IYATOMI, H.  et al. An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm. **Computerized Medical Imaging and Graphics,** v. 32, n. 7, p. 566-579,  2008. ISSN 0895-6111.

KOVÁCS, Z. L. **Redes neurais artificiais**. 4ª. São Paulo: Editora Livraria da Fisica, 2002.  ISBN 8588325144.

KUSUMOPUTRO, B.; ARIYANTO, A. Neural network diagnosis of malignant skin cancers using principal component analysis as a preprocessor. Neural Networks Proceedings. IEEE World Congress on Computational Intelligence. , 1998, Anchorage, AK p.310-315. https://doi.org/10.1109/ijcnn.1998.682283

MANOUSAKI, A. G.  et al. A simple digital image processing system to aid in melanoma diagnosis in an everyday melanocytic skin lesion unit. A preliminary report. **International journal of dermatology,** v. 45, n. 4, p. 402-410,  2006. ISSN 1365-4632.

MANTESE, S. A. O.  et al. Carcinoma basocelular-Análise de 300 casos observados em Uberlândia-MG Basal cell Carcinoma-Analysis of 300 cases observed in Uberlândia-MG, Brazil. **An Bras Dermatol,** v. 81, n. 2, p. 136-42,  2006.  https://doi.org/10.1590/S0365-05962006000200004

MELLI, R.; GRANA, C.; CUCCHIARA, R. Comparison of color clustering algorithms for segmentation of dermatological images. Medical Imaging, 2006,   International Society for Optics and Photonics. p.61443S-61443S-9.

MENZIES, S. W.  et al. Short-term digital surface microscopic monitoring of atypical or changing melanocytic lesions. **Archives of dermatology,** v. 137, n. 12, p. 1583,  2001.  https://doi.org/10.1001/archderm.137.12.1583

MIOT, H. A.; PAIXÃO, M. P.; WEN, C. L. Teledermatology: past, present and future. **Anais brasileiros de dermatologia,** v. 80, n. 5, p. 523-532,  2005. ISSN 0365-0596.

NEVES, E.; OSELAME, G. Software para acompanhamento de lesões dermatológicas por imagem fotográfica. XXIII Congresso Brasileiro em Engenharia Biomédica, 2012, Porto de Galinhas - PE.  Anais do XXIII Congresso Brasileiro em Engenharia Biomédica. p.932-935.

OSELAME, G. B.; NEVES, E. B.; SANCHES, I. J. Máquinas contra o câncer. **Ciência Hoje,** v. 306, p. 24-27, 2013.

PAROLIN, A.; HERZER, E.; JUNG, C. R. Semi-automated diagnosis of melanoma through the analysis of dermatological images. Graphics, Patterns and Images (SIBGRAPI), 2010 23rd SIBGRAPI Conference on, 2010,   IEEE. p.71-78. https://doi.org/10.1109/sibgrapi.2010.18

PERIN, A.  et al. Use of biophotogrammetry for assessment of trunk flexibility. **Revista Brasileira de Medicina do Esporte,** v. 18, n. 3, p. 176-180,  2012. ISSN 1517-8692.

POPIM, R. C.  et al. Câncer de pele: uso de medidas preventivas e perfil demográfico de um grupo de risco na cidade de Botucatu. **Ciênc saúde coletiva,** v. 13, n. 4, p. 1331-6,  2008.

REZENDE, M. L.; DE PAULA CORRÊA, M.; DUARTE, E. M. Estimativa do custo do tratamento de câncer de pele tipo melanoma no Estado de São Paulo–Brasil. **An Bras Dermatol,** v. 84, n. 3, p. 237-43,  2009. https://doi.org/10.1590/S0365-05962009000300004

REZZE, G. G.  et al. Structural correlations between dermoscopic features of cutaneous melanomas and histopathology using transverse sections. **The American journal of dermatopathology,** v. 28, n. 1, p. 13-20, 2006. ISSN 0193-1091.

RIPKA, W.  et al. Biophotogrammetry model of respiratory motion analysis applied to children. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, 2012,   IEEE. p.2404-2407. https://doi.org/10.1109/embc.2012.6346448

ROCHA, F. P.  et al. Marcadores e fatores de risco para queratoses actínicas e carcinomas basocelulares: um estudo de caso-controle Risk markers and risk factors for actinic keratosis and basal cell carcinoma: a case-control study. **An Bras Dermatol,** v. 79, n. 4, p. 441-454,  2004. https://doi.org/10.1590/S0365-05962004000400006

ROCHA, R. C. C. **Rastreamento de câncer de pele em um município de colonização pomerana no estado do Espírito Santo, 1997-2002; Skin cancer screening in pomerana settling county in State of Espirito Santo, 1997-2002**. 2005.   Universidade de São Paulo. Faculdade de Saúde Pública

ROSADO, L. **Sistema automático para diagnóstico de lesões cutâneas baseado em imagens dermoscópicas**. 2009.   Dissertação de mestrado em Engenharia Biomédica, Universidade Técnica de Lisboa, Portugal

RUBEGNI, P.  et al. Automated diagnosis of pigmented skin lesions. **International Journal of Cancer,** v. 101, n. 6, p. 576-580,  2002. ISSN 1097-0215.

SOARES, H. B. **Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte**. 2008.  Tese  Universidade Federal do Rio Grande do Norte, Natal.

SOBIERANSKI, A. C.  et al. Metodologia computacional para aplicaçao da regra abcd na avaliaçao de lesoes pigmentadas. X Congresso Brasileiro de Informática em Saúde, 2007, Porto de Galinhas.

SOYER, H. P.  et al. Dermoscopy of pigmented skin lesions. **EJD,** v. 11, n. 3, p. 270-276,  2001. ISSN 1167-1122.

TAOUIL, K.; ROMDHANE, N. B. Automatic segmentation and classification of skin lesion images. Distributed Frameworks for Multimedia Applications, 2006. The 2nd International Conference on, 2006, IEEE. p.1-12. https://doi.org/10.1109/dfma.2006.296918

VACARI, D.  et al. Evaluation of pelvis slope and flattening on children gymnasts by biophotogrammetry technique. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, 2012,   IEEE. p.1948-1951. https://doi.org/10.1109/embc.2012.6346336

VON ZUBEN, I. P. F. J.; ATTUX, R. R. Redes Neurais com Funções de Ativação de Base Radial. **DCA/FEEC/Unicamp**, p. 1-38,  2001.