34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

On the Role of Texture and Color in the Classification of Dermoscopy Images

Jorge S. Marques[1] Catarina Barata[1] Teresa Mendonça[2]

Abstract— **This paper addresses the detection of melanoma lesions in dermoscopy images, using texture and color features. Although melanoma detection has been studied in several works, using different types of texture, color and shape features, it is not always clear what is the role of each set of features and which features are most discriminative. This papers aims at clarifying the role of texture and color features. Furthermore, the proposed systems is based on features which can be easily implemented and tested by other resarchers. It is concluded that both types of features achieve good detection scores when used alone. The best results (SE=94.1%, SP=77.4%) are achieved by combining them both.**

I. Introduction

Skin cancer is a major source of mortality in modern societies. Although there are several types of skin cancer, melanoma is by far the most aggressive. The survival rate for melanoma lesion strongly depends on its early detection and removal. Therefore, a large effort has been done to develop efficient screening procedures. Dermoscopy is a wide spread imaging technique which allows the magnification of the skin lesion and the observation of the inner layers of the skin through the use of a microscope and special illumination [1]. This technique allows a more detailed inspection of the skin lesion and, in some cases a more accurate diagnosis, since several dermoscopic structures become visible [2].

The interpretation of dermoscopy images requires a long training of medical specialists. Several medical procedures were developed to guide the specialists such as the Menzies method [3], the ABCD rule [4] and the 7 point checklist [5]. For example, the ABCD rule takes 4 criteria into account: asymmetry (A), border (B), number of colors (C) and differential structures (D). Each of these criteria receives a score and the final score associated to a lesion is the weighted sum of the four scores. Although these methods help the medical doctor to extract useful information from the image and to combine it into a malignancy degree, the final decision remains difficult and subjective in many cases.

These difficulties led to the development of computer aided systems for the diagnosis of skin lesions [6], [7], [8], [9], [10], [11], [12], [14]. Two approaches have been adopted. Some research groups try to mimic the tasks performed by specialists and to extract the same type of scores in an automatic way [14]. This requires the extraction of very subtle information namely number of colors and differential structures such as regression regions, blue-white veil, vascular networks, pigment network, dots, globules and streaks (Fig. 1).

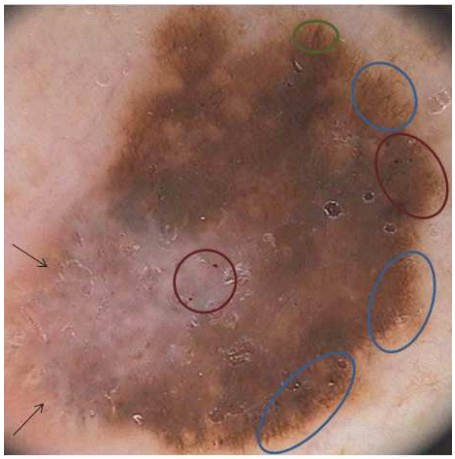

Fig. 1. Dermoscopy image: differential structures. Atypical Pigment Network (blue ellipses); Irregular Dots (red ellipses); Atypical Streaks (green circle); Regression Region (black arrows)

The majority of works adopts a pattern recognition approach. They extract a large number of features from each dermoscopy image (texture, color, shape features) and train a classifier to discriminate melanomas from other types of lesions [6], [7], [8], [9], [10], [11], [12]. Very interesting results have been achieved by different authors while using distinct types of features and classifiers. Between the most interesting works is the one of Ganster et al. [6] that uses features which aim to characterize the color and shape of each lesion and a kNN classifier to distinguish between melanoma and benign nevi. This group achieved a SE=87% and a SP=92% in a large dataset with more than 5000 lesions and 100 melanomas. Rubegni et al. [7] also obtain interesting results with their system that uses an ANN as classifier to distinguish between melanoma and other melanocytic lesions. They use a large set of features that discriminate shape, color and texture and achieved a SE=96% and a SP=93% in a database of 588 images (37% of which were melanomas). The works of Celebi et al. [11] and Iyatomi et al. [12] are also significant. The first group proposes a system for diagnosing melanomas using a set of shape, color and textural features and SVM as the learning algorithm. They obtained a SE=93% and a SP=92% in a database with

*This work was supported by FCT in the scope of projects PTDC/SAUBEB/ 103471/2008 and PEst-OE/EEI/LA0009/2011.

[1]J. Marques and C. Barata are with Instituto Superior Tecnico and Institute for Systems and Robotics, Lisbon, jsm at isr.ist.utl.pt and ana.c.fidalgo.barata@ist.utl.pt

[2]T. Mendonça is with the Faculty of Sciences of University of Porto, tmendo at fc.up.pt

Fig. 2. Block diagram of the CAD system

90 melanomas and 474 nevi. The system developed by the second group is currently available in the Internet. They use features which describe the color, degree of symmetry, border and texture of a given lesion and an ANN do classify a given lesion.

Previous works, use all the features considered as relevant for the decision. However, in most cases no attempt is made to assess the importance of each type of feature. A notable exception is the work of Seidenari et al. [10], which aims to determine the importance of color distribution in melanocytic lesions, achieving a SE=88% and a SP=86%. Furthermore, it is not always clear how the features are computed.

This paper aims to achieve two goals: i) describing a reproducible systems which can be easily implemented by other researchers and ii) assessing the importance of texture and color features in the final decision. It is shown that both features are relevant. Good results can be obtained by using texture features or color features only. However improved results are achieved by combining them both (SE=94.1%, SP=77.4%).

II. Overall Description

The proposed systems aims to detect melanoma lesions in dermoscopy images and to discriminate them from other types of lesions. Fig. 2 shows the block diagram. First, the skin lesion is segmented to separate the lesion from healthy skin. Then, a set of features is extracted from the lesion region. Finally, a binary classifier is trained to discriminate melanomas from other types of lesions. An adaptive threshold algorithm was used in the first block (see details in [13]). Then texture and color features are extracted as explained in section 3. Finally, a classifier is used to obtain the final decision.

III. Image Features

Two types of features are considered in this study: i) texture features which convey information about the differential structures (pigment network, dots, streaks, etc) present in the lesion and ii) color features which contain information about the color distribution and number of colors in the skin lesion. We decided to consider color and texture descriptors as random variables and to characterize them by their normalized histograms.

A. Texture Features

The histograms of gradient magnitude and phase are powerful descriptors of the image texture and have been sucsessfully used in several image analysis problems (e.g., see [15]). We have considered both the histogram of the gradient magnitude and phase, computed for all the lesion pixels. First, we convert the RGB dermoscopy image into a gray level image. This is done by selecting the color channel with the highest entropy i.e., the most informative channel.

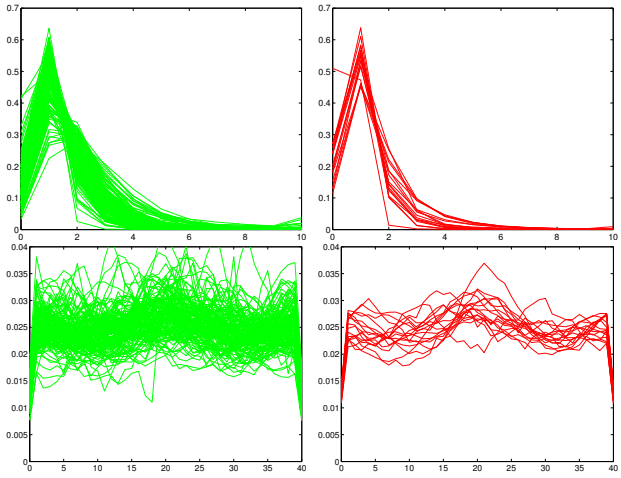

Fig. 3. Texture features for non-melanomas (green) and melanomas (red): histograms for the gradient magnitude (1st row) and orientation (2nd row)

To clarify this issue, let $p_i(k)$ denote the probability of the $i-th$ color channel being equal to k at a lesion pixel x. This distribution can be easily obtained by computing the histogram of each color channel inside the lesion. The entropy is defined by,

$$S(i) = -\sum_k p_i(k) \log p_i(k) \,, \tag{1}$$

and we select the channel with highest entropy.

In order to compute the image gradient, we first filter the gray level image using a Gaussian filter with $\sigma = 2$, and then compute the gradient vector at each point $g(x) = [g_1(x)\ g_2(x)]^T$ using Sobel masks. The gradient magnitude and orientation are then computed as usual

$$\|g(x)\| = \sqrt{g_1^2(x) + g_2^2(x)} \,, \qquad \theta(x) = \tan^{-1}\left(\frac{g_2(x)}{g_1(x)}\right) \,. \tag{2}$$

The gradient magnitude and orientation are then characterized by their histograms, using $M_a = 10$ and $M_\theta = 40$ bins, respectively.

Fig. 3 shown the histograms for all the lesions considered in this study (see details of the database in section IV). Although the histograms of melanomas look similar to the ones associated to other lesions, a good discrimination is still possible, as shown in section IV.

B. Color Features

Color perception is a response of the human brain to the electromagnetic radiation arriving at the human retina. Fortunately, the trichromatic theory of Young and Helmholtz suggests that color can be represented by three coefficients characterizing the response of three types of photosensitive cells (known as cones) in the retina. Several color representations have been proposed but it is not clear which of them performs best in medical image analysis. In this paper, we

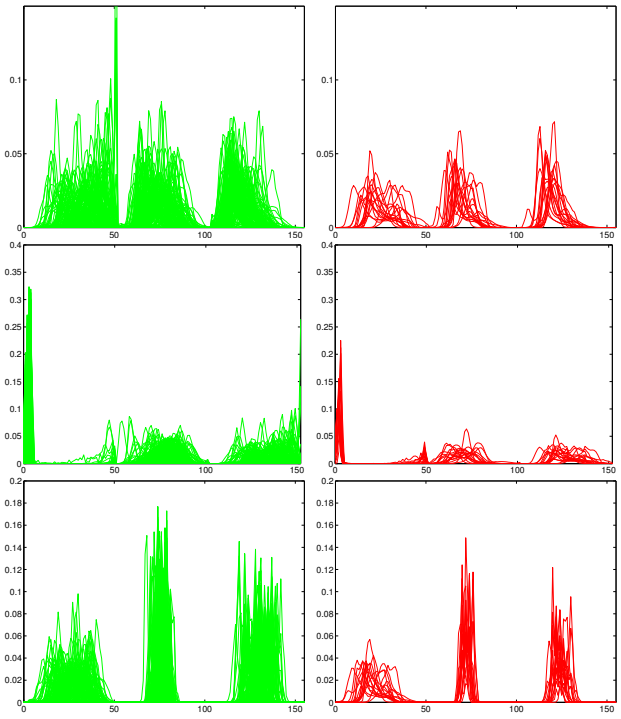

Fig. 4. Color features for non-melanomas (green) and melanomas (red): histograms RGB (1st row), HSV (2nd row) and L*a*b* (3rd row) coefficients

will consider three color spaces: RGB, HSV and L*a*b*. RGB is the most popular color representation but it has two main disadvantages: perceptual color differences cannot be reliably measured in RGB space and RGB coefficients do not provide an intuitive description of color. The HSV coordinates are much closer to a perceptual description of color into hue, saturation and value components and the L*a*b* color space was designed to provide a good correlation between perceptual difference of colors and measured color distance. The L*a*b* coordinates are the most obvious choice for the problem we wish to solve.

We have characterized the color content of the skin lesion by three histograms (one histogram per component) which approximate the probability distribution of each color component. We could have used the joint histogram of the three components since they are not statistically independent. However, the joint descriprion involves the computation of a volume of probabilities which is much more demanding from the computational and storage points of view. Fig. 4 shows the color histogram for the three different color spaces.

C. Multiple regions

Previous sections make two simplifying assumptions: i) all the regions inside the lesion play the same role; ii) the lesion is homogeneous and can be described by a global feature vector (histogram) which represents the average properties. The first assumption is too restrictive since we know that regions close to the lesion border play an important role (e.g., the smooth or abrupt transition between the lesion and the healthy skin are taken into account in the ABCD rule).

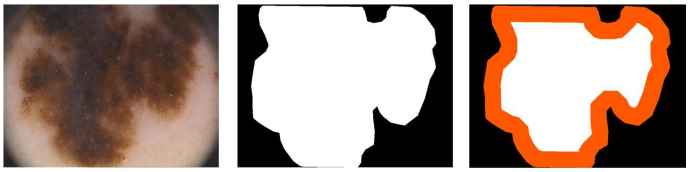

Fig. 5. Segmentation: (left) Original image; (mid) Segmented image; (right) Two regions discriminated; border(orange) and inner (white)

We have therefore tested a multiple region approach in which we consider two regions: the inner region and the border region. The inner region is obtained by eroding the segmented image using a disk of radius r; r is chosen to be $1/10$ of the lesion diameter. The boundary region is obtained by dilating the segmented region, using the same structured element, and then removing the pixels of the inner region. A descriptive example can be seen on Fig.5. The color and texture histograms described in the previous section are computed for each of the two regions and they are concatenated to obtain an extended feature vector.

IV. Experimental Results and Discussion

The system described in section II was evaluated with a database of 163 dermoscopy images (17 melanomas, 146 nevi), kindly provided by Hospital Pedro Hispano. These images were obtained by specialists during clinical routine using a magnification of 20×. They have a resolution of 573×765 pixels and are available in bitmap or jpeg formats.

In order to avoid segmentation errors, the output of the first block was checked by a specialist and the segmentation errors were corrected. We have evaluated the performance of the melanoma detection system for each type of feature namely: T1) gradient magnitude and T2) gradient orientation histograms; C1) RGB, C2) HSV and C3) L*a*b* histograms. We have also considered all texture features T=(T1, T2), all color features C=(C1, C2, C3), the best texture and color feature (T1,C2) and all texture and color features All=(T1, T2, C1, C2, C3).

Since the data set is small we have tested the system using a leave-one-out approach. This means that for each set of features we trained 163 classifiers, using all the feature vectors except one which is used for testing. The final score is obtained by averaging the 163 individual scores. In addition, since the two classes are unbalanced we have repeated the melanoma features belonging to each training set, until we obtain the same number of examples for both classes.

The statistics of all these tests are provided in table 1. The conclusions are very interesting. We obtain good results using texture features only (SE=82.4%, SP=76.7%). The best texture features are the histograms of gradient amplitude. These results are rather surprising since the histograms of melanomas and non-melanomas shown in Fig. 3 look similar. We also obtain good results using color features only (SE=82.4%, SP=74.7%). The HSV color space proved to be slightly better than the other color representations.

TABLE I

SYSTEM EVALUATION USING TEXTURE FEATURES ($T = (T1, T2)$), COLOR FEATURES ($C = (C1, C2, C3)$) AND BOTH ($All = (T, C)$)

	1 Region			2 Regions		
Features	SE	SP	ACC	SE	SP	ACC
T1	82.4	76.7	77.3	76.5	71.2	71.8
T2	64.7	42.5	44.8	76.5	50.0	52.8
T	82.4	69.9	71.2	70.6	69.9	69.9
C1	82.4	68.5	69.9	82.4	**74.0**	**74.8**
C2	82.4	74.7	75.5	82.4	70.5	71.8
C3	82.4	71.9	73.0	82.4	67.8	69.3
C	82.4	75.3	76.1	82.4	70.5	71.8
(T1,C2)	**94.1**	**77.4**	**79.1**	82.4	70.5	71.8
All	88.2	75.3	76.7	**88.2**	70.5	72.4

TABLE II

CONFUSION MATRIX: TN, FP (1ST LINE), FN AND TP (2ND LINE).

	0	1
0	113	33
1	1	16

This means that perceptually inspired color space performed slightly better in this task. Furthermore, we obtain an additional performance gain when we use all texture and color features. However, the best results are achieved when we use the best texture features (T1) and the best color features (C2) leading to SE=94.1%, SP=77.4%. The use of multiple regions with separate color and texture features does not bring further improvements.

The confusion matrix associated to the best classifier is shown in table II. All melanomas (except one) were detected. Fig. 6 shows one example of each of the four classification cases: TN, FP (1st line), FN and TP (2nd line).

V. CONCLUSION

This paper describes a pattern recognition system for the detection of melanomas in dermoscopy images. We tried to answer to following question: which features, texture or color, convey the most relevant information? It has been concluded that both features achieve very interesting results (SE=82.4%, SP=76.7% and SE=82.4%, SP=74.7%, respectively) if used alone. If we combine both types of features the performance of the system increases (SE=94.1%, SP=77.4%).

We stress that we have used standard texture and color features. Therefore, the proposed system can be easily implemented by any researcher wishing to test this approach and improve it using additional features.

Future work should extend these results by using a larger database of images and additional image features. Shape and symmetry features can be easily included in the system and may further improve these results.

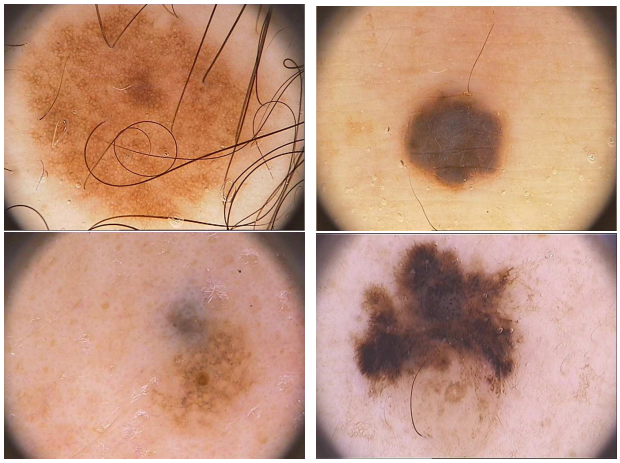

Fig. 6. Examples of the 4 classification cases: TN, FP (1st line) FN and TP (2nd line)

Acknowledgement

We thank Dr. Jorge Rozeira from the Hospital Pedro Hispano for kindly providing the database of dermoscopy images and for segmenting these lesions.

REFERENCES

[1] Dermoscopy Tutorial, www.dermoscopy.org/atlas/base.htm
[2] J. Mayer, Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma, Medical Journal Australia, 206-210, 1997.
[3] Scott W. Menzies, A Method for the diagnosis of primary cutaneous melanoma using surface microscopy, Dermatologic Clinics, Vol. 19, 299-305, 2001.
[4] W. Stolz et al., ABCD rule of Dermatoscopy: a new practical method for early recognition of malignant melanoma, European Journal of Dermatology, Vol. 4, 521-527, 1994.
[5] G. Argenziano et al., Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis, Archive of Dermatology, Vol. 134, 1563-1570, 1998.
[6] H. Ganster et al., Automated melanoma recognition, IEEE Transactions on Medical Imaging, Vol.20,233-239, 2001.
[7] P. Rubegni et al., Automated diagnosis of pigmented skin lesions, International Journal of Cancer, Vol. 101, 576-580, 2002.
[8] A. Blum et al., Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on the analysis of 837 melanocytic lesions, British Journal of Dermatology, Vol. 151, 1029-1038, 2004.
[9] M. Burroni et al., Dysplastic naevus vs. in situ melanoma: digital dermoscopy analysis, British Journal of Dermatology, Vol. 152, 679-684, 2005.
[10] S. Seidenari et al., Pigment Distribution in melanocytic lesions images: a digital parameter to be employed for computer-aided diagnosis, Skin Research Technology, Vol. 11, 236-241, 2005.
[11] M. Celebi et al., A methodological approach to the classification of dermoscopy images, Computerized Medical Imaging and Graphics, Vol. 31, 362-372, 2007.
[12] H. Iyatomi et al., An improved Internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm, Computerized Medical Imaging and Graphics, Vol. 32,566-579, 2008.
[13] M. Silveira, et al., Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images, IEEE Journal of Selected Topics in Signal Processing, special Issue on Digital Image Processing Techniques for Oncology, Vol. 3, Number 1, pp. 25-35, 2009.
[14] G. Di Leo et al., A software tool for the diagnosis of melanomas, In: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences, IEEE Computer Society, 1818-1823, 2010
[15] N. Dalal, B. Triggs, Histograms of Oriented Gradients for Human Detection, IEEE Int. Conf. Computer Vision and Pattern Recognition, 2005.