

Research Article

Zhiying Xu*, Fatima Rashid Sheykhahmad, Noradin Ghadimi*, Navid Razmjoojy

Computer-aided diagnosis of skin cancer based on soft computing techniques

<https://doi.org/10.1515/med-2020-0131>

received April 07, 2020; accepted August 12, 2020

Abstract: Skin cancer is a type of disease in which malignant cells are formed in skin tissues. However, skin cancer is a dangerous disease, and an early detection of this disease helps the therapists to cure this disease. In the present research, an automatic computer-aided method is presented for the early diagnosis of skin cancer. After image noise reduction based on median filter in the first stage, a new image segmentation based on the convolutional neural network optimized by satin bowerbird optimization (SBO) has been adopted and its efficiency has been indicated by the confusion matrix. Then, feature extraction is performed to extract the useful information from the segmented image. An optimized feature selection based on the SBO algorithm is also applied to prune excessive information. Finally, a support vector machine classifier is used to categorize the processed image into the following two groups: cancerous and healthy cases. Simulations have been performed of the American Cancer Society database, and the results have been compared with ten different methods from the literature to investigate the performance of the system in terms of accuracy, sensitivity, negative predictive value, specificity, and positive predictive value.

Keywords: skin cancer, image segmentation, feature extraction, feature selection, convolutional neural networks, classification, satin bowerbird optimization

1 Introduction

Skin cancer is a high-prevalence disease in the United States, which occurs in the largest part of the body, i.e., the skin. Skin cancer usually develops on the outer layer of the skin, which may initially appear as a swelling, bump or other parts of the skin. Melanoma is considered one of the most malignant and deadly skin cancers worldwide. Although only about 1% of skin cancer is related to melanoma, it is the main reason of death among other skin cancer diseases [1]. In ref. [2], it has been reported that in 2020 there will be 1,00,350 new cases and 6,850 deaths due to melanoma in the United States. Researchers have found that skin cancer increases the likelihood of other cancers and hence early diagnosis of skin cancer is very important, which can significantly prevent death from this fatal cancer [3]. The two major problems in this regard are as follows:

- (1) In most cases, skin lesions become malignant due to lack of attention to the skin lesions on their body surface or lack of access to experienced dermatologists.
- (2) In many cases, skin lesions are misdiagnosed by physicians because of the high similarity of their characteristics.

Melanoma and Clark, for example, are two very similar skin lesions, except that melanoma is a malignant and deadly cancer and Clark is a benign skin lesion. Therefore, providing a method for the diagnosis of melanoma at an early stage is very useful and valuable [4]. In the last two decades, many studies have been performed on the rapid and accurate diagnosis of melanoma by dermoscopy images with diagnostic accuracy between 70% and 95% [5–9]. Recently, the use of machine vision and artificial intelligence as non-destructive tools in medical applications has been increasing. Besides, the significance of image processing in medical applications helps the physicians and radiologists to reduce the complexity and increase the early detection speed for disease diagnosis. One of the beneficial tools for medical cancer diagnosis is the artificial neural networks (ANNs). The ANN is a widely

* Corresponding author: Zhiying Xu, Yuanpei College, Shaoxing University, Shaoxing, Zhejiang, 312000, China, e-mail: xuzhy@usx.edu.cn

* Corresponding author: Noradin Ghadimi, Young Researchers and Elite Club, Ardabil Branch, Islamic Azad University, Ardabil, Iran, e-mail: noradin1364@gmail.com

Fatima Rashid Sheykhahmad, Navid Razmjoojy: Young Researchers and Elite Club, Ardabil Branch, Islamic Azad University, Ardabil, Iran

used methodology in artificial intelligence, which is inspired by the human brain interactions between synapses and neurons. The ANN method is a good black box-based tool for classification of the nonlinear problems with the least attempts [10–16]. In recent years, a new kind of neural network has been proposed based on deep learning, which is known as convolutional neural network (CNN). The CNN is often employed for image or speech analyses in machine learning. After the application of CNN in image processing, several researchers started to work on using CNN as a tool in medical image processing. For example, Sreelatha et al. [9] proposed a technique for melanoma diagnosis from dermoscopy images by the Gradient and Feature Adaptive Contour (GFAC) model. The study also used pre-processing and noise reduction to make the process faster and more accurate. The proposed GFAC model is a noise-free method. The method was then applied to the PH2 dataset. The method was then compared with different methods in the literature to indicate its effectiveness. Hekler et al. [17] proposed another skin cancer classification method based on artificial intelligence. The study applied dermoscopy images to classify it into five divisions. The method was based on the deep learning method to train a single CNN. The method was applied in a collection of 13 German university hospitals. The final results of the method were good.

Tschandl et al. [18] presented a classification method using CNN by training 7,895 dermoscopy images. The method was then performed to a set of 2,072 test cases, and the results were validated by some experts.

Tan et al. [19] proposed an optimized method for skin cancer detection based on an improved version of particle swarm optimization (PSO) algorithm. The PSO was used to optimize the feature extraction of the dermoscopy images. The method was then applied to

some different benchmarks to show the PSO superiority in terms of performance toward some well-known methods.

The satin bowerbird optimization (SBO) algorithm is a new meta-heuristic algorithm that was first proposed by Musavi et al. [20]. The SBO algorithm was inspired by the method of making special nests by the male bowerbird to attract the females. In this study, an optimized comprehensive methodology is presented for skin cancer diagnosis in the dermoscopy images. The main contribution of the method is to use the CNN optimized by the SBO, which increases the accuracy of the network compared with the classic gradient descent method. The flowchart of the presented system is shown in Figure 1.

As shown in Figure 1, after receiving the input dermoscopy image, a preprocessing including median filtering has been applied to enhance the image quality by eliminating a part of the existing noises. In the next step, a new image segmentation based on the aforementioned optimized CNN has been applied to segment the skin-like area. In the next step, some features of the segmented area have been extracted for reducing the data complexity and easier classification. Here, an optimized feature selection has been used to select the useful features. Finally, after extracting the useful data from the image, a support vector machine (SVM) has been used for classifying the images into the following two groups: cancerous and healthy.

2 Materials and methods

All the required materials and methods that are used for skin cancer diagnosis in this research are as follows.

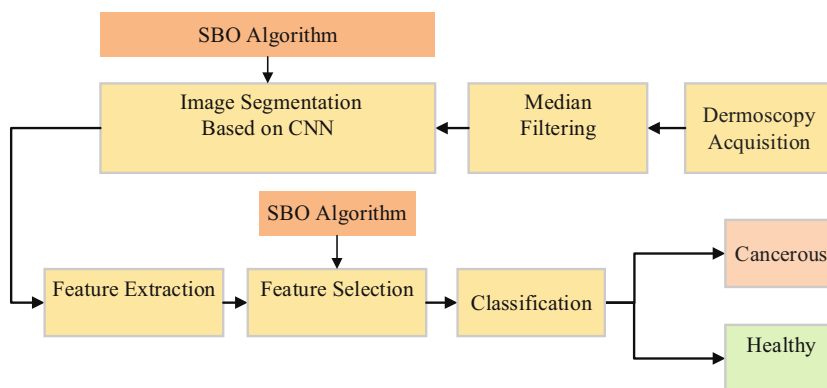


Figure 1: The flowchart of the proposed skin cancer diagnosis system.

2.1 CNN

The CNN is a mostly used method based on deep learning in which multiple layers are trained in new ways. These networks are a kind of multilayer neural network designed for two-dimensional data such as image. Different parts of the image have been implemented as inputs to the neural network layers, which are hierarchical, and each layer is extracted by applying digital filters. Convolution layers convolve the image by using different kernels. The convolution process has the following three important features:

1. Sharing weights for the identical feature to decrease the parameter number.
2. Local learning connection.
3. Invariance according to the position of the object.

By considering the advantages provided by the convolution process, some research studies in the literature used it as an alternative to fully integrated layers to speed up the learning process. In general, the CNN contains three layers, including the convolutional layer, the pooling layer, and the fully connected layer, where different layers perform various tasks. The CNN architecture has several convolution layers that convolve the input image with filters whose output coefficients can be trained. These filters move over the image. Each separate filter creates a set of features, which consequently gives us the use of n filters of the set of features. Each filter has weights called w , which can be trained as filter coefficients. During the training, the network is sequenced, and the output obtained from the convolution operation is computed with a number that we consider to be a bias number and is stored in the feature plane. The pooling layer is usually followed by the convolution layer, which can be used to reduce the dimension of network features and parameters. Like the convolution layer, the pooling layer also uses adjacent pixels to compute. There are different strategies for pooling, but the most common way is to use the maximum pooling method. In the pooling method,

blocks of the size 2×2 are considered and by moving these blocks over the image through 4 pixels, their maximum value is selected and transferred to the next layer that the number of features has been kept but its size has been reduced. Therefore, in the final step, we use fully connected layers, which actually convert the 2D feature into a one-dimensional vector. This layer accounts for 77% of the CNN parameters. Figure 2 shows the arrangement of the employed CNN.

The CNN has two levels of training: the feed-forward level and the backpropagation (BP) level. In the feed-forward level, the original image is injected to the network by applying a dot multiplication with the parameters of the neurons to provide convolutional operations on the layers.

Afterward, the output of the network is evaluated. For network training (set the network parameters), the output is employed to evaluate the network error rate. Then, the error rate of the network based on BP is calculated. In this stage, the chain rule is employed for evaluating the gradient of the parameters and the parameters change based on the error impact on the network. After updating the parameters, feed-forward is performed until the correct number of training. Here, local feature extraction is adopted for achieving the regional features of the input image. The learning is applied here to achieve some number of kernel matrices to extract the main characteristics of the dermoscopy skin cancer images. The present research adopts the BP technique for optimizing the weights of network connections. Sliding window is employed as a vector for implementing the convolution so that the weights and the dot product have been added up. For applying the activation function, rectified linear unit (ReLU) is used, i.e., $f(x) = \max(x, 0)$ [21]. For decreasing the output scale, Max pooling is used. In the present study, the subsequent layer of the sliding grid is considered as the highest value. After initializing the CNN, an optimization technique is required for error minimization of the actual output and the estimated output by varying the internal weights. Here, the BP algorithm is

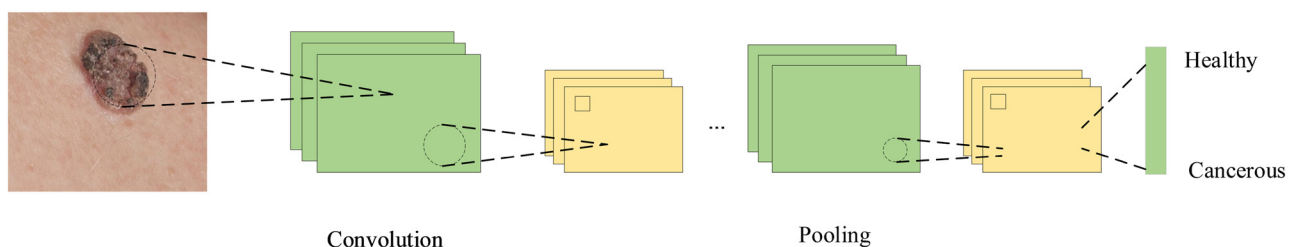


Figure 2: The structure of a CNN.

used for this purpose [6,22]. The main part of BP is to adopt the gradient descent algorithm for minimizing. The gradient descent is a minimization procedure on the cross-entropy loss [23]. The considered cost function for the explained cases is as follows:

$$L = \sum_{j=1}^N \sum_{i=1}^M -d_j^{(i)} \log z_j^{(i)}, \quad (1)$$

where $d_j = (0, \dots, 0, \underbrace{1, \dots, 1}_k, 0, \dots, 0)$ represents a vector with the desired output and z_j determines the softmax function of the m th class by the following equation:

$$z_j^{(i)} = \frac{e^{f_j}}{\sum_{i=1}^M e^{f_i}}, \quad (2)$$

where N represents the sample quantity.

Therefore, by modifying the cost function based on the weight penalty with a coefficient θ ,

$$L = \sum_{j=1}^N \sum_{i=1}^M -d_j^{(i)} \log z_j^{(i)} + \frac{1}{2} \theta \sum_K \sum_L W_{K,L}^2, \quad (3)$$

where W_k describes the weight of connection, L represents the total number of layer connections, and K is the number of layer l connections.

Although CNN has optimal arrangement for classification, most of the layouts are achieved experimentally that reduces its accuracy.

In recent years, some methods have been introduced to modify them based on bio-inspired techniques [24,25]. Metaheuristics are some kinds of optimization methods that have been inspired by different phenomena and have better efficiency in finding a global optimum value in less time. Recently, several types of metaheuristics have been introduced [26]. For instance, genetic algorithm [27] that simulates Darwin's principles of selection to find the optimal formula, grass fibrous root optimization algorithm that simulates the fibrous root behavior [28], butterfly optimization algorithm [29] that simulates the butterflies' migration from the cold areas to the warmer places in the cold seasons, and teaching-learning-based optimization algorithm [30] that simulates the relations between teaching and learning. In this study, another new metaheuristic method, called SBO algorithm, is employed for improving the CNN efficiency.

2.2 SBO algorithm

The bowerbirds are some kinds of interesting Australian birds which have a unique way for mating. In the mating season, the adult male bower builds his special bower by

using his leathers, sparkling objects, flowers, fruits, and branches along with dramatic gestures to attract females. These parameters (the beauty of the bower and the male dramatic gestures) attract the females to the bower. These parameters form the main structure of the SBO algorithm [20]. The steps of the SBO algorithm are briefly described in the following.

2.2.1 Initializing

Like any population-based metaheuristic, the SBO algorithm was initialized with a set of random population. The population in the SBO is the positions of the bowers such that each of them is an n dimensional vector of parameters of the problem that should be optimized. The bower parameters form the variables of the optimization problem. The bower's attractiveness is a combination of these parameters. The initial population of the SBO can be defined as follows:

$$W_h = (w_1, w_2, \dots, w_m), \quad (4)$$

where W_h describes the h th solution and (w_1, w_2, \dots, w_m) is the population of the solution.

The probability of fitness evaluation is mentioned as the attractiveness of the bower. Based on the bower probability theory, the male stain is selected by a female bowerbird. Similarly, after selecting a bower by the assigned probability of the male, it mimics the bower construction. Therefore

$$\text{Prob}_i = \frac{\text{fit}_i}{\sum_{n=1}^{NB} \text{fit}_n}, \quad (5)$$

$$\text{fit}_i = \begin{cases} \frac{1}{1 + f(x_i)}, & f(x_i) \geq 0 \\ 1 + |f(x_i)|, & f(x_i) < 0, \end{cases} \quad (6)$$

where $f(x_i)$ describes the cost function value of the i th position.

2.2.2 Elitism

The best solution of each stage is considered as elitism. Basically, each male satin bowerbird constructs his bower based on his natural instinct. Each satin bowerbird uses its instinct on its similarity of every other bird in the mating season to construct and beautification of his bower. In other words, although each male satin bowerbird uses material for decorating its bower, it benefits its experience as a key factor to attract lot of

attention to his specific bower. In a word, we can say that experience has a significant effect on both construction of bower and dramatic gestures, which makes elder males to have more potential to attract more attention to their bower. In the algorithm, the position of the best bower is considered as iteration's elite which has the capability of affecting other positions.

2.2.3 Position updating

All of the variations at the bowers in every cycle of algorithm are performed based on the following equation:

$$W_{hj}^{\text{new}} = W_{hj}^{\text{old}} + \lambda_j \left(\frac{W_{ij} + W_{\text{elite},j}}{2} \right) - W_{hj}^{\text{old}}, \quad (7)$$

where W_i defines the target solution in the current iteration, the term i is achieved based on the roulette wheel mechanism, i.e., the better solution will have more chance to be selected as W_i , W_{hj} describes the j th member of W_h , W_{elite} describes the position of the elite stored in the cycles, λ_j represents the attraction power in the target bower (solution) that is achieved by the following equation:

$$\lambda_j = \frac{\alpha}{1 + o_i}, \quad (8)$$

where the term α stands for the greatest size of step using the target solution and o_i describes the probability obtained by Prob_i .

2.2.4 Mutation

While building the bower, the male satin bowerbird might be attacked or even totally ignored by other animals. During the bower construction, stronger males destroy the weaker males' bower or steal their materials. Therefore, arbitrary changes have been adopted at the end of each cycle with a specific probability. This modification is implemented to W_{hj} with a specific probability. A normal distribution (L) has been adopted with variance α^2 and average of W_{hj}^{old} as follows:

$$w_{hj}^{\text{new}} \sim L(w_{hj}^{\text{old}}, \alpha^2), \quad (9)$$

$$L(w_{hj}^{\text{old}}, \alpha^2) = w_{hj}^{\text{old}} + (\alpha \times L(0, 1)). \quad (10)$$

The proportion of space width is represented as a value of α that is evaluated in equation (16).

$$\alpha = \gamma \times (\text{Var}_{\text{max}} - \text{Var}_{\text{min}}), \quad (11)$$

where γ describes the variance ratio among lower and upper ranges, and var_{max} and var_{min} represent the upper and the lower limits assigned to the variable.

The population that is achieved after modifications at the end of every cycle and the old population is evaluated. The populations are then combined and sorted after evaluating, and the new population has been generated. Once the termination condition is satisfied, the algorithm is stopped.

2.3 Median filtering

Medical images are often damaged by noise when acquired and passed on. The purpose of image noise reduction methods is to eliminate such noise while maintaining the essential features of the image as much as possible. Dermoscopy imaging is a widely used medical imaging process, as it is economically, relatively safe, and adaptable. However, one of its major disadvantages is the poor quality of images due to spot noise. Spot noise is a disgrace, as it undermines image quality and affects single interpretation and recognition operations. Consequently, point refining is a key stage for the function's extraction, analyzing and recognizing medical images. There are several useful methods for noise removal from the medical images. Median filter due to its specificity has the most application in medical image noise removal [31,32]. The main idea behind median filtering is to present an $m \times n$ neighborhood to assemble all neighborhoods in the ascending order, select the median value of the ordered numbers, and replace the central pixel. This strategy is modeled by the following equation:

$$y_{(m,n)} = \text{median}\{x_{(i,j)}, (i, j) \in C\}, \quad (12)$$

where C represents the centered neighborhood around location (m, n) of the image.

In this research, a median filter is adopted for digital noise removal in the skin cancer dermoscopy image, in which a filter mask with a size of 3×3 is employed.

3 Medical skin cancer diagnosis based on optimized CNN

The main contribution of this study is to adopt the SBO algorithm for optimal justification of the number of

hyperparameters in CNN for the skin cancer diagnosis application. To avoid system error, the minimum range is considered to be 2 and the maximum range is considered equal to the size of sliding window. The reason for selecting 2 as the minimum range is that the allowed minimum value for the maximum pooling is such that there is no lower size. It is important to note that the inequality constraint of this problem is that the value of the sliding window should be smaller than the input data. In this study, the population for bowerbirds is considered 200. Table 1 indicates the architecture of the suggested CNN.

As mentioned earlier, the cost function is the half-value precision of CNN. After initializing, the position of each bowerbird has been updated based on the SBO algorithm. Once the stopping criteria is reached, the optimization process will be stopped. In this study, the weights and the biases of the CNN are chosen for optimization, i.e.,

$$W = \{w_1, w_2, \dots, w_p\}, \quad (13)$$

$$A = \{a_1, a_2, \dots, a_A\}, \quad (14)$$

$$w_n = \{w_{1n}, w_{2n}, \dots, w_{Ln}\}, \quad (15)$$

$$\begin{aligned} b_n &= \{b_{1n}, b_{2n}, \dots, b_{Ln}\} \\ l &= 1, 2, \dots, L \end{aligned} \quad (16)$$

$$n = 1, 2, \dots, N,$$

where N describes the overall number of optimization individuals, L represents the overall number of layers, l determines the layer index, n represents the number of bowerbirds, and w_{in} represents the value of the weight in the i th layer. The optimized configuration of the SBO-based CNN is shown in Figure 3.

Table 1: Architecture of the suggested CNN

Layer #	Layer name	Properties
1	Input layer	Input image patch size: $256 \times 256 \times 3$
2	Convolutional	Blocks of the size: 11×11
3	ReLU	—
4	Max pooling	Pool size 2×2
5	Convolutional	Blocks of the size: 7×7
6	ReLU	—
7	Max pooling	Pool size 2×2
8	Convolutional	Blocks of the size: 3×3
9	ReLU	—
10	Dropout	Dropout ratio: 0.6
11	Fully connected	1×256
12	Softmax	—

The error between the desired output and the network output is calculated by the following equation:

$$E = \frac{1}{Ns} \sum_{i=1}^{Ns} \sum_{j=1}^k (d_{ji} - o_{ji})^2, \quad (17)$$

where k and Ns represent the output layers quantity and training samples, respectively, and o_{ji} and d_{ji} describe the CNN output and the desired output.

Since the gradient descent has been trapped into the local minimum in some cases, the SBO algorithm has been adopted [33–35]. Figure 4 shows the results of cancer diagnosis for some examples by the SBO-based CNN.

The confusion matrix for the system is 2×2 as the American Cancer Society (ACS) database is selected and divided into two classes: the first 40% plus the second 60% with noise (ACS1) and the second 60% plus the first 40% with noise (ACS2). This matrix has been achieved by adding two confusion matrices on the aforementioned two categories of the dataset. We note a good percentage of correctly classified ACS1 and ACS2. The number of correctly classified ACS2 at a first glance looks quite lower than ACS1, which is due to the small number of high-quality images in ACS2 compared with that of ACS1 (Table 2).

4 Feature extraction and selection based on SBO

The process of extracting more precise information from the cancer area is called feature extraction, i.e., the purpose of feature extraction in this step is to use this trend as a method for skin cancer detection after segmenting the skin cancer area. For more clarification, the original image is a raw datum with enormous information. This reason increases the complexity of image processing and is time-consuming. A good method here is to use feature extraction for simplifying the process. Several features have been introduced for the extraction of image features. In this study, statistical features, geometric features, and texture features are adopted, which are listed in Table 3.

Nevertheless, some of the explained features have more effect and some others have less effect with low information from the image. In this study, an optimized procedure is used for better selection of the useful features from the image. The optimal feature selection here is performed by SBO. The fitness function of the feature selection is considered as follows:

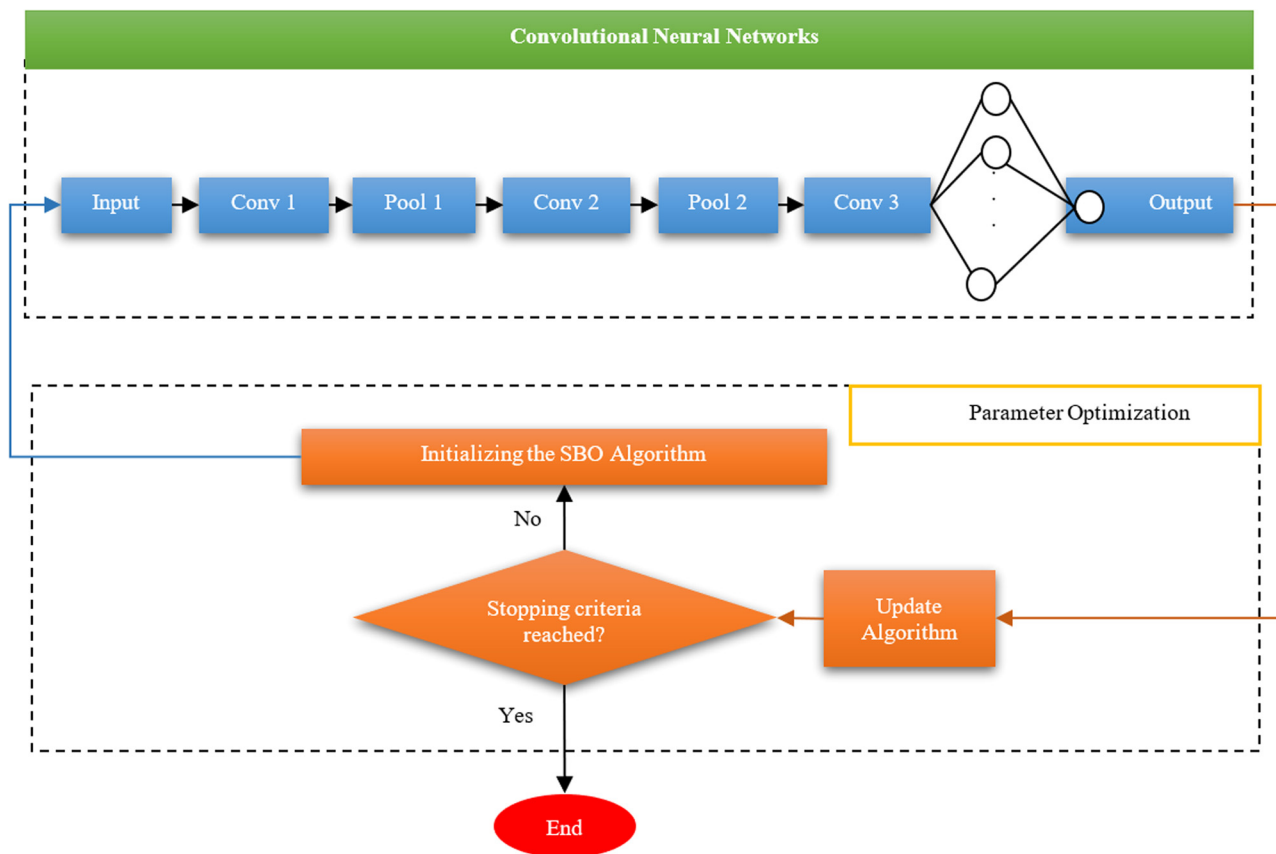


Figure 3: The optimized structure of the SBO-based CNN.

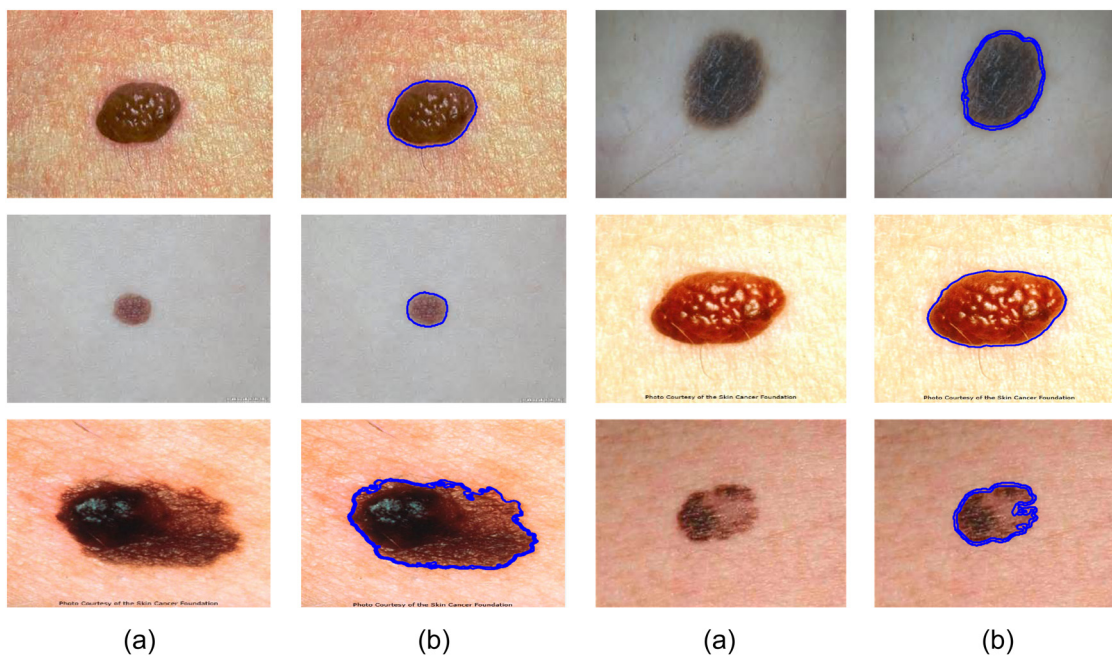


Figure 4: The results of cancer diagnosis for some examples by the SBO-based CNN. (a) before process, (b) after process.

Table 2: Confusion matrix for the system

	ACS1	ACS2
ACS1	570	9
ACS2	7	320

$$\text{Fitness} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{\sqrt{((\text{TN} + \text{FP}) \times (\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FN}))}} \quad (18)$$

where the main parts of the function are false positive (FP), false negative (FN), true positive (TP), and true negative (TN). The main idea based on the proposed method is to minimize the fitness value in equation (18) by optimal selection of the features. After optimal feature extraction of the medical images, a classification method is required to determine the cancerous and the healthy cases. To understand how the features can help easier classification, some significant example features for both cancerous and healthy images are given in Table 4.

5 Image classification

In this study, SVM has been employed for feature classification. The method includes a set of points in the n -dimensional space, which indicates and sorts the

class boundaries and can be changed by replacing with one of two cases. SVM selects the optimal decision surface by the following formula:

$$y = \text{sgn} \left(\sum_{i=1}^N y_i \alpha_i K(x, x_i) + b \right), \quad (19)$$

where x_i represents the training set vector number i , x describes a d -dimensional test set vector, N defines the training set numbers, $K(x, x_i)$ describes a kernel function, y describes a class label in the range $[-1, 1]$, and $\alpha = \{\alpha_1 \dots \alpha_N\}$, and b determines the model parameters. Due to the high potential of SVM, it is adopted for classifying the skin cancer dermoscopy image into two categories of cancerous and healthy groups.

6 The database

There are several datasets for skin cancer detection and the validation of different methods. This research adopts the ACS dataset. This dataset contains 68 pairs of XLM and TLM images that are obtained by the same Nevoscope device. All the images have been resized into 256×256 pixels to reduce the system computational complexity [36]. Figure 5 shows some examples of the used dataset from ACS.

Table 3: Adopted features in this study

Parameter	Equation	Parameter	Equation
Perimeter	$\sum_{i=1}^M \sum_{j=1}^N B_p(i, j)$	Correlation	$\sum_{i=1}^M \sum_{j=1}^N \frac{\rho(i, j) - \mu_r \mu_c}{\sigma_r \sigma_c}$
Area	$\sum_{i=1}^M \sum_{j=1}^N p(i, j)$	Mean	$\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N p(i, j)$
Solidity	Area/Convex area	Entropy	$-\sum_{i=1}^M \sum_{j=1}^N p(i, j) \log p(i, j)$
Elongation	$\frac{2\sqrt{\text{Area}}}{a\sqrt{b}}$	Variance	$\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (p(i, j) - \mu)^2$
Rectangularity	$\frac{\text{Area}}{b \times a}$	Standard deviation	$\text{variance}^{\frac{1}{2}}$
Irregularity index	$4\pi \times \text{Area}/\text{Perimeter}^2$	Invariant moments	$\varphi_1 = \eta_{20} + \eta_{02}$ $\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$ $\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \mu_{03})^2$
Form factor	$\frac{\text{Area}}{a^2}$	Energy	$\sum_{i=1}^M \sum_{j=1}^N p^2(i, j)$
$E_{\text{centricity}}$	$2a^{-1}(a^2 - b^2)^{0.5}$	Homogeneity	$\sum_{i=1}^M \sum_{j=1}^N \frac{p(i, j)}{1 + i - j }$
Contrast	$\sum_{i=1}^M \sum_{j=1}^N p^2(i, j)$		

B_p describes the length of external side for the boundary pixel, MN represents the image size, $p(i, j)$ defines the value of the pixels intensity at location (i, j) , a is the major axis and b is the minor axis, μ describes the mean value, and σ stands for the standard deviation.

Table 4: Some significant example features for both cancerous and healthy images

Parameter	Cancerous				Healthy			
	#1	#2	#3	#4	#1	#2	#3	#4
Correlation	0.9915	0.9886	0.9872	0.9505	0.9905	0.9898	0.9864	0.9779
Area	0.3847	0.4492	0.5133	0.9389	0.3988	0.4266	0.4996	0.6086
Energy	0.6419	0.6348	0.5746	0.6185	0.6538	0.6320	0.6193	0.6510
Form factor	0.1205	0.1811	0.2253	0.6428	0.1096	0.1328	0.1973	0.1308
Eccentricity	0.1215	0.1886	0.1872	0.1505	0.1205	0.1798	0.1864	0.2779
Homogeneity	0.9974	0.9964	0.9953	0.9843	0.9972	0.9968	0.9955	0.9934
Contrast	0.1480	0.2018	0.2635	0.8817	0.1591	0.1820	0.2496	0.3704

7 Simulation results

This section validates and discusses about the results of the proposed skin cancer diagnosis system. The method has been simulated based on MATLAB software version R2018a with a 32 GB Ram, Core i7Intel processor and two SLI GeForce Titan GPUs. The method is performed for the ACS skin cancer dataset for performance analysis of the system. The procedure of the method is given as follows:

1. Apply median filtering for noise elimination of the input dermoscopy image.
2. Apply optimized CNN based on the SBO algorithm for cancer area segmentation.
3. Extract important features from the segmented image.
4. Apply feature selection based on the SBO algorithm.
5. Classify cancerous and healthy cases based on the SVM classification methodology.

For the proposed SVM classifier, 75% of data is considered for training and 25% is used for validation and testifying the method. Training stage for the network is considered by 8,000 iterations. The training step is repeated 15 times for achieving a robust analysis. The analysis is applied based on five performance indexes, including accuracy, specificity, positive predictive value (PPV), negative predictive value (NPV), and sensitivity, which are formulated as follows:

$$\text{Accuracy} = \frac{\text{correctly detected cases}}{\text{total cases}}, \quad (20)$$

$$\text{Specificity} = \frac{\text{correctly detected healthy skin cases}}{\text{total healthy skin cases}}, \quad (21)$$

$$\text{PPV} = \frac{\text{correctly detected skin cancer cases}}{\text{detected skin cancer cases}}, \quad (22)$$

**Figure 5:** Some examples of the used dataset from ACS.

Table 5: Efficiency analysis of the proposed method compared with the other state-of-art techniques

Method	Performance metric				
	Accuracy	Sensitivity	Specificity	NPV	PPV
LIN [43]	0.89	0.92	0.90	0.93	0.82
AlexNet [41]	0.83	0.84	0.62	0.85	0.67
VGG-16[40]	0.87	0.91	0.87	0.91	0.79
Spotmole [39]	0.69	0.84	0.60	0.86	0.59
Ordinary CNN	0.82	0.82	0.80	0.87	0.76
ResNet-50 [38]	0.83	0.87	0.80	0.84	0.72
ResNet-101 [38]	0.86	0.85	0.78	0.90	0.76
Inception-v3 [42]	0.85	0.86	0.66	0.73	0.64
MED-NODE texture descriptor [37]	0.79	0.64	0.87	0.80	0.77
MED-NODE color descriptor [37]	0.75	0.77	0.74	0.84	0.66
Proposed method	0.95	0.95	0.92	0.96	0.87

$$NPV = \frac{\text{correctly detected healthy skin cases}}{\text{detected healthy skin cases}}, \quad (23)$$

$$\begin{aligned} \text{Sensitivity} \\ = \frac{\text{correctly detected skin cancer cases}}{\text{Total skin cancer cases}}. \end{aligned} \quad (24)$$

The final results are compared with ten different methods from the literature in terms of efficiency that contain a method [37] using semi-supervised system, some methods based on CNN, ResNet [38], a commercial tool [39], VGG-16[40], AlexNet [41], Inception-v3 [42], and LIN [43]. Table 5 indicates the efficiency results of the proposed method compared with the other state of art techniques.

Figure 6 shows the radar plot of the classification rate. As shown in Table 2 and Figure 6, the presented procedure has the highest accuracy compared with the other methods from the literature.

8 Conclusions

In the present study, an automatic computer-aided methodology was proposed for the skin cancer detection from dermoscopy images. To achieve higher accuracy for the diagnosis, the input images were first denoized based on a median filter. Afterward, a new optimized method was

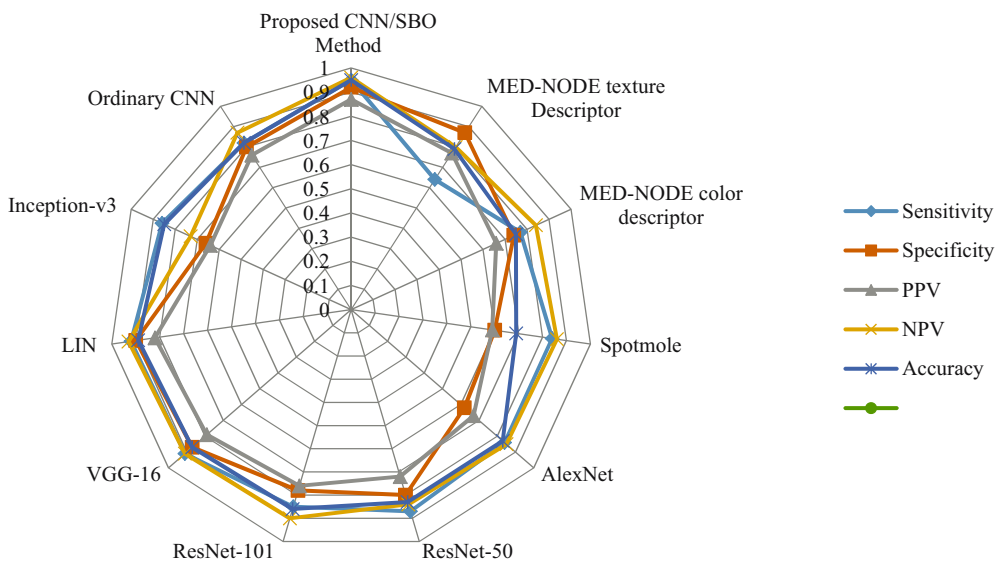


Figure 6: The Radar plot of the classification rate of skin cancer classification based on the presented technique and the other methods from the literature and its efficiency is indicated by the confusion matrix.

presented for cancer area segmentation using a developed CNN based on the SBO algorithm to classify the cancerous region from the background. Then, numerous features were extracted from the segmented images. To simplify the classification process, an optimal method based on the SBO algorithm was adopted for optimal feature selection and decreasing the order of features. To do the final classification, the extracted features from the images were classified based on an optimized SVM classifier using the SBO algorithm. The design was then applied to the ACS database and was compared with ten state-of-art methods to indicate the system efficiency. Final results showed that by analyzing accuracy, specificity, NPV, sensitivity, and PPV, the presented technique has better performance than the compared methods. The future work will be about using an improved model of the proposed method to achieve good results for breast cancer detection.

Conflict of interest: Authors state no conflict of interest.

References

- [1] Razmjoo N, Ashourian M, Karimifard M, Estrela V, Loschi H, do Nascimento D, et al. Computer-aided diagnosis of skin cancer: a review. *Curr Med Imaging Rev.* 2020.
- [2] Society AC. *Cancer Facts and Figures 2020.* Atlanta: American Cancer Society; 2020; available from: <https://www.aimatmelanoma.org/about-melanoma/melanoma-stats-facts-and-figures/>.
- [3] Zhang L, Gao HJ, Zhang J, Badami B. Optimization of the convolutional neural networks for automatic detection of skin cancer. *Open Med.* 2020;15(1):27–37.
- [4] Narayanamurthy V, Padmapriya P, Noorasafin A, Pooja B, Hema K, Nithyakalyani K, et al. Skin cancer detection using non-invasive techniques. *RSC Adv.* 2018;8(49):28095–130.
- [5] Razmjoo N, Mousavi BS, Soleymani F, Khotbesara MH. A computer-aided diagnosis system for malignant melanomas. *Neural Comput Appl.* 2013;23(7–8):2059–71.
- [6] Razmjoo N, Sheykhahmad FR, Ghadimi N. A hybrid neural network–world cup optimization algorithm for melanoma detection. *Open Med.* 2018;13(1):9–16.
- [7] Celebi ME, Codella N, Halpern A, Shen D. Guest editorial skin lesion image analysis for melanoma detection. *IEEE J Biomed Health Inform.* 2019;23(2):479–80.
- [8] Mukherjee S, Adhikari A, Roy M. Malignant Melanoma Detection Using Multi Layer Perceptron with Optimized Network Parameter Selection by PSO. In: Mandal J, Sinha D, Bandopadhyay J, editors. *Contemporary advances in innovative and applicable information technology. Advances in intelligent systems and computing, vol. 812.* Singapore: Springer.
- [9] Sreelatha T, Subramanyam M, Prasad MG. Early detection of skin cancer using melanoma segmentation technique. *J Med Syst.* 2019;43(7):190.
- [10] Moallem P, Razmjoo N. A multi layer perceptron neural network trained by invasive weed optimization for potato color image segmentation. *Trends Appl Sci Res.* 2012;7(6):445.
- [11] Fan X, Sun H, Yuan Z, Li Z, Shi R, Ghadimi N. High voltage gain DC/DC converter using coupled inductor and VM techniques. *IEEE Access.* 2020 Jun 16;8:131975–87.
- [12] Ye H, Jin G, Fei W, Ghadimi N. High step-up interleaved dc/dc converter with high efficiency. *Energy Sources Part A Recovery Util Environ Eff.* 2020 Jan 24;1–20.
- [13] Such FP, Madhavan V, Conti E, Lehman J, Stanley KO, Clune J. Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. *arXiv preprint arXiv:171206567.* 2017.
- [14] Yuan Z, Wang W, Wang H, Ghadimi N. Probabilistic decomposition-based security constrained transmission expansion planning incorporating distributed series reactor. *IET Generat Transm Distrib.* 2020;14.17:3478–87.
- [15] Hagh MT, Ghadimi N. Radial basis neural network based islanding detection in distributed generation. *Int J Eng Trans A Basic.* 2014;27:1061–70.
- [16] Leng H, Li X, Zhu J, Tang H, Zhang Z, Ghadimi N. A new wind power prediction method based on ridgelet transforms, hybrid feature selection and closed-loop forecasting. *Adv Eng Inform.* 2018 Apr 1;36:20–30.
- [17] Hekler A, Utikal JS, Enk AH, Hauschild A, Weichenthal M, Maron RC, et al. Superior skin cancer classification by the combination of human and artificial intelligence. *Eur J Cancer.* 2019;120:114–21.
- [18] Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. *JAMA Dermatol.* 2019;155(1):58–65.
- [19] Tan TY, Zhang L, Neoh SC, Lim CP. Intelligent skin cancer detection using enhanced particle swarm optimization. *Knowl Syst.* 2018;158:118–35.
- [20] Moosavi SHS, Bardsiri VK. Satin bowerbird optimizer: a new optimization algorithm to optimize ANFIS for software development effort estimation. *Eng Appl Artif Intell.* 2017;60:1–15.
- [21] Koehler F, Risteski A. Representational Power of ReLU Networks and Polynomial Kernels: Beyond Worst-Case Analysis. *arXiv preprint arXiv:180511405.* 2018.
- [22] Meng Q, Liu T, Su C, Niu H, Hou Z, Ghadimi N. A single-phase transformer-less grid-tied inverter based on switched capacitor for PV application. *J Control Autom Electr Syst.* 2020 Feb 1;31(1):257–70.
- [23] Van Merriënboer B, Bahdanau D, Dumoulin V, Serdyuk D, Warde-Farley D, Chorowski J, et al. *Blocks and fuel: Frameworks for deep learning.* arXiv preprint arXiv:150600619. 2015.
- [24] Martens J, Sutskever I. Learning recurrent neural networks with hessian-free optimization. *Proceedings of the 28th International Conference on Machine Learning (ICML-11).* Bellevue, WA, USA; 2011.
- [25] Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. *Advances in neural information processing systems.* 2007. p. 153–60.
- [26] Aghajani G, Ghadimi N. Multi-objective energy management in a micro-grid. *Energy Rep.* 2018;4:218–25.
- [27] Ghadimi N. Genetically tuning of lead-lag controller in order to control of fuel cell voltage. *Sci Res Essays.* 2012;7(43):3695–701.

- [28] Akkar HA, Mahdi FR. Grass fibrous root optimization algorithm. *Int J Intell Syst Appl.* 2017;11(6):15.
- [29] Arora S, Singh S. Butterfly optimization algorithm: a novel approach for global optimization. *Soft Comput.* 2019;23(3):715–34.
- [30] Kaveh A, Bakhshpoori T. Teaching-Learning-Based Optimization Algorithm. In: *Metaheuristics: Outlines, MATLAB Codes and Examples.* Springer, Cham; 2019. p. 41–9.
- [31] Anoraganingrum D. Cell segmentation with median filter and mathematical morphology operation. *Proceedings of the 10th International Conference on Image Analysis and Processing, Venice, Italy; 1999.* p. 1043–6.
- [32] Loupas T, McDicken W, Allan PL. An adaptive weighted median filter for speckle suppression in medical ultrasonic images. *IEEE Trans Circuits Syst.* 1989;36(1):129–35.
- [33] Zhang L, Suganthan PN. A survey of randomized algorithms for training neural networks. *Inf Sci.* 2016;364:146–55.
- [34] Jaddi NS, Abdullah S. Optimization of neural network using kidney-inspired algorithm with control of filtration rate and chaotic map for real-world rainfall forecasting. *Eng Appl Artif Intell.* 2018;67:246–59.
- [35] Emary E, Zawbaa HM, Grosan C. Experienced gray wolf optimization through reinforcement learning and neural networks. *IEEE Trans Neural Netw Learn Syst.* 2018;29(3):681–94.
- [36] ACS. Skin Cancer Database. USA: ACS; 2019 [cited 2019]; available from: <https://www.cancer.org/>.
- [37] Giotis I, Molders N, Land S, Biehl M, Jonkman MF, Petkov N. MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert Syst Appl.* 2015;42(19):6578–85.
- [38] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV; 2016.* p. 770–8.
- [39] Munteanu C, Coocelea S. Spotmole – melanoma control system. 2009.
- [40] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:14091556.* 2014.
- [41] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems.* 2012.
- [42] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV; 2016.* p. 2818–26.
- [43] Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors.* 2018;18(2):556.