IET Computer Vision

Special Section: Selected Papers from the 19th International Conference on Image Analysis and Processing (ICIAP) 2017

IET Journals
The Institution of Engineering and Technology

Evaluation of Levenberg–Marquardt neural networks and stacked autoencoders clustering for skin lesion analysis, screening and follow-up

ISSN 1751-9632
Received on 9th April 2018
Accepted on 18th July 2018
E-First on 24th August 2018
doi: 10.1049/iet-cvi.2018.5195
www.ietdl.org

Francesco Rundo[1] ✉, Sabrina Conoci[1], Giuseppe L. Banna[2], Alessandro Ortis[3], Filippo Stanco[3], Sebastiano Battiato[3]

[1]*STMicroelectronics, ADG Central R&D, Catania, Italy*
[2]*Medical Oncology Department, Cannizzaro Medical Hospital, Catania, Italy*
[3]*DMI IPLAB, University of Catania, Catania, Italy*
✉ *E-mail: francesco.rundo@st.com*

Abstract: Traditional methods for early detection of melanoma rely on the visual analysis of the skin lesions performed by a dermatologist. The analysis is based on the so-called ABCDE (Asymmetry, Border irregularity, Colour variegation, Diameter, Evolution) criteria, although confirmation is obtained through biopsy performed by a pathologist. The proposed method exploits an automatic pipeline based on morphological analysis and evaluation of skin lesion dermoscopy images. Preliminary segmentation and pre-processing of dermoscopy image by SC-cellular neural networks is performed, in order to obtain *ad-hoc* grey-level skin lesion image that is further exploited to extract analytic innovative hand-crafted image features for oncological risks assessment. In the end, a pre-trained Levenberg–Marquardt neural network is used to perform *ad-hoc* clustering of such features in order to achieve an efficient nevus discrimination (benign against melanoma), as well as a numerical array to be used for follow-up rate definition and assessment. Moreover, the authors further evaluated a combination of stacked autoencoders in lieu of the Levenberg–Marquardt neural network for the clustering step.

1 Introduction

The early detection of skin cancer, specifically, the so-called melanoma, is one of the main issues addressed by medical oncologists and dermatologists, as the probability of full oncological remission is strongly correlated to early detection. A robust and efficient approach for nevus discrimination is currently under investigation by physicians and bio-medical engineers with the aim to develop a non-invasive 'Point of Care (PoC)' for real-time skin cancer detection. Both oncologists and dermatologists make use of a heuristic approach to evaluate skin lesions by the visual inspection of dermoscopy images, also known as ABCDE (Asymmetry, Border irregularity, Colour variegation, Diameter, Evolution). The ABCDE assessment of the nevus is based on physician experience and background. Clearly, this strategy suffers from clinician subjectivity such as low sensibility and specificity; often it is required a nevus invasive biopsy to confirm the diagnosis. In order to address this issue, this paper proposes an automatic skin lesion pipeline based on the analysis of dermoscopy images in order to discriminate between benign lesions and malignant ones, with the objective to have a good trade-off between sensibility and specificity. The method has been successfully validated on a dermoscopy image dataset of PH2 open database [1]. Several approaches have been proposed in the past for automatic and robust detection of the skin cancer, including statistical hand-crafted features or soft-computing approaches based on machine learning algorithms. In [2] the authors review several recent methods for melanoma detection based on the exploitation of image features, neuro-fuzzy approaches, clustering methods (K-means, support vector machine – SVM etc.) and some methods based on the study of melanocytes distribution on the skin histopathologic images, as well as based on probabilistic analysis (e.g. Bayesian classifier). The overall analysis reports some very promising methods based on the usage of neural networks or adaptive thresholding analysis for the classification of skin lesions. Among other, methods based on pattern recognition approaches combined with classical statistical dermoscopy, image features pointed out acceptable results [2]. In [1, 2] the authors show a method for classifying skin cancer that relies on global and local features combined with different classification systems such as SVM, artificial neural network (ANN), K-nearest neighbour, Naive-Bayes algorithm comparing pros and cons of the analysed approaches. Conoci *et al.* in [3] propose a set of hand-crafted image features combined with a feed-forward neural network system (NNS). An interesting approach is proposed in [4], where some hand-crafted image features are combined with a deep-learning algorithm able to extract learned features. An SVM engine is then used for performing final classification by means of a scores-based approach. The results reported in [4] show a limited sensibility/specificity for the analysed image dataset. Other approaches [5, 6] make use of 'standard' image features coupled with a classification engine (i.e. SVM or K-nearest neighbour etc.) with limited results. The proposed method is based on a combined approach of *ad-hoc* customised hand-crafted image features and a neural network system (i.e. we evaluated both a Levenberg-Marquardt Neural Network and a stacked autoencoder) which properly trained. The presented pipeline is able to perform a robust and efficient classification. The remaining part of the paper is organised as follows. Section 2 provides a description of the proposed pipeline, by detailing the pre-processing steps and the exploited features, as well as the Levenberg–Marquardt (LV) based approach. Section 3 describes the approach based on the stacked autoencoders, which provides a further improvement of the performances. The achieved results are commented in Section 4, which also provides insights for future works.

$$
\begin{aligned}
C\frac{\mathrm{d}x_{ij}(t)}{\mathrm{d}t} &= -x_{ij}(t) \\
&+ \sum_{C(k,l)\in N_r(i,j)} A(i,j;k,l)y_{kl}(t) \\
&+ \sum_{C(k,l)\in N_r(i,j)} B(i,j;k,l)u_{kl}(t) \\
&+ \sum_{C(k,l)\in N_r(i,j)} C(i,j;k,l)x_{kl}(t) \\
&+ IN_r(i,j) = \{C(k,l), (\max(|k-i|,|l-j| < r\} \\
&k \in [1,m], l \in [1,n] \\
y_{ij}(t) &= 0.5\left(\left|x_{ij}(t)+1\right| - \left\|x_{ij}(t)-1\right\|\right)
\end{aligned} \tag{1}
$$

2 Proposed pipeline: system description

The overall scheme of the processing pipeline is synthetically sketched in Fig. 1. The output of the system consists of an array containing both the representation of morphologic features applied and the skin classification lesion expressed in terms of probability of oncological malignant progression. The following subsections describe the details of each involved step.

2.1 Dermatoscope and SC-CNNs pre-processing block

Fig. 2 shows a typical dermoscopy image of a nevus acquired by means of a medical optical dermoscope. The RGB image provided by the dermatoscope is first converted into YCbCr format [7], then the method is applied only to the luminance component of the source RGB dermoscopy image $D(x,y)$. The luminance component, denoted as $Y_{D(x,y)}$, is fed as 'input' and 'state' to a state-controlled cellular neural network (SC-CNN) with a two-dimensional (2D) matrix structure and size equal to the source image $Y_{D(x,y)}$ of size m x n.

The classical cellular neural network (CNN), introduced by Chua and Yang [8], can be defined as a high speed local interconnected computing array of analogue processors. The CNN processing is defined through the instructions provided by the so-called *cloning templates* [8]. Each cell of the CNN array may be considered as a dynamical system that is arranged into a topological structure, usually a 2D or 3D grid. The CNN cells interact with each other within their neighbourhood defined by a heuristic radius [8]. Each CNN cell has an *input*, a *state* and an *output* which is a functional mapping of the state, usually by means of *PWL* (PieceWise Linear) function. The CNN can be implemented with analogue discrete components or *VLSI* (Very Large Scale Integration) technology so that it is able to perform high speed 'near real-time' computations. Some stability results and considerations about the dynamics of the CNNs can be found in [8, 9]. Arena *et al.* [9] introduced an updated version of the CNN model called '*State Controlled Cellular Neural Network* (*SC-CNN*)', as it directly explicates the dependency of the dynamic evolution of the cell to the 'state' of the single cell. We refer to SC-CNNs in the following mathematical formulations. By assigning each normalised grey-level of the input source image (i.e. $Y_{D(x,y)}$ of size m x n) to each cell of the SC-CNNs (input and state of each cell), several image processing tasks may be performed according to the defined cloning templates instructions [8]. Equation (1) defines the state equation of a (m x n) SC-CNNs.

$N_r(i,j)$ represents the neighbourhood of each cell $C(i,j)$ with radius r. The terms in (1) represent the state x_{ij}, the output y_{ij}, and the input u_{ij} of the cell $C(i,j)$, the cloning templates $A(i,j;\ k,l)$, $B(i,j;k,l)$, $C(i,j;k,l)$, and the constant bias I. The proposed SC-CNNs block can be used for several pre-processing operations such as lesion segmentation, edge detection, noise reduction, pixels averaging and so on [9]. In this paper, we propose the usage of SC-CNNs for skin lesion pre-processing with a novel *ad-hoc* cloning templates setup. Robust segmentation of the nevus is another challenging task since there are several issues related to noise removal (hairs, angiomas etc.) without significant distortion of the original region of interest (ROI) of the lesion image. The approach proposed in this paper is evaluated by using dermoscopy images provided by PH2 database in which segmented mask for each medical image is provided [7].

The SC-CNNs pre-processing pipeline is detailed in Fig. 3. It performs a *transient* pre-processing of grey-level converted lesion image $Y_{D(x,y)}$, by using the cloning templates defined in (2) [7, 10, 11]

$$
A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 0.25 & 0.25 \\ 0.25 & 3 & 0.25 \\ 3 & 0.25 & 0.25 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad I = 0.7 \tag{2}
$$

The cloning templates in (2) are useful to configure SC-CNNs in order to perform *ad-hoc* adaptive time-transient increasing of grey-level $Y_{D(x,y)}$ image contrast, reducing the image distortion produced by the typical gel used by a dermatologist during medical dermoscopy.

The dermoscopy grey-level image processed by SC-CNNs is denoted as $Y_{\text{SC-CNN}(x,y)}$. Figs. 4*a–c* show the output images related

Fig. 1 *Proposed feed-forward skin lesion analysis pipeline*

Fig. 2 *Typical dermoscopy image of a skin nevus*

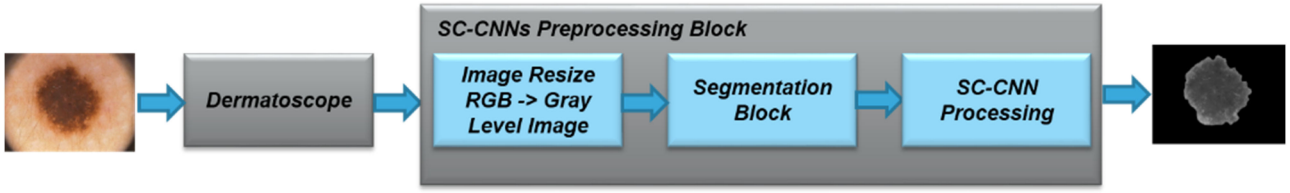


Fig. 3 *Pre-processing block based on SC-CNNs*

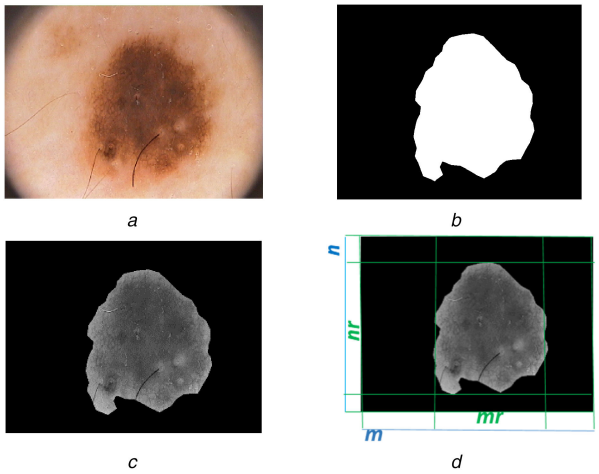


Fig. 4 *Image preprocessing details*

(*a*) Original RGB dermoscopy image, **(*b*)** Image-mask provided by PH^2 database, **(*c*)** SC-CNNs pre-processed segmented grey-level image, **(*d*)** Some geometric parameters for the segmented SC-CNNs grey-level image

to the SC-CNN pre-processing for a dermoscopy example image, as well as the image-mask provided in the PH2 database [1].

2.2 Hand-crafted image features and SC-CNNs hand-crafted blocks

At this level, the pre-processed grey-level dermoscopy image referred as $Y'_{\text{SC-CNN}(x,y)}$, is further processed by using an *ad-hoc* morphological-heuristic set of hand-crafted features useful to reproduce the well-known *ABCDE* rule [12]. The image features are re-scaled via a logarithmic function, with the aim to reduce the range of their values. Some of them are classical statistical indicators, whereas others are heuristically defined encoding the *ABCDE* rule, with high precision and robustness. Furthermore, some of the proposed hand-crafted image features try to analyse the melanocytes distribution on the skin epidermis [12]. We denote with *nr* and *mr* the dimension of the bounding box enclosing the signal (Fig. 4*d*). The grey-level intensity of the single image pixel is denoted by $p(i,j)$, whereas $\Theta(p'(i,j))$ is a measure of the frequency rate of the pixel $p'(i,j)$ in the image. In the following, the overall list of employed features F_i is reported

$$F_1 = \log\left(\frac{1}{m\cdot n}\sum_{i=1}^{mr}\sum_{j=1}^{nr} p(i,j)\right) \tag{3}$$

$$F_2 = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}(|p(i,j)-F_1|)\right) \tag{4}$$

$$F_3 = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}(p(i,j)-F_1)^2\right) \tag{5}$$

$$F_4 = \sqrt{|F_3|} \tag{6}$$

$$F_5 = \log\left(\frac{\pi}{2}\cdot\frac{1}{m\cdot n}\sum_{i=1}^{m}\sum_{j=1}^{n}(|p(i,j)-F_1|)\right) \tag{7}$$

$$F_6 = -\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\Theta(p'(i,j))\cdot\log(\Theta(p'^{(i,j)}))\right) \tag{8}$$

$$F_7 = \log\left(\sqrt{|F_3-(F_6)^2|}\right) \tag{9}$$

$$F_8 = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{|p(i,j)-F_1|}{F_4}\right)^3\right) \tag{10}$$

$$F_9 = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right|\right)\cdot\left(\left|j-\frac{n}{2}\right|\right)\left(\frac{p(i,j)-F_1}{F_4}\right)^4\right) \tag{11}$$

$$F_{10} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right|\right)\cdot\left(\left|j-\frac{n}{2}\right|\right)\left(\frac{|p(i,j)-F_1|}{F_4}\right)^5\right) \tag{12}$$

$$F_{11} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right|\right)\cdot\left(\left|j-\frac{n}{2}\right|\right)\left(\frac{p(i,j)-F_1}{F_4}\right)^6\right) \tag{13}$$

$$F_{12} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right|\cdot\left|j-\frac{n}{2}\right|\cdot p(i,j)\right)\right) \tag{14}$$

$$F_{13} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right|\cdot\left|j-\frac{n}{2}\right|\cdot(p(i,j))^2\right)\right) \tag{15}$$

$$F_{14} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right|\cdot\left|j-\frac{n}{2}\right|\cdot(p(i-j)\cdot(i-j)^2)\right)\right) \tag{16}$$

$$F_{15} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m-k}\sum_{j=1}^{n-k}\left(\left|i-\frac{m}{2}\right|\cdot\left|j-\frac{n}{2}\right|\cdot(|p(i,j)-p(i+k,j+k)|)\right)\right) \tag{17}$$

$$F_{16} = \log\left(\frac{1}{mr\cdot nr}\sum_{i=1}^{m-k}\sum_{j=1}^{n-k}\left(\left|i-\frac{m}{2}\right|\cdot\left|j-\frac{n}{2}\right|\right.\right.$$
$$\left.\left.((|p(i,j)-F_1|)\cdot(|p(i+k,j+k)+\ -F_1|))\right)\right) \tag{18}$$

$$F_{17} = \log\left(m\cdot n\cdot\frac{1}{6}\cdot\frac{1}{mr\cdot nr}\cdot\left((F_8)^2+\left(\frac{1}{4}\cdot(F_9-3)^2\right)\right)\right) \tag{19}$$

The first set of features (from F_1 to F_5) is useful to provide a classical statistic representation of the ROI of the previously segmented input image. Specifically, it corresponds to the statistical study in logarithmic scale of the lesion pixel distribution $p(i,j)$ with respect to statistical indices such as mean, weighted mean, variance or standard deviation.

Regarding the feature F_6, it is a modified logarithmic scaled computation of the entropy applied to the lesion image: this is aimed to obtain an entropic characterisation (degree of order) of the temporal evolution of the lesion (component '*E*' of ABCDE rule).

Features F_7–F_{13} represent the weighted computation logarithmic scaled of moments with many orders, useful to

characterise the intensity distribution of the pixels of the analysed lesion image, along with several directions, with the aim to make more evident the gradient of such intensity (i.e. the '*C*' component of the ABCDE rule).

Features F_{14}–F_{16} allow to analytically define the so called melanocytic distribution over all the ROI of the lesion. They are also useful to contribute to the definition of the contours and the borders of the lesion.

Note that the feature F_{17} is a modified version of the '*Jarque–Bera index*', which is able to point-out *kurtosis* and *skewness* of time-series and often applied in the field of financial markets. In the proposed work, we adapted the Jarque–Bera index to the purpose related to the 2D analysis of the *kurtosis* and *skewness* features for the analysed nevus. We used $k = 5$ in (17) and $k = 3$ in (18) (see (20)) . The feature F_{18} is a modified version of the so-called '*Cosine similarity*', used to analyse the intrinsic similarity of the nevus. The features from F_{19} to F_{21} have been computed as results of further SC-CNNs post-processing performed in the 'SC-CNNs Hand Crafted Block'. The SC-CNNs programmed via '*edge detection*' cloning templates [8, 9] is used to detect edges of the grey-level segmented image $Y'_{\text{SC-CNN}(x,y)}$, producing the image $Y_{\text{SC-CNN}(x,y)}{}^{pp}$. The latest set of hand-crafted features, related to the above described image $Y_{\text{SC-CNN}(x,y)}{}^{pp}$, are defined as follows:

$$F_{19} = \log\left(\frac{1}{m \cdot n}\sum_{i=1}^{mr}\sum_{j=1}^{nr} p(i,j)\right) \tag{21}$$

$$F_{20} = \log\left(\pi \cdot \left(\frac{nr-1}{2}\right)^2\right) \tag{22}$$

$$F_{21} = \frac{\min(mr, nr)}{\max(mr, nr)} \tag{23}$$

where the parameters *nm*, *nr*, *n and m* are defined as shown in Fig. 4*d*.

2.3 Low-voltage (LV) network clustering system

After we have collected the numerical values corresponding to morphological-heuristic and hand-crafted features, we are ready to perform a skin lesion classification needed for oncological risk assessment of the analysed nevus, by providing a classification of the lesion as 'benign nevus' or 'melanoma' (skin cancer). Based on the above classification output, the algorithm provides a follow-up rate, according to an *ad-hoc* heuristic risk assessment. In particular, a time-rate of periodic medical check-up of the analysed lesion is provided, in order to perform a monitoring of the temporal evolution of the lesion by means of the numerical sequence of features $F_1, \ldots, F_{21}$.

The normalised set of the proposed innovative image features $F_1, \ldots, F_{21}$ is used as input of a feed-forward *Levenberg–Marquardt* (LV) optimised neural network, previously pre-trained with a training set containing skin lesions examples with different morphological features (colour, border, geometry, irregularities etc.), related to either benign and malignant (melanoma) examples. We designed and trained that LV feed-forward neural network with one hidden layer with dimension is $21 \times 25 \times 2$. In the proposed system, we used a Levenberg–Marquardt network with two output neurons, thus the output of the classifier is a 2D vector which values lies in the compact set $[0,1] \times [0,1]$. Output values higher than 0.5 will be approximated with 1, whereas values lower than 0.5 will be approximated with 0. Therefore, the final output vectors will be equal to [1,0] or [0,1], while the pair values [0,0] and [1,1] are discarded as they represent the invalid state. This design choice is motivated by the fact that we aim to extend the proposed system by allowing the model to output a fine-grained score of benignity or malignancy (or suspected malignancy) useful for biomedical characterisation of the so-called dysplastic nevus, which are the most difficult to be classified by dermatologists and oncologists. Moreover, the system has been also evaluated by considering only one output neuron, by approximating with 0 (benign) or 1 (malign) the classifier output, without significant performance differences. The exploited neural network is trained with efficient optimised Levenberg–Marquardt back-propagation learning algorithm [9, 13] with stop criteria correlated to learning mean square error. Several works in the literature have proposed the usage of ANN or CNNs in order to address the task of early detection of skin cancer by the analysis of dermoscopy images [14–16]. In our experiments, in order to improve the generalisation capability of the neural network as well as for avoiding the issue of local convergence of training algorithm, we have defined the dataset by introducing selected dermoscopy images with different patterns (colours, borders, irregularities, benign lesions, melanoma etc.). Then we split the dataset into train and test/validation sets. In our experiments, we have used MATLAB with Image and Neural Network toolboxes. The used dermoscopy images for the dataset have been taken partially from PH2 database [1] and partially from other medical database provided us by oncologists. All the source images have been resized to 768×576 pixels by using the same algorithm proposed in [17] as extended in [10].

2.4 Autonomous diagnosis system

The numerical output of the LV neural network, together with hand-crafted features, is also used to define the follow-up rate of the analysed nevus, by using heuristic threshold-based rules. We define a set of thresholds for the proposed mainly discriminative hand-crafted image features (i.e. the features F_{12}, F_{17}, F_{21}) by means of averaging computation of the values heuristically computed during the training phase of the LV neural network according to the oncologist advice

$$\text{Th}_{F_{12}} = 12.99, \quad \text{Th}_{F_{17}} = 104.00, \quad \text{Th}_{F_{21}} = 6.49$$

The above thresholds are used to define an *ad-hoc* follow-up rate according to nevus diagnosis performed by LV neural clustering system. Specifically, the autonomous diagnosis system for 'melanoma diagnosis' suggests 'Contact Physician as soon as possible' in the case that features F_{12}, F_{17}, F_{21} are higher than related thresholds, otherwise it extends the follow-up to '1 month' (at least one of the above features is less of the corresponding thresholds) since the exanimated nevus does not show high malignant medical indices according to oncologist's consultation, although the diagnosis is 'melanoma' with high probability. As per 'melanoma diagnosis', in case of 'benign nevus' diagnosis performed by LV neural clustering system, the autonomous diagnosis system suggests 'Follow-up rate >= 1 year' in the case of the main features (F_{12}, F_{17}, F_{21}) are lower than the corresponding thresholds, while it reduces the time-range 'Follow-up rate >= 6 months' in the case of at least one of the main features has a value greater than the corresponding thresholds.

3 Deep learning extended approach

We have extended the proposed algorithm by replacing the LV neural networks with recent deep learning architecture in order to evaluate further improvements to the overall performance of the proposed approach. We propose a power Stacked Autoencoder

$$F_{18} = \log\left(\frac{1}{mr \cdot nr} \cdot \left(\sum_{i=1}^{\text{round}(m/2)}\sum_{j=1}^{n}\left(\left|i-\frac{m}{2}\right| \cdot \left|j-\frac{n}{2}\right|\right) \cdot \frac{|p(i,j)-p(i+1,j)|}{\sqrt{(p(i,j))^2+(p(i+1,j))^2}}\right)\right.$$
$$\left. + \left(\sum_{i=1}^{m}\sum_{j=1}^{\text{round}(n/2)}\left(\left|i-\frac{m}{2}\right| \cdot \left|j-\frac{n}{2}\right|\right) \cdot \frac{|p(i,j)-p(i,j+1)|}{\sqrt{(p(i,j))^2+(p(i,j+1))^2}}\right)\right) \tag{20}$$

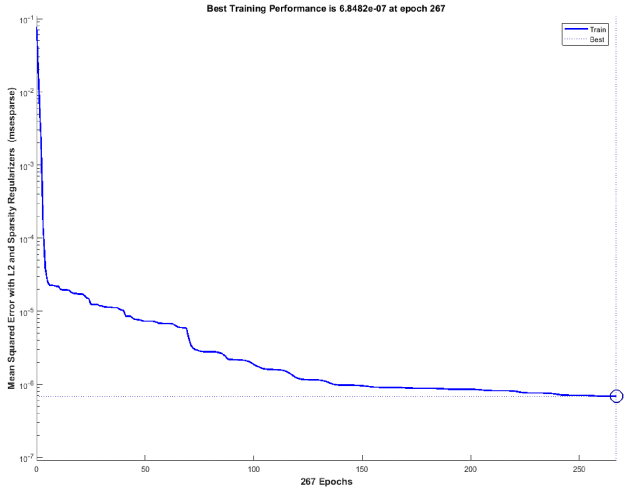


Fig. 5 *Dynamic of the mean squared error during the training phase*

Table 1 Benchmarks comparison for the proposed pipeline

Method	Performance indicators		
	Sensibility, %	Specificity, %	*C*
LV approach	97	*95*	*0.038*
first 13 features	93	92	0.074
stacked autoencoders	*98*	**98**	**0.020**
global method (colour)	90	89	0.104
global method (textures)	93	78	0.130
local method (colour)	93	84	0.106
local method (textures)	88	76	0.168
global features C6 – kNN classifier	**100**	71	0.116
global features C1 – kNN classifier	88	81	0.149
global features C6 – SVM	84	78	0.183
global features C6 – AdaBoost	96	77	0.117
local features BoF	*98*	79	0.100

Best results are highlighted in bold.

based framework composed by two autoencoders sequentially grouped with a SoftMax layer needed for clustering the features computed by each autoencoder into two different classes: [0.5, 1] for malignant nevus and [0, 0.5] foe benign ones. Fig. 5 shows the dynamic of the mean squared error during the training phase. To avoid overfitting, we can define a regularised empirical risk, where the regularisation imposes a degree of sparseness on the derived encodings. We used a common sparsity constraint based on the classic Kullback–Leibler divergence approach [11, 18, 19]. Each autoencoder has 20 neurons in the hidden layer. The first autoencoder takes the proposed hand-crafted features as input data, whereas the second autoencoder takes the features defined by the first autoencoder.

Finally, the SoftMax layer performs a clustering of the second encoded features into two basic classes as above described (benign/malignant).

The proposed approach shows very promising results as it is able to improve the overall performance. Indeed, the sensibility is 98%, and specificity is 98% with the same image dataset (PH2). Future works will be devoted to the investigation of further deep learning approaches with the aim to improve the performances.

4 Experimental results and future works

The proposed method has been validated by using the skin lesions database named 'PH2', kindly provided within the ADDI project [1, 20]. The PH2 database consists of 176 dermoscopy images carefully classified in the related project webpage [1, 20]. The performance of the proposed pipeline has been compared with algorithms proposed in the ADDI project [1, 20] and validated with

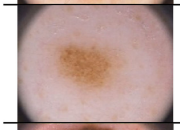
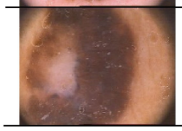
PH² image	*Proposed pipeline classification*	*Classification in PH² database*	*Follow-up rate*
	Non Melanoma (Benign Nevus)	Common Nevus (Benign Nevus)	>= 6 months
	Non Melanoma (Benign Nevus)	Common Nevus (Benign Nevus)	>= 1 year
	Non Melanoma (Benign Nevus)	Dysplastic Nevus (Benign Nevus)	>= 6 months
	Melanoma (Cancer lesion)	Melanoma (Cancer lesion)	Contact physician as soon as possibile

Fig. 6 *Some instances of classified nevus*

PH2 database. In order to perform a fair comparison of the proposed method with respect to other approaches, we evaluated our approach by computing the same benchmark indicators computed in [1, 20]. Namely, the Sensibility (SE), the Specificity (SP) and the cost function 'C'

$$C = \frac{c_{10}(1 - \mathrm{SE}) + c_{01}(1 - \mathrm{SP})}{c_{10} + c_{01}} \qquad (24)$$
$$c_{10} = 1.5c_{01}, \quad c_{01} = 1$$

Equation (24) defines c_{10} as the weight coefficient of incorrect classification of a melanoma as benign nevus (a false negative FN), while c_{01} indicates the weight coefficient for incorrect classification of benign nevus as melanoma (a false positive FP). From values reported in (24), results clear that the system considers a missed classification of melanoma more dangerous with respect to a wrong classification of benign nevus as melanoma. Table 1 reports the performance benchmark comparison [1, 20].

Table 1 shows the comparison between the evaluated methods. For each metric, the best results are highlighted in bold whereas the second best results are underlined. The comparison between benchmarks shows the very promising performance of the proposed methods with respect to others proposed in the literature. An innovative combination between *ad-hoc* hand-crafted image features with supervised neural networks drastically improves the ability of the overall discrimination system in identifying the key nevus features for a robust classification of benign lesions against high suspected ones or melanoma cancer (see Fig. 6 for results). Several works in the literature, in order to improve the ability of the pipelines, increases the sensibility reducing the over-all checks of the skin lesion image statistics. The drawback of poor specificity of the pipeline means more nevus biopsy for the patients.

We reached only 93% of sensitivity with 92% of specificity with only first 13 features of the overall set of hand-crafted features herein proposed. These results confirm that we have strongly improved the discrimination capability of the proposed approach extending the hand-crafted features as herein reported by means of more accurate melanocytes mathematical characterisation.

Experiments have been conducted with the aim to reduce the number of the employed features. However, the reduction in the number of features resulted in lower performances. These experiments confirm the need of all the proposed features.

We further improved the performances obtained with the LV network by exploiting a stacked autoencoder architecture. In particular, the improvements have been obtained by replacing the LV with the stacked autoencoders in the same pipeline described above.

The time performance of the proposed method is acceptable as the proposed pipeline is able to analyse a single nevus in about 2.5

s (we tested the pipeline in a PC Intel Core i5 3.10 GHz @ 64 bit with 4 Gbyte of RAM). The proposed pipeline can be ported easily from the MATLAB environment to an embedded platform based on STM32 [21]. We proposed a very efficient method for skin nevus analysis and oncological classification both for screening and follow-up of the exanimated skin lesion. We are extending skin lesions image dataset in order to cover all the specific skin lesion features with the aim to perform a more accurate clustering, further including the class of 'suspected' nevus, that represents one of the main issues for dermatologists and oncologists. Preliminary results and comparisons with analytic thresholds-based strategy have been done in [12] without any strong improvement. Future works will be devoted to evaluating other recent deep architectures.

5 References

[1] Mendonça, T., Ferreira, P.M., Marques, J.S., *et al.*: 'PH2 – a dermoscopic image database for research and benchmarking'. 35th Int. Conf. the IEEE Engineering in Medicine and Biology Society, 3–7 July 2013, Osaka, Japan

[2] Binu Sathiya, S., Kumar, S.S., Prabin, A.: 'A survey on recent computer-aided diagnosis of melanoma'. 2014 Int. Conf. Control, Instrumentation, Communication and Computational Technologies (ICCICCT)

[3] Conoci, S., Rundo, F, Petralia, S., *et al.*: 'Advanced skin lesion discrimination pipeline for early melanoma cancer diagnosis towards PoC devices'. IEEE Proc. the Circuit Theory and Design European Conf. (ECCTD), 4–6 September 2017, Catania

[4] Majtner, T., Yildirim-Yayilgan, S., Hardeberg, J.Y.: 'Combining deep learning and hand-crafted features for skin lesion classification'. Sixth Int. Conf. Image Processing Theory, Tools and Applications (IPTA), 2016

[5] Jamil, U., Khalid, S., Usman Akram, M.: 'Dermoscopic feature analysis for melanoma recognition and prevention'. Sixth Int. Conf. Innovative Computing Technology (INTECH), 2016

[6] Rashad, M.W., Takruri, M.: 'Automatic non-invasive recognition of melanoma using support vector machines'. BioSMART Conf., 2016

[7] Gonzalez, R.C., Woods, R.E.: '*Digital image processing*' (Prentice-Hall, New York, NY, USA, 2018, 4th edn.)

[8] Chua, L.O., Yang, L.: 'Cellular neural networks: theory', *IEEE Trans. Circuits Syst.*, 1988, **35**, (10), pp. 1257–1272

[9] Arena, P., Baglio, S., Fortuna, L., *et al.*: 'Dynamics of state controlled CNNs'. IEEE Proc. of Int. Symp. Circuits and Systems, ISCAS'96, Atlanta, GA, USA, 1996

[10] Battiato, S., Rundo, F., Stanco, F.: 'ALZ: adaptive learning for zooming digital image'. IEEE Proc. of Int. Conf. Consumer and Electronics, Las Vegas, NV, USA, 2007, pp. 1–2

[11] Lee, H., Ekanadham, C., Ng, A.Y.: 'Sparse deep belief net model for visual area V2', *Adv. Neural. Inf. Process. Syst.*, 2007, **7**, pp. 873–880

[12] Rundo, F., Banna, G.L.: 'A Method of analyzing skin lesions, corresponding system, instrument and computer program product'. EU Registered Patent App. N. 102016000121060, 29 November, 2016

[13] Hagan, M.T., Menhaj, M.: 'Training feed-forward networks with Marquardt algorithm', *IEEE Trans. Neural Netw.*, 1994, **5**, (6), pp. 989–993

[14] Fukushima, K.: 'Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position', *Biol. Cybern.*, 1980, **36**, (4), pp. 93–202

[15] Fridan, U., Sarı, İ., Kumrular, R.K.: '*Classification of skin lesions using ANN*' Medical Technologies National Congress (TIPTEKNO), Antalya, Turkey, 2016

[16] Xie, F., Fan, H, Li, Y., *et al.*: 'Melanoma classification on dermoscopy images using a neural network ensemble model', *IEEE Trans. Med. Imaging*, 2017, **36**, (3), pp. 849–858

[17] Battiato, S., Gallo, G., Stanco, F.: 'A new edge-adaptive zooming algorithm for digital images'. Proc. Signal Processing and Communications SPC, 2000, pp. 144–149

[18] Bengio, Y.: 'Learning deep architectures for AI', *Found. Trends Mach. Learn.*, 2009, **2**, (1), pp. 1–127

[19] Titsias, M.K.: 'Variational learning of inducing variables in sparse Gaussian processes'. Int. Conf. Artificial Intelligence and Statistics, 2009a, vol. 12, pp. 567–574

[20] Barata, C., Ruela, M., Francisco, M., *et al.*: 'Two systems for the detection of melanomas in dermoscopy images using texture and color features', *IEEE Syst. J.*, 2013, **99**, pp. 1–15

[21] STM32 32-bit ARM Cortex MCUs: Available at http://www.st.com/en/microcontrollers/stm32-32-bit-arm-cortex-mcus.html?querycriteria=productId=SC1169, accessed August 2018