

Research Article

Long Zhang[#], Hong Jie Gao^{*}, Jianhua Zhang[#], Benjamin Badami

Optimization of the Convolutional Neural Networks for Automatic Detection of Skin Cancer

<https://doi.org/10.1515/med-2020-0006>

received February 27, 2019; accepted October 3, 2019

Abstract: Convolutional neural networks (CNNs) are a branch of deep learning which have been turned into one of the popular methods in different applications, especially medical imaging. One of the significant applications in this category is to help specialists make an early detection of skin cancer in dermoscopy and can reduce mortality rate. However, there are a lot of reasons that affect system diagnosis accuracy. In recent years, the utilization of computer-aided technology for this purpose has been turned into an interesting category for scientists. In this research, a meta-heuristic optimized CNN classifier is applied for pre-trained network models for visual datasets with the purpose of classifying skin cancer images. However there are different methods about optimizing the learning step of neural networks, and there are few studies about the deep learning based neural networks and their applications. In the present work, a new approach based on whale optimization algorithm is utilized for optimizing the weight and biases in the CNN models. The new method is then compared with 10 popular classifiers on two skin cancer datasets including DermIS Digital Database Dermquest Database. Experimental results show that the use of this optimized method performs with better accuracy than other classification methods.

Keywords: Computer-aided diagnosis; Skin cancer detection; Convolutional neural networks; Whale optimization algorithm; Image segmentation

***Corresponding author: Hong Jie Gao**, Department of medical equipment, People's hospital of Zhengzhou University, Zhengzhou, 450001, China, E-mail:532136292@qq.com

Long Zhang, Department of medical equipment, People's hospital of Zhengzhou University, Zhengzhou, 450001, China

Jianhua Zhang, Institute of Medical Engineering Technology and Data Mining, Zhengzhou University, Zhengzhou, 450001, China

Benjamin Badami, University of Georgia, Athens, USA

1 Introduction

Skin cancer involves abnormal changes in the outer layer of the skin. This cancer is the most prevalent cancer in the world and contains about 75% of the world's cancer. Although most people with skin cancer are healed, it is still a major concern due to its high prevalence [1]. Most skin cancers grow only locally and invade adjacent tissues, but some of them, particularly melanoma (cancer of the pigment cells), which is the rarest type of skin cancer, may spread through the circulatory system or lymphatic system and reach the farthest points of the body [2]. Melanoma forms the highest percentage of probability among different types of skin cancer [3-7]. On average, 4740 males and 2490 females died in 2019 due to melanoma [8]. Melanoma is more prevalent in some areas, especially in western regions and countries. According to the findings, the diagnosis of melanoma in the initial stages can significantly reduce the mortality due to this cancer; but since the diagnosis of this disease at an early stage, even by specialists, is difficult, it will be very helpful to provide a method to early diagnosis of the melanoma or skin cancer [9-12].

In recent years, with the advancement of technology, particularly artificial intelligence, suitable methods have been developed for this issue. In the meantime, image processing techniques are progressing as successful methods [9, 13-16]. The application of image processing and computer vision for automatically identifying the patterns like cancer from images reduces human errors and increases the speed of detection. In addition, the importance of medical image processing can be considered as it helps physicians and radiologists to more easily diagnose the disease, thus protecting the patient against irreparable risks that will come about. Artificial Neural Networks (ANNs) are one of the popular methods used in image processing. ANN is inspired by the intricate structure of the human brain, in which millions of neurons (cells) communicate with each other (synapse) to solve problems or store information. These networks are a collection of different models that are proposed by mathematicians and engineers to simulate a part of brain function [15-21].

The system is made up of a large number of extravagant processing elements called neurons that work together to solve a problem. Learning in natural systems occurs adaptively; this means that there is a change in the synapse as a result of learning. Recently, important developments are proposed based on new kinds of neural networks for analyzing visual systems. CNN is a trail of deep neural networks which is usually used on image or speech analyzes in machine learning [22-24].

In addition to different applications of CNN's in image processing, they have especially promising performance in different medical image problems like lesion classification [25], breast cancer [26], tumor diagnosis [26], brain analysis [27], panoptic analysis [28], and MR images fusion [29]. In the mentioned examples of CNN applications, the image has to be first divided into a lot of small superpixels and then the methods have to be performed on all of the superpixels. From the literature, it is observed that using CNN models improves the diagnosis system performance [30]. A part of the training step in the neural networks is to find the optimal solution to fit the target problem based on internal weights which is usually established by the back propagation (BP) algorithm. BP is a classic method that evaluates the error on each training pairs and adjusts the neurons weights to fit the desired output [31]. The error minimization is established by the gradient descent algorithm as it minimizes the cross-entropy loss in the image. This is a complicated optimization problem which needs high cost for solving it.

Recently, the utilization of meta-heuristics in different applications is extensively increasing. One of these applications can be to use them for the cross-entropy loss minimization [32]. In the recent years, several kinds of meta-heuristic algorithms have been introduced. In 2016, Mirjalili and Lewis proposed a new meta-heuristic method called whale optimization algorithm [33]. The whale optimization algorithm is an inspiration of the bubble net hunting strategy of the humpback whales. Despite being new, it has good results for different applications [34-38]. Here, this algorithm is employed for the cross-entropy loss minimization of skin cancer images to improve the method efficiency. In the present study, the whale optimization algorithm is employed for the diagnosis of cancer images. The main purpose here is to optimize the weights of any layer of CNN. The proposed optimization algorithm shows suitable improvements about optimal training of the CNN.

The main structure of the paper is given in the following. Section "Materials and Methods" describes comprehensive explanations about materials and methods

including convolutional neural networks (CNN) and whale optimization algorithm (WOA).

In Section "The Proposed WOA based CNN", the new optimized convolutional neural network based on a whale optimization algorithm is presented. Section "Dataset Description" briefly presents the dataset which is considered for performance analysis.

Section "Implementation Results" investigates the experimental results by performing a comparison between the proposed method and 10 popular cancer detectors, and the conclusions of the paper are given in section "Conclusions".

2 Materials and Methods

In the following, a general description about the CNN and how to optimize them will be described.

2.1 Convolutional neural networks

Here, the membranous neurons respond to the motive in bounded areas called *receptive fields*. The receptive field for each neuron partially overlaps until the visual field is tiled. The reply to the every single neuron to the motive can be mathematically approximated by a convolution operation [39,40].

Convolutional neuron layers contain the most principal part of a CNN. For classification applications like image classification, multiple 2D matrices can be considered as the input and the output of the convolutional layer. It is important to note that there is no restriction about the equality in the number of the input and the output matrices.

In this step, local feature extraction has been applied to extract the regional characteristics of the original image. The main purpose of the learning procedure is to obtain some kernel matrices to get better prominent features to be utilized in image classification. The BP algorithm can be used here for optimizing the network connection weights. The convolution in this layer is performed by a sliding window. Afterward, a vector has been generated based on the sliding window and the dot product and the weights are added up.

Then, an activation function is utilized for each neuron which is often a rectified linear unit (ReLU), with a function $f(x) = \max(x, 0)$ [41]. This process has been implemented on the original image. For more scale reduction of the output, another process called max pooling has

been employed; here, only the highest value is reported to the subsequent layer of the sliding grid. After initializing the structure of a CNN, an optimization method will be required to fit the target problem based on internal weights. This process is usually applied by the BP algorithm. In BP, the error on each training pairs is evaluated and then it is employed to adjust the weights of the neurons to fit the desired output [6, 31]. BP uses a gradient descent algorithm for the error minimization. The gradient descent is a method based on minimizing the cross-entropy loss as the fitness function [42]. The proposed fitness function is given in the following.

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

where, $d_j = (0, \dots, 0, \underbrace{1}_{k}, 0, \dots, 0)$ describes the desired output vector and z_j is the obtained output vector of the m^{th} class.

The softmax function is illustrated in the following formula:

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

where, N is the number of samples.

The function L can be modified by the weight penalty to include a γ value, to keep the values of the weights from getting larger:

$$L = \sum_{j=1}^N \sum_{i=1}^M -d_j^{(i)} \log z_j^{(i)} + \frac{1}{2} \gamma \sum_k \sum_L W_{k,l}^2 \tag{3}$$

where, W_k describes the connection weight, k in layer l and L and K are the total number of layers and the layer l connections, respectively.

Although CNN has been introduced as a powerful classification tool, designing an optimal structure for its layout is a significant problem: most of the designed layouts are based on trials and errors.

Recently, there are some new works which have introduced modifications based on meta-heuristic algorithms [43, 44]. Fig. 1 shows a simple skin cancer detection using ordinary CNN.

In the fig. 1 the convolution layer evaluates the output of the neurons that are connected to the local area at the input. The calculation is performed by the point multiplication between the weights of each neuron and the area they are connected to (the activation mass). The main purpose of the pooling layer is also to subsample the input image to reduce the computational load, memory, and the number of parameters (over fitting). Reducing the size of the input image also causes the neural network to be less sensitive to image displacement (independent of the position).

The WOA is a new stochastic optimization method which is derived from the hunting process of whales [33]. Like any evolutionary method, WOA starts with a random population set (candidate solutions) to search and find the global optimum (maximum or minimum) solution for the problem. The algorithm continues to improve and update the solution based on its structure until the optimum value is satisfied.

The basic difference between the WOA and other meta-heuristic methods is how the WOA rules develop and update the solution.

The WOA is inspired from the whale’s trap and attack hunting strategy; the use of bubbles in spiral movement around the prey with which the trap is formed is given the name “*bubble-net feeding behavior*”. The behavior of the bubble-net feeding process is shown in Fig. 2. From the

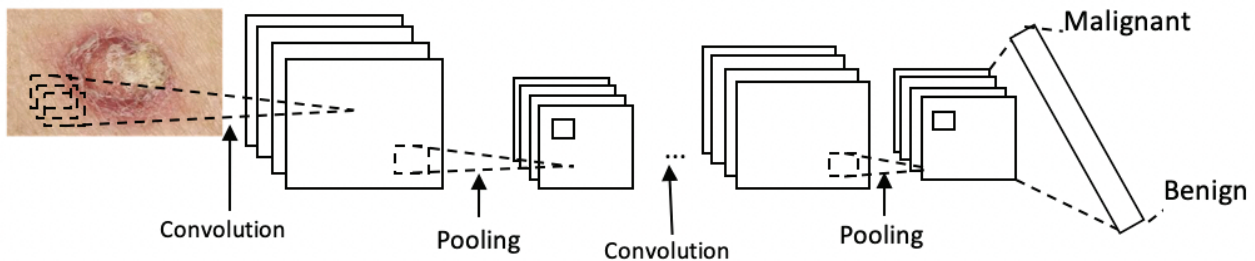


Figure 1: A simple skin cancer detection using ordinary CNN

figure 2, it is clear that the humpback whale first creates bubbles around the prey. This process is performed by spiral motion of the whale. Afterward, it attacks the prey. This process comprises the main contributor of the WOA. The explained created bubble-net system is mathematically defined as follows:

$$X(t+1) = \begin{cases} X^*(t) - AD & p < 0.5 \\ D'e^{bl}\cos(2\pi t) + X^*(t) & p \geq 0.5 \end{cases} \quad (4)$$

$$D' = |CX^*(t) - X(t)| \quad (5)$$

$$A = 2ar - a \quad (6)$$

$$C = 2r \quad (7)$$

where, p and r are random constants and are bounded between $[0, 1]$, l is a random constant in the interval $[-1, 1]$, t illustrates the present iteration, D' describes the distance for the i^{th} whale from the prey (the best solution), b defines the logarithmic shape of the spiral motion, a is a linear descent from 2 to 0 over the iteration.

In the above equation, the first term models the encircling process and the second term models the bubble-net process. In WOA, these two terms comprise the exploitation and exploration terms of the algorithm [33].

As it is mentioned before, the WOA starts with a random population. The solutions are then updated in each iteration in order for mathematical modeling of the bubble net hunting and the prey encircling. Here, to ensure the convergence of the algorithm, the best solution improves the position of the agents when $|X| > 1$. Otherwise, the best solution will play the rule of the pivot point. In the following, the general pseudo code for the WOA is given.

Start

Initializing the whale population X_i

Initializing A, C , and a

Evaluate the cost value of each agent

X^ shows the best solution in the current iteration*

Apply WOA:

$t=1$

while $t \leq \text{max iteration}$:

for all the agents

if $|A| \leq 1$ **then**

Update the position of the agents

else if $|A| \geq 1$ **then**

Find a random search agent X_{rand}

Update the position of the agents

end if

end for

Update A, C , and a

*Update X^**

$t = t + 1$

end while

return X^*

End

2.2 The Proposed WOA based CNN

In the present study, a different strategy is utilized specifying the number of hyper-parameters; not only the most appropriate hyper-parameters for the CNN at the moment can be considered, it can also take into account the time to run each moment.

As mentioned before, the primary purpose of this research is to design an optimization based technique for skin cancer classification. The main idea here is to utilize an optimized method to improve the system accuracy. Candidate solutions in the proposed optimized classification problem are a sequence of integers. In this method, at first, the minimum (min) and the maximum (max) limitations for the algorithm is determined to prevent the system errors. In this problem, max describes the size of the sliding window and min is 2. Here, the constant 2, presents the minimum value that is acceptable for the max pooling where no lower size exists.

The other point which should be considered is that the value of the input data should be greater than the sliding window. Afterward, a group of solutions is randomly generated. In this problem, the initial population is set to 150, where the hyper-parameters settings of the CNN are described by the individuals, occurring within 10 integer values. The search agent vector for the proposed CNN is shown in Fig. 2.

Afterward, the solutions are evaluated. Here, the half-value precision for the proposed optimized CNN is considered as the cost function on a skin cancer validation process. It is important to know that the general strategy has high computational cost; each member of the population describing the CNN requires training on the skin

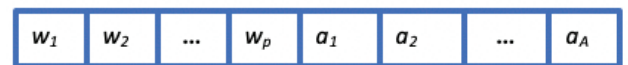


Figure 2: The search agent vector assigning of WOA on the CNN

cancer dataset with applying back propagation algorithm for 1500 iterations.

After initial population generating and evaluation of the initial cost, the position of the search agents are updated based on the parameters like prey encircling and bubble net hunting and the process repeats until the stop criteria are achieved.

The designed optimized system was tested on the DermIS and the Dermquest databases based on the minimization of the MSE value for validating and testing. In the following, more explanations are described. Weights and biases are two important parameters which are used for optimizing the structure of the CNN. Therefore, in this research, these two features have been selected for optimizing, such that:

$$W = \{w_1, w_2, \dots, w_p\} \tag{8}$$

$$A = \{a_1, a_2, \dots, a_A\} \tag{9}$$

$$w_n = \{w_{1n}, w_{2n}, \dots, w_{Ln}\} \tag{10}$$

$$b_n = \{b_{1n}, b_{2n}, \dots, b_{Ln}\} \tag{11}$$

$$l = 1, 2, \dots, L$$

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

where, A and L are the total number of agents and the total number of layers respectively, l describes the layer index, n describes the number of the agent, and w_{in} describes the value of the weight in layer i . In other words, the total parameters for optimizing can be described by the following vector:

$$W_n = \{W, A\} \tag{12}$$

Fig.2. shows these assignments.

A simplified measured error between the reference and the system output is given below.

$$E = \frac{1}{T} \sum_{i=1}^T \sum_{j=1}^k (d_{ji} - o_{ji})^2 \tag{13}$$

where, T describes the training samples number, k is the output layers number, d_{ji} and o_{ji} are the desired value and output value of the CNN respectively. Gradient descent contains the main part of the ordinary BP algorithm; this technique can be trapped easily into the local minimum. This shortcoming can lead to wrong results in some complicated pattern recognition problems [45-47].

Another advantage of using WOA than the BP for Error minimization is that the WOA based method does not require the backward phase as a high computational cost process. Fig. 3 shows the flowchart diagram of the proposed method.

2.3 Dataset Description

Two different dermoscopy databases have been employed for testifying and analyzing the proposed method:

1) DermIS Digital Database [48]: This is an image atlas on different kinds of skin cancers with differential diagnoses which was launched for medical image processing applications. This database is the largest online information service available on the Internet.

2) Dermquest Database [49]: This is an online medical atlas for dermatologists and dermatologist-based healthcare

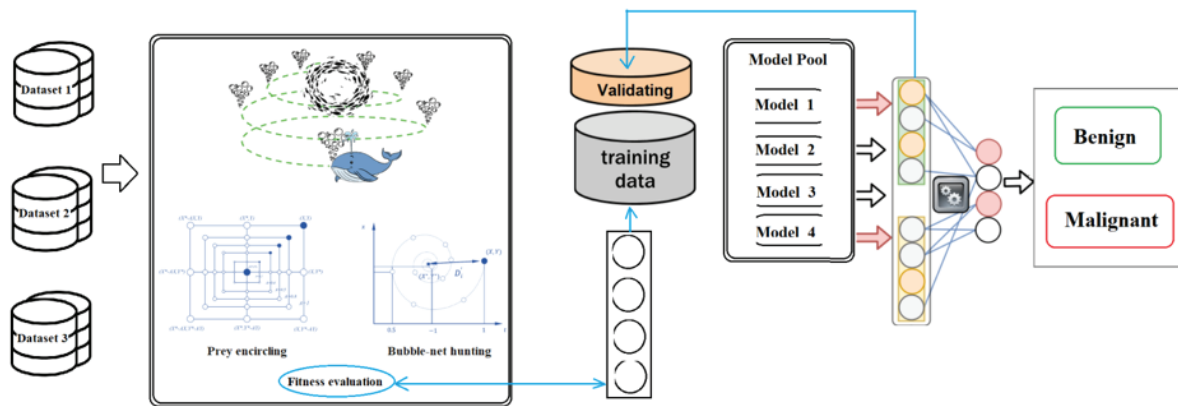


Figure 3: The WOA based framework for the structure of the convolutional neural network

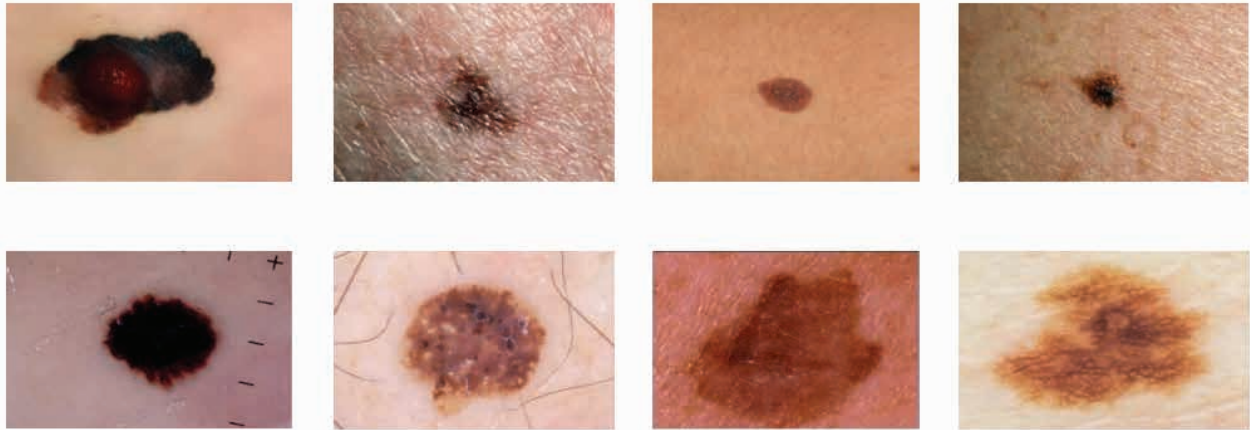


Figure 4: Some examples of the DermIS and the Dermquest databases

professionals. All images of this database are reviewed and approved by famous international editorial boards. It uses an extensive number of dermatologists and includes over 22,000 clinical images.

Some examples of the databases are shown in Fig. 4.

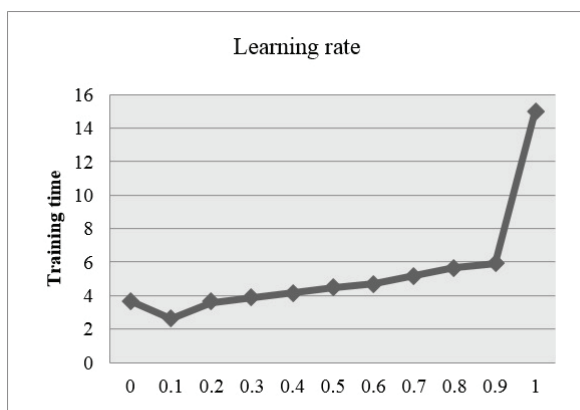
3 Implementation Results

Experimental simulations were implemented using Matlab R2017® software on a Intel Core i7-4790K processor with 32 GB of RAM, and two NVIDIA GeForce GTX Titan X GPU cards with scalable link interface (SLI). The proposed simulations were implemented on two of the standard skin cancer databases to analyze the system performance.

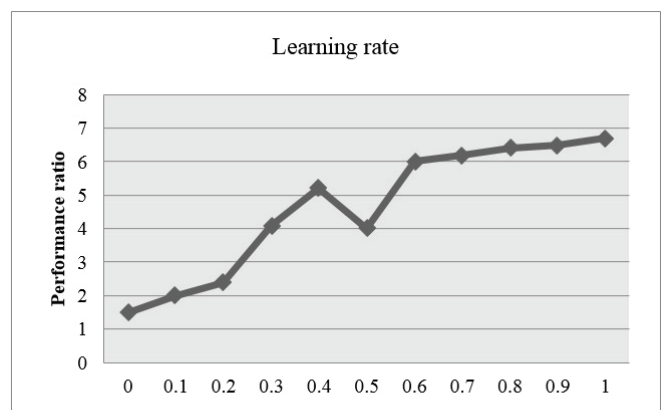
70% of data was utilized as the training set and 10% for validating set. The remaining 20% was used as test

sets. This division is known as the Pareto principle (80/20 rule) that states that, for many events, roughly 80% of the effects come from 20% of the causes [50]. For determining what images should be included in the training, validating, or testing part, they were selected randomly. For fair image processing, all images of the datasets were resized to 640×480 . The proposed CNN was trained by the WOA method. In the presented experiment (Fig. 5), it is obvious that the learning rate varied between 0.2 and 0.9, since the radius and the number of neuron cells are different, almost 100% of training pixels will be included in the prototype neurons.

The best case is to select a neural network with smallest neuron volume. Based on [51], it is possible to select a proper learning rate by the performance ratio. Fig. 5 shows that by increasing the learning rate, both the performance ratio and training time will increase.



(A)



(B)

Figure 5: (A) learning rate vs. training time, (B) learning rate vs. performance ratio

However performance ratio is significant, for making a trade-off between performance ratio and training time, the learning rate is selected to be 0.9.

As explained before, Dermis and the Dermquest are employed as two most applicable databases for testifying the proposed method.

30,000 iterations were performed to train the proposed network. For making a correct and independent

analysis of the images, the training step was repeated 60 times and the final results were described based on the mean values.

For testifying the performance of the proposed system, five performance metrics were employed and are defined as follows.

Table 1: Comparison of the performance metrics for skin cancer detection

Method	Performance Metric				
	Sensitivity	Specificity	PPV	NPV	Accuracy
Proposed CNN/WOA Method	0.95	0.92	0.84	0.95	0.91
MED-NODE texture Descriptor[56]	0.64	0.87	0.76	0.79	0.78
MED-NODE color descriptor [56]	0.76	0.74	0.66	0.83	0.75
Spotmole [55]	0.84	0.59	0.58	0.85	0.69
AlexNet [57]	0.84	0.61	0.67	0.85	0.82
ResNet-50 [59]	0.86	0.80	0.71	0.84	0.83
ResNet-101 [59]	0.85	0.77	0.75	0.89	0.85
VGG-16[58]	0.90	0.86	0.79	0.90	0.86
LIN [60]	0.91	0.89	0.80	0.92	0.88
Inception-v3 [61]	0.84	0.65	0.64	0.72	0.84
Ordinary CNN	0.83	0.81	0.77	0.88	0.83

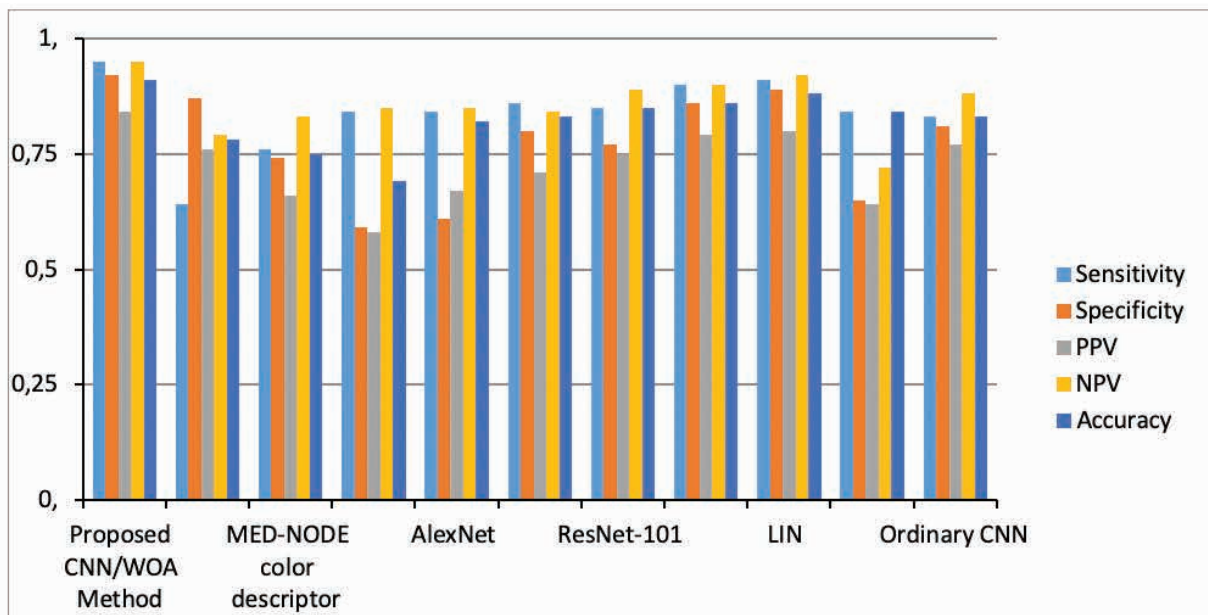


Figure 6: Distribution of classification performance of the methods for skin cancer detection

$$\text{Sensitivity} = \frac{\text{Number of correctly detected skin cancer cases}}{\text{Total number of skin cancer cases}} \quad (14)$$

$$\text{Specificity} = \frac{\text{Number of correctly detected healthy skin cases}}{\text{total number of healthy skin cases}} \quad (15)$$

$$\text{PPV} = \frac{\text{Number of correctly detected skin cancer cases}}{\text{Number of detected skin cancer cases}} \quad (16)$$

$$\text{NPV} = \frac{\text{Number of correctly detected healthy skin cases}}{\text{Number of detected healthy skin cases}} \quad (17)$$

$$\text{accuracy} = \frac{\text{Number of correctly detected cases}}{\text{total number of cases}} \quad (18)$$

There are different research works which have been introduced for skin cancer detection [52-54]. Each of these methods have their own difficulties and shortcomings. Introducing all of these methods is not possible. Therefore 10 methods have been selected for comparison with our proposed method.

The method of [55] is based on a commercial tool. The method of [56] is about a framework based on the semi-supervised system. For a fair comparison, automatically extracted descriptors of this method are employed. Some deep learning based systems like Ordinary CNN, AlexNet [57], VGG-16[58], ResNet [59], LIN [60], and Inception-v3 [61] are also utilized for this comparison. Table 1 illustrates a performance comparison between the proposed system and the aforesaid methods.

As can be observed, the CNN/WOA method is most accurate when compared with the other 10 aforesaid methods.

This is due to the combination of the CNN with the whale optimization algorithm. Applying this optimization algorithm for the CNN allows it to escape from the local minima. This gives a global minimum for the BP problem in the CNN and addresses better performance for the proposed method.

The results show the effect of using the WOA optimization algorithm on the deep learning framework.

For more clarification, the distribution of classification performance of the above table is shown in Fig. 6 as a bar chart.

The proposed detection system here has two classes: the background region and the cancerous region. The input layer of the proposed CNN/WOA network contains $3 \times n$ pixel feature vectors which describe the information

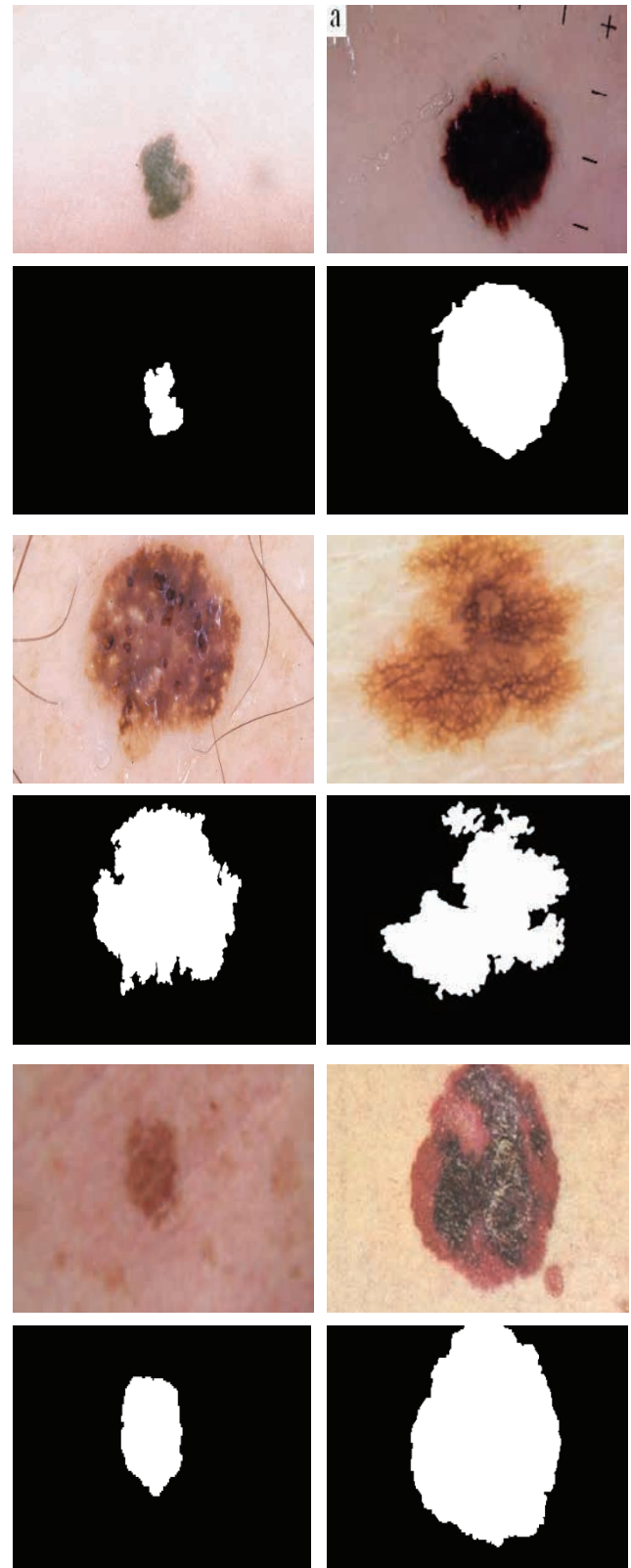


Figure 7: Sample skin cancer detection results: first and third columns: input images, second and fourth column: detected masks based on optimized CNN/WOA method.

about R, G, and B of the image. As it is explained before, rectified linear unit (ReLU) is utilized as the activation function of the network. The output layer presents a two labeled image including 0 (background region) or 1 (cancerous region).

Fig. 7 show the results of some samples of the process of the skin cancer detection system detection using the proposed CNN/WOA method. In the figure, the first and the third columns show the original images and the second and the fourth column show the detected masks based on optimized CNN/WOA method. Experimental results show the high efficiency of the presented method for the diagnosis of the skin cancer regions.

4 Conclusions

In this paper, a new method is proposed for skin cancer detection. The proposed method uses a meta-heuristic based algorithm for optimization of the Convolutional neural network in training the biases and the weights of the network based on back propagation. To do so, the half-value precision is considered for the proposed optimized CNN as the cost function on a skin cancer validation process which includes a simplified measured error between the reference and the system output. In this study, a recently introduced algorithm called whale optimization algorithm is utilized for minimizing the error rate of the learning step for the Convolutional neural network. The proposed method is called CNN/WOA. The proposed method is then tested on images from two well-known databases including DermIS Digital Database and Dermquest Database and compared with 10 number of popular classification methods. Final results show the accuracy prominence of the proposed system toward the compared classifiers.

Conflict of interest statement: Authors state no conflict of interest

References

- [1] Force, U. P. S. T., Behavioral counseling to prevent skin cancer: US Preventive Services Task Force recommendation statement. *JAMA-Journal of the American Medical Association* 2018; 319(11); 1134-1142
- [2] Al-Jamal, R. a. T., Cassoux, N., Desjardins, L., Damato, B., Konstantinidis, L., Coupland, S. E., et al., The pediatric choroidal and ciliary body melanoma study a survey by the european ophthalmic oncology group. *Ophthalmology* 2016
- [3] Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., and Smith, J. R., Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images, in *International Workshop on Machine Learning in Medical Imaging*, 2015; 118-126
- [4] Dalila, F., Zohra, A., Reda, K., and Hocine, C., Segmentation and classification of melanoma and benign skin lesions. *Optik-International Journal for Light and Electron Optics* 2017; 140; 749-761
- [5] Razmjoo, N., Mousavi, B. S., Soleymani, F., and Khotbesara, M. H., A computer-aided diagnosis system for malignant melanomas. *Neural Comput Appl* 2013; 23(7-8); 2059-2071
- [6] Razmjoo, N., Sheykhahmad, F. R., and Ghadimi, N., A hybrid neural network-world cup optimization algorithm for melanoma detection. *Open Medicine* 2018; 13(1); 9-16
- [7] Silveira, M., Nascimento, J. C., Marques, J. S., Marçal, A. R., Mendonça, T., Yamauchi, S., et al., Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE Journal of Selected Topics in Signal Processing* 2009; 3(1); 35-45
- [8] Siegel, R. L., Miller, K. D., and Jemal, A., *Cancer statistics, 2019*. CA: a cancer journal for clinicians 2019
- [9] Cohen, V. M., Pavlidou, E., DaCosta, J., Arora, A. K., Szyszko, T., Sagoo, M. S., et al., Staging uveal melanoma with whole-body positron-emission tomography/computed tomography and abdominal ultrasound: Low incidence of metastatic disease, high incidence of second primary cancers. *Middle East African journal of ophthalmology* 2018; 25(2); 91
- [10] Ghadimi, Noradin. "An adaptive neuro-fuzzy inference system for islanding detection in wind turbine as distributed generation." *Complexity* 21.1 (2015): 10-20
- [11] Kulkarni, A. and Mukhopadhyay, D., SVM Classifier Based Melanoma Image Classification. *Research Journal of Pharmacy and Technology* 2017; 10(12); 4391-4392
- [12] Narasimhan, K. and Elamaran, V., Wavelet-based energy features for diagnosis of melanoma from dermoscopic images. *International Journal of Biomedical Engineering and Technology* 2016; 20(3); 243-252
- [13] Ali, Y. and Hamed, S., Early breast cancer detection using mammogram images: a review of image processing techniques. *Biosciences Biotechnology Research Asia* 2015; 12(1); 225-234
- [14] Razmjoo, N., Mousavi, B. S., and Soleymani, F., A real-time mathematical computer method for potato inspection using machine vision. *Computers & Mathematics with Applications* 2012; 63(1); 268-279
- [15] Razmjoo, N., Mousavi, B. S., and Soleymani, F., A hybrid neural network Imperialist Competitive Algorithm for skin

- color segmentation. *Math Comput Modell* 2013; 57(3); 848-856
- [16] Moallem, P. and 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
- [17] Mirjalili, S., Genetic Algorithm. in *Evolutionary Algorithms and Neural Networks*, ed: Springer, 2019; 43-55
- [18] Such, F. P., Madhavan, V., Conti, E., Lehman, J., Stanley, K. O., and Clune, J., Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. *arXiv preprint arXiv:1712.06567* 2017
- [19] Abedinia, O., Amjady, N., and Ghadimi, N., Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm. *Computational Intelligence* 2018; 34(1); 241-260
- [20] Ghadimi, N., Akbarimajid, A., Shayeghi, H., and Abedinia, O., Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting. *Energy* 2018; 161; 130-142
- [21] Razmjoo, N. and Ramezani, M., Training wavelet neural networks using hybrid particle swarm optimization and gravitational search algorithm for system identification. *International Journal of Mechatronics, Electrical and Computer Technology* 2016; 6(21); 2987-2997
- [22] Ghadimi, Noradin. "A new hybrid algorithm based on optimal fuzzy controller in multimachine power system." *Complexity* 21.1 (2015): 78-93
- [23] Jadhav, A. R., Ghontale, A. G., and Shrivastava, V. K., Segmentation and Border Detection of Melanoma Lesions Using Convolutional Neural Network and SVM. in *Computational Intelligence: Theories, Applications and Future Directions-Volume I*, ed: Springer, 2019; 97-108
- [24] Ronneberger, O., Fischer, P., and Brox, T., U-net: Convolutional networks for biomedical image segmentation, in *International Conference on Medical image computing and computer-assisted intervention*, 2015; 234-241
- [25] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., and Greenspan, H., GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification. *arXiv preprint arXiv:1803.01229* 2018
- [26] Zhou, Y., Xu, J., Liu, Q., Li, C., Liu, Z., Wang, M., et al., A Radiomics Approach with CNN for Shear-wave Elastography Breast Tumor Classification. *IEEE Transactions on Biomedical Engineering* 2018
- [27] Bernal, J., Kushibar, K., Asfaw, D. S., Valverde, S., Oliver, A., Martí, R., et al., Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artificial intelligence in medicine* 2018
- [28] Zhang, D., Song, Y., Liu, D., Jia, H., Liu, S., Xia, Y., et al., Panoptic segmentation with an end-to-end cell R-CNN for pathology image analysis, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018; 237-244
- [29] Zhang, L., Yin, F., and Cai, J., A Multi-Source Adaptive MR Image Fusion Technique for MR-Guided Radiation Therapy. *International Journal of Radiation Oncology• Biology• Physics* 2018; 102(3); e552
- [30] Sudharshan, P., Petitjean, C., Spanhol, F., Oliveira, L. E., Heutte, L., and Honeine, P., Multiple instance learning for histopathological breast cancer image classification. *Expert Systems with Applications* 2019; 117; 103-111
- [31] Roy, K., Mandal, K. K., and Mandal, A. C., Ant-Lion Optimizer algorithm and recurrent neural network for energy management of micro grid connected system. *Energy* 2019; 167; 402-416
- [32] Blanco, R., Cilla, J. J., Malagón, P., Penas, I., and Moya, J. M., Tuning CNN Input Layout for IDS with Genetic Algorithms, in *International Conference on Hybrid Artificial Intelligence Systems*, 2018; 197-209
- [33] Mirjalili, S. and Lewis, A., The whale optimization algorithm. *Advances in Engineering Software* 2016; 95; 51-67
- [34] Oliva, D., El Aziz, M. A., and Hassanien, A. E., Parameter estimation of photovoltaic cells using an improved chaotic whale optimization algorithm. *Applied Energy* 2017; 200; 141-154
- [35] Kaveh, A. and Ghazaan, M. I., Enhanced whale optimization algorithm for sizing optimization of skeletal structures. *Mechanics Based Design of Structures and Machines* 2017; 45(3); 345-362
- [36] Ahadi, Amir, Noradin Ghadimi, and Davar Mirabbasi. "An analytical methodology for assessment of smart monitoring impact on future electric power distribution system reliability." *Complexity* 21.1 (2015): 99-113
- [37] El Aziz, M. A., Ewees, A. A., and Hassanien, A. E., Whale Optimization Algorithm and Moth-Flame Optimization for multilevel thresholding image segmentation. *Expert Systems with Applications* 2017; 83; 242-256
- [38] Trivedi, I. N., Pradeep, J., Narottam, J., Arvind, K., and Dilip, L., Novel adaptive whale optimization algorithm for global optimization. *Indian Journal of Science and Technology* 2016; 9(38)
- [39] Schmidhuber, J., Deep learning in neural networks: An overview. *Neural networks* 2015; 61; 85-117
- [40] Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., and Adeli, H., Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput Biol Med* 2018; 100; 270-278
- [41] Koehler, F. and Risteski, A., Representational Power of ReLU Networks and Polynomial Kernels: Beyond Worst-Case Analysis. *arXiv preprint arXiv:1805.11405* 2018
- [42] 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:1506.00619* 2015
- [43] Martens, J. and Sutskever, I., Learning recurrent neural networks with hessian-free optimization, in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, 2011; 1033-1040
- [44] Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H., Greedy layer-wise training of deep networks, in *Advances in neural information processing systems*, 2007; 153-160
- [45] Zhang, L. and Suganthan, P. N., A survey of randomized algorithms for training neural networks. *Information Sciences* 2016; 364; 146-155
- [46] Jaddi, N. S. and Abdullah, S., Optimization of neural network using kidney-inspired algorithm with control of filtration rate and chaotic map for real-world rainfall forecasting. *Engineering Applications of Artificial Intelligence* 2018; 67; 246-259

- [47] Emary, E., Zawbaa, H. M., and Grosan, C., Experienced gray wolf optimization through reinforcement learning and neural networks. *IEEE transactions on neural networks and learning systems* 2018; 29(3); 681-694
- [48] Xu, H. and Mandal, M., Epidermis segmentation in skin histopathological images based on thickness measurement and k-means algorithm. *EURASIP Journal on Image and Video Processing* 2015; 2015(1); 18
- [49] Database, D. (2019). Dermquest Database Available: <https://www.derm101.com/dermquest/>
- [50] Zhu, Q. and Xiang, H., Differences of Pareto principle performance in e-resource download distribution: An empirical study. *The Electronic Library* 2016; 34(5); 846-855
- [51] Sui, C., Kwok, N. M., and Ren, T., A restricted coulomb energy (rce) neural network system for hand image segmentation, in *2011 Canadian Conference on Computer and Robot Vision*, 2011; 270-277
- [52] Deepa, S. and Devi, B. A., A survey on artificial intelligence approaches for medical image classification. *Indian Journal of Science and Technology* 2011; 4(11); 1583-1595
- [53] Celebi, M. E., Wen, Q., Iyatomi, H., Shimizu, K., Zhou, H., and Schaefer, G., A state-of-the-art survey on lesion border detection in dermoscopy images. *Dermoscopy Image Analysis* 2015; 97-129
- [54] Loescher, L. J., Janda, M., Soyer, H. P., Shea, K., and Curiel-Lewandrowski, C., Advances in skin cancer early detection and diagnosis, in *Seminars in oncology nursing*, 2013; 170-181
- [55] Munteanu, C. and Coooclea, S., *Spotmole—melanoma control system*. ed, 2009
- [56] Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F., and Petkov, N., MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert systems with applications* 2015; 42(19); 6578-6585
- [57] Krizhevsky, A., Sutskever, I., and Hinton, G. E., Imagenet classification with deep convolutional neural networks, in *Advances in neural information processing systems*, 2012; 1097-1105
- [58] Simonyan, K. and Zisserman, A., Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* 2014
- [59] Jalili, Aref, and Noradin Ghadimi. "Hybrid harmony search algorithm and fuzzy mechanism for solving congestion management problem in an electricity market." *Complexity* 21.S1 (2016): 90-98
- [60] Li, Y. and Shen, L., Skin lesion analysis towards melanoma detection using deep learning network. *Sensors* 2018; 18(2); 556
- [61] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z., Rethinking the inception architecture for computer vision, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016; 2818-2826