Semi-Automated Diagnosis of Melanoma Through the Analysis of Dermatological Images

Alessandro Parolin, Eduardo Herzer
Universidade do Vale do Rio dos Sinos
Av. Unisinos, 950. São Leopoldo, Brazil, 93022-000
{*alessandro.parolin,eduardoherzer*}*@gmail.com*

Cláudio R. Jung
Universidade Federal do Rio Grande do Sul - PPGC
Av. Bento Gonçalves, 9500. Porto Alegre, Brazil, 91501-970
crjung@inf.ufrgs.br

***Abstract*—Melanoma is the deadliest kind of skin cancer, but it can be 100% cured if recognized early in advance. This paper proposes a non-invasive automated skin lesion classifier based on digitized dermatological images. In the proposed approach, the lesion is initially segmented using snakes guided by an edge map based on the Wavelet Transform (WT) computed at different resolutions. A set of features is extracted from lesion pixels, and a probabilistic classifier is used to identify melanoma lesions. The detection rate of the proposed system can be adjusted to control the tradeoff between false positives and false negatives, and experimental results indicated that a false negative rate of 1.89% can be achieved, in a total accuracy rate of 82.55%.**

***Keywords*-melanoma classification; image processing; feature reduction**

I. Introduction

Melanoma is the deadliest form of skin cancer, and its incidence rate has been increasing over the latest years throughout the world [1], [2]. This cancer can be totally cured if diagnosed early. If not, however, it can spread through the body (metastasis) and be fatal. The survival rate of the patient is strongly related to the stage of the disease [3].

Early diagnosis of melanoma is not an easy task, since the disease can be very similar to other skin lesions in its early stages. Therefore, the diagnosis relies mostly on the knowledge and experience of a physician and his/her observation skills. According to [4], a dermatologist is incorrect in 1 of every 3 diagnosis through the simple visual analysis of a cutaneous lesion. To guarantee a precise diagnosis, other exams are necessary, such as biopsies, which are expensive and painful [3]. Another problem towards early diagnosis of melanoma is the frequency that people, especially in third world countries, go under periodic revisions with dermatologists.

Considering the facts presented before, this paper proposes an automated diagnostic system of cutaneous lesions through the analysis of digitalized images to help a physician, acting like a second opinion. Many papers have already proposed solutions to this problem, but, most of them uses only neural networks to classify the lesion, as in [5]–[9]. In this paper, we propose a novel approach using a probabilistic classifier, which allows us to determine the confidence rate of the results, and that can be fine tuned to control the compromise between false positives (FP) and false negatives (FN).

Automated classifiers, besides working as physician assistants, could be also used in teledermatology, which is a recent field in dermatology. Through teledermatology, there is no need of a patient being in the same place as the physician. For instance, a patient can go to smaller health care outposts, where a trained employee could take a picture of the lesion and upload it to a server, where a physician can access the picture some other time and return a diagnosis. A detailed explanation about teledermatology can be found in [10]. In this cases, an automated diagnosis system could be used for preselecting images that have high probability of being melanoma, setting their priority of being verified by the specialist higher than the others.

II. Related Work

Automated diagnosis has been getting a lot of attention from medical and academic communities due to the increasing rate of melanoma. Many papers have already proposed solutions to the problem presented in this one. In [9], two sets of images were evaluated, one containing 66 images and the other containing 300. Eleven features were extracted from each image: asymmetry index, irregularity, average red, green and blue inside the tumor, colorbin, local and average color for RGB and area and elevation (manually estimated). The features were used in a feedforward neural network, trained with backpropagation algorithm. Tests were executed using half of the database for training and the other half for testing. The results show that the neural network, when tested with the smaller database achieved a success rate of 92%, while with the larger database, the success rate drops to 85%.

In [8] two approaches are compared: neural networks and discrimant analysis. Different databases were composed, containing approximately 30 images each and the classifiers were validated using the “leaving-one-out” method. The features extracted from each image were: two irregularity indexes and thiness ratio. Besides, statistical data related

to different color spaces, such as RGB, Lab, IHS were also extracted, totalizing 20 features. In some cases, PCA (Principal Component Analysis) was applied to reduce the dimensionality from the database. The experimental results presented that both classifiers achieved high success rates, from 85% to 100%.

In [1], 1258 dermoscopy images have been analyzed and 428 features were extracted from each of them. The lesion segmentation was performed automatically by using a threshold method for an initial estimative of the tumor area followed by a region growing algorithm. From the segmented lesion, 140 features related to color, 80 to symmetry, 32 to border and 176 to texture were extracted. The classification was performed by an Artificial Neural Network (ANN) trained with a backpropagation algorithm, and the results were evaluated by a leave-one-out cross-validation method. The inputs for the ANN were selected through the method of hypothesis test of Wilks. The best results achieved a sensitivity of 85.9% and specificity of 86.0%, when 72 features were selected from the whole group of features and using an ANN with 6 hidden neurons. The system was made available on the internet for physicians and dermatologists.

Christensen et al. [2] proposed an approach aiming to achieve at least 64% of accuracy (the same rate of dermatologists, when inspecting lesions at naked-eye). They applied a closing operation to reduce the influence of hair, and used an adaptive threshold method to segment the lesions. The features extracted from the images for classification were: border irregularity, area index, best-fit ellipse index (melanomas have greater tendency than benign lesions to have a ellipse-shaped form), standard deviation of mean radius, roundness and Heywood circularity index. Other features related to the color of the lesion were also obtained. Each feature was statistically evaluated to determine if it contributed in a positive or negative way for the accuracy of the system. Then, all the features were summed together in order to create a single score which would describe if the lesion was benign or malignant. The system achieved an accuracy of 77%, and although it is not clear in the paper, it seems that an online image database (Dermis) was used.

Despite the existence of different approaches for automatic lesion classification, there are several aspects that influence the performance of the classifier, such as the initial segmentation of the lesion (which may be manual or automatic), the features that are used in the classifier, the type of classifier, and the database used for testing (number of samples, controlled environment for acquiring the images, etc.). This work presents a novel approach for automatic classification of skin lesion that presents contributions in different stages of the process. First, the automatic segmentation of the lesion contour is obtained using snakes guided by edge maps based on the WT at different resolutions in a coarse-to-fine manner, getting a rougher estimate of the contour initially and refining it later on. Then, a new way to compute the symmetry of the lesion is introduced, and combined with other features commonly used for classification (such as color, irregularity, etc.). Finally, a Bayesian classifier is employed, providing a statistical framework to evaluate the confidence of classification results (and also allowing to control the compromise between FP and FN). The proposed approach is described next.

III. The Proposed Approach

The proposed system architecture can be seen in Figure 1. A digital image (photograph) of the lesion is given to the system, which segments the lesion automatically. Then, eleven key features are extracted, and the Fukunaga-Koontz Transform (MDA-FKT) is applied for feature reduction. The features are used in a classifier, which will determine if the lesion is melanoma or not based on its training using a specific database. Each step will be detailed in the next subsections.

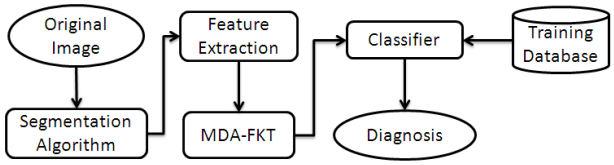


Figure 1. System architecture

A. Image Segmentation

The image segmentation step consists on isolating the lesion from the image background. There are several techniques to solve this problem, such as thresholding, region growing, snakes and split and merge [11]. These techniques are usually applied to solve problems involving images in many fields, such as cardiac, ankle cartilage and skin lesion images [12]–[14] and many others. In this paper, several techniques were analyzed.

We have chosen the snake segmentation technique, due to its widespread use in medical imaging [12], [13], and flexibility to get the contour accurately. This technique was proposed by [15], and it consists on a curve $c(s,t) = (x(s,t), y(s,t))$, which moves dynamically until it finds the contour of the object. Therefore, this model deforms the curve in order do minimize the its energy functional, which is given by:

$$E = \int_0^1 \left[\frac{1}{2} \left(\alpha \left| c^{'}(s) \right|^2 + \beta \left| c^{''}(s) \right|^2 \right) + \kappa E_{ext}\left(c(s)\right) \right] ds, \quad (1)$$

where α is the tension of the snake, β is the rigidity of the snake, κ is the weight of the external force, and $c(s)$ is the final curve. The functional energy can be divided in two parts: the internal and external energies. The former term of the integral is the internal energy, which represents the curve itself, and the latter term is the external force E_{ext}, which consider the image data, guiding the curve to the boundary.

The snake has some convergence problems due its weak external force. To solve this issue, the Gradient Vector Flow Snake (GVF snake) was used, proposed by [16]. The GVF snake has a behavior like the traditional snake, but it solves the small attraction basin of traditional snakes using a new external force: the Gradient Vector Flow Field (GVF field). On the other hand, GVF snakes may be strongly affected by noise, which can push the snake away from the actual contour. In this work, an edge map based on the WT is employed, due to the inherent denoising of the wavelet decomposition.

To obtain the initial contour, the color image is initially converted to grayscale and blurred using a Gaussian filter (with standard deviation 10 and size 15×15), to minimize the presence of hair and small skin signs. Then, Otsu's automatic threshold [11] is applied to the blurred image, and pixels lower that the threshold are retrieved (it is assumed that the lesion is darker than the skin). However, shadows may degrade the thresholding procedure (and they tend to appear in the boundaries of the image). A connected components algorithm is applied, resulting in a set of blobs. The largest blob is then retrieved as the initial lesion region if it is larger than a minimum area threshold T_A (set experimentally to 200, based on the resolution of the images in our database), and if it is not connected to the boundary (since shadow regions may yield larger connected sets). If this largest set is not validated, the initial threshold is reduced by a fixed amount (0.05, set experimentally), and the process repeats iteratively.

The boundary of the validated blob is used as the initial contour required for snakes segmentation. To evolve this contour, the non-decimated wavelet transform described in [17] is used, resulting in a set of detail images in the horizontal and vertical directions $W^h_{2^j}$ and $W^v_{2^j}$, for $j = 1, ..., N$, and a low-pass image S_{2^N} (here, N is the coarsest scale). As described in [17], the wavelet used to compute the WT is similar to the derivative of a Gaussian, so that $W^h_{2^j}$ and $W^v_{2^j}$ act as edge detectors, with magnitude

$$M_{2^j}[n,m] = \sqrt{(W^h_{2^j}[n,m])^2 + (W^v_{2^j}[n,m])^2}. \quad (2)$$

For larger values of j, M_{2^j} provides a coarse edge map presenting only the most relevant objects in the image, but also suppressing noise. In our approach to evolve the snake, the GVF field [16] is computed with the coarsest edge map M_{2^N}, and the initial contour is evolved. When it converges, the resulting contour is used as input for another snake evolution, using a finer edge map M_{2^j} $(j < N)$ to build the GFV field, and the process is repeated until a scale $N_0 < N$ is reached. In all examples, we used just two scales ($N = 6$ and $N_0 = 4$), based on the resolution of our database. Also, we have made an exhaustive analysis of the snake parameters α , β and κ, and set these parameters respectively to 0.1, 0.1, 0.9 in all experiments.

Figure 2 shows an example of the segmentation procedure. The original color image is shown in Figure 2(a), and the blurred grayscale version is shown in Figure 2(b). The thresholding procedure is illustrated in Figure 2(c), and it can be seen that the lesion was not separated from the background (due to uneven illumination). However, such separation can be achieved after two reductions of the threshold (Figure 2(d)). The final contour obtained with wavelet-based magnitudes and snakes is shown in Figure 2(e).

B. Feature Extraction

The feature extraction step was based on the ABCDE rule [3], [18], which is widely used by dermatologists to visually diagnose melanoma. This rule describes that malignant melanoma lesions are characterized by their asymmetry, border irregularity, color variance, diameter (greater than 5mm) and evolution over time. Since the images evaluated in this paper were selected from online databases[1], it was not possible to determine their diameter in centimeters (since the images do not have any scale), neither their evolution. However, as it can be seen in [7], [9], there are other relevant features that can be extracted from the images. Thus, based on the ABCDE rule and in other similar papers, the features extracted from the images were: irregularity, asymmetry index, color variance (RGB), relative chromaticity (RGB) and average color (Lab color space). The approaches to extract each feature can be seen below.

1) Asymmetry Index: It represents how asymmetric the lesion is. In this paper, we propose the following approach to detect the asymmetry of a skin lesion. An asymmetry index is calculated with respect to every axis that passes through the lesion centroid, and the lesion asymmetry is defined by the smallest value of asymmetry, considering all the axis. More specifically, consider the lesion centroid $\boldsymbol{c}$ and its perimeter, containing N pixels $\boldsymbol{p}_i$ (where N depends on the size of the lesion and the resolution of the image). For every point $\boldsymbol{p}_i$, its Euclidean distance d_i to the centroid is computed through:

$$d_i = ||\boldsymbol{p}_i - \boldsymbol{c}||, i = 1, \ldots, N. \quad (3)$$

To obtain the asymmetry A_j of the lesion around a point $\boldsymbol{p}_j$ (i.e., around an axis that passes through $\boldsymbol{c}$ and $\boldsymbol{p}_j$), we compute the average of the distances[2] d_i between the points on the left and the right of $\boldsymbol{p}_j$:

$$A_j = \frac{2}{N} \sum_{i=1}^{N/2} |d_{j-i} - d_{j+i}|, j = 1, \ldots, N. \quad (4)$$

and the lesion asymmetry index is given by the smallest value of A_j:

$$A = \min_j A_j. \quad (5)$$

[1]Avaliable at: http://www.dermis.net and http://www.dermnet.com.

[2]In Equation (4), the vector containing distances d_i is extended periodically, so that $d_{-1} = d_N$, $d_{-2} = d_{N-1}$, etc.

(a) Original image (b) Blurred version (c) Binary image (d) Binary image after second thresholding (e) Segmented lesion

Figure 2. Steps of the segmentation procedure.

2) Irregularity: This feature describes the irregularity index of the lesion border. According to [7], it can be calculated by:

$$I = \frac{N^2}{4\pi A}, \tag{6}$$

where A denotes the area of the lesion, and N denotes the perimeter of the lesion, both in pixels.

3) RGB Color Variance: This feature describes the color variance among all pixels from the lesion. This feature is calculated for each color component in the RGB color space. According to [6] and [7], melanoma is characterized by the mixture of the following colors: tan, brown, red and black, which implies in high color variation in the RGB channels. The variance is given by:

$$\sigma_c^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^c - \mu_{RGB}^c)^2, c \in \{R, G, B\}, \tag{7}$$

where x_i^c is the value of the component c from pixel i in the RGB color space, μ_{RGB}^c is the mean value of the same color component and n is the number of pixels in the lesion.

4) Lab Mean Values: As demonstrated in [6], melanoma can be very similar to Dysplastic Nevi (a benign skin lesion, which is a potential beginning of melanoma) in its early stages, due to the small variation of brightness, color and chroma between them. According to the authors, the Lab colors are the most efficient way to distinguish those lesions. To obtain this feature, the image is converted to the Lab color scheme and then the average of the Lab components is calculated through:

$$\mu^c = \frac{1}{n}\sum_{i=1}^{n} y_i^c, c \in \{L, a, b\} \tag{8}$$

where y_i^c is the value of the component c from the pixel i in the Lab color space, and again n is the number of pixels in the lesion.

5) Relative Chromaticity: According to [7], this feature reduces the variation of light, printing and digitalization of the image, and it also equalizes the variations of individual human color. This feature is calculated for each component in the RGB color space, and it is given by:

$$CR_c = \frac{\mu_{RGB}^c}{\sum_{c\in\{R,G,B\}} \mu_{RGB}^c} - \frac{\nu_{RGB}^c}{\sum_{c\in\{R,G,B\}} \nu_{RGB}^c}, \tag{9}$$

where μ_{RGB}^c are the RGB mean values within the lesion, and ν_{RGB}^c are the mean values computed in healthy skin parts around the lesion.

As it can be seen in Equation (9), it is necessary a small portion of the healthy skin from the image to compute this feature. Since the images used here were selected from online databases, most images contains logos, which are not always positioned in the same place of the images. Therefore, it should be necessary to develop an algorithm that could detect a part of the healthy skin without any kind of symbol overlapping it and under good illumination. Since the process of selecting a small portion of the image is not complicated and can be easily done manually by a non-specialist, this step was not automated.

C. Image Classification

The features extracted from each image were used in a Bayesian classifier. This kind of classifier works with probabilities, which allows a greater flexibility towards the results. In our problem, we have just two classes ω_1 and ω_2, that represent melanoma and not melanoma, respectively. The Bayesian decision rule is based on the posterior probability that a feature vector $\boldsymbol{x}$ belongs to a class ω_i:

$$P(\omega_i|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\omega_i)P(\omega_i)}{p(\boldsymbol{x})}, \tag{10}$$

where $p(\boldsymbol{x}|\omega_i)$ is the probability density function (PDF) for the class ω_i, and $p(\boldsymbol{x}) = p(\boldsymbol{x}|\omega_1)P(\omega_1) + p(\boldsymbol{x}|\omega_2)P(\omega_2)$ is the PDF for the whole distribution of feature vectors. The class ω_i is chosen if $P(\omega_i|\boldsymbol{x}) > P(\omega_j|\boldsymbol{x})$, $\forall j \neq i$. For a two-class problem, this reduces to selecting ω_i if and only if $P(\omega_i|\boldsymbol{x}) > 0.5$.

One interesting property of the Bayesian classifier given in Equation (10) is the association of a cost to wrong selections. In particular, in the problem of classifying skin lesions into melanoma or not, FNs (i.e., the classification of melanoma lesions as not melanoma) are dangerous, since they may not be promptly analyzed by the physician. On the other hand, FPs (i.e., the classification of non melanoma lesions as a melanoma) are not that bad, since a thorough analysis by the dermatologist could provide the correct diagnosis.

This cost can be computed by a loss function λ_{ij}, that provides the cost of selecting a class ω_i when the correct one was actually ω_j. To avoid penalizing correct classifications it

is desirable to have $\lambda_{ij} = 0$ if $i = j$, and $\lambda_{ij} > 0$ otherwise. With the introduction of the cost functions, the boundary of the Bayesian classifier changes accordingly [19]. More specifically, the class ω_1 is selected over ω_2 if and only if:

$$\frac{P(\omega_1|\boldsymbol{x})}{P(\omega_2|\boldsymbol{x})} > \frac{\lambda_{12}}{\lambda_{21}}. \tag{11}$$

The selection of λ_{21} and λ_{12} depends on how strongly one wants to avoid FNs (at the cost of possibly increasing the number of FPs). This issue is analyzed in the next section.

Another challenging task is the determination of the underlying PDFs for the two classes of the problem. For sakes of simplicity, we assumed a multivariate Gaussian model for each of the two classes.

D. Data Dimensionality Reduction

To characterize a multivariate Gaussian distribution, the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ are needed. Although their computation is straightforward, the number of samples required to compute them reliably increases as the dimensionality of the problem grows. To avoid this kind of problem, several techniques to reduce the dimensionality of the problem have been proposed, and the main approaches are Principal Component Analysis (PCA) and the Fisher's Linear Discriminant Analysis (FLDA). PCA is useful to find linear combinations of features that contain large variance, not accounting for the individual distribution of each class. In FLDA, the main goal is to reduce a multidimensional feature to a scalar through a linear transformation that maximally separates class patterns according to the Fisher Criterion, being optimal only for two Gaussian distributions with equal covariance [19].

In this work, we explore the Multiple Discriminant Analysis based on the Fukunaga-Koontz Transform (MDA-FKT), that maximizes the Bhattacharyya distance between two Gaussian distributions with same mean but possibly different covariance matrices [20]. Another advantage of the MDA-FKT over the FLDA is that it allows to reduce a d-dimensional feature vector to a f-dimensional feature vector, for $1 \leq f < d$, whereas FLDA allows only the reduction to a scalar (in a two-class problem).

As in [20], the first step is to reformulate the classification problem to generate two distributions with the same mean. This is done by defining the intraclass space Ω_I and the extraclass space Ω_E through

$$\begin{aligned} \Omega_I &= \{\boldsymbol{x}_i - \boldsymbol{x}_j | L(\boldsymbol{x}_i) = L(\boldsymbol{x}_j)\}, \\ \Omega_E &= \{\boldsymbol{x}_i - \boldsymbol{x}_j | L(\boldsymbol{x}_i) \neq L(\boldsymbol{x}_j)\}, \end{aligned} \tag{12}$$

where $L(\boldsymbol{x})$ is the label of sample $\boldsymbol{x}$. Since both distributions have zero mean, it was shown in [20] that the optimal subspace with dimension f that maximizes the Bhatacharyya distance is characterized by f generalized eigenvectors of the pair of covariance matrices $\boldsymbol{\Sigma}_I$, $\boldsymbol{\Sigma}_E$ (related to Ω_I and Ω_E, respectively) corresponding to the largest values $\lambda + 1/\lambda$, where λ are the generalized eigenvalues.

Once the original data $\boldsymbol{x}_i$ are reduced to f-dimensional feature vectors $\boldsymbol{y}_i$ using the MDA-FKT approach, the resulting vectors are used to obtain the corresponding Gaussian distributions of the melanoma and non-melanoma classes, characterized the the mean vectors and covariance matrices. Although these parameters can be obtained using traditional estimators, the presence of outliers can degrade such estimates. In fact, they can be obtained in a more robust manner using the algorithm presented in [21]. In their approach, the covariance matrix is created through the remotion of outliers, which are detected by projecting the data in directions that maximized the kurtosis coefficient.

E. Image preprocessing

In order to increase the accuracy of the proposed system, an image enhancement procedure [22] was applied as a preprocessing stage. This algorithm is based on an adaptive modification of wavelet coefficients that stretches low-contrast edges and shrinks noise-related coefficients. In [22], it was shown that such enhancement algorithm improves the visual quality of skin lesions for manual classification by the physician (an example of lesion enhancement is shown in Figure 3), and in this work we propose to use edge-enhanced images to improve classification results using semi-automatic techniques.

(a) Original image (b) Enhanced Lesion

Figure 3. Example of lesion enhancement.

IV. Experimental Results

In the study presented in this paper, 290 images (139 non melanomas and 151 melanomas), obtained from an online database, were evaluated. The tests were separated in two parts: segmentation and classification.

The segmentation procedure is just a preliminary stage needed for the classification procedure. The evaluation of segmentation results was not based on a precise comparison of the contour generated by the proposed approach and some kind of ground truth. Instead, just a visual inspection was performed to detect if the lesion contour was correctly delineated using the proposed approach, neither leaving much of the lesion outside the boundary, nor presenting healthy skin portions within the segmented region.

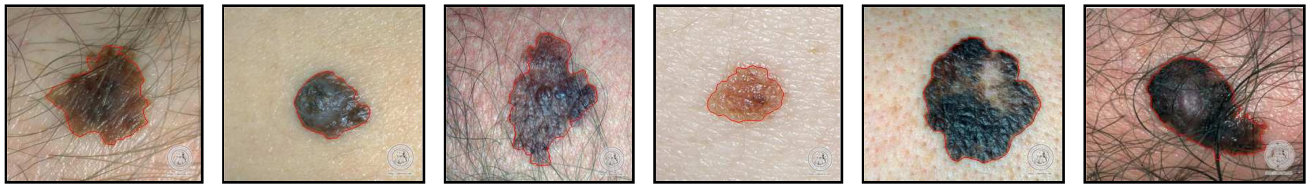


Figure 4. Examples of lesions and segmentation results.

In fact, since the segmentation is the input for the classification procedure, the result of the segmentation algorithm is implicitly evaluated based on the results of the classification. The classification tests were validated with the 10-fold-cross validation technique, which, according to [23], provides more reliable results than dividing the database in two sets: training and testing. We ran the 10-fold-cross validation technique twenty times, obtaining the confusion matrix from the mean values.

As explained in the previous section, the Bayesian classifier was evaluated using a minimum acceptable threshold T_p, and also using the cost-based approach. The results of each step is discussed with more details in this section.

A. Segmentation Results

Despite the fact that the segmentation results were not validated by a specialist, it can be seen that for most of the input images in the database, the segmentation did not include healthy skin or left behind significant portions of the lesion, as shown in Figure 4. In fact, from the 290 images used in this work, the system could segment automatically 85.86% images (the correctness of the segmentation results was evaluated visually).

Figure 5. Example of manual initialization (left) and final segmentation results (right).

The remaining 41 images that could not be segmented automatically usually had a pellucid stain into the lesion. This errors occurred due to a bad automatic initialization of the snake. However, the problem of having a bad initialization in snakes segmentation can be overcame by manually drawing a polygonal line that roughly approximates the lesion boundary, which is neither difficult nor time consuming. Thus, all the images which could not be segmented automatically were successful segmented after a small user intervention. Figure 5 illustrates an example of manual initialization (green polygon) and the corresponding segmentation (red curve).

B. Classification Results

To evaluate the classification results, confusion matrices were built to show the average number of false positives (FP), false negatives (FN), true positives (TP) and true negatives (TN), given by:

$$\text{TP} = \frac{P_{cc}}{N_p}, \quad \text{FN} = 1 - \text{TP}, \tag{13}$$

$$\text{TN} = \frac{N_{cc}}{N_n}, \quad \text{FP} = 1 - \text{TN}, \tag{14}$$

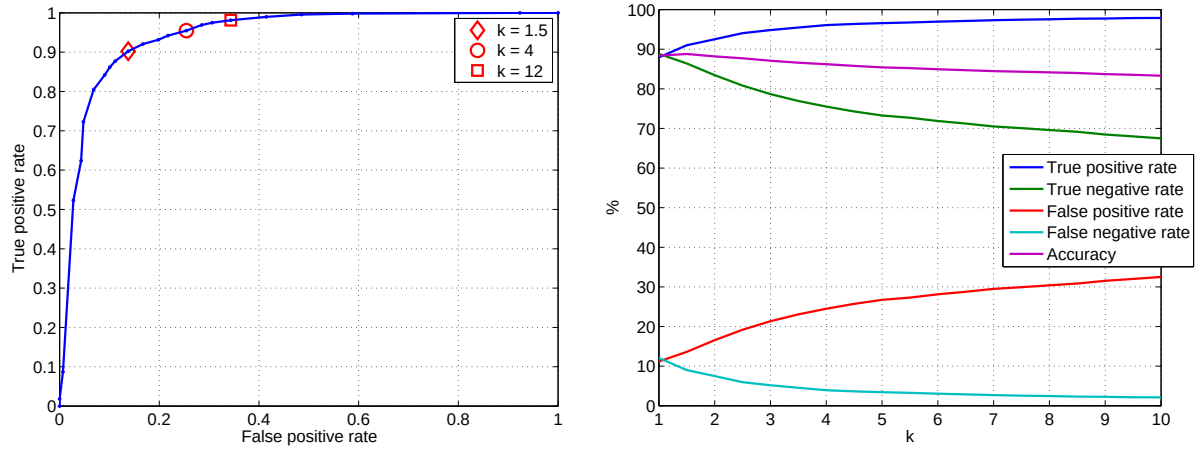
where P_{cc} and N_{cc} are number of positive and negative instances correctly classified, respectively, N_p and N_n are the total number of positive and negative instances being evaluated, respectively.

For all analyzed technique, the 10-fold-cross validation procedure was applied 20 times, and the mean results were retrieved. Table I shows the results obtained using the Bayesian classifier according to condition (11) using $\lambda_{12} = \lambda_{21}$, i.e., assigning the same penality for FPs and FNs. More specifically, Table I presents the confusion matrices for 5 different classification approaches: using the full 11-dimension feature vector (Multivariate), using Fisher's Linear Discriminant Analysis (FLDA) to reduce the feature vectors to a single scalar, the MDA-FKT approach reducing to a single scalar (MDA-FKT1), the MDA-FKT approach reducing to 2-dimensional feature vector (MDA-FKT2), and the MDA-FKT approach reducing to 2-dimensional feature vector applied to enhanced versions of the lesions (MDA-FKT2 + Enhancement).

As it can be seen, both FLDA and MDA-FKT approaches present some improvement over the multivariate approach with all feature vectors, both in terms of FPs and FNs. Comparing the approaches that reduce the feature vector to a scalar (FLDA and MDA-FKT1), it can be observed that FLDA produces slightly better results, particularly with respect to the number of FNs. However, when MDA-FKT is used to reduce the feature vectors to two dimensions (MDA-FKT2), the results are better than FLDA w.r.t. to both FPs and FNs, and an additional gain is obtained when

		Predicted									
		Multivariate		FLDA		MDA-FKT1		MDA-FKT2		MDA-FKT2 + Enhancement	
		Mela	Non-mela	Mela	Non-mela	Mela	Non-mela	Mela	Non-mela	Mela	Non-mela
Actual	Mela	130.45	20.55	130.95	20.05	131.90	19.10	131.55	19.45	133.35	17.65
	Non-Mela	18.35	120.65	18.00	121.00	31.35	119.65	16.55	122.45	15.95	123.05

Table I
COMPARISON OF DIFFERENT CLASSIFICATION APPROACHES USING $\lambda_{12} = \lambda_{21}$



(a) (b)

Figure 6. (a) ROC curve for the MDA-FKT2 approach with enhancement. (b) Results assigning higher costs to FNs, where $k = \lambda_{21}/\lambda_{12}$.

this last approach is combined with the image enhancement procedure [22].

C. Assigning penalities to errors

Another set of experiments was performed by assigning a higher cost to FNs, by setting $\lambda_{21} > \lambda_{12}$ in Equation (11). As $k = \lambda_{21}/\lambda_{12} \geq 1$ increases, the number of FNs tends to decrease, at the cost of increasing the number of FPs. Figure 6(b) shows the relation between the classification error types (FPs and FNs) as a function of k, along with the total accuracy[3]. Figure 6(a) shows the receiver operating characteristic (ROC) curve for the proposed approach (MDA-FKT2 + Enhancement) with varying values for k, and the compromise between FPs and FNs can be observed explicitly.

As it can be observed in Figure 6, the FN rate decreases at the cost of rapidly increasing the FP rate as k increases, which causes the global accuracy rate to be decreasing. The selection of the best value for k is context-dependent, and relates to the tradeoff between FPs and FNs. Figure 6(a) highlights three values for k, to illustrate such tradeoff. The FP rates for $k = 1.5$, $k = 4$ and $k = 12$ are, respectively, 13.78%, 25.43% and 34.35%, and the FN rates are 9.77%, 4.54% and 1.89%.

[3]The accuracy is the rate of total samples (either positives or negatives) that were correctly classified.

V. CONCLUSIONS

This paper proposed a quasi-automated system for diagnosing melanoma lesions based on the analysis of digitalized images. In our approach, a novel segmentation algorithm based on snakes and multiple wavelet gradients was proposed to obtain the lesion contour. A set of features is then computed for the lesion (and healthy skin portions), that are evaluated using a Bayesian classifier. Variants of the classifier were employed, using MDA-FKT for dimensionality reduction, using robust estimation of the class parameters, using costs to reduce the number of false negatives, and using an enhancement algorithm as a preprocessing stage.

As the experimental results have shown, the proposed approach can achieve an accuracy of 88.41 % with a false negative rate of 11.47%, using MDA-FKT2 approach plus the image enhancing preprocessing, indicating its use as a pre-screening for tele-dermatology applications. Besides, associating a cost to the FN instances (setting its cost twelve times the cost of FP), we were able to reduce the false negative rate to 1.89% of the non-melanoma samples, but at the cost of decreasing the total accuracy to 82.55%.

An objective comparison with competitive approaches is difficult, since they are applied to different databases. However, it should be noticed that the approach described in [1] achieved TP and TN rates of 85.9% and 86.0%, respectively. Also, the technique described in [2] apparently used the same database as the one used in our work, achieving an

accuracy of 77%. As it was presented in the last section, our system achieved a TP rate of 88.31% and a TN rate of 88.52% using only 2 features (MDA-FKT2+Enhancement), yielding an accuracy of 88.41%.

As future work, we intend to use other databases to train and validate the classifiers, obtained with a standardized procedure for image acquisition (which would probably increase the accuracy of the proposed system). Another point for improvement is the study and development of other features that better discriminate melanomas for other lesions, and the use of other distributions (rather than the Gaussian) to model the classes melanoma and non-melanoma.

REFERENCES

[1] H. Iyatomi, H. Oka, Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka, and K. Ogawa, "An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm," *Computerized Medical Imaging and Graphics*, vol. 32, no. 7, pp. 566–579, October 2008.

[2] J. H. Christensen, M. B. T. Soerensen, Z. Linghui, S. Chen, and M. O. Jensen, "Pre-diagnostic digital imaging prediction model to discriminate between malignant melanoma and benign pigmented skin lesion," *Skin Research and Technology*, vol. 16, 2010.

[3] V. T. Devita, S. Hellman, and S. A. Rosenberg, *Cancer: Principles & Practice of Oncology*. Lippincott Williams & Wilkins, 2001.

[4] A. A. Marghoob, L. D. Swindle, C. Z. M. Moricz, F. A. S. Negron, B. Slue, A. C. Halpern, and A. W. Kopf, "Instruments and new technologies for the in vivo diagnosis of melanoma," *Journal of the American Academy of Dermatology*, vol. 49, no. 5, pp. 777 – 797, 2003.

[5] M. Barzegari, H. Ghaninezhad, P. Mansoori, A. Taheri, Z. Naraghi, and M. Asgari, "Computer-aided dermoscopy for diagnosis of melanoma," *BMC Dermatology*, vol. 5, no. 1, p. 8, 2005.

[6] F. Ercal, A. Chawla, W. Stoecker, H.-C. Lee, and R. Moss, "Neural network diagnosis of malignant melanoma from color images," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 9, pp. 837–845, Sep 1994.

[7] B. Kusumoputro and A. Ariyanto, "Neural network diagnosis of malignant skin cancers using principal component analysis as a preprocessor," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 1, May 1998, pp. 310–315 vol.1.

[8] I. Maglogiannis, S. Pavlopoulos, and D. Koutsouris, "An integrated computer supported acquisition, handling, and characterization system for pigmented skin lesions in dermatological images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 1, pp. 86–98, March 2005.

[9] Z. Zhang, R. H. Moss, and W. V. Stoecker, "Neural networks skin tumor diagnostic system," in *Proceedings of the International Conference on Neural Networks and Signal Processing*, vol. 1, Dec. 2003, pp. 191–192 Vol.1.

[10] C. Ebner, G. Gabler, C. Massone, R. Hofmann-Wellenhof, G. P. Lozzi, E. Wurm, and H. P. Soyer, "Mobile teledermatology coming of age," *e & i Elektrotechnik und Informationstechnik*, vol. 123, pp. 148–151, 2006.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (2nd Edition)*. Prentice Hall, January 2002.

[12] V. B. Medina, R. Valdes, O. Yanez-Suarez, M. Garza-Jinich, and J.-F. Lerallut, "Automatic initialization for a snakes-based cardiac contour extraction," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 3, 2000, pp. 1625–1628 vol.3.

[13] J. Tang, S. Millington, S. Acton, J. Crandall, and S. Hurwitz, "Ankle cartilage surface segmentation using directional gradient vector flow snakes," in *Proceedings of the IEEE International Conference on Image Processing*, vol. 4, Oct. 2004, pp. 2745–2748 Vol. 4.

[14] L. Chen, X.-L. Zhang, Z. Wang, and L.-T. Fang, "Contour extraction based on improved gvf snake model in asar images," in *Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition*, vol. 1, Nov. 2007, pp. 301–304.

[15] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. V1, no. 4, pp. 321–331, January 1988.

[16] C. Xu and J. Prince, "Gradient vector flow: A new external force for snakes," in *Proceedings of Computer Vision and Pattern Recognition (CVPR '97)*. San Juan, Puerto Rico: IEEE, June 1997, pp. 66–71.

[17] S. G. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, 1992.

[18] R. J. Friedman, D. S. Rigel, and A. W. Kopf, "Early detection of malignant melanoma: The role of physician examination and self-examination of the skin," *CA Cancer J Clin*, vol. 35, no. 3, pp. 130–151, 1985.

[19] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley-Interscience, November 2000.

[20] S. Zhang and T. Sim, "Discriminant subspace analysis: A Fukunaga-Koontz approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 10, pp. 1732–1745, 2007.

[21] D. Pena and F. J. Prieto, "Multivariate outlier detection and robust covariance matrix estimation," *Technometrics*, vol. 43, pp. 286–310(25), 2001.

[22] C. R. Jung and J. Scharcanski, "Sharpening dermatological color images in the wavelet domain," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 4–13, Feb. 2009.

[23] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, October 1999.