

Original Research Paper

New and Efficient Features for Skin Lesion Classification based on Skeletonization

¹Youssef Filali, ¹Hasnae El Khoukhi, ¹My Abdelouahed Sabri, ¹Ali Yahyaouy and ²Abdellah Aarab

¹Department of Computer Science, Faculty of Sciences Dhar-Mahraz, USMBA, Fez, Morocco

²Department of Physics, Faculty of Sciences Dhar-Mahraz, USMBA, Fez, Morocco

Article history

Received: 28-01-2019

Revised: 18-03-2019

Accepted: 31-08-2019

Corresponding Author:

Youssef Filali

Department of Computer Science, Faculty of Sciences Dhar-Mahraz, USMBA, Fez, Morocco

Tel: +212614352283

Email: youssef.filali1@usmba.ac.ma
Filaliucf@gmail.com

Abstract: This paper presents a new approach to detect and classify skin lesions for melanoma diagnosis with high accuracy. Skin lesion detection is based on an image decomposition into two components using the Partial Differential Equation (PDE). The first component that sufficiently preserves the contour is thus exploited to have an adequate segmentation of image lesion while the second component provides a good characterization of the texture. Moreover, to improve the classification accuracy, new and powerful features extracted by skeletonization of the lesion are presented. These features are compared and combined with well-known features from the literature. Features engineering was applied to select the most relevant features to be retained for the classification phase. The proposed approach was implemented and tested on a large database and gave a good classification accuracy compared to recent approaches from the literature.

Keywords: Image Processing, Skin Cancer, PDE Multiscale Decomposition, Texture, Shape and Color Analysis, Features Engineering, Skeletonization

Introduction

Skin Cancer (SC) is known to be the most dangerous cancer to cause a high mortality rate. This cancer is the result of an abnormal production of melanocyte cells, mainly due to excessive exposure to the sun. These cells generate skin pigmentation and are also the main cause of producing the substance of melanin. In addition, some non-melanoma lesions, such as the "nevus", can evoke a melanoma lesion.

If diagnosed and treated early, SC is almost always curable. Computer-Aided Diagnosis (CAD) systems are a tool to help dermatologists at the early detection and diagnosis of SC. The main challenge in using image processing is to be able to correctly distinguish between a melanoma and non-melanoma lesion (Vasconcelos *et al.*, 2015). In addition to image acquisition, the main steps of such system are Pre-processing, Segmentation, Features Extraction and Classification:

Skin Lesion (SL) images acquired from macroscopic or dermoscopic device usually contain artifacts like hair, shadows and skin lines, which requires a pre-processing to improve the contrast, smooth the image and reduce noise. Several of Linear, Non-linear and Morphological filters are widely used in this step. Linear filter remain

effective only for a specific processing, non-linear filters can be used to eliminate artifact like shadows and hair while preserving the borders of the lesion and Morphological filters are effective for lesions having a very low contrast (Filho *et al.*, 2015).

Segmentation plays an important role, requires more attention and attracts many researchers. It has a straight impact on the result of the classification. Indeed, poor segmentation will lead to a poor Region Of Interest (ROI) detection and automatically to poor classification result even if the best classifier is used but a good segmentation will certainly lead to a better ROI detection and to high classification rate even with a very basic classifier. Several algorithms of segmentation has been proposed in literature to identify and detect the skin lesion (Pennisi *et al.*, 2016; Oliveira *et al.*, 2016) as Edge-based segmentation or Region-based segmentation (Victor and Ghalib, 2017), Threshold segmentation (Garnavi *et al.*, 2011; Korotkov and Garcia, 2012) and active contour method (Chopra and Dandu, 2012; Zhou *et al.*, 2010). Other methods derived from the data analysis like K-means, Artificial Intelligence (AI), Artificial Neural Networks (ANN) have been tried in order to an improve the segmentation results. They are

sometimes combined with meta-heuristics like Genetic Algorithms (GAs) (Filali *et al.*, 2017a; Sabri *et al.*, 2016).

In the Features Extraction (FE) step, we try to better describe the lesion in order to keep the maximum amount of useful information. The features adopted are generally inspired by the clinical approaches used by dermatologists in their diagnosis. The Asymmetry, Border, Color and Diameter (ABCD) rules that consist of calculating the Asymmetry of the lesion in terms of shape and color and the irregularity of the edge, variation of the color and the diameter of the lesion have been widely used (Vasconcelos *et al.*, 2015; Barata *et al.*, 2018; Nezhadian and Rashidi, 2017). These features can contain appearance, color, texture, shape parameters (Leguizam, 2015; Ferris *et al.*, 2015). The shape of the lesion is calculated either by evaluating the compactness or by the spatial moments that give more details on the geometry of the lesion. The texture is defined through statistical measures, spectral characteristic, or co-occurrence matrix and the color of the lesion is determined from intensities information that contains the lesion. However, all proposed primitives in the literature remains insufficient and the classification accuracy is not very high, especially for melanoma, that is why high-level object descriptors has become vital.

The final step is the classification, which uses the extracted features for the recognition and interpretation of the type of lesions. Supervised classification remains the most used type such as Support Vector Machine (SVM), ANN, Decision Tree (DT) and Logistic Regression (LR) (Deepa and Aruna Devi, 2011; Dreiseitl *et al.*, 2001; Mathew and Sathyakala, 2015).

CAD approaches for the diagnosis of skin lesions commonly use the previously described steps from pre-processing to lesion classification through segmentation and feature extraction. In this paper, we propose to use an image decomposition based on an algorithm derived from PDE, which has been proven both for image segmentation and for content retrieval of textured images (Filali *et al.*, 2018; 2017b). This algorithm yields two components. The first, denoted geometric component can be attributed to low frequencies and enhances both the shape and the boundaries of the objects in the image, while the second, denoted texture component, contains the rest of the information as texture and noise. The geometric component is thus exploited to have a suitable segmentation while the texture component provides a good characterization of the texture.

This work also proposes new and powerful form features that represent judiciously both the shape of the lesion and its contour. These features more precisely describe the skeleton pattern of the segmented lesion and thus provide useful and robust

indications for classification and can therefore be used in related medical research.

The rest of the paper is organized as follows. Section 2 gives a description of the proposed approach. The results and metric evaluations are presented in section 3. Finally, section 4 is dedicated to a conclusion and some future directions.

The Proposed Approach

Flowchart of our proposed approach is depicted in Fig. 1. The identification and classification of the lesion involve four main steps:

Decomposition image, segmentation, features engineering and classification. The first step concerns the use of the PDE decomposition image that leads to separate the original image into two components. The first component is well structured and contains the geometric information of the image and the second component, contains oscillating information as textures and noise. The object component is used to perform the segmentation. Then we proceed to the skeletonization of the segmented lesion followed by the extraction of nine proposed features to better characterize the skeleton. The texture component allows evaluating the texture characteristic. The various features thus extracted are concatenated and selected at the features engineering step. The final step is the classification of the lesion using the select features into melanoma or non-melanoma.

Segmentation

Before extracting color, shape and texture information on the lesion, it is necessary to proceed to a separation of the lesion from its surrounding healthy part of the skin. This step is, however, essential to properly characterize the discriminant information of the lesion shape and edge, which depends closely on the border detection precision.

Many approaches have been developed to provide a segmentation tool. The difficulties encountered during this operation are actually several and varied, they concern both the brightness variations encountered in the image, the presence of artifacts and the variability of colors and textures. The presence of texture in skin cancer images is the major problem that encountered during the segmentation step.

This creates a poor segmentation; thus, incorrect identification of the lesion will negatively influence the results of classification step. One way to overcome these difficulties can be achieved through the application of the decompositions of the image using mathematical formulations derived from partial differential equations.

PDE decompositions have proved to be an interesting tool for breaking down the image into several components so that each has additional information.

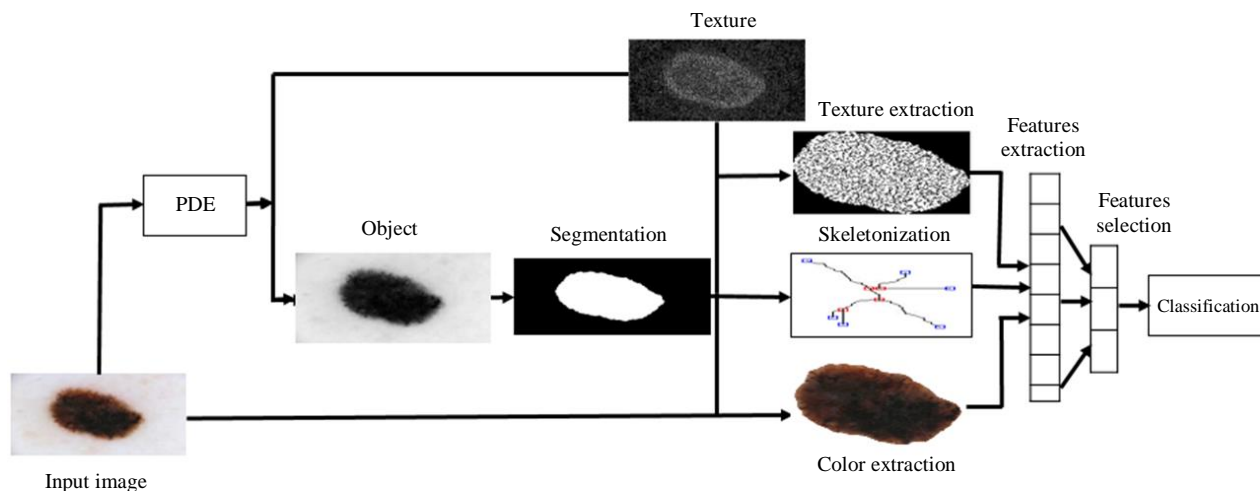


Fig. 1: Flowchart of the proposed approach

Multiscale Image Decomposition

Among the models of PDE decomposition, we have retained the decomposition model proposed by Aujol (Aujol *et al.*, 2005; Aubert and Aujol, 2004), which consists in separating the image into two components. The first component, denoted u , is well structured and contains the geometric information of image. The second component, denoted v , contains oscillating information such as textures and noise.

This decomposition is given by minimizing a functional F :

$$F_{\lambda}(u, v) = J(u) + \frac{1}{2\lambda} \|f - u - v\|_2^2 \quad (1)$$

where, u and v are respectively the object and texture components. λ is the regularization coefficient according to the image area.

Figure 2 shows the decomposition result of a melanoma image.

It is very clear from the decomposition in Fig. 2 that the object component (Fig. 2b) contains only the geometrical information of the lesion without any textures. Therefore, the segmentation is performed on the object component, which allowed us to avoid the problems encountered by the presence of artifacts (contrast, hair, blood vessels and skin lines). Thus, we do not need complex algorithms for segmentation; just Otsu's algorithm is sufficient.

A segmentation example of the object component using Otsu algorithm is presented in Fig. 3. The segmentation evaluation will be presented in section 3.1.

Skeletonizing

Skeletonizing an image aims to provide a wired representation of objects while preserving the essential information carried by each object in the simplest form. At the beginning its application concerned the character recognition (Arcelli *et al.*, 1985) then it was extended to other research area namely biometrics, road extraction, medical imaging for the detection of blood vessels, or extraction of bone architecture (Zhu *et al.*, 2014; Beumier and Neyt, 2017). There is currently a wide variety of skeletonizing methods; the most common are based on thinning, on distance maps, or on hybrid techniques, but each of them gives a different result (Saha *et al.*, 2016; De Oliveira *et al.*, 2009):

- **Thinning:** The skeleton is obtained by iteratively peeling the boundary layers of the object by removing single points using morphological operations (points that can be removed without changing the topology of the shape) until having a skeleton homotopic to the object and not necessarily centered
- **Distance maps:** The principle consists in computing the distance map inside the object, to find the local maximums and then reconnect them to have the desired skeleton. The resulting skeleton is not necessarily homotopic, thin and centered
- **Hybrid techniques:** The hybrid methods have also been introduced recently to take advantage of the two approaches previously mentioned

In our proposed approach, the skeletonization is performed based on thinning technique. Fig. 4 presents the skeletonization result of the segmented lesion.

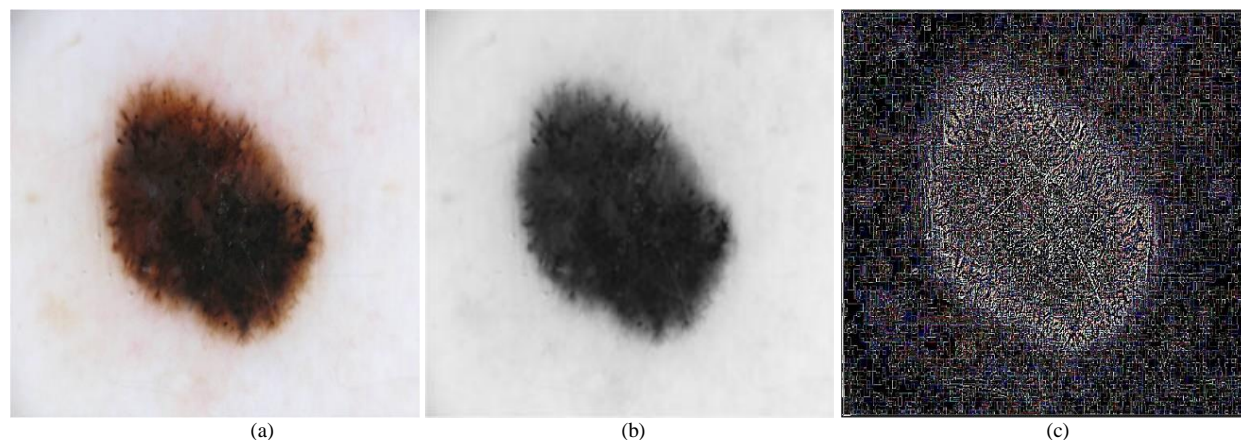


Fig. 2: Image decomposition: (a) original image; (b) object component; (c) texture component

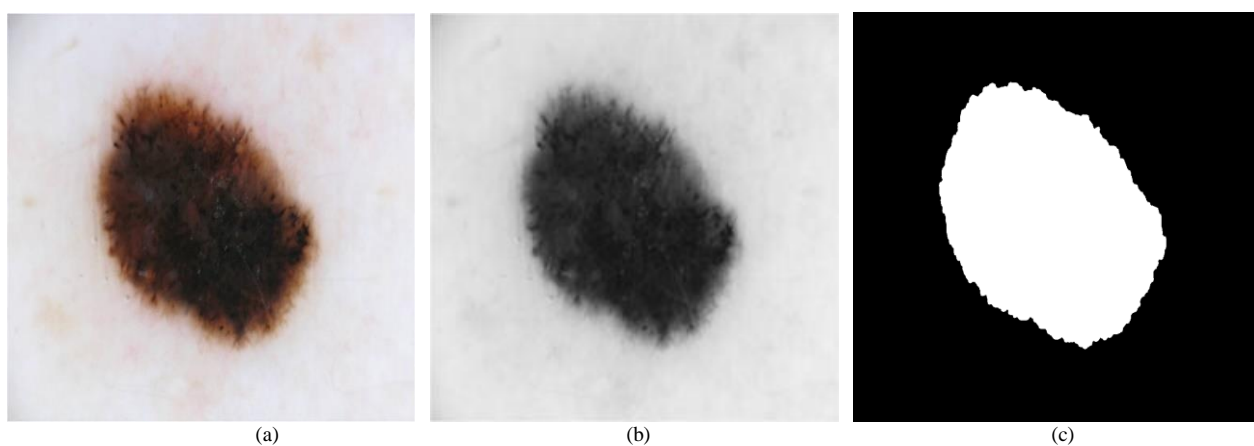


Fig. 3: Skin lesion segmentation: (a) original image; (b) object component; (c) segmentation result performed on (b) using Otsu

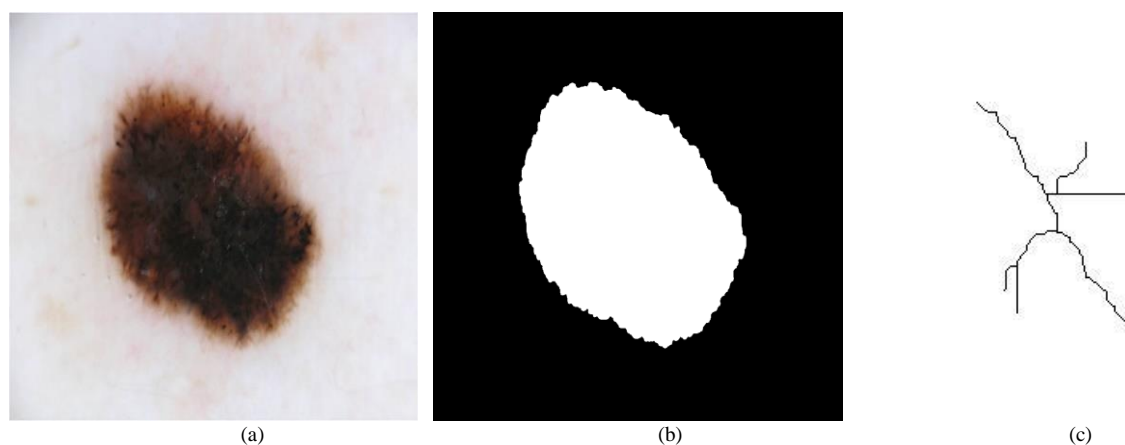


Fig. 4: Lesion skeletization: (a) original image; (b) segmented lesion; (c) skeleton of the lesion

Features Extraction

In order to achieve a classification with a good score, it is necessary to extract appropriate features from the image data. Thus, the need for high-level

selective object descriptors has become crucial. With the presence of noise and the loss of spatial information and depth, it is essential to collect information from all the primitives (colors, shape and texture) for more efficient analysis. In this context, we

use (i) skeleton features of the lesion obtained from the object component, (ii) textures present in the lesion obtained by projection of the segmented object component on the texture component (iii) colors of the lesion obtained after projection of the segmented object on the original image.

Skeleton Features

The skeleton topology is used to well describe the shape of the lesion while keeping the geometrical properties and then to obtain a good classification rate. (Fig. 5) the skeletonization that we can extract the endpoints and branches.

From the endpoints and branches, our new descriptor based on the skeleton of the lesion will consist of nine features:

- Number of Endpoint: An active pixel that has exactly one active neighbor pixel
- Number of Branch-points: Points where two or more branches meet
- Number of Sub-branches: The ending part of a derived branch
- Size of the skeleton: Number of all the skeleton pixels
- Length of the skeleton image
- Width of the skeleton image
- R1: Report between width and length of the skeleton
- Maxleng: Maximum length between two endpoints
- Minleng: Minimum length between two endpoints

From Fig. 6 we can conclude that the skeleton shape of melanoma and non-melanoma is very different. Melanoma skeleton contains several branches and several endpoints.

Texture Features

In order to quantify the texture present in a lesion, a set of statistical texture descriptors are used, namely:

Contrast:

$$Cont = \sum_{n=0}^{G-1} n^2 \left\{ \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j) \right\}, |i, j| = n \quad (2)$$

G is the number of gray levels used.

This measure of contrast or local intensity variation will favour contributions from $P(i, j)$ away from the diagonal.

Correlation:

$$Corl = \frac{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i \times j\} \times P(i, j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y} \quad (3)$$

μ_x, μ_y, σ_x and σ_y are the means and standard deviations of P_x and P_y .

Correlation is a measure of gray level linear dependence between the pixels at the specified positions relative to each other.

Energy (Angular Second Moment (ASM)):

$$ASM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{P(i, j)\}^2 \quad (4)$$

ASM is a measure of homogeneity of an image. A homogeneous scene will contain only a few gray levels, giving a GLCM with only a few but relatively high values of $P(i, j)$. Thus, the sum of squares will be high.

Entropy:

$$Entropy = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j) \times \log(P(i, j)) \quad (5)$$

Inhomogeneous scenes have low first order entropy, while a homogeneous scene has a high entropy.

Inverse Difference Moment (IDM):

$$IDM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i, j)^2} P(i, j) \quad (6)$$

IDM is also influenced by the homogeneity of the image.

Mean:

$$Mean(\mu) = \sum_N \frac{1}{N} P(i, j) \quad (7)$$

It provides average Color value in the image.

Variance:

$$E_i = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - \mu)^2 P(i, j) \quad (8)$$

μ is the mean value of $P(i, j)$.

This feature puts relatively high weights on the elements that differ from the average value of $P(i, j)$.

Standard Derivation (Stddev):

$$Stddev(\sigma) = \sqrt{\left(\frac{1}{N} (P(i, j) - E_i)^2 \right)} \quad (9)$$

N is the total number of pixels in the image.

The standard deviation is the square root of the variance of the distribution.

Root Mean Square (RMS):

$$RMS = \sqrt{\frac{1}{n \times m} \sum_{j=1}^n \sum_{i=1}^m [P(i, j)]^2} \quad (10)$$

RMS is defined as the some sort of average or sum (or integral) of square of the error n, m is the dimensions of the image.

Kurtosis:

$$Kurtosis = \frac{1}{\sigma^3} \sum_{i=0}^N (i - \mu)^3 P(i, j) \quad (11)$$

It is measure the peakness or flatness of a distribution relative to a normal distribution.

Skewnes:

$$Skewnes = \frac{1}{\sigma^3} \sum_{i=0}^N (i - \mu)^3 P(i, j) \quad (12)$$

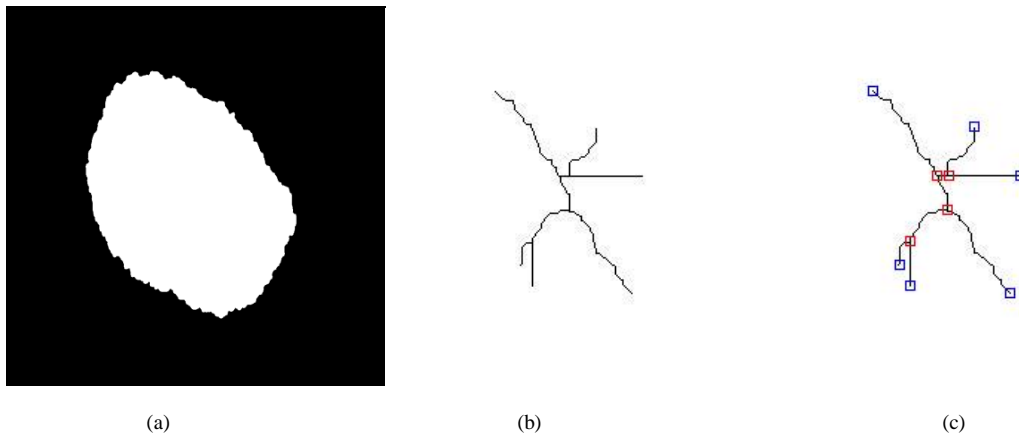


Fig. 5: Example of the extracted features from skeleton; (a) segmentation result; (b) skeleton of the lesion; (f) skeleton features: branches in red and endpoints in bleu

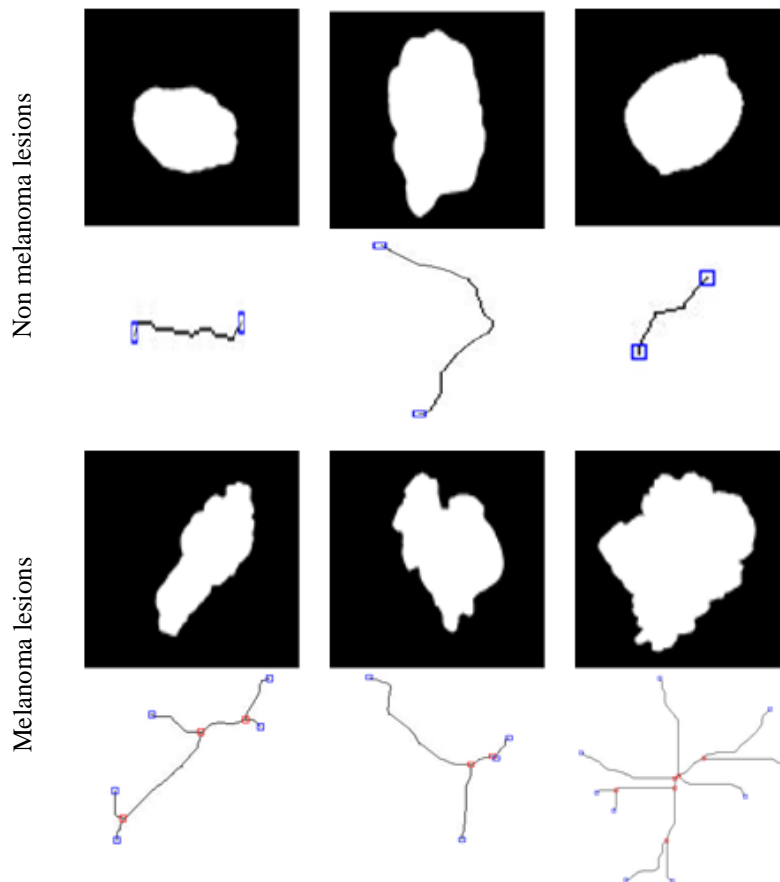


Fig. 6: Skeletization of non-melanoma and melanoma lesions

It is the measurement of the inequality of the intensity level distribution about the mean.

Color Features

Dermatologists can also diagnosis the difference between two types of lesions (melanoma and non-melanoma) just by calculating the number of colors that exist, which are 6 colors (white, red, light brown, dark brown, blue-gray and black). To calculate the number of colors that exist in the lesion, we proceed as follows:

$$N_{Bc} = \sum (I_R < 0.2 \& I_G < 0.2 \& I_B < 0.2) > 0) * 100 / N_T \quad (13)$$

$$N_{Rd} = \sum (I_R > 0.8 \& I_G < 0.2 \& I_B < 0.2) > 0) * 100 / N_T \quad (14)$$

$$N_{Bg} = \sum \left(I_R < 0.2 \& I_G > 0.32 \& I_G < 0.72 \& I_B > 0.34 \& I_B < 0.74 \right) > 0) * 100 / N_T \quad (15)$$

$$N_w = \sum (I_R > 0.8 \& I_R > 0.8 \& I_B > 0.8) > 0) * 100 / N_T \quad (16)$$

$$N_{Lb} = \sum \left(I_R > 0.6 \& I_R < 1 \& I_G > 0.32 \& I_G < 0.72 \& I_B > 0.05 \& I_B < 0.45 \right) > 0) * 100 / N_T \quad (17)$$

$$N_{Db} = \sum \left(I_R > 0.2 \& I_R < 0.6 \& I_G > 0.06 \& I_G < 0.46 \& I_B > 0 \& I_B < 0.33 \right) > 0) * 100 / N_T \quad (18)$$

where, N_T is the number of pixels in the lesion. I_R , I_G and I_B are consecutively the intensities of pixels in Red, Green and Blue components.

Features Engineering

In classification challenges, feature engineering is considered as an essential step, which help in increasing the classification accuracy.

First, as the features are not always in the same range of values we first perform a normalization of the data. Among normalization methods we adopt the standard score transformation (z-score):

$$Z = \frac{X - \mu}{\sigma} \quad (19)$$

Z is the normalized value for each element of X (features matrix), where μ is the mean (average) and σ is the standard deviation from the mean.

Second, to improve the classification rate, we need to select the relevant features and eliminate redundant features for a better data representation. ReliefF algorithms, Correlation-based Features Selection (CFS), Recursive Feature Elimination (RFE) and Chi2 method are the most used algorithms in features selection (Chandrashekar and Sahin, 2014; Oliveira, 2014; Nasir *et al.*, 2017):

- ReliefF: This algorithm determines the nearest neighbors from several samples token randomly from the data set. For each feature's value is compared with those of the nearest neighbors, then the score of every feature are then updated. The purpose is to estimate the quality of the attributes according to their values and to distinguish between samples that are close to each other
- Correlation based Feature Selection (CFS): This algorithm tries to determine a set of features that are well correlating individually with the classes but with a small inter-correlation between each other
- Recursive Feature Elimination (RFE): This algorithm is based on the multivariate mapping method, that build a model and choose either the worst or the best performing. The idea of RFE is to take initially all the features of a large region then to progressively exclude features that not contribute in discriminating classes
- Chi2 method: This method calculates the weights of the features using the Chi2 statistic. It measures the independence between a feature and a class. It begins with a high level of significance for all features for discretization and each feature is sorted according to its values. Then we perform the classification of the features

Classification

Several classification techniques have been developed. Certainly each of them has its own weaknesses and their use depends on the types of data to classify, but it is also possible to find a good compromise and to use simple methods that lead to high performance with a fast computing time. In our work, the choice of a classification algorithm to use is based on our previous work (Filali *et al.*, 2018) where the SVM with quadratic kernel gives the best result compared with other classifiers and kernels. The SVM is a supervised learning. It finds the optimal separation between classes by using a hyper plane with the maximum margin.



Fig. 7: Example of images from database: melanoma in the first line and non-melanoma in the second line

Experimentation and Results

Database

The database used to implement and evaluate our proposed approach is composed of macroscopic and dermoscopic skin lesion images. It contains 1000 images collected from two databases where 392 are "melanoma" and 608 are "non-melanoma" images (<http://dermis.net/dermisroot/en/home/index.htm> at dermis.net; <https://challenge.kitware.com>). The two databases are classified by experts and contain the segmentation ground-truth. Figure 7, we present an example of melanoma and non melanoma images from the database. As is customary and to evaluate our proposed approach, the dataset is randomly divided into training and test sets using k-fold cross validation (5 fold in this study). That preserves the fairness of performance of our proposed approach.

Performance Measure

To evaluate our proposed approach, the performance measure used in this paper are Sensitivity (TP rate), specificity (TN rate) and precision:

$$\text{Sensitivity} = TP / (TP + FN) \quad (20)$$

$$\text{Specificity} = TN / (FP + TN) \quad (21)$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (22)$$

Where:

- TP = The number of true positives
- TN = The number of true negatives
- FP = The number of false positives
- FN = The number of false negatives

Results and Discussion

As it were mentioned above, the result of the segmentation, features extraction, features selection and classification result are discussed in this section.

Segmentation

In order to evaluate the contribution of the PDE image decomposition to the segmentation in terms of precision, we carry out a preliminary study concerning a random number of images.

The segmentation is performed on the object component that contain only the object shape without any noise or texture. Otsu algorithm is used as a segmentation algorithm in our approach because of its simplicity, the effectiveness of the result and time of execution. Table 1 presents the average of sensitivity and specificity of the segmentation of the whole database in regard of the database ground of truth.

From results presented in Table 1, we can say that the segmentation of the object component obtained by the use of the PDE decomposition is much better than segmenting the original skin lesion image.

Features Extraction

After performing a good segmentation, we use the skeleton of the segmented lesion and extract skeleton features as described in section 2.1 which are the size of the skeleton, length, width, the report between length and width, maximum and minimum length between two endpoints and the number of sub-branches. It has been noted that the skeleton of a melanoma cancer is more developed and has more branches and sub-branches than the skeleton of a non-melanoma cancer which is smaller and has fewer branches (Red are branches, blue are endpoints) (Fig. 5 and 6).

An evaluation of the use of only the skeleton features is first conducted in comparison with other features such as the number colors and textural features of the lesion as presented in sections 2.2 and 2.3. Table 2 presents results of the classification accuracy using our proposed features in comparison with color and texture features.

From the results in Table 2, it is clear that the use of the new skeleton features increases the classification accuracy.

In the next section, we will concatenate all the features as described in section 2 and make a Features-Engineering to keep only the relevant features and delete redundancy.

Features Engineering

To improve the classification rate, we will combine all the features presented in section 2: textural, skeleton and colors features. Then, we will make a features selection to keep only the best features, since not all features are relevant.

The result of features selection using different algorithms is presented as follows:

- Chi2: 4 features selected; **Number of branches, Report of Wight/length**, Number of colors and Contrast
- RFE: 5 features selected; **Skeletal length, Number of endpoints**, IDM, RMS and Light brown color
- ReliefF: 6 features selected; **Sub branch of the skeletal, number of branches, Report of Wight/length**, IDM, Light brown color and Number of colors
- CFS: 6 features selected; **Skeleton size, Number of endpoints, number of sub-branches of the skeleton, Wight of the skeleton**, Number of colors and RMS

The average of the classification accuracy using SVM and features selected by each of the selection algorithms presentend above is presented in Table 3. The best classification accuracy is by the features selected by ReliefF (skeletal Sub branch, number of branches, Report of Wight/length, IDM, Light brown color and Number of colors) imply that the new features based on skeletization are very important and can improve the SL classification results.

In the following sections, the classification will be done using only the six features selected by the ReliefF algorithm.

Classification Result

In this section, our proposed approach is evaluated first by presenting the confusion matrix, which shows the

performance of our classification approach when predicting the skin lesion type using cross-validation. Then, a comparative study between our proposed approach with recent approaches from literature will be conducted.

In Table 4 the confusion matrix of the proposed approach using 5 folds in cross validation with 96.42% in sensitivity and 96.54% in specificity demonstrate the effectiveness of the new used features.

Table 5 presents the classification accuracy of our proposed approach in comparison with recent approaches proposed respectively by Immagulate and Vijaya (2015), Dalila *et al.* (2017), Waheed (2017) and Victor and Ghalib (2017) based on the same dataset. In fact, in (Immagulate and Vijaya, 2015) authors used color and texture features as a FE and SVM as a classifier. They got an average of 85.00% accuracy rate. Dalila *et al.* (2017) authors extracted geometrical properties, texture and colors features and used two classifiers K-Nearest Neighbor (KNN) and ANN and the classification accuracy was 89.50%. Waheed (2017) the features used are colors in HSV and RGB space and texture using GLCM and the SVM as a classifier. The classification accuracy was 91.20%. In another hand, the proposed approach in (Korotkov and Garcia, 2012) uses area, mean, variance and the standard deviation as a features extraction and SVM as classifiers got 93.50% accuracy rate. Finally, our proposed approach gives the best result and exceeds all other approaches with a 96.50% accuracy rate.

Table 1: Average results of sensitivity and specificity of the segmentation using Otsu algorithm

	Without PDE	Using PDE
Sensitivity	88.30	93.20
Specificity	87.60	92.35

Table 2: Results of classification using features resulting from the skeleton, texture and color

	Color	Texture	Skeleton
Accuracy (%)	65.03	82.02	87.00

Table 3: Classification accuracy average using SVM for each of features obtained by FS algorithms

	Chi2	RFE	ReliefF	CFS
Accuracy (%)	85.26	92.80	96.50	94.00

Table 4: Confusion matrix of the support vector machine with ReliefF as a features selection

	Melanoma	Non-melanoma
Melanoma	378	21
Non-melanoma	14	587

Table 5: Results of the proposed approach compared to recent approaches from literature

	Immagulate and Vijaya (2015)	Dalila <i>et al.</i> (2017)	Waheed (2017)	Korotkov and Garcia (2012)	Proposed approach
Accuracy (%)	85.00	89.50	91.2	93.50	96.50

Conclusion

Skin lesion is considered as the most dangerous cancer. If diagnosed and treated earlier, it is almost always curable. Computer-Aided Diagnosis (CAD) is widely used to automatically diagnosis a skin lesion as melanoma or not. The process is often done in three basic steps; segmentation, features extraction and classification. A good segmentation implies an exact identification of the lesion, which influences the quality of the extracted characteristics and therefore the accuracy of the classification. The presence of texture in the skin lesion images makes the segmentation very difficult.

In this article, we first propose to add a preprocessing step based on a multiscale image decomposition based on an algorithm derived from PDE. This algorithm gives two components as a result. The first one can be attributed to the low frequency which contains only the lesion geometry and boundaries while the second component contains only texture. The geometric component is thus exploited to have a suitable segmentation while the second component provides a good characterization of the texture.

The second proposed contribution in this paper is to provide new descriptors of the lesion's shape based on the skeleton pattern, which provide useful and robust features for classification and can thus be used in related medical research.

The experimentations conducted, justified the use of a pretreatment by the PDE to segment the image and to better identify the lesion on the one hand and on the other hand, showed that the new proposed features based on the skeletonization allowed a clear improvement of the classification rate compared to other approaches from literature.

In future work, will be concentrated on improving more the segmentation; to reduce the processing time caused by pretreatment, or to develop a fast alternative method.

Acknowledgement

This research work has been funded by the laboratories LIIAN and LESSI and the Faculty of Sciences, University Sidi Mohamed Ben abdellah, Fez, Morocco.

Author's Contributions

Youssef Filali: Developpement, expermimentations testing, validation and writing the manuscript.

Hasnae El Khoukhi: Expermimentations validation and proofreading.

My Abdelouahed Sabri: Concept developpement, experimentations testing, validation and proof reading.

Ali Yahyaouy: Concept developpement and proof reading.

Abdellah Aarab: Concept developpement, experimentations validation and proof reading.

Ethics

This paper is original with unpublished material. The corresponding author confirms that this manuscript has not been published elsewhere and that no ethical issues are involved.

References

- Arcppli, C. G. Sanniti and D.I. Baja, 1985. A width-independent fast thinning algorithm. *IEEE Trans. Patt. Anal. Machine Intell.*, 7: 463-474.
DOI: 10.1109/TPAMI.1985.4767685
- Aubert, G. and J.F. Aujol, 2004. Modeling very oscillating signals. Application to image processing. *Applied Math. Optim.*, 51: 163-182.
DOI: 10.1007/s00245-004-0812-z
- Aujol, J.F., G. Aubert, L. Blanc-Féraud and A. Chambolle, 2005. Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imag. Vis.*, 22: 71-88.
DOI: 10.1007/s10851-005-4783-8
- Barata, A.C.F., E.M. Celebi and J. Marques, 2018. A survey of feature extraction in dermoscopy image analysis of skin cancer. *IEEE J. Biomed. Heal. Inform.* DOI: 10.1109/JBHI.2018.2845939
- Beumier, C. and X. Neyt, 2017. Camera motion compensation from T-junctions in distance map skeleton. *Proceedings of the International Conference on Systems, Signals and Image Processing*, May 22-24, IEEE Xplore Press, Poznan, Poland, pp: 1-5.
DOI: 10.1109/TWSSIP.2017.7965613
- Chandrashekar, G. and F. Sahin, 2014. A survey on feature selection methods. *Comput. Electr. Eng.*, 40: 16-28. DOI: 10.1016/j.compeleceng.2013.11.024
- Chopra, A. and B.R. Dandu, 2012. Image segmentation using active contour model. *Int. J. Comput. Eng. Res.*, 2: 819-822.
- Dalila, F., A. Zohra, K. Reda and C. Hocine, 2017. Segmentation and classification of melanoma and benign skin lesions. *Optik (Stuttg)*, 140: 749-761.
DOI: 10.1016/j.ijleo.2017.04.084
- De Oliveira, J.R., R. Aparecida, F. Romero and S. Carlos, 2009. Image skeletonization method applied to generation of Topological Maps. *Proceedings of the 6th Latin American Robotics Symposium*, Oct. 29-30, IEEE Xplore Press, Valparaiso, Chile, pp: 1-6.
DOI: 10.1109/LARS.2009.5418319
- Deepa, S.N. and B. Aruna Devi, 2011. A survey on artificial intelligence approaches for medical image classification. *Indian J. Sci. Technol.*, 4: 1583-1595.
DOI: 10.17485/ijst/2011/v4i11/30291

- Dreiseitl, S., L. Ohno-Machado, H. Kittler, S. Vinterbo and H. Billhardt *et al.*, 2001. A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *J. Biomed. Inform.*, 34: 28-36. DOI: 10.1006/jbin.2001.1004
- Ferris, L.K., J.A. Harkes, B. Gilbert, D.G. Winger and K. Golubets *et al.*, 2015. Computer-aided classification of melanocytic lesions using dermoscopic images. *J. Am. Dermatol.*, 73: 769-76. DOI: 10.1016/j.jaad.2015.07.028
- Filali, Y., M.A. Sabri and A. Aarab, 2017a. An improved approach for skin lesion analysis based on multiscale decomposition. *Proceedings of the International Conference on Electrical and Information Technologies*, Nov. 15-18, IEEE Xplore Press, Rabat, Morocco, pp: 15-18. DOI: 10.1109/EITech.2017.8255250
- Filali, Y., A. Ennoui, M.A. Sabri and A. Aarab, 2017b. Multiscale approach for skin lesion analysis and classification. *Proceedings of the International Conference on Advanced Technologies for Signal and Image Processing*, May 22-24, IEEE Xplore Press, Fez, Morocco, pp: 1-6. DOI: 10.1109/ATSIP.2017.8075545
- Filali, Y., A. Ennoui, M.A. Sabri and A. Aarab, 2018. A study of lesion skin segmentation, features selection and classification approaches. *Proceedings of the International Conference on Intelligent Systems and Computer Vision*, Apr. 2-4, IEEE Xplore Press, Fez, Morocco, pp: 1-4. DOI: 10.1109/ISACV.2018.8354069
- Filho, M., Z. Ma and J.M.R.S. Tavares, 2015. A review of the quantification and classification of pigmented skin lesions: From dedicated to hand-held devices. *J. Med. Syst.*, 39: 1-12. DOI: 10.1007/s10916-015-0354-8
- Garnavi, R., M. Aldeen, M.E. Celebi and A. Bhuiyan, 2011. Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels. *Computerized Med. Imag. Graph.*, 35: 105-115. <http://dermis.net/dermisroot/en/home/index.htm@dermis.net>
<https://challenge.kitware.com>
- Immagulate and M.S. Vijaya, 2015. Categorization of non-melanoma skin lesion diseases using support vector machine and its variants. *Int. J. Med. Imag.*, 3: 34-40. DOI: 10.11648/j.ijmi.20150302.15
- Korotkov, K. and R. Garcia, 2012. Computerized analysis of pigmented skin lesions: A review. *Artif. Intell. Med.*, 56: 69-90. DOI: 10.1016/j.artmed.2012.08.002
- Leguizam, D.N., 2015. Computerized diagnosis of melanocytic lesions based on the ABCD method. *Proceedings of the 41th Latin American Computing Conference*, Oct. 19-23, IEEE Xplore Press, Arequipa, Peru, pp: 1-12. DOI: 10.1109/CLEI.2015.7360029
- Mathew, S. and D. Sathyakala, 2015. Segmentation of skin lesions and classification by neural network. *Int. J. Adv. Res. Electron. Commun. Eng.*, 4: 192-201.
- Nasir, M., M. Attique Khan, M. Sharif, I.U. Lali and T. Saba *et al.*, 2017. An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. *Microsc. Res. Tech.*
- Nezhadian, F.K. and S. Rashidi, 2017. Melanoma skin cancer detection using color and new texture features. *Proceedings of the Artificial Intelligence and Signal Processing Conference*, Oct. 25-27, IEEE Xplore Press, Shiraz, Iran, pp: 1-5. DOI: 10.1109/AISP.2017.8324108
- Oliveira, R.B., 2014. Features selection for the classification of skin lesions from images. *Proceedings of the 6th European Conference on Computational Fluid Dynamics*, Jul. 20-25, Barcelona, Spain, pp: 1-2.
- Oliveira, R.B., J.P. Papa, A.S. Pereira and J.M.R.S. Tavares, 2016. Computational methods for pigmented skin lesion classification in images: Review and future trends. *Neural Comput. Applic.*, 29: 613-636. DOI: 10.1007/s00521-016-2482-6
- Pennisi, A., D.D. Bloisi, D. Nardi, A. Rita and C. Mondino *et al.*, 2016. Computerized medical imaging and graphics skin lesion image segmentation using Delaunay triangulation for melanoma detection. *Comput. Med. Imag. Graph.*, 52: 89-103. DOI: 10.1016/j.compmedimag.2016.05.002
- Sabri, M.A., A. Ennoui and A. Aarab, 2016. Automatic estimation of clusters number for K-means. *Proceedings of the 4th IEEE International Colloquium on Information Science and Technology*, Oct. 24-26, IEEE Xplore Press, Tangier, Morocco, pp: 450-454. DOI: 10.1109/CIST.2016.7805089
- Saha, P.K., G. Borgefors and G. Sanniti, 2016. A survey on skeletonization algorithms and their applications. *Patt. Recognit. Lett.*, 76: 3-12. DOI: 10.1016/j.patrec.2015.04.006
- Vasconcelos, M.J.M., L. Rosado and M. Ferreira, 2015. A new risk assessment methodology for dermoscopic skin lesion images. *Proceedings of the IEEE International Symposium on Medical Measurements and Applications*, May 7-9, IEEE Xplore Press, Turin, Italy, pp: 570-575. DOI: 10.1109/MeMeA.2015.7145268
- Victor, A. and M.R. Ghalib, 2017. Automatic detection and classification of skin cancer. *Int. J. Intell. Eng. Syst.*, 10: 444-451. DOI: 10.22266/ijies2017.0630.50

- Waheed, Z., 2017. An efficient machine learning approach for the detection of melanoma using dermoscopic images. Proceedings of the International Conference on Communication, Computing and Digital Systems, Mar. 8-9, IEEE Xplore Press, Islamabad, Pakistan, pp: 316-319. DOI: 10.1109/C-CODE.2017.7918949
- Zhou, H., G. Schaefer, M.E. Celebi, H. Iyatomi and K.A. Norton *et al.*, 2010. Skin lesion segmentation using an improved snake model. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology, Aug. 31-Sept. 4, IEEE Xplore Press, Buenos Aires, Argentina, pp: 1974-1977. DOI: 10.1109/IEMBS.2010.5627556
- Zhu, S., J. Yang and X. Zhu, 2014. A Thinning Model for Handwriting-Like Image Skeleton. In: Computer Engineering and Networking, Wong, W.E. and T. Zhu (Eds.), Springer, Cham, pp: 687-693.