STATISTICAL IMAGE PROCESSING FOR THE DETECTION OF DERMOSCOPIC CRITERIA

Valerio De Vita[1], *Giuseppe Di Leo*[2], *Gabriella Fabbrocini*[1], *Alfredo Paolillo*[2], *Paolo Sommella*[2]

[1] Dermatology Research - University Federico II of Naples, Via S. Pansini 5 80131 Naples, ITALY, gafabbro@unina.it
[2] DIIn University of Salerno Via Ponte Don Melillo, 84084 Fisciano (SA), ITALY, {gdileo,apaolillo,psommella}@unisa.it

Abstract: An image based system implementing a well-known diagnostic method is disclosed for the automatic detection of melanomas as support to clinicians. The software procedure is able to recognize automatically the skin lesion within the digital image, measure morphological and chromatic feature, carry out a suitable classification for the detection of structural dermoscopic criteria provided by the 7-Point Check List. Experimental results about the adoption of statistical techniques applied to the border detection, feature extraction and classification as well as the resulting diagnostic score are described with reference to a large image set.

Key words: *Biomedical instrumentation, image processing, diagnosis and classification,*

1. INTRODUCTION

Malignant melanoma is nowadays one of the leading cancer in the world with a rapidly increasing incidence observed in Australia, America and Europe. Fortunately the curability of skin cancer is very high, if it is treated surgically early enough. To this aim, epiluminescence microscopy (ELM, also known as dermoscopy) has become an established noninvasive tool for improving the early detection of melanoma accuracy (10% to 30% higher sensitivity, [1] when compared with the clinical diagnosis by the naked eye). Based on the interpretation of the features inspected by dermoscopy, three different diagnostic models (Pattern Analysis, ABCD-rule and ELM 7 point checklist, [2]-[4]) have become more widely accepted by clinicians.

At the same time there has been much scientific endeavor aimed at obtaining an improved and consistent differentiation between benign and malignant melanocytic skin lesions by means of digital dermoscopy analysis.

Computerized dermoscopy image analysis, in fact, adds a quantitative evaluation to the "clinical eye observation" and can be used to improve biopsy decision-making ([5]). A very interesting summary of the main researches about the digital dermoscopy (in terms of recording system, calibration, image datasets, methods and approaches) is reported in [6]. As results of this survey, high accuracy may be achieved by computer aided diagnostic systems employing statistics obtained from low-levels feature and parameters. Nevertheless, it is not likely that the digital system will completely substitute the expert in dermoscopy.

The automated system should be in fact integrated by higher level features based on a particular diagnostic scheme in order to gain greater clinical acceptance. More precisely, the software diagnostic system should be able to reproduce the expertise of a well-trained dermatologist and support the clinician in his/her visual inspection and diagnosis according to the dermoscopic methods previously introduced.

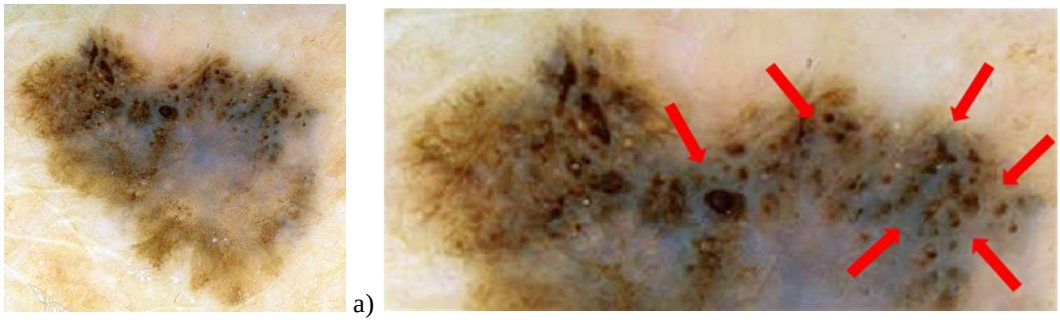
Starting from these considerations, the authors have tackled the problem of defining suitable image processing algorithm implementing the *7-Point Check List*. The list of dermoscopic structures as defined by the diagnostic method and the associated score are reported in Table 1 (for the diagnose of melanoma, a total score not lower than 3 has to be achieved summing up the partial scores).

A preliminary study about the image processing techniques for the extraction of the pigmented lesion (from healty skin) and the detection of chromatic features was reported in [7]. Further studied have been led to the introduction of a common framework [8] for the automatic detection of dermoscopic criteria, where the inputs of the Computer Aided System were ELM images and the processing algorithms were derived from the clinical knowledge gained by expert dermatologists (well-trained in the 7-Point Check List application).

This paper investigates the image processing algorithms developed for the automatic detection of *Irregular Dots/Globules*, which are defined as black, brown or blue round structures irregularly distributed within the lesion (see details reported in Figure 1). About these dermoscopic structures, statistical approaches are proposed as suitable solutions for the identification and classification sub-tasks. Finally experimental results are reported and analyzed with respect to a large image set of pigmented lesions.

Table 1 Dermoscopic criteria and scores according to the 7 Point Check List method

ELM criterion	Score
Atypical pigment network	2
Blue-whitish Veil	2
Atypical vascular pattern	2
Irregular streaks	1
Irregular pigmentation	1
Irregular dots/globules	1
Regression structures	1



a) b)

Figure 1 Example of lesion (a) characterized by Irregular Dots and Globules (b)

2. METHODS

According to the scheme reported in Figure 2, the software procedure developed for the automatic detection of Irregular Dots/Globules is organized into three main stages: the *boundary detection*, which allows the pigmented lesion to be extracted from the surrounding healthy skin, the *feature extraction*, which aims to measure morphological and chromatic features related to the dermoscopic structures of interest, and the *feature classification*, through which the detection of the criterion introduced in the 7-Point Check List is achieved. In the following subsections, the statistical techniques and algorithms are introduced correspondingly to the main stages of the automatic procedure.

3.1 Boundary detection

Boundary detection is a critical problem in ELM images because the transition between the lesion and the surrounding skin is smooth and hard to detect accurately, even for a trained dermatologist. The algorithm proposed for the skin lesion border extraction consists of three steps:

(i) color to monochrome image conversion;

(ii) image binarization using an adaptive threshold;

(iii) border identification, based on a blob-finding algorithm.

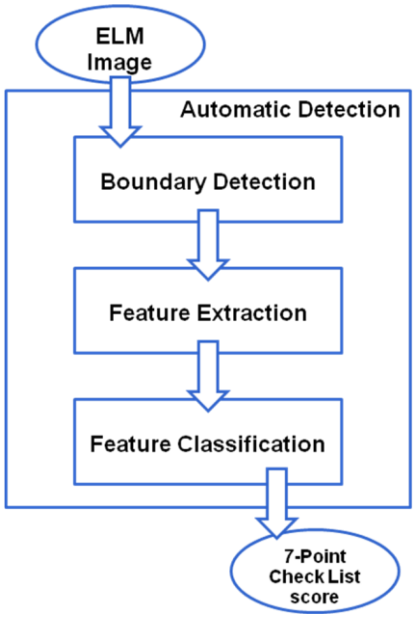


Figure 2 Scheme of the proposed procedure for the automatic detection of Irregular Dots/Globules

In the first step, 3 different monochrome images are obtained from the source image (RGB standard color) corresponding to the red, green and blue planes. For each component (see Figure 3.a), two modes (classes) are typically evident in the pixel intensity histogram (as depicted in Figure 3.b) corresponding respectively to the pigmented lesion (the image foreground) and the surrounding skin (the image background).

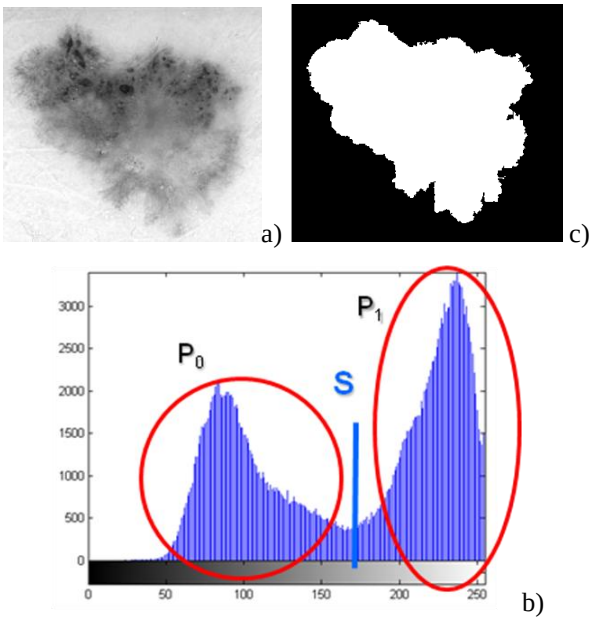
Then, the algorithm introduced by Otsu [9] is adopted to select the optimum threshold S* for each histogram, thus allowing the image background and foreground to be detected. In more details, the adaptive algorithm aims to maximize the between-class variance σ_B and minimize the intra-class variance σ_W:

$$S^* = ArgMax\{\frac{\sigma_B(S)}{\sigma_W(S)}\} \quad (1)$$

$$\sigma_B^2(S) = P_0(S)*[m_o(S)-m]^2 + P_1(S)*[m_1(S)-m]^2 \quad (2)$$

$$\sigma_W^2(S) = P_0(S)*\sigma_0^2(S) + P_1(S)*\sigma_1^2(S)^2 \quad (3)$$

In the previous equations m is the image intensity mean value, whereas P_i, m_i and σ_i are respectively the distribution, mean and standard deviation of the intensity class C_i



a) c)

b)

Figure 3 Example of Boundary detection: a) image conversion; b) intensity histogram; c) binary mask.

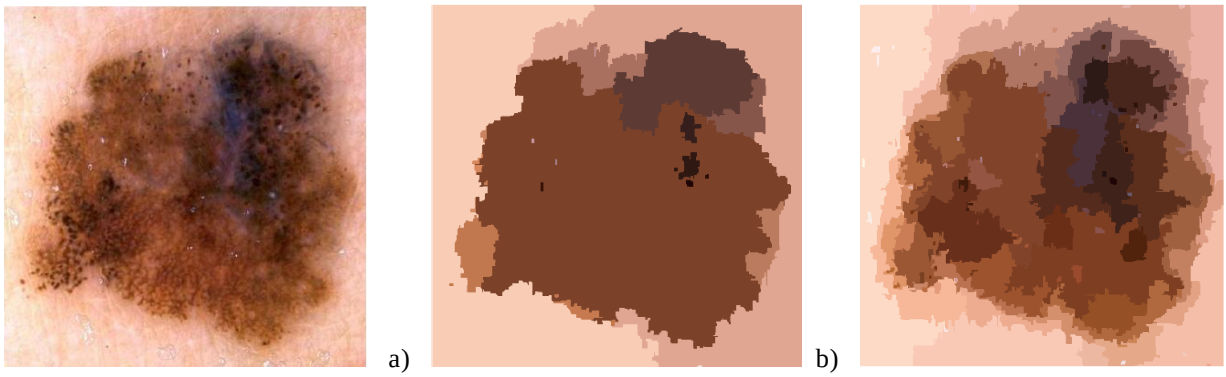


Figure 4 Segmentation using Statistical Region Merging: a) ELM Image; b) lesion map resulting from Q=32; c) lesion map resulting from Q=256;

resulting from the *S* histogram bin:

$$P_0(S) = \sum_{k=1}^{S} \frac{f_k}{N}, \qquad P_1(S) = \sum_{k=S+1}^{L} \frac{f_k}{N} \tag{4}$$

where N is the number of the image pixels, L is the number of histogram bins and f_k the number of pixels associated with k intensity value.

Since the proposed approach has been experimentally revealed to be more sensitive to surrounding skin (the image background), the thresholding result corresponding to the wider skin lesion area (the image foreground) has to be considered as binary mask for next processing. An example of result is shown in Figure 3.c.

Finally, a simple blob-finding algorithm is adopted to extract the contour of the lesion from the binary mask: the tracking algorithm collects and sorts out the edges of the black-white image into an ordered list. At this point, the border is superimposed on the color ELM image and displayed for visible inspection to the diagnostician.

3.2 Feature extraction

The dermoscopic criterion of interest is characterized both by chromatic and morphological structures. Thus, once the lesion is localized, several first order features are extracted and measured by means of techniques grouped into the following macro-categories:

- color segmentation
- structural analysis

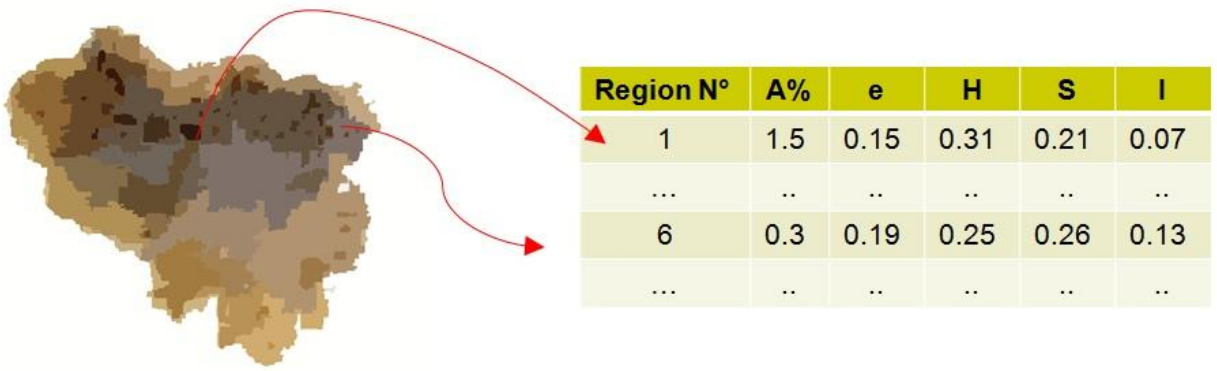
The proposed *color segmentation* is based on Statistical Region Merging (SRM), a recent technique [10] belonging to the region growing and merging group. The method models segmentation as an inference problem, in which the image is treated as an observed instance I of an unknown theoretical image I^*, whose statistical (true) regions are to be disclosed. This method is typically adopted for its simplicity, computational efficiency, and excellent performance without the use of quantization or color space transformations.

In more details, each pixel of the true image I* can be modeled as a set of Q independent random variables whereas the statistical regions represent theoretical objects sharing a common homogeneity property:

- inside any statistical region the pixels have the same expectation for each color channel (for example Red, Green and Blue or Hue, Saturation and Intensity);
- the expectation of adjacent regions are different for at least one color channel.

Given the homogeneity property the ideal segmentation of the observed image I relies on the frontiers between the statistical regions which are connecting pixels with differences in their color expectation. Figure 4 depicts an example of color segmentation for the ELM image performed through the SRM: each region is displayed according to its mean RGB values (averaged on pixels constituting the region). A similar result is also held for the HSI color space.

The parameter Q allows to quantify the statistical complexity of I^*, the generality of the model and finally control the coarseness of the segmentation. For the application of interest (detection of small round-shaped areas) a fine level of color segmentation is required which can be achieved by considering Q=256.



Region N°	A%	e	H	S	I
1	1.5	0.15	0.31	0.21	0.07
…	..	..	..	..	..
6	0.3	0.19	0.25	0.26	0.13
…	..	..	..	..	..

Figure 5 Example of fine color segmentation (SRM) and corresponding feature extraction

As you can easily note in the example reported in Figure 5, the darkest segments may be deeply investigated to seek for the *structures* which represents Irregular Dots and Globules. More in details, a statistical analysis based on the histogram of the SRM image has been performed by considering and ordering the statistical regions with respect to the increasing value of Intensity value (within a suitable range for Hue component).

Moreover the following morphological measurements are also considered to extract round items from the inside of the lesion:

- relative dimension *A%*, defined as the number of the region pixels with respect to the lesion area;
- eccentricity *e* of the ellipse that has the same second-order moments as the region; it is computed as the ratio of the distance between the foci of the ellipse and its major axis length with value between 0 and 1 (the degenerate cases corresponding respectively to a circle and a line segment).

3.3 Feature classification

About the classification of chromatic and/or morphological features determined and measured by applying the techniques previously described, a statistical approach is proposed based on Test Hypothesis.

In Figure 6 an example of the feature extraction is reported, where the darkest round items are high lightened as well as the main symmetry axes of the lesions. According to the corresponding definition provided by the diagnostic method, the Dots and Globule within the lesion can be considered as irregular if their spatial distribution is not uniform. In the opposite case, i.e if the observed (spatial) round-shaped items were randomly scattered within the lesion, the number of dots and globules in each of 4 quadrants (as resulted from the drawing of the main lesion axes) could be modeled according to the Binomial Distribution. Then, by considering the random distribution of *N* round objects as Null Hypothesis of a bilateral Test, the thresholds $k_{1,min}$, $k_{1,MAX}$ and k_2 can be jointly adopted for estimating the irregularity of Dots and Globules:

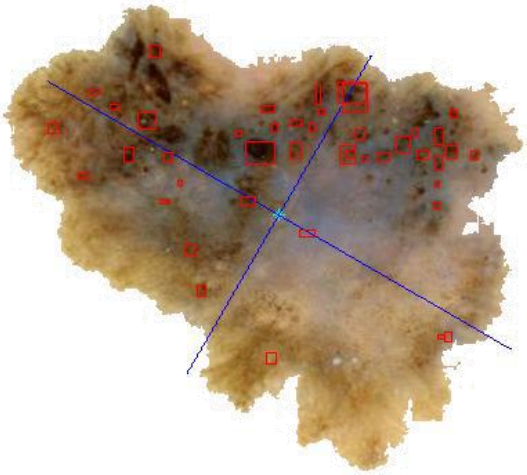


Figure 6 Result of Feature Extraction: detection of round-shaped items within the lesion area

$$\sum_{k=0}^{k_{1,\min}}\binom{N}{k}*\left(\frac{1}{4}\right)^{k}*\left(\frac{3}{4}\right)^{N-k}+\sum_{k=k_{1,MAX}}^{N}\binom{N}{k}*\left(\frac{1}{4}\right)^{k}*\left(\frac{3}{4}\right)^{N-k}\leq\alpha \quad (5)$$

$$2*\sum_{k=0}^{k_2}\binom{N}{k}*\left(\frac{1}{2}\right)^{k}*\left(\frac{1}{2}\right)^{N-k}\leq\alpha \quad (6)$$

where α is the accepted risk of Type I Error.

According to the proposed approach, if the paucity or plenty of objects is observed in any quadrant (Eq. 5) and/or couple of quadrants (Eq.6), the Null Hypothesis is refused and the lesion is classified as featured by Irregular Dots and Globules.

4. RESULTS

In order to develop and test the automatic procedure for the diagnosis of pigmented skin lesions, images of benign and malignant lesions were collected and stored in a *database* (200 cases were extracted from a dermoscopy atlas [11], whereas about 100 images were available from a screening activity carried out at the Dermatology Section of University "*Federico II*" in Naples). For each image, the corresponding clinical and histological analysis (when available) has been considered as well as the 7-Point Check List score computed by a group of three expert physicians. About the image quality, all the pictures are 24-bit RGB color images in JPEG) format with dimensions ranging from 700x447 to 2272x1520 pixels. The lesions are imaged completely with healthy skin visible at margins. Moreover, as image pre-processing for artifact removal, the strategy based on mathematical morphology has been adopted as suggested in [12].

4.1 Boundary detection

The proposed technique based on the adaptive thresholding has been compared with an unsupervised approach based on the Statistical Region Merging (*SRM*) algorithm, which was revealed as the most effective method [13] for contour detection in dermoscopy images of pigmented skin lesions. Comparison has taken into account 120 dermoscopy images (60 invasive malignant melanoma and 60 benign) randomly selected from the starting dataset.

As a ground truth for the evaluation of the border detection error, a manual border was obtained by selecting a number of points on the lesion border, connecting these points by a second-order B-spline and finally filling the resulting closed curve. More in detail, three dermatologists were asked to select the points on the lesion border, then the corresponding binary images were combined according to a majority policy to achieve the manual border (image pixels resulted at least twice as inner points have been considered as forming the lesion in the ground truth binary image).

Using the dermatologist-determined borders, the automatic borders resulting from the Adaptive Thresholding and SRM have been compared using the metric suggested in [14]. Here, the percentage border error is given by:

$$BorderError = \frac{(AutomaticBorderXORManualBorder)}{Area(ManualBorder)} * 100\% \quad (7)$$

where *AutomaticBorder* is the binary image obtained by filling the computer detected border, *ManualBorder* is the binary image previously described, the exclusive-OR operation gives the pixels for which the *AutomaticBorder* and *ManualBorder* disagree, and Area(*I*) denotes the number of pixels in the binary image *I*.

Table 2 shows the mean and standard deviation border error for the automated methods considered. Although the error rates increase in the melanoma group (due to the presence of higher border irregularity and color variegation in these lesions), the proposed approach has achieved the best results (lowest error values) in terms of both accuracy (mean) and consistency (standard deviation). An example of automatic contour extraction for a melanoma is reported in Figure 7 where the resulting borders are compared with the manual border.

4.2 Feature extraction and classification

About the feature extraction and classification, a *Training* and *Test Set* have been suitably selected from the reference database. In more details, 150 digital images has been adopted to develop the automatic detection of *Irregular Dots/Globules*, whereas the remaining 137 images have been adopted as Test Set to verify the software procedure. As result, the Training and Test Set share the same case distribution with respect to the criterion of interest: respectively 45 and 39 skin lesions are characterized by Irregular Dots/Globules.

Based on the image properties of the Training Set, suitable thresholds have been determined through ROC curves for the quantities introduced in the feature extraction stage (maximum region dimension *A%*, and eccentricity *e*, range for *I* component) as well as the classification (minimum number N_0 of round-shaped items to perform the statistical test and the risk α). The classification results from physicians have been taken into account: 3 expert dermatologists were asked to inspect the results from the feature extraction in order to set the classification attribute for each image.

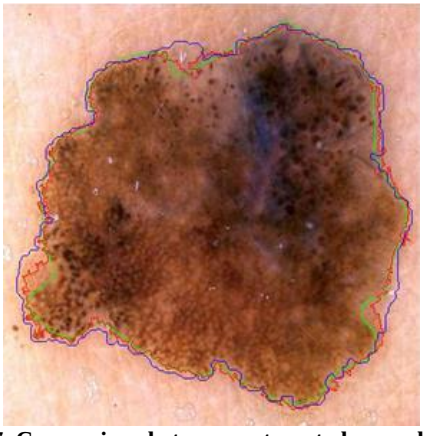


Figure 7 Comparison between automated procedures for border detection: ground truth (green line), adaptive thresholding (red line), unsupervised approach (blue line)

Table 2 Experimental results about the proposed automatic procedure for border extraction

Diagnosis	Statistic	Adaptive Thresholding	Unsupervised Approach
Benign	Mean	7,4 %	8,5 %
	SD	2,3 %	3,3 %
Melanoma	Mean	10,0 %	13,1%
	SD	6,2 %	8,7 %
All	Mean	8,7 %	10,8%
	SD	4,8 %	6,9 %

The verification of the proposed approach with respect to the Test Set has resulted in 35 skin lesions where Irregular Dots/Globules have been correctly revealed and 15 false detections. Table 3 summarizes the corresponding results (both for Training and Test Set) in terms of *sensibility* and *specificity*, intended as ratios of corrected decisions and the total number of cases where Irregular Dots/Globules are respectively present and absent. As you can easily note, the overtraining has been avoided (similar performance of the classifier for the two Image Sets). Example of correct classifications are reported in Figure 8 both for Irregular Dot/Globules and pigmented lesions where dark round object are not significant.

5. CONCLUSION

The paper has described the statistical techniques adopted for the automatic detection of a dermoscopic criterion in digital images according to the 7-Point Check List diagnostic method. Namely, the procedures for the detection of Irregular Dot/Globules have been be tested with respect to a quite extensive metrological characterization (performance of each classifier estimated in terms of the sensibility and specificity). The approach will be extended and suitably integrated with all the procedures constituting the image-based measurement system for the diagnosis of melanoma. To this aim future investigation will be addressed to: *i)* determine the correlation existing among the seven dermoscopic criteria and *ii)* compute a confidence level for each intermediate classification, in order to achieve more accurate diagnostic results (improve sensibility and specificity of the software system as whole).

REFERENCES

[1] J. Mayer, "Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma", Med. Journal Aust Vol. 167, pp. 206-210, 1997.

[2] H. Pehamberger, A. Steiner, K. Wolff, "In vivo epiluminescence microscopy of pigmented skin lesions: Pattern analysis of pigmented skin lesions", J. Am. Acad. Dermatol. Vol. 17, pp. 571-583, 1987.

Table 3 Performance of Classifiers (Testing Set)

Image Set	Sensibility	Specificity
Training Set	91%	87 %
Test Set	90%	85 %

[3] W. Stolz, O. Braun Falco et alii, “ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma”, European Journal of Dermatology, Vol. 4, pp. 521-527, 1994

[4] G. Argenziano et alii, “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”, Arch Dermatol. Vol. 134, pp. 1563-1570, 1998.

[5] M. Burroni, R. Corona, G. Dell'Eva et alii, “Melanoma computer-aided diagnosis: reliability and feasibility study”, Clin. Cancer Res. Vol. 10, pp. 1881-1886, 2004.

[6] A. Blum, I. Zalaudek, “Digital image analysis for diagnosis of skin tumors”, Semin Cutan Med Surg., Vol. 27, pp.11-15, 2008.

[7] G. Di Leo, G. Fabbrocini et alii, “ELM image processing for melanocytic skin lesion based on 7-point checklist: a preliminary discussion” Proc. of the 13th IMEKO TC-4 Symposium, Athens, Greece, Vol.2, pp. 474-479, 2004.

[8] G. Di Leo, G. Fabbrocini et alii, "Automatic Diagnosis of Melanoma: a Software System based on the 7-Point Check-List", Proceedings of the 43rd Annual Hawaii International Conference on System Sciences, January 5-8, 2010, Computer Society Press, 2010.

[9] Otsu, “A threshold selection method from gray-level histogram”, IEEE Trans. on System Man Cybernetics, Vol. SMC-9, N° 1, pp. 62-66, 1979.

[10] R. Nock, F. Nielsen, “Statistical Region Merging”. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 26 N° 11, pp. 1452-1458, 2004;

[11] G. Argenziano, H.P. Soyer, V. De Giorgi et al., “Interactive Atlas of Dermoscopy”, EDRA Medical Publishing & New Media, Milan, Italy, 2002.

[12] P. Schmid, J. Guillod , J.P. Thiran,"Towards a computer-aided diagnosis system for pigmented skin lesions", Computerized Medical Imaging and Graphics, Vol N° 27, (2003), pp. 65-78.

[13] M.E. Celebi et alii, “Border detection in dermoscopy images using statistical region merging”, Skin Research and Technology Vol 13, N° 4, pp.347-353, 2008;

[14] G.A. Hance et alii, “Unsupervised color image segmentation with application to skin tumor borders”, IEEE Eng. in Medicine and Biology, Vol 15, N°1, pp. 104-111, 1996.

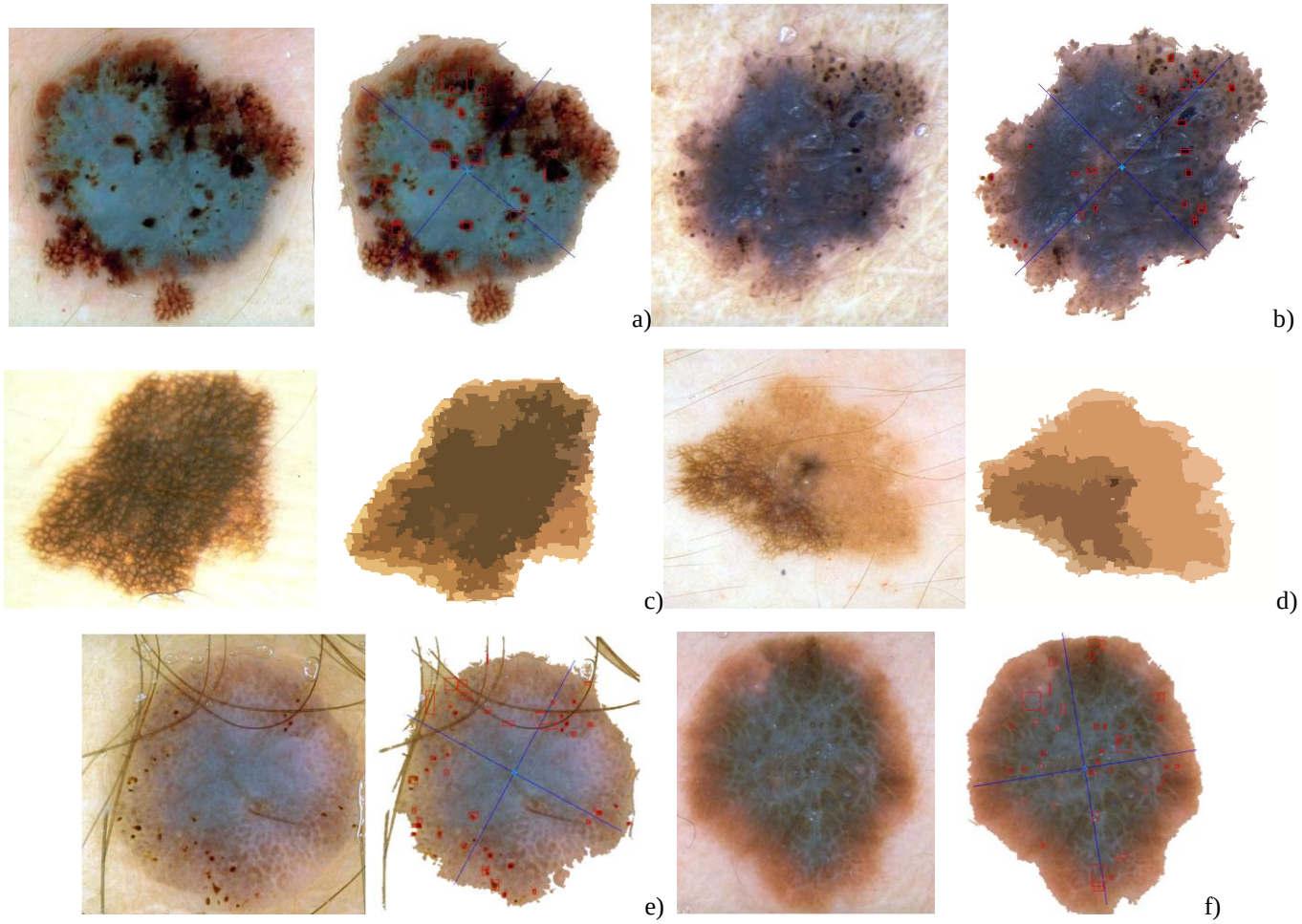


Figure 8 Example of classification results: a)-b) correct detection of Irregular Dot/Globules (partial score=1); c)-d): correct analysis of lesion with texture (partial score=0); e) false detection (regular dots/globules); f) false detection (erroneous round-shaped items).