METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.pg.gda.pl

DIAGNOSIS OF MALIGNANT MELANOMA BY NEURAL NETWORK ENSEMBLE-BASED SYSTEM UTILISING HAND-CRAFTED SKIN LESION FEATURES

Michał Grochowski, Agnieszka Mikołajczyk, Arkadiusz Kwasigroch

Gdańsk University of Technology, Faculty of Electrical and Control Engineering, G. Narutowicza 11/12, 80-222, Gdańsk, Poland (✉ *michal.grochowski@pg.edu.pl*, +48 58 347 2904, *agnieszka.mikolajczyk@pg.edu.pl*, *arkadiusz.kwasigroch@pg.edu.pl*)

Abstract

Malignant melanomas are the most deadly type of skin cancer, yet detected early have high chances of successful treatment. In the last twenty years, the interest in automatic recognition and classification of melanoma dynamically increased, partly because of appearing public datasets with dermatoscopic images of skin lesions. Automated computer-aided skin cancer detection in dermatoscopic images is a very challenging task due to uneven sizes of datasets, huge intra-class variation with small interclass variation, and the existence of many artifacts in the images. One of the most recognized methods of melanoma diagnosis is the ABCD method. In the paper, we propose an extended version of this method and an intelligent decision support system based on neural networks that uses its results in the form of hand-crafted features. Automatic determination of the skin features with the ABCD method is difficult due to the large diversity of images of various quality, the existence of hair, different markers and other obstacles. Therefore, it was necessary to apply advanced methods of pre-processing the images. The proposed system is an ensemble of ten neural networks working in parallel, and one network using their results to generate a final decision. This system structure enables to increase the efficiency of its operation by several percentage points compared with a single neural network. The proposed system is trained on over 5000 and tested afterwards on 200 skin moles. The presented system can be used as a decision support system for primary care physicians, as a system capable of self-examination of the skin with a dermatoscope and also as an important tool to improve biopsy decision making.

Keywords: decision support, diagnostics, image processing, artificial neural networks, ensemble of neural networks, melanoma malignant.

1. Introduction

Malignant melanomas are the most deadly type of skin cancer, yet detected early have high chances of successful treatment. Regular skin examinations are very important, especially for people who belong to a group with a higher risk of melanoma *e.g.*: Caucasian race, with blue or green eyes, having light hair, with family melanoma history, in advanced age. Although skin lesions are easily visible to an unaided eye, early-stage melanomas are often not discerned due to their similarity to the benign ones. On the other hand, many benign lesions are unnecessary

Article history: received on Apr. 30, 2018; accepted on Oct. 13, 2018; available online on Mar. 29, 2019, DOI: 10.24425/mms.2019.126327.

examined by biopsy – still, they are always worth examining just in case. The skin can be examined either by health care professionals themselves or by specialized dermatologists. The most popular way of diagnosing skin cancer is dermatoscopy, with the use of popular methods like the ABCD criterion, 7-point checklist, Menzies method [1–3].

In the last twenty years, the interest in automatic computer-aided detection and classification of skin cancer significantly increased, partly because of appearing public datasets with dermatoscopic images of skin lesions. The most popular datasets are the Dermofit Image Library consisting of 1300 images of skin lesions, the MoleMap dataset (over 32 000 images) [4], the ISIC Archive (over 13 000) [5], and the database from ISBI melanoma recognition challenge (around 1300 images) [6].

Researches tried to resolve the skin cancer diagnosis problem using two different types of approaches:

- based on hand-crafted features, with a simple classifier, usually preceded by advanced preprocessing and lesion segmentation;
- based on complex deep neural networks, which are taught to extract the features and to classify the inputs during the training process.

In the first case, the greatest emphasis is put on the selection and extraction of reasonable features, often based on the methods proposed by doctors or medical specialists. Such an approach requires the use of appropriate pre-processing methods, since if poorly performed they will result in a poor extraction of features that may result in an incorrect diagnosis. In the second approach, pre-processing is usually limited to basic operations or even entirely omitted due to the use of convolutional filters that already provide graphical operations on images with weights that are adjusted by an optimization algorithm during the training. What is more, the features describing the mole are automatically selected during the training. In this case, the greatest attention is paid to choosing the right architecture and preparing the data. It is worth noticing that this kind of solution may lead to finding important visual features that are hard to notice by just the human eye or even features that were not considered by the specialists before, but also to recognizing dangerous "false features", for example associating the occurrence of curly hair on the picture with malignancy. A few researchers have addressed the problem of adversarial attacks, for example, the researchers in [7] performed an experiment, in which they proved that changing even only one pixel in the image can fool the CNN (*Convolutional Neural Network*) and significantly change the given output. Their attack ended with success in most cases: the decision of the CNNs changed after modifying the value of one pixel in the image.

An advantage of the hand-crafted selection of features, as it is for example in the ABCD rule, and their later use by the classifier, is that the medical specialists feel that they oversee the decision-making process. The researchers in [8] detect melanoma with the Otsu thresholding method, extract features like asymmetry, border, colour and diameter, and directly apply the ABCD Stoltz rule without any kind of neural classifier. Directly applied medical methods have strong disadvantages: in this kind of approach the final results may be different for a rotated image or different colour balance and it may even lead to an incorrect final diagnosis. In [9] the authors proposed a system based on five different asymmetry features, one border feature and four colour features (without the lesion pattern feature) reaching 0.90 sensitivity and 0.72 specificity. In [10–11] the authors presented respectively the summary of ISBI 2016 and 2017 Skin Lesion Analysis towards Melanoma Detection Challenge, making this the largest standardized and comparative study for melanoma diagnosis in dermatoscopic images to date (2016: 79 submissions from a group of 38 participants, 2017: 46 finalized submissions and approximately 50 attendees). According to the paper, the best approaches in each edition reached: 80.4% AUC and 91.1% AUC. Analysis of the classi?cation task performed by the authors demonstrates that ensembles

of deep learning NNs led to the highest performance. In addition, collaborative fusions of all participant systems outperformed any single system alone. The researchers in papers [12-13] focused on analyzing which kind of features are really important in melanoma recognition. They noticed that most of the melanomas show irregularity in two axes, consist of 3 or more colours and are characterized by atypical networks or streaks, dots and globules, and that skin lesion classification has a strong similarity with a research field called fine-grained image classification (big intra-class variation with small inter-class variation) [13].

This paper proposes a new approach to automatic malignant melanoma diagnostics based on the modified ABCD dermatoscopic criteria of Stoltz. We shed a new light on the well-known ABCD rule by proposing a set of hand-crafted features essential for the diagnosis. Along with the algorithm for automatic feature extraction, we present a specially designed neural system with two layers of neural networks. The obtained results are competitive to the results reported in other publications. We would like to emphasize that we do not aim to give a broad description of pre-processing and skin lesion segmentation stages because both these steps can be replaced with other customized methods.

The paper is organized as follows: first, a simplified introduction to the ABCD Stoltz rule of dermatoscopic melanoma recognition is presented. Then, a brief description of the pre-processing methods and the process of hair reduction from dermatoscopic images is given. Next, a set of features: symmetry, border, colour and structure is proposed along with an explanation of its application. Based on the extracted features, the neural network diagnosis system was taught to classify the skin lesions and the architecture of designed classifier is presented. The preliminary results are presented in [14].

The presented system can be used as a decision support system for primary care physicians and as a system capable of self-examination of the skin with a dermatoscope and also as an important tool to improve biopsy decision making.

2. Diagnosis of malignant melanoma

2.1. Fundamentals

The ABCD Stolz is a list of criteria being used for dermatoscopic differentiation between melanoma malignant lesions and benign moles. The rules are based on analysis of the following features of the skin moles: Asymmetry (A), Border (B), Colour (C) and Differential structural components (D). Each feature is carefully analyzed: symmetry is checked in two axes (A); border is divided into 8 parts and checked if they have smooth, even edges without noticeable structures; colours in the mole are counted; types of existing structures and textures are counted (see the example in Fig. 1).

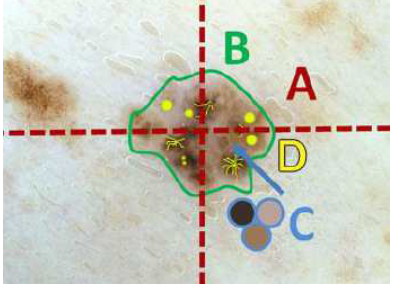


Fig. 1. ABCD dermatoscopic features.

Branched streaks and dots are counted only when more than two are clearly seen, yet the presence of a single globule is sufficient to be counted as being present in a lesion. Other structures defined in a different method proposed by Menzies are blue-white veil, multiple brown dots, pseudopods, radial streaming, scar-like depigmentation, peripheral black dots/globules, multiple blue/grey dots and broad pigment network [13]. The atypical network is closely associated with early melanoma [12]. After a wide analysis of the presence of a skin lesion, the clinician assigns points to each observed feature. The final diagnosis is based on the number of calculated points during the process of mole analysis. We have carefully analyzed the method and proposed a new extended set of features (Subsection 2.3).

2.2. Pre-processing

The purpose of the system is to detect and recognize malignant melanoma on a dermatoscopic photograph of a skin lesion. The main goal of pre-processing is to increase the image quality by reducing noise, deleting black frames around the picture and other unwanted elements such as hair, light reflections, gel bubbles. The presence of such elements makes the further steps of automatic diagnosis impossible. Examples of the images hard to pre-process are presented in Fig. 2. In this section, we briefly describe all the steps carried out in the pre-processing stage, although we do not aim to give a broad description of the pre-processing and skin lesion segmentation stages. Both steps can be replaced with other customized methods.

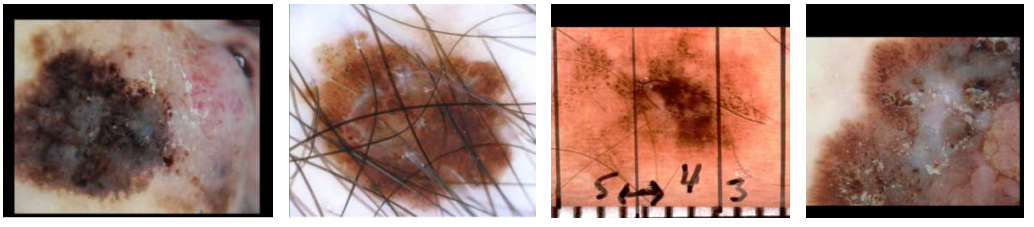


Fig. 2. Examples of images hard to pre-process.

One of the most frequently mentioned problems during segmentation of the image of skin lesion is the occurrence of hair on photographs [12, 15–16]. The visible occurrence of dense hair causes severe disturbance of the analysis process. With the above-mentioned problems in mind, we developed, prepared and implemented a method of hair detection and reduction based on a combination of reliable and traditional image processing techniques.

We divided the process of hair reduction into 3 steps: hair detection, image estimation, and hair replacement. In the first step, we perform the hair detection algorithm. The process of hair detection begins with filtration of the image: specifically, each of the RGB layers is filtered independently, with the use of a convolution matrix designed to detect corners. During the first phase the selected matrix is used twice: in its original form and after rotation by 180 degrees (Fig. 3). Using an original filter as well as its rotated form, we are able to detect corners in two different directions. Each of the RGB channels of the basic image (Fig. 4a) is filtered separately with two filters depicted in Fig. 3, resulting in receiving two binary images per each RGB layer.

Afterwards, all of the individual images were binarily summed up and all small elements were deleted resulting in a binary mask of the hair (Fig. 4b).

As soon as the hair detection process had been accomplished, we moved to the process of estimating the area under the hair. We opted for a fast algorithm which will enable blurring

-5	-5	3
-5	0	3
3	3	3

3	3	3
3	0	-5
3	-5	-5

Fig. 3. The filters applied to hair detection.

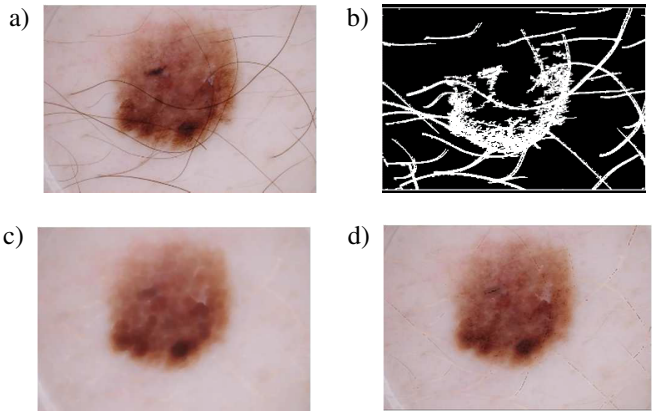


Fig. 4. The hair reduction process: a) original image; b) segmentation hair mask; c) blurred image without hair; d) final effect.

the image in such a way that the hair will disappear, but the overall colour palette will remain unchanged. To achieve this we performed a morphological closing operation on the original photograph. Having both, the hair binary image (Fig. 4b) and the blurred photograph (Fig. 4c) enabled us to replace the hair from the original picture (Fig. 4a) with the estimation of the under-hair area. Hence, all pixels in the original image marked by the binary mask of hairs were replaced by pixels from the blurred image. The final result, achieved by the described above steps is presented in Fig. 4d.

The algorithm enables to perform a notable hair reduction without significant interference in the skin lesion appearance. In Fig. 5 other results of the hair removing process are shown. After the hair reduction, the procedure moved to the next step of the pre-processing method. In the proposed system, firstly the contrast of the image is enhanced, then the number of colours in the picture is reduced and the photograph is symmetrically blurred.

After these operations, the image is morphologically opened in order to delete the noise. Finally, the picture is converted to the black-and-white form for the purpose of the next procedure: segmentation of the lesion.

Next, we performed automated segmentation of the skin mole from the dermatoscopic images. The quality of the binary mask influences all the following steps and the final decision. Two different but complementary methods perform the image segmentation. The first method (Method I) is characterized by lower quality in recognizing brighter fragments of the mole, however, it works well while determining darker areas. The effect of this method is depicted in Fig. 6. On the other hand, the second method (Method II) works well in recognizing brighter elements of the mole (Fig. 7).

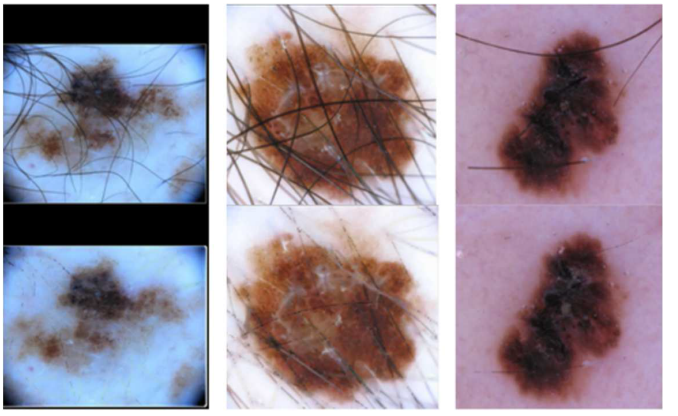


Fig. 5. Examples of hair reduction.

Method I:

1. Perform an erosion on the pre-processed image.
2. Perform the threshold operation on the received image.

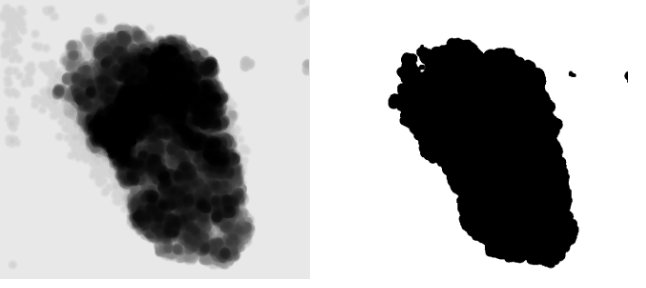


Fig. 6. Segmentation – Method I.

Method II:

1. Take the original input image, saved in the RGB colour space and:
   a) Add the R-layer to Layer B and subtract the G-layer (R + B – G), to create Picture1 (Fig. 7a);
   b) Subtract Layer R and Layer B from Layer G (G – R – B), to create Picture2 (Fig. 7b).
2. Subtract Picture2 from Picture1 (Picture1 – Picture2).
3. Enhance contrast in the image in order to make it more clearly visible (Fig. 7c).
4. Reverse colours, the skin should be bright (Fig. 7d).
5. Perform the threshold operation (Fig. 7e).

Combining the advantages of both methods results in eliminating their disadvantages and makes possible to obtain a more versatile method, which works well even in the case of multicolour features. Binary images from both methods are summed up resulting in the final segmentation mask that needs only simple final improvements: small objects are removed, holes filled and edges of the mole are smoothed (Fig. 7f).

The final result of the segmentation method, on an example of 8 different moles, is presented in Fig. 8 where the contour of segmented lesion defines the lesion border.

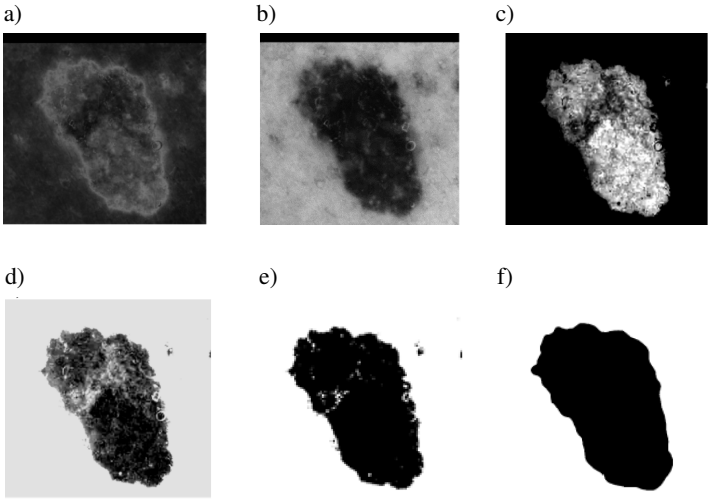


Fig. 7. Segmentation – Method II (the steps), and the final result.

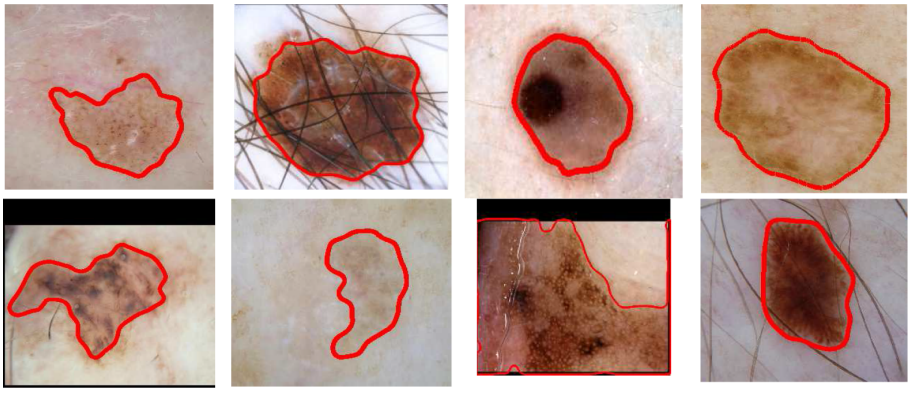


Fig. 8. Examples of border detection.

2.3. Feature extraction

The previous steps (pre-processing and segmentation) are carried out in order to extract the most important features of melanoma malignant skin lesions: symmetry, border, colour and structures. The feature analysis carried out in [12] showed that most of melanomas show irregularities in two axes, consist of 3 or more colours, have an atypical network or streaks, dots and globules. We propose a new set of features based on the ABCD rule.

Asymmetry. In the original ABCD method symmetry is checked in 2 axes. It turns out that such an approach may end with different results depending on the rotation of the input image. Instead, we propose to compare the shape of a lesion to the perfectly symmetric shape of a circle in order to avoid errors due to the selection of wrong axes. The estimated radius of the circle and the centre of the lesion mass are calculated based on the detected binary image of a mole (see Fig. 9).

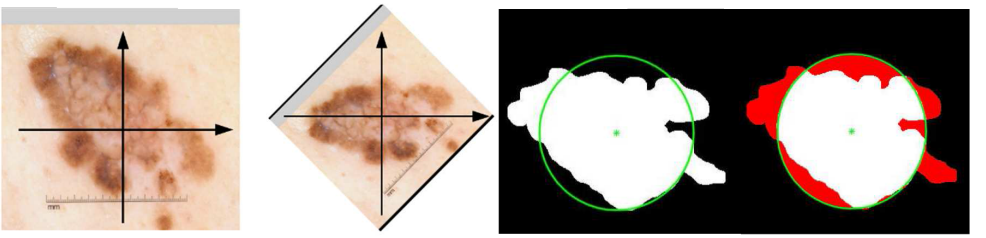


Fig. 9. An image with a highlighted asymmetry feature.

The asymmetry feature A is defined as follows:

$$A = \frac{\sum_{i=0}^{k}\sum_{j=0}^{l} \neg p_{i,j} + \sum_{i=0}^{k}\sum_{j=0}^{l} q_{i,j}}{\pi r^2}, \tag{1}$$

where: i, j – the pixel coordinates, k, l – the image dimensions, $p_{i,j}$ – the binary value of a pixel (i, j) inside a circle, $q_{i,j}$ – the binary value of a pixel (i, j) outside a circle, r – the radius [px].

In other words, the asymmetry feature is defined by the sum of the pixels defining the skin lesion outside the circle, and the pixels defining the skin inside the circle (red pixels in Fig. 9).

Border. The benign moles have smooth, even edges without noticeable structures. According to the ABCD rule, it is suggested to start the analysis of border from dividing it into 8 segments and then to evaluate them separately. Again, such a solution might lead to various final results according to differently divided segments during the automatic analysis, hence it was improved. The border feature B is defined as follows:

$$B = \frac{O * r}{2 * P}, \tag{2}$$

where: O – the perimeter of a circle [px], r – the radius [px], P – the area of a circle [px].

For the perfectly smooth border, the B feature would be equal to one. Every deformation of the contour will cause an increase of this coefficient.

Colour. The colour is a very important feature which may significantly improve recognition of a malignant skin lesion. The ABCD rule distinguishes 6 colours: light brown, dark brown, black, white, blue and red, and suggests counting the number of visible colours in the mole. However, in some cases (like a different type of lighting or colour balance), the described method can lead to overestimation or underestimation of the result and opinions may differ among the doctors. Based on this remark, we propose to apply a larger amount of automatically calculated features.

The six-colour features are proposed as: C_1 – the maximum change in colours averaged on R, G and B channels; C_2 – the size of area with dominating red colour; C_3 – the size of area with dominating green colour; C_4 – the size of area with dominating blue colour, C_5 – the maximum change of brightness in the image; C_6 – the variety in colours.

Before extracting the colour features, all unwanted elements that disturb the diagnosis process, like skin and hair, are removed.

The first colour feature C1 is defined as follows:

$$C_1 = \frac{R_{\max} - R_{\min} + G_{\max} - G_{\min} + B_{\max} - B_{\min}}{3}, \tag{3}$$

where: $R_{\max}$, $G_{\max}$, $B_{\max}$, $R_{\min}$, $G_{\min}$, $B_{\min}$ are respectively the maximum and minimum R, G, B values.

Next, three colour features determine the sizes of areas with dominating red (C2), green (C3) and blue (C4) colours. The value of each pixel is checked in the red, green and blue channel of an RGB image. For dominating red colour one should check if the value of a pixel in R channel is higher than or equal to the same pixel in other layers and if yes, include it in the sum. Then, in order to normalize the measure, all summed up pixels are divided by the total number of pixels within the image. The process of calculating features C3 and C4 is carried out in the same way.

The next feature C5 determines the maximum change of brightness in the black-and-white image:

$$C_5 = \frac{C_{\max} - C_{\min}}{2}, \tag{4}$$

where: $C_{\max}$, $C_{\min}$ are the values of brightest and darkest pixels found in the picture, respectively.

The last feature (C6) is calculated with the use of the k-means algorithm. The image is divided into five-colour clusters. The clustering process is repeated three times in order to avoid local minima (Fig. 10). Next, there is calculated the mean value of a colour for all the groups, and also the mean value of the colour for each group. For each group, the squared difference between the mean colour of all groups as well as the mean colour of a selected group are calculated. Finally, all five values are summed up and divided by the area of the lesion being analyzed.

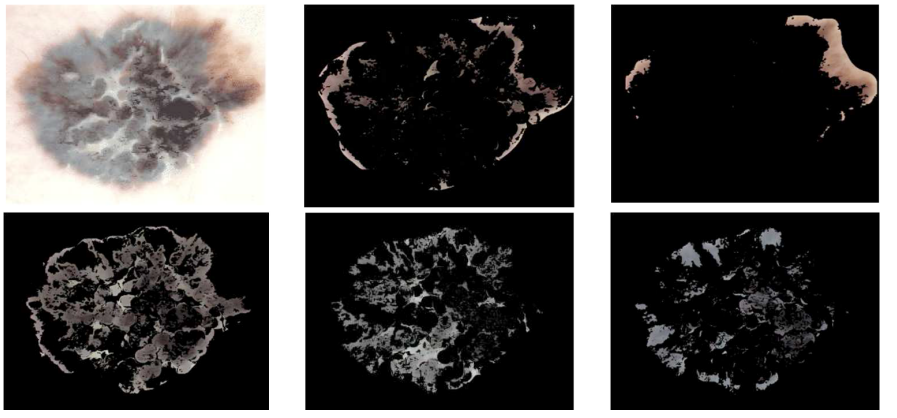


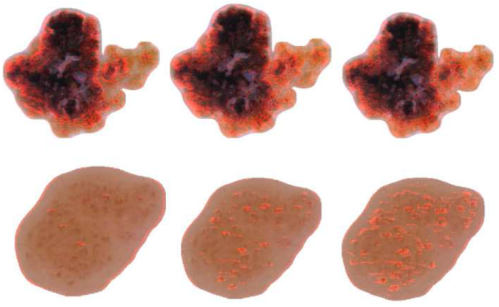
Fig. 10. K-means clustering: an original image and 5 corresponding clusters.

Different structural components. The last step in the ABCD method is analysis of different structural components of the skin lesion. There are many possible types of existing structures and they are difficult to distinguish even for experts in the field. We decided to use methods that enable to detect the general occurrence of the structures on the skin mole, without analyzing their types. The features indicate the density of the texture visible in the image.

The first method is based on filtration of the image with a specially prepared 6×6 digital convolutional filter, designed to detect corners and edges (Fig. 11b). This operation enables to calculate three structure features (D1, D2, D3). Each of the RGB layers is separately filtered by four filters, the original one (Fig. 11b) and the versions of the original filter being rotated by 90°, 180° and 270°. Next, we performed the threshold operation on all of the filtered images. Once we performed all of the threshold operations, we have received 12 binary images: 4 images per each of the RGB channels. We grouped the binary masks into 3 colour groups with 4 binary masks per each group. Next, we summed up the binary images inside the groups. The regions in each of the RGB channels remaining after filtration (pixels equal to 1 on the binary image) were summed up

and then divided by the area of the lesion. This method enables to detect all kinds of corners: all possible abnormalities from an even structure, a flat structure like radial streaming, an atypical network (Fig. 11a).

a)



b)

-2.5	-1.5	0	1.5	2.5	4.5
-2.5	1.5	2.5	4.5	0	-1.5
2.5	2.5	0	-0.5	-1.5	-2.5
4.5	0.5	-0.5	-1.5	-2.5	-2.5
2.5	0.5	0	-1.5	-2.5	-2.5
2.5	0.5	-0.5	-1.5	-2.5	-2.5

Fig. 11. The structure feature based on the convolutional filtering method: a) examples of the extracted features; b) the applied filter.

Other important elements of the structure that need to be searched for, are various types of circular spots, blobs, and drops. First, for the purpose of detecting the mentioned features, we used the *Maximally Stable Extremal Regions* (MSER) method [18]. Again, after using the MSER method, all of the found regions are summed up and divided by the area of the lesion. With this method the occurrences of peripheral black dots and globules as well as blue or grey dots are checked and labelled as D4 features. As an example, the determined MSER regions on malignant and benign skin lesions are shown in Fig. 12a. In the case of a malignant lesion, several of these areas have been detected, but in the case of a benign mole, such areas have not occurred.

a) b)

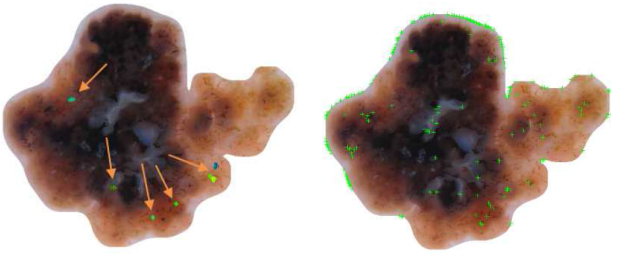


Fig. 12. The structure feature based on a) the MSER method; b) the Harris Corner Detector.

The last feature (D5) is calculated by using the Harris Corner Detector developed in 1988 [17]. The Harris detector enables to detect combined corners and edges. In order to calculate the feature, the Harris detecting method is applied to the image of the lesion. The detected pixels are counted and summed up, then divided by the area of the lesion and stored as a new feature (D5). The new feature covers all kinds of clearly outstanding structures on the mole, such as pseudopods, radial streaming, and branched streaks. An example of the image detected with the Harris method is presented in Fig. 12b. The melanoma malignant is numerously covered with markers, and at the same time, a normal mole has only a small number of markers on its edge.

3. Decision support system

The proposed decision support system is based on feedforward neural networks, which have proved to be well suited for the classification of such problems [18–22].

The described in Subsection 2.3 features of asymmetry (A), border (B), colour (C1–C6) and different structural components (D1–D5) extracted during the pre-processing phase constitute the inputs to the decision support system. However, analysis of NN efficiency working on various training and testing input sets showed that the asymmetry (A) and colour with dominating red (C2) features significantly lowered the final AUC and accuracy score. Hence, it was decided to exclude those features from the set of inputs that were fed to the neural network. Excluding any other feature ended with a lower efficiency.

The best neural network structures have been determined by analyzing possible combinations of the number of neurons and kinds of activation functions in the hidden layer, as well as the activation function in the output neuron on different datasets. Moreover, the neural networks being analyzed were initialized with different initial values of weights and biases.

The final structure, a feedforward neural network [23] with ten neurons with a sigmoid activation function in the hidden layer and a linear neuron at the network output has been selected.

Thorough research has shown that neural networks initialized with different initial conditions of weights and biases, make improper decisions when analyzing different images. However, they have similar classification effectiveness over the whole dataset. It was decided to take advantage of this fact and to develop a structure based on many different neural networks (the best among the analyzed), cooperating together in order to achieve the best classification results [24]. In this regard, the decision support system consists of ten neural networks working in parallel (each ANN with 10 neurons in the hidden layer with a logsig activation function, 1 output linear neuron) and serving as a kind of voting system, and one network using their results to generate the final decision (10 neurons in the hidden layer, 1 output neuron, all neurons with a logsig activation function). The output values greater than 0.5 indicate that the skin lesion is malignant, otherwise it is treated as a benign one. Such a system structure made it possible to achieve much better results than those obtained with a single neural network. The structure of the proposed system is shown in Fig. 13.

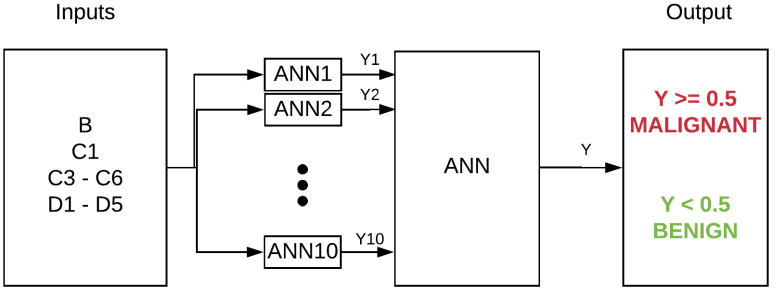


Fig. 13. The structure of the decision support system. ANN1-ANN10 structures: 10 neurons in the hidden layer with a logsig activation function, 1 output linear neuron; ANN structure: 10 neurons in the hidden layer, 1 output neuron, all neurons with a logsig activation function.

4. Results

All networks were trained using the Bayesian regularization algorithm [25]. The database used for the training consisted of 4759 benign and 976 malignant moles, while the test database contained 100 images of each type. In order to investigate the generalization ability of the network, the k-fold cross-validation procedure was applied [26]. The dataset was divided into five different representative sets of images.

The results obtained with each individual neural network are depicted in Figs. 14–18 in the form of ROC curves and gathered in Table 1. The coloured thin lines in the figures represent the results obtained with a single neural network which is part of the decision support system, whereas the thick black lines show the efficiency of the system as a whole. Table 1 contains the detailed information on AUC (*area under the curve*), accuracy, sensitivity, and specificity over 5

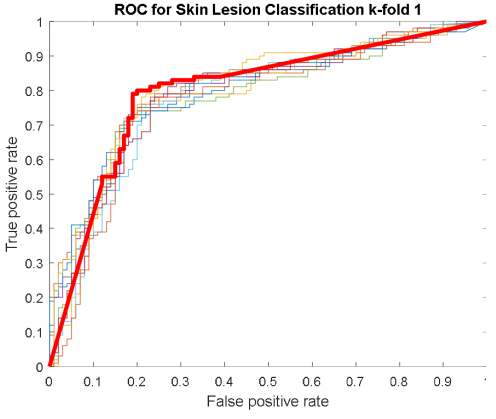


Fig. 14. Single neural network ROC curves for skin lesion classification for k-fold 1.

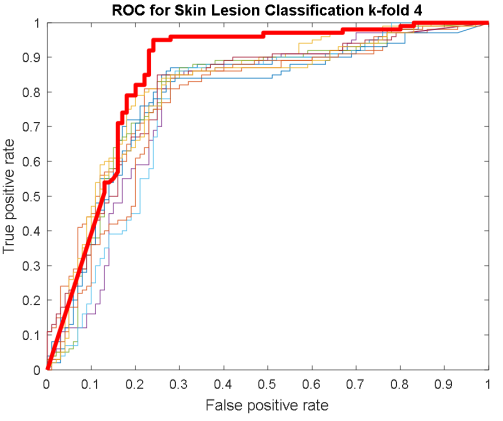


Fig. 15. Single neural network ROC curves for skin lesion classification for k-fold 4.

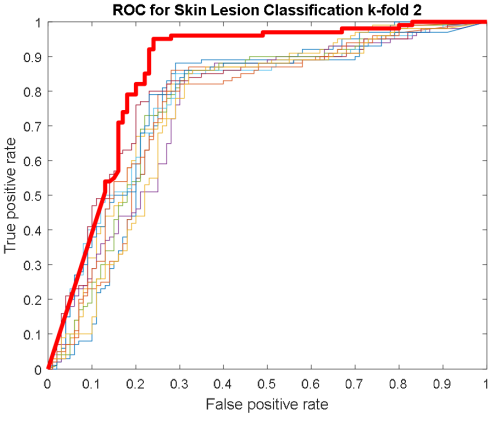


Fig. 16. Single neural network ROC curves for skin lesion classification for k-fold 2.

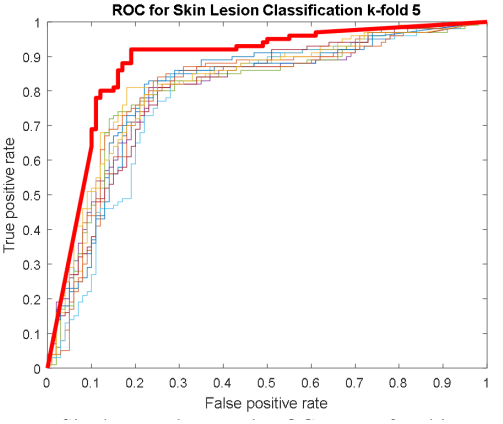


Fig. 17. Single neural network ROC curves for skin lesion classification for k-fold 5.

different folds of the dataset [27]. The average values of these measures, calculated from all folds are gathered in the last row.

Table 1. The average results calculated for 10 single NNs.

K-FOLD	AUC	ACCURACY	SENSITIVITY	SPECIFICITY
1	79.04	76.70	73.50	79.90
2	76.99	77.00	83.90	70.10
3	78.66	77.60	82.50	72.70
4	79.26	78.50	84.00	73.00
5	80.54	78.60	81.10	76.10
AVG	78.90	77.68	81.00	74.36

Table 2 and Fig. 19 show the results of the decision system being an ensemble of the neural networks described above, obtained for 5 testing folds. It should be noted that applying an ensemble of the networks increased all the indices for all testing folds (Figs. 14–18) and, in particular, the AUC by 6.02%, accuracy by 6.72%, sensitivity by 9.80% and specificity by 3.64% on average.

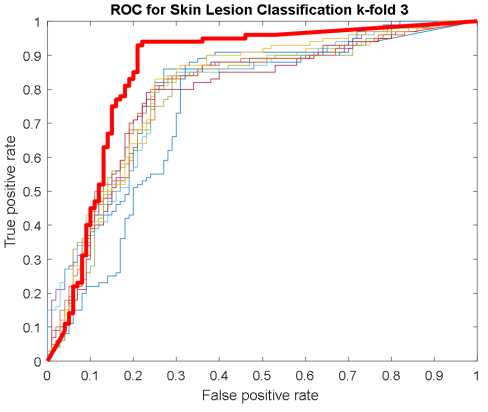


Fig. 18. Single neural network ROC curves for skin lesion classification for k-fold 3.

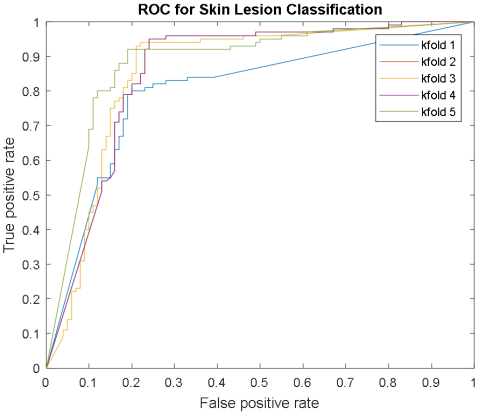


Fig. 19. Neural network ensemble-based decision support system ROC curves for skin lesion classification for all testing folds.

Table 2. The results of the NN ensemble-based decision support system.

K-FOLD	AUC	ACCURACY	SENSITIVITY	SPECIFICITY
1	80.02	79.50	80.00	79.00
2	85.53	85.00	94.00	76.00
3	85.49	86.00	94.00	78.00
4	85.53	85.00	94.00	76.00
5	88.02	86.50	92.00	81.00
AVG	84.92	84.40	90.80	78.00

5. Summary and concluding remarks

In this paper, we presented an intelligent system supporting the diagnosis of malignant melanoma using an extended ABCD rule. As stated in the introduction we presented a novel method of hair reduction based on convolutional filtering as well as the skin lesion segmentation method. Moreover, based on thorough analysis we proposed to extend the well-known ABCD criteria with additional hand-crafted features that helped to improve further lesion classification. We presented a way of automatic extraction of the mentioned features and performed its wide analysis.

Finally, we proposed an NN-based decision support system consisting of ten neural networks working in parallel and one network using their results to generate the final decision on the lesions being analyzed. The results are very encouraging: the system reached high classification results of 85% AUC, 84% accuracy, 91% sensitivity, and 78% specificity, on the average.

The achieved results are comparable to those obtained with deep learning techniques [28–31] but in contrast to these very black-box methods, the process of inferring in the character of the lesions is based on medical procedures and is much more transparent for specialist doctors, what is an advantage.

Acknowledgments

This work was supported by the Ministry of Science and Higher Education, Poland (Diamond Grant #DI2016020746).

References

[1] Nachbar, F., Stolz, W., Merkle, T., Cognetta, A.B., Vogt, T., Landthaler, M., *et al.* (1994). The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *J. Am. Acad. Dermatol.*, 30, 551–559.

[2] Johr, R.H. (2002). Dermoscopy alternative melanocytic algorithms – the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist. *Clin. Dermatol.*, 20, 240–247.

[3] Henning, J.S., Dusza, S.W., Wang, S.Q., Marghoob, A.A., Rabinovitz, H.S., Polsky, D., *et al.* (2017). The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *J. Am. Acad. Dermatol.*, 56, 45–52.

[4] MoleMap, *https://molemap.net.au*. (Jan. 2018).

[5] ISIC Archive, *https://isic-archive.com*. (Jan. 2018).

[6] ISBI 2017, *http://biomedicalimaging.org/2017/challenges*. (Jan. 2018).

[7] Su, J., Vargas, D.V., Kouichi, S. (2017). *One pixel attack for fooling deep neural networks*. ArXiv171008864 Cs Stat.

[8] Gopinathan, S., Rani, S.N.A. (2016). The melanoma skin cancer detection and feature extraction through image processing techniques. *Orthopedics*, 5, 11.

[9] Jafari, M.H., Samavi, S., Karimi, N., Soroushmehr, S.M.R., Ward, K., Najarian, K. (2016). Automatic detection of melanoma using broad extraction of features from digital images. *Eng. Med. Biol. Soc. EMBC 2016 IEEE 38th Annu. Int. Conf. On, IEEE*, 1357–1360.

[10] Gutman, D., Codella, N.C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., *et al*. (2016). *Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC).* ArXiv Prepr ArXiv160501397

[11] Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., *et al*. (2017). *Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)*. ArXiv Prepr ArXiv171005006.

[12] Jamil, U., Khalid, S., Akram, M.U. (2016). Dermoscopic feature analysis for melanoma recognition and prevention. *Innov. Comput. Technol. INTECH 2016 Sixth Int. Conf. On, IEEE*, 290–295.

[13] Ge, Z., Demyanov, S., Bozorgtabar, B., Abedini, M., Chakravorty, R., Bowling, A., *et al*. (2017). Exploiting local and generic features for accurate skin lesions classification using clinical and dermoscopy imaging. Biomed. *Imaging ISBI 2017 2017 IEEE 14th Int. Symp. On, IEEE*, 986–990.

[14] Mikołajczyk, A., Kwasigroch, A., Grochowski, M. (2017). Intelligent system supporting diagnosis of malignant melanoma. *In Polish Control Conference Springer, Cham*., 828–837.

[15] Fiorese, M, Peserico, E, Silletti, A. (2011). VirtualShave: automated hair removal from digital dermatoscopic images. *Eng. Med. Biol. Soc. EMBC 2011 Annu. Int. Conf. IEEE*, 5145–5148.

[16] Toossi, M.T.B., Pourreza, H.R., Zare, H., Sigari, M.H., Layegh, P., Azimi, A. (2013). An effective hair removal algorithm for dermoscopy images. *Skin Res. Technol*., 19, 230–235.

[17] Harris, C., Stephens, M. (1988). A combined corner and edge detector. *Alvey Vis. Conf*., 15, Manchester, UK, 10–5244.

[18] Erkaymaz, O., Ozer, M., Perc, M. (2017). Performance of small-world feedforward neural networks for the diagnosis of diabetes. *Applied Mathematics and Computation*, 15, 311, 22–28.

[19] Grochowski, M., Wąsowicz, M., Mikołajczyk, A., Ficek, M., Kulka, M., Wróbel, M., Jędrzejewska-Szczerska, M. (2019). Machine Learning System For Automated Blood Smear Analysis. *Metrol. Meas. Syst*., 26(1).

[20] Qian, G., Zhang, L. (2018). A simple feedforward convolutional conceptor neural network for classification. *Applied Soft Computing*, 1, 70, 1034–41.

[21] Frery, A.C., Rangayyan, R.M., Azevedo-Marques, P.M., Ramos, H.S. (2018). Evaluation of Deep Feedforward Neural Networks for Classification of Diffuse Lung Diseases. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. *22nd Iberoamerican Congress, CIARP 2017, Valparaíso, Chile*, 10657, 152, Springer.

[22] Mirchandani, M.S., Pendse, M., Rane, P., Vedula, A. (2018). Plant disease detection and classification using image processing and artificial neural networks. *Plant Disease*, 5(06).

[23] Bebis, G., Georgiopoulos, M. (1994). Feed-forward neural networks. *IEEE Potentials*, 13(4), 27–31.

[24] Zhou, Z.H., Wu, J., Tang, W. (2002). Ensembling neural networks: many could be better than all. *Artificial intelligence*, 1, 137(1–2), 239–263.

[25] Williams, P.M. (1995). Bayesian regularization and pruning using a Laplace prior. *Neural computation*, 7(1), 117–143.

[26] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *InIjcai*, 14(2), 1137–1145.

[27] Zhu, W., Zeng, N., Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG proceedings: health care and life sciences*, Baltimore, 14, 19, 67.

[28] Kwasigroch, A., Mikołajczyk, A., Grochowski, M. (2017). Deep neural networks approach to skin lesions classification – A comparative analysis. *Methods Models Autom. Robot. MMAR 2017 22nd Int. Conf. On, IEEE*, 1069–1074.

[29] Kwasigroch, A., Mikołajczyk, A., Grochowski, M. (2017). Deep convolutional neural networks as a decision support tool in medical problems – malignant melanoma case study. *Pol. Control Conf., Springer*, 848–856.

[30] Yu, Z., Jiang, X., Zhou, F., Qin, J., Ni, D., Chen, S., Lei, B., Wang, T. (2018). Melanoma Recognition in Dermoscopy Images via Aggregated Deep Convolutional Features. *IEEE Transactions on Biomedical Engineering*.

[31] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115.