

A robust system for melanoma diagnosis using heterogeneous image databases

Khaled Taouil¹, Zied Chtourou¹, Nadra Ben Romdhane²

¹UR CMERP National Engineering School of Sfax, Sfax, Tunisia;

²Faculty of Sciences of Sfax, Sfax, Tunisia.

Email: Khaled.taouil@cmerp.net

Received 5 April 2010; revised 8 May 2010; accepted 15 May 2010.

ABSTRACT

Early diagnosis of melanoma is essential for the fight against this skin cancer. Many melanoma detection systems have been developed in recent years. The growth of interest in telemedicine pushes for the development of offsite CADs. These tools might be used by general physicians and dermatologists as a second advice on submission of skin lesion slides via internet. They also can be used for indexation in medical content image base retrieval. A key issue inherent to these CADs is non-heterogeneity of databases obtained with different apparatuses and acquisition techniques and conditions. We hereafter address the problem of training database heterogeneity by developing a robust methodology for analysis and decision that deals with this problem by accurate choice of features according to the relevance of their discriminative attributes for neural network classification. The digitized lesion image is first of all segmented using a hybrid approach based on morphological treatments and active contours. Then, clinical descriptions of malignancy signs are quantified in a set of features that summarize the geometric and photometric features of the lesion. Sequential forward selection (SFS) method is applied to this set to select the most relevant features. A general regression network (GRNN) is then used for the classification of lesions. We tested this approach with color skin lesion images from digitized slides data base selected by expert dermatologists from the hospital "CHU de Rouen-France" and from the hospital "CHU Hédi Chaker de Sfax-Tunisia". The performance of the system is assessed using the index area (Az) of the ROC curve (Receiver Operating Characteristic curve). The classification permitted to have an Az score of 89,10%.

Keywords: Melanoma; Computer Aided Diagnosis System;

Segmentation; Feature Selection; Classification; Generalized Regression Neural Network

1. INTRODUCTION

Melanoma is the most deadly form of skin cancer. The World Health Organization estimates that more than 65000 people a year worldwide die from too much sun, mostly from malignant skin cancer [1].

The five-year survival rate for people whose melanoma is detected and treated before it spreads to the lymph nodes is 99 percent. Five-year survival rates for regional and distant stage melanomas are 65 percent and 15 percent, respectively [2]. Thus the curability of this type of skin cancer depends essentially on its early diagnosis and excision.

The ABCD (asymmetry, border, colour and dimension) clinical rule is commonly used by dermatologists in visual examination and detection of early melanoma [3]. The visual recognition by clinical inspection of the lesions by dermatologists is 75% [4]. Experienced ones with specific training can reach a recognition rate of 80% [5].

Several works has been done on translating knowledge of expert physicians into a computer program. Computer-aided diagnosis (CAD) systems were introduced since 1987 [6]. It has been proved that such CAD systems can improve the recognition rate of the nature of a suspect lesion particularly in medical centres with no experience in the field of pigmented skin lesions [7,8]. For these systems to be efficient, the shots of the suspected lesion have to be taken using the same type of apparatuses than the one used for the learning database [9] and with identical lighting and exposure conditions. This could be very challenging in the majority of cases.

In order to overcome the lack of standardization in stand alone CADs and to provide an open access to dermatologists, web-based melanoma screening systems were proposed [10,11]. These systems have to consider the heterogeneity in databases collected in different cen-

tres.

This work describes an enhanced CAD system that addresses the problem of robustness of such tools under the use of different databases.

2. PROPOSED CAD SYSTEM

The proposed software combines automated image segmentation and classification procedures and is designed to be used by dermatologists as a complete integrated dermatological analysis tool. CAD systems in melanoma detection are usually based on image processing and data classification techniques. Five steps are generally needed: data acquisition, pre-processing, segmentation, feature extraction and classification (Figure 1).

2.1. Data Acquisition

The main techniques used for this purpose are the epiluminescence microscopy (ELM, or dermoscopy), transmission electron microscopy (TEM), and the image acquisition using still or video cameras. The use of commercially available photographic cameras is also quite common in skin lesion inspection systems, particularly for telemedicine purposes [12].

2.2. Segmentation of Lesion Images

Image segmentation is the most critical step in the entire process. It consists of the extraction of the region of interest (ROI) which is the lesion. The result of segmentation is a mask image. This mask is the base for the computation of several shape and colour features.

The computer has a great difficulty in finding lesion edge accurately. This task alone has formed the basis of much research [13]. The difficulty of segmentation is due to low contrast between the lesion and the surrounding skin and irregular and fuzzy lesion borders. Artefacts (light reflections, shadows, overlapping hair,

etc) can also give a false segmentation result. Some works rely on the physician to outline the suspicious area [14].

We use a hybrid segmentation approach based on two steps. The first consists in applying morphological pre-processing filters to facilitate the extraction of the approximate region of the lesion from the safe skin. The second consists in applying active contour method on the approximate mask to have the final contour of the lesion [15].

Active contours or snakes are curves defined within an image domain that can move under the influence of internal forces coming from within the curve itself and external forces computed from the image data. Snakes were introduced by Kass *et al.* [16]. Snakes are parameterized curves:

$$v(s) = [x(s), y(s)], s \in [0, 1] \quad (1)$$

This curves move through the spatial domain of an image to minimize the functional energy [17]:

$$E_{snake} = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s)) ds \quad (2)$$

$$E_{int}(v(s)) = \alpha |v'(s)|^2 + \beta |v''(s)|^2 \quad (3)$$

$$E_{ext}(x, y) = -|\nabla I(x, y)|^2 \quad (4)$$

where:

- $v(s)$ is a set of coordinates to form a snake contour.
- $v'(s)$ and $v''(s)$ denote the first and second derivatives of $v(s)$ with respect to s .
- α and β are weighting parameters that control respectively the snake's tension and rigidity.
- $\nabla I(x, y)$ is the gradient of grey-level image I .

A snake that minimizes E_{snake} must satisfy the Euler equation.

$$\alpha v''(s) - \beta v''''(s) - \nabla E_{ext} = 0 \quad (5)$$

The internal force F_{int} discourages stretching and bending while the external potential force F_{ext} pulls the snake toward the desired image edges.

To find a solution to (4), the snake $v(s)$ is made dynamic by adding the parameter of time t to the equation of the curve that becomes:

$$v_t(s, t) = \alpha v''(s, t) - \beta v''''(s, t) - \nabla E_{ext} \quad (6)$$

indicating how the snake must be modified at the instant $t+1$ according to its position at the instant t .

When $v(s, t)$ stabilizes, we achieve a solution of (6).

2.3. Features Extraction from Lesion Images

To characterize the different types of lesions we consider a parametric approach. In such approach, the skin lesion is resumed in a vector of features which dimension depends on the number of extracted primitives. We use

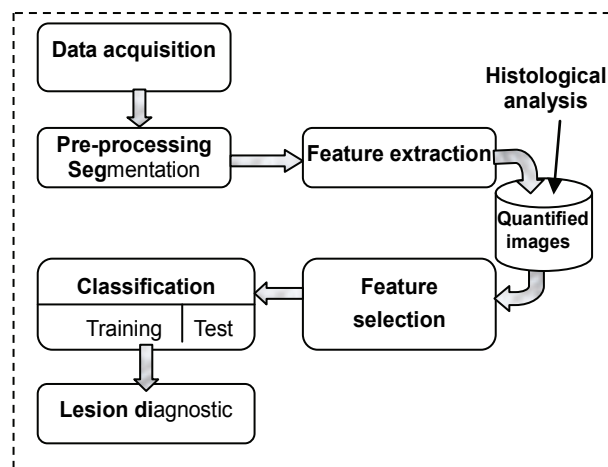


Figure 1. CAD system in melanoma detection.

quantitative parameters from the descriptions of dermatologists based on the ABCD rule to model clinical signs of malignancy.

The preliminary developed set of parameters underwent a series of tests to evaluate their robustness when quantifying multiple shots of the same lesion acquired under different lighting conditions with different apparatuses as is the case when using different slides from heterogeneous databases [18]. Our set of parameters was besides used by [19] to develop an automatic recognition of melanoma which reached a correct classification rate of 79.1%.

The most important clinic signs that were kept to characterize melanoma and the different lesions are the irregularity of the contour, the asymmetry of colour and shape, as well as the heterogeneity of the colour. We classify parameters in two categories: geometric and photometric parameters.

2.3.1. Geometric Parameters

Geometric parameters are extracted from the binary shapes obtained after image segmentation. These parameters permit to characterize the shape of the lesion, its elongation and the regularity of its contour. All these parameters are standardized and independent of translation, zoom and rotation effects and therefore compensate for rigid transformations introduced by the optics conditions and the scene selection and framing.

2.3.2. Photometric Parameters

Photometric parameters are calculated from true colour and binary images. These parameters permit to describe homogeneity and symmetry of the colour as well as the deviation between the mean colour of the lesion and the mean colour of the surrounding safe skin. We tested different colour spaces representations calculated from the red green blue components. We reflect the correlation between the level of a colour component of a pixel and its position on the digital image which is supposed to be independent of lighting conditions and spectral ensitivity of the camera sensor.

2.4. Feature Selection

Feature selection allows choosing the most relevant parameter subset to perform the classification step. This subset must contain the more robust and the most discriminative primitives [20]. Three criteria's must be fixed: the assessment method of the variables set relevance, the research procedure to follow and the stopping criteria of the selection. In this work we report the use of a sequential forward selection with a stopping criteria based on the minimum error generated by the classifier.

2.4.1. Sequential Forward Selection

The SFS [21] is an ascending research method (bot-

tom-up) of the set of most discriminative parameters from an initial set of parameters (E_i) with:

$$E_i = \{p_j, j = 1, 2, \dots, N\}$$

$$E_{SFSn} \subset E_i, n \leq N$$

For this method one parameter p_j is added at a time to the E_{SFS} subset. If the assessment criterion is an artificial neural network, for each step, we insert one by one remaining parameters p_j of E_i in E_{SFS} and we calculate the corresponding classification error (Err) with:

$$Err = \frac{1}{q} \sum_{i=1}^q (d_i - a_i)^2 \quad (7)$$

with q : total image number of the training database.

d_i : desired output.

a_i : real output.

Initially, the subset $E_{SFS0} = \varnothing$;

For every step, parameter p_j that will be selected is the one for which the new E_{SFS} subset permits to minimize the classification error:

$$Err(E_{SFS} U_{p_j}) \leq Err(E_{SFS} U_{p_j'}) \quad (8)$$

Thus, the first selected parameter is the most discriminative one of the initial set of parameters. The selection of parameters stops when while adding a new parameter to E_{SFS} , the classification error increases.

2.5. Classification of Lesion Images

After having summarized information contained in the different images of our databases in vectors of parameters we use a classifier based on a general regression network (GRNN) [22]. GRNN network is much faster to train than a multilayer perceptron network (MPN). GRNN gave better recognition rates than MPN for melanoma classification [23]. The architecture of this network is illustrated on **Figure 2**. The GRNN is composed of four layers: the input layer, the first intermediate

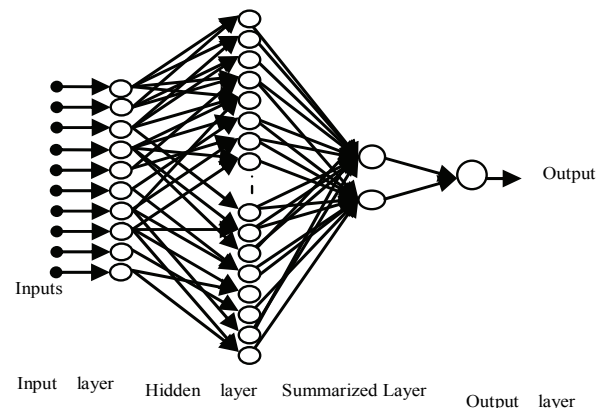


Figure 2. Architecture of the used GRNN.

layer constituted of radial units, the second intermediate layer constituted of summarized units and the output layer.

3. ASSESSMENT OF THE CLASSIFICATION

The performance of our CAD system is evaluated in term of sensitivity and specificity. These measures are defined as follow:

$$\text{Sensitivity} = \frac{\# \text{ True Positives}}{\# \text{ True Positives} + \# \text{ False Negatives}} \quad (9)$$

$$\text{Specificity} = \frac{\# \text{ True Negatives}}{\# \text{ True Negatives} + \# \text{ False Positives}} \quad (10)$$

With #true-Positive and #False-Negative corresponds respectively to the number of malign lesions well classified and badly classified. #True-Negative and #False-Positive corresponds respectively to the number of benign lesions well classified and badly classified.

A ROC curve consists in representing the value of the sensitivity according to (1-specificity) [24]. The area under the ROC curve or area index (Az) represents the probability to correctly identify the image with anomaly when an image with anomaly and an image without are presented simultaneously to the observer.

We use two databases of image lesions whose malign or benign nature is perfectly known after histological analysis. The first one has been collected in CHU Rouen France with the collaboration of the research laboratory PSI-INSA Rouen and has been supported by the French National League against Cancer. This database was digitized in true colours by a 35 mms slides Nikon LS-1000S scanner. It was used in previous works [19,25] and [26,27]. We divide this database in a training (B0) and test (B1) sets used in the first assessment of classification.

The second image database has been collected in Tunisia from the dermatology service of CHU Hédi CHAKER in Sfax. Images were digitized in true colours with a HP Scanjet 3570c scanner (cf. **Table 1**).

Our approach for efficiency assessment of the developed tool has been achieved in three steps:

3.1. First Assessment of the System

For the first step, we evaluate the diagnosis results of our system while using two sets of images (B0 for training and B1 for test) from the same CHU Rouen image database.

Table 1. Distribution of images databases.

Data-base	Set	Class	Nbr of BL	Nbr of ML	Total
CHU Rouen	B0	Training	59	27	86
	B1	Test	34	15	49
CHU Sfax	B2	Test	31	20	51
	Total		124	62	186

3.2. Comparison of Our System's Diagnosis with Dermatologist's Visual Diagnosis

This step consists in comparing diagnostic results of our system with the opinion of four expert dermatologists for the same test database (B1). Dermatologists are part of the dermatology service of the CHU Sfax. We asked every dermatologist to give his diagnosis for each lesion.

3.3. Third Assessment of the System

For this step, we evaluate our system while using the second image database (B2) collected at CHU Sfax. This test has been done while using the artificial neural network having the best recognition rate according to the first assessment.

4. RESULTS AND DISCUSSION

4.1. Results of the Segmentation Step

For image lesion segmentation, we propose a hybrid method that combines the advantages of morphological treatments, histogram thresholding and active contours techniques.

First the contrast of the gray level original image is enhanced using top-hat and bottom-up filtering (**Figure 3**).

The extraction of the image mask is based on the detection of regional minima of the complementary image of the contrast enhanced image. The detection of these regions requires the application of a threshold. This threshold is obtained with histogram thresholding using Otsu method [28]. We apply then a morphological opening on the obtained image. Lakes of the resulting image are eliminated by filling holes. The approximate zone or approximate mask of the lesion is finally obtained following a labelling, conservation of the biggest element.

We initialize a snake at the approximate boundary of the safe skin (cf. **Figure 4**). The snake begins with the calculation of a field of external forces over the image domain. The forces drive it toward the boundary of the lesion. The process is iterated until it matches the contour of the lesion. We superpose the obtained contour on the original color image.

Figures 5 and 6 show some examples of segmentation results of lesions collected from both CHU Rouen and CHU Sfax databases.

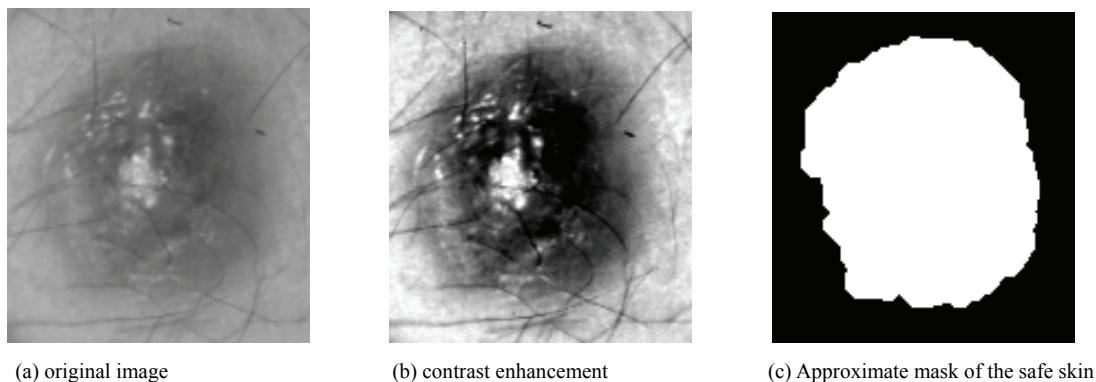


Figure 3. Extracting the approximate mask of the safe skin.

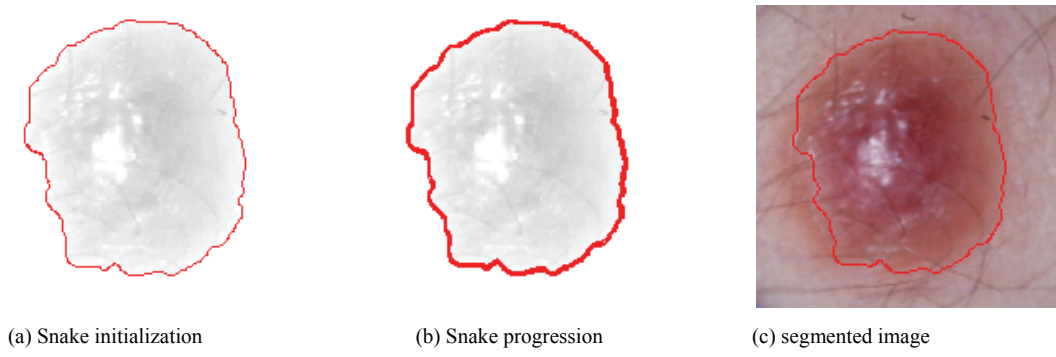


Figure 4. Application of the active contour on the approximate mask.

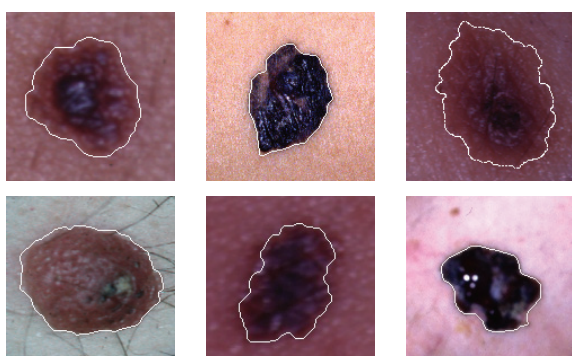


Figure 5. Segmentation results of lesions collected from CHU Rouen database.

4.2. Results of Features Selection

For features extraction, a set of 68 parameters are extracted for every lesion. Through correlation and robustness study, a set of 42 parameters have been kept. To find the most discriminative ones for the classification step, we apply the sequential forward selection (SFS) method. Training and test databases of images have been randomly selected.

For the SFS method, the assessment of parameter selection is based on the comparison of the error generated by the general regression neural network (GRNN) for the dif-

ferent subsets of selected variables.

The chosen set of variables is the one that generates the minimal error. We pursued research until the selection of all parameters. Then we chose the smallest subset gotten with the minimal error.

The result of this selection method is illustrated on **Figure 7**. It illustrates the variation of the classification's mean square error (MSE) according to the number of included parameters, during trainings and tests of the GRNN.

According to the test curve, we note that the set of the

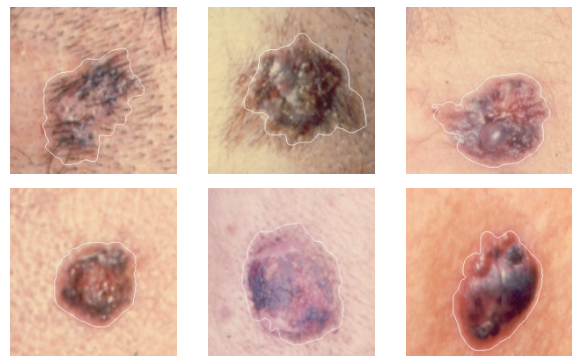


Figure 6. Segmentation results of lesions collected from CHU Sfax database.

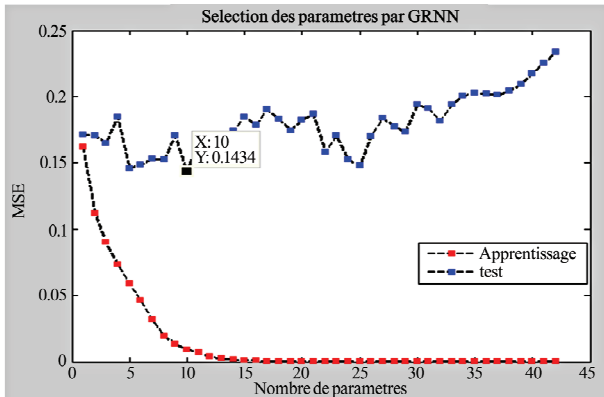


Figure 7. Parameter selection using SFS method.

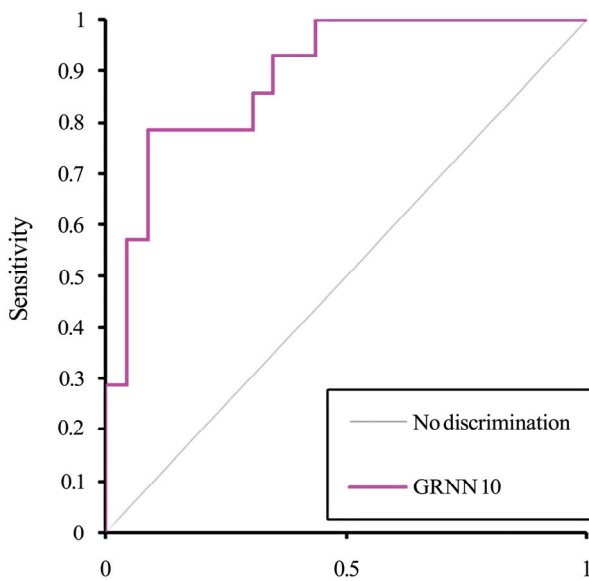


Figure 8. The ROC curve obtained using GRNN classifier and the selected parameters.

Table 2. Order of the selected parameters using SFS method.

Rg	N°	Description
01	24	scG symmetry of the green component
02	09	rmoyr mean of the normalized red in the lesion
03	03	DeltaD expanse of the distance to the center
04	35	Beta spherical coordinate beta
05	23	scR symmetry of the red component
06	01	comp compactness of the shape
07	39	RB Ratio between lesion and safe skin for the red
08	29	gamma1B Fisher coefficient for the blue
09	16	sigmamoyLb Standard deviation of normalized bleu in the safe skin
10	06	scB symmetry of the blue component

most discriminative parameters permitting to minimize the MSE of test database are gotten after the insertion of the first ten parameters (MSE = 0.1434).

The selection by the SFS method permitted a reduction of 76.19% of the total number of parameters. The list of parameters selected is presented in Table 2.

4.3. Results of the Classification Step

The classification is based on parameters kept in features selection. Figure 8 illustrates results of the classification while using the 10 most discriminative parameters selected by the SFS method.

The performances of classification using B1 set of images extracted from CHU Rouen database are based on the comparison of the value of the area index (Az) of the ROC curve. We obtain a value of Az equal to 89.1%. The recognition rate of the system is 92.05%.

To validate the efficiency of our system, we compare the obtained classification results with the diagnosis of four dermatologists from CHU Sfax. We asked every dermatologist to give his diagnosis for every lesion of the same test set B1. Results of their diagnoses are given in Table 3.

According to these results, we notice that the mean value of sensitivity provided by dermatologists is equal to the mean percentage of visual recognition of the true positives by dermatologists that is 75% [4].

The recognition rate of our CAD system (92.05%) exceeds that obtained by dermatologists. It even exceeds that of experimented trained ones which is about 80% [5].

Results of the second assessment of the system are given in Table 4. We note that even when using a test set of image lesions selected randomly from a different database the recognition rate our system is 90.15%. It remains better than the one of the visual diagnosis of experimented dermatologists.

5. CONCLUSIONS

In this paper we have described the different steps used in the CAD system that we propose for melanoma detection. To make this tool useful by the dermatologist community outside specialized centres, each stage of processing had to be automatic and robust to different conditions of acquisition and apparatus. The system segments and extracts parameters of description of the lesion. These parameters are normalized and used as inputs for the neural network classifier which decides if the lesion is suspicious. We have also described the different steps used for the evaluation of our system. This evaluation had proved the robustness of our system when using different databases in training and test. This property makes it a suitable and an efficient candidate for use in a context of a telemedicine dermatological application.

Table 3. Recognition rate dermatologists (D).

RR / D	D A	D B	D C	D D	
FP	12	0	12	8	mean
VP	9	8	14	14	
FN	6	7	1	1	
VN	22	34	22	26	
Sensitivity	60	53,3	93,3	93,3	74,9
Specificity	64,7	100	64,7	76,5	76,4
RR	62,3	76,6	79,0	84,9	75,7

- FP (False Positive)
 - TP (True Positive)
 - FN (False Negative)
 - TN (True Negative)
 - RR: Recognition Rate

Table 4. Assessment of the system with database B2.

TR	Base2
FP	03
VP	18
FN	02
VN	28
Sensitivity	90
Specificity	90,3
Recognition rate	90,15

6. ACKNOWLEDGEMENTS

We sincerely thank:

- Dermatology Services of CHU Rouen, France and CHU Hédi Chaker, sfax, Tunisia.
- Laboratory PSI (Perception Systems Information), INSA Rouen, France.
- Research Unit: Sciences and Technologies of Image and Telecommunications, High Institute of Biotechnology, Sfax.

REFERENCES

- [1] (2006) Solar ultraviolet radiation: Global burden of disease from solar ultraviolet radiation. Environmental Burden of Disease Series, World Health Organization, **13**.
- [2] (2009) Cancer Facts and Figures. American Cancer Society, Atlanta.
- [3] Friedman, R.J. (1985) Early detection of malignant melanoma: the role of the physician examination and self examination of the skin. *Cancer Journal for Clinicians*, **35**, 130-151.
- [4] Schmid-Saugeon, P., Guillod, J. and Thiran, J.P. (2003) Towards a computer-aided diagnosis system for pigmented skin lesions. *Computerized Medical Imaging and Graphics*, **27**, 65-78.
- [5] Blanzieri, E. (2000) Exploiting classifier combination for early melanoma diagnosis support. *Proceedings of the 11th European Conference on Machine Learning*, Barcelona, 31 May-2 June 2000, 55-62.
- [6] Cascinelli, N., Ferrario, M., Tonelli, T., *et al.* (1987) A possible new tool for clinical diagnosis of melanoma. *Journal of the American Academy of Dermatology*, **16**, 361-367.
- [7] Ercal, F., Chawla, A., Stoecker, W.V., Lee, H.C. and Moss, R.H. (1994) Neural network diagnosis of malignant melanoma from color images. *IEEE Transactions on Biomedical Engineering*, **41(9)**, 837-845.
- [8] Ganster, H., Pinz, A., Röhner, R., *et al.* (2001) Automated melanoma recognition. *IEEE Transactions on Medical Imaging*, **20(3)**, 233-238.
- [9] Burrioni, M. (2004) Melanoma computer-aided diagnosis: Reliability and feasibility study. *Clinical Cancer Research*, **10**, 1881-1886.
- [10] Guillod, J., Schmid-Saugeon, P., Décaillet, F., *et al.* (2003) An open internet platform to distributed image processing applied to dermoscopy. *Studies in Health Technology and Informatics*, **95**, 107-112.
- [11] Oka, H., Hashimoto, M., Iyatomi, H. and Tanaka, M. (2004) Internet-based program for automatic discrimination of dermoscopic images between melanoma and Clark nevi. *British Journal of Dermatology*, **150(5)**, 1041.
- [12] Maglogiannis, I. and Charalampos, N.D. (2003) Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transaction on Information Technology in Biomedicine*, **13(5)**, 721-733.
- [13] Hall, P.N., Claridge, E. and Morris Smith, J.D. (1995) Computer screening for early detection of melanoma—Is there a future? *British Journal of Dermatology*, **132**, 325-338.
- [14] White, R., Rigel, D.S. and Friedman, R.J. (1991) Computer applications in the diagnosis and prognosis of malignant melanoma. *Dermatologic Clinics*, **9**, 695-702.
- [15] Taouil, K. and Romdhane, N.B. (2006) A new automatic approach for edge detection of skin lesion images. *International Conference on Information & Communication Technologies: From Theory to Applications*, Damascus.
- [16] Kass, M., Witkin, A. and Terzopoulos, D. (1987) Snakes: Active contour models. *International Journal of Computer Vision*, **1**, 321-331.
- [17] Xu, C.Y. and Prince, J.L. (1998) Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, **7(3)**, 359.
- [18] Taouil, K., Bouhlel, M.S., Elloumi, M. and Kamoun, L. (2002) Quantification des caractéristiques de mélanomes en vue d'une classification. *JTEA'2002*, Sousse, 179-185.
- [19] Zagrouba, E. and Barhoumi, W. (2004) A preliminary approach for the automated recognition of malignant melanoma. *Image Analysis and Stereology Journal*, **23(2)**, 121-135.
- [20] Taouil, K. and Romdhane, N.B. (2006) Automatic segmentation and classification of skin lesion images. *The 2nd International Conferences on Distributed Frameworks for Multimedia Applications*, Pulau Pinang.
- [21] Pudil, P., Novoviova, J. and Kittler, J. (1994) Floating search methods in feature selection. *Pattern Recognition Letters*, **15**, 1119-1125.
- [22] Specht, D.F. (1991) A general regression neural network. *IEEE Transactions on Neural Networks*, **2**, 568-576.
- [23] Romdhane, N.B., Taouil, K., Boudaya, S. Turki, H. and Bouhlel, M.S. (2007) Sélection des Variables et Classification par Réseaux de Neurones des Lésions Derma-

- tologiques. *4th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia.
- [24] Tokan, F., Türker, N. and Yıldırım, T. (2006) ROC analysis as a useful tool for performance evaluation of artificial neural networks. *Artificial Neural Networks – ICANN 2006*, **4132**, 923-931.
- [25] Zagrouba, E. and Barhoumi, W. (2005) An accelerated system for melanoma diagnosis based on subset feature selection. *Journal of Computing and Information Technology*, **13(1)**, 69-82.
- [26] Vannoorenberghe, P., Colot, O. and de Brucq, D. (1999) Dempster-shafer's theory as an aid to color information processing application to melanoma detection in dermatology. *Proceedings of the 10th International Conference on Image Analysis and Processing*, Venice, 774-779.
- [27] Lefevre, E., Colot, O., Vannoorenberghe, P. and de Brucq, D. (2000) Knowledge modeling methods in the framework of evidence theory: An experimental comparison for melanoma detection Systems. *2000 IEEE International Conference on Man and Cybernetics*, Nashville, **4**, 2806- 2811.
- [28] Sezgin, M. and Sankur, B. (2004) Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, **13(1)**, 146-165.