Research Article

Navid Razmjooy, Fatima Rashid Sheykhahmad, Noradin Ghadimi*

A hybrid neural network – world cup optimization algorithm for melanoma detection

Abstract: One of the most dangerous cancers in humans is Melanoma. However, early detection of melanoma can help us to cure it completely. This paper presents a new efficient method to detect malignancy in melanoma via images. At first, the extra scales are eliminated by using edge detection and smoothing. Afterwards, the proposed method can be utilized to segment the cancer images. Finally, the extra information is eliminated by morphological operations and used to focus on the area which melanoma boundary potentially exists. To do this, World Cup Optimization algorithm is utilized to optimize an MLP neural Networks (ANN). World Cup Optimization algorithm is a new meta-heuristic algorithm which is recently presented and has a good performance in some optimization problems. WCO is a derivative-free, Meta-Heuristic algorithm, mimicking the world's FIFA competitions. World cup Optimization algorithm is a global search algorithm while gradient-based back propagation method is local search. In this proposed algorithm, multi-layer perceptron network (MLP) employs the problem's constraints and WCO algorithm attempts to minimize the root mean square error. Experimental results show that the proposed method can develop the performance of the standard MLP algorithm significantly.

Keywords: World Cup Optimization Algorithm, Melanoma, Cancer, Tumors, Artificial Neural Network

*Corresponding author: **Noradin Ghadimi,** Young Researchers and Elite club, Ardabil Branch, Islamic Azad University, Ardabil, Iran, Tel. +989147028949, Noradin1364@gmail.com
Navid Razmjooy, Fatima Rashid Sheykhahmad, Young Researchers and Elite club, Ardabil Branch, Islamic Azad University, Ardabil, Iran

1 Introduction

In the preceding years, skin cancer has become one of the most common cancers in the world. In addition, modeling of skin cancer due to its fine-scale geometry and the complex surface has become a difficult case study.

Skin cancer can be easily diagnosed visually; however, there are a lot of specific aspects of the skin which can better assess by non-invasive imaging methods [1].

For the past 30 years, Melanoma rates have increased in the United States. However, the melanoma lifetime risk is 1 in 50 for whites, about 1 in 1,000 for blacks and 1 in 200 for Hispanics. Rising melanoma rates have motivated practitioners to detect lesions in their curable, early phase.

By detecting the skin cancer in early stages, it can be cured. However, when advanced, it spreads to other parts of the body, becoming harder to treat and often fatal [2].

In the melanoma detection process, architectural and cellular characteristics can be utilized to determine the malignancy of the skin tissue if the melanocytes are identified correctly.

The clinical characteristics of melanoma detection include Asymmetry, irregular Borders, more than one or uneven distribution of Color, or a large (greater than 6mm) Diameter. The Evolution of moles is also a critical factor [3, 4]. These characteristics were first introduced by the American Cancer Society as the ABCD rule to provide a standard and easily remembered guideline for the patient to use in self-examination for MM. Physicians can detect melanoma by using the ABCD rule. For analyzing the ABCD score, the criteria are assigned semi-quantitatively [5]. Each of the criteria is then multiplied by a given weight factor to calculate a total dermoscopy score. The ABCD rule works appropriately for thin melanocytic wounds. The ABCD rule has about 59% to 88% accuracy in diagnosing melanoma, but biopsy is needed for more precise diagnosis [5, 6].

The first step in achieving image characteristics for melanoma detection is to diagnose and localize the lesions in the image. Automated melanoma detection systems are based on using one imaging modality (like dermoscopy),

computer algorithms and mathematical models to predict if a skin lesion is a melanoma [7].

In 1999, Xu et. al. proposed a method based on converting the color images into the intensity dimension on which the lesion boundaries were then developed by using a nonlinear sigmoid function [8]; they were applied then Double-thresholding to localize the boundary edges, which were then checked with a closed elastic curve to get a smooth lesion boundary.

In 2001, Ganster et al. synthesize dynamic thresholding, global thresholding and 3-D color clustering along with a fusion technique to characterize a lesion; they achieved 96% performance on a set of 4000 images [9].

In 2004, Zagrouba and Barhoumi motivated by the desire to classify skin lesion from color images; they employed fuzzy classifier after noise removing to detect the melanoma and achieved 79.1% accuracy for correct classify of lesions [10].

Orientation sensitive Fuzzy c-mean [9], Density-Based Spatial Clustering of Application with Noise [11], and JSEG [12] are the other examples of implementing the clustering algorithms in melanoma detection.

In 2004, Zouridakis, et al. [13] developed a new automatic melanoma detection technique based on size difference of two image modalities: TLM and XLM. The XLM imaging modality captures only surface pigmentation.

In 2011, Fassihi et al. used coefficients of wavelet decomposition to extract image characteristics. Melanoma classification is carried out by utilizing the mean and variance of the wavelet coefficients of the input images as the input of neural network [14]. Final results show about 90% accuracy in the distinction between benign and melanoma.

In the melanoma detection, researchers proposed employing back-propagation neural network to model unstructured problems due to its ability to map complex non-linear relationships between input and output variables.

Unfortunately, the back-propagation algorithm is known as a local search algorithm which uses gradient descent to iteratively develop the weights and biases in the neural network [15-20]. A significant drawback of the gradient descent technique is that Easy trapped in local minimum and slow convergence

In this paper to compensate this drawback, world cup optimization (WCO) algorithm has been used to find the optimal values for weights and biases in the back-propagation algorithm.

WCO a new proposed swarm-based metaheuristic algorithm [21]. This algorithm imitates the social leadership and hunting behavior of grey wolves in nature.

Because of its metaheuristic feature, it can search for optimal solutions in different directions in order to minimize the chance of trapped in a local minimum and increment the convergence speed.

2 Filtering

In Biomedical imaging, performing some kinds of noise and over-segmentation reduction on the considered image is often desirable which makes easy the next processing steps.

The median filter is a nonlinear digital filtering technique which is often employed to remove noise from an image or signal. This process is a pre-processing step to improve the results of later processing (in this paper detect of melanoma parts of an image). Median filtering is one of the most utilized methods in medical imaging because, under certain conditions, it preserves edges while removing noise. In case, the median filter replaces a pixel by the median of all pixels in its neighborhood as below:

$$y[m,n] = median\{x[i,j], (i,j) \in \omega\} \qquad (1)$$

where ω is a neighborhood centered around the location (m, n) in the image.

Median filter considers the pixels in the image in turn and looks at their neighbors to make a decision which is representative of its surroundings or not. Median filter gets evaluated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then placing the pixel being considered with the middle pixel value [22].
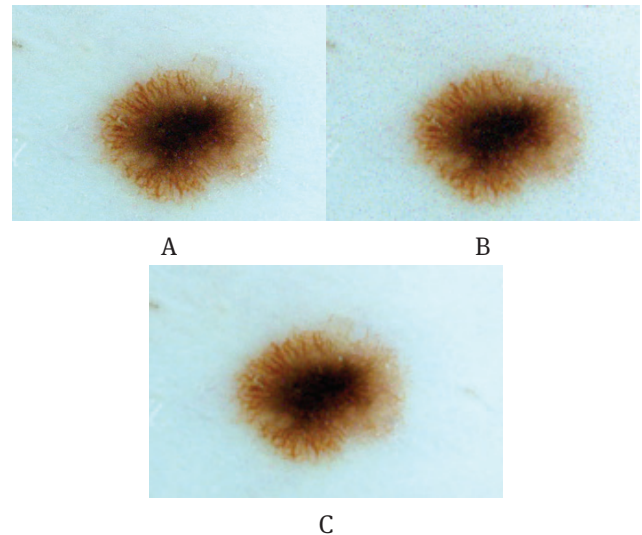

Figure 1: Image noise reduction: (A) input image, (B) apply salt and pepper noise to image and (C) Filtered image

In this paper, a median filter is applied to the image to assigns to each pixel over a neighborhood of a given size. This filter decreases the small structures affections, like noise, hair, and scale lines on the segmentation result. The employed neighborhood of the median filter depends on the image resolution. In this research, 9×9 neighborhood is utilized for images by the size of 256×256 pixels to show a complete melanoma.

3 Supervised classification of the melanoma

Supervised classification is the technique which is often utilized for the quantitative analysis of biomedical imaging. The purpose of supervised classification in melanoma detection is to divide all the pixels of the input image into two classes (Melanoma and not melanoma classes). By using supervised classification, we categorize examples of the information classes (i.e., melanoma type) of interest in the image. Melanoma color is one of the considered cases which can become a classification issue. In addition, the purpose of melanoma color pixel classification is to decide whether a color pixel is a melanoma color or not. Good Melanoma color pixel classification should make coverage of all various melanoma types. Such a mentioned problem can be evaluated by artificial neural networks which have been proven as an efficient tool for pattern classification purposes where decision rules are hidden in highly complex data and can be learned only from examples. The image is then classified by attempting the performance for each pixel and decides about which of the signatures being similar most; figure 2 shows the steps of classification.

4 Artificial neural network

Artificial Neural Networks (ANNs) are relatively crude electronic models based on the neural structure of the brain. Natural neurons receive signals through synapses located on the dendrites or membrane of the neuron [23]. When the received signals get strong enough, the neuron is activated and emits a signal through the axon. This signal might be sent to another synapse and might activate other neurons.

From the practical point of view, ANNs are just parallel computational systems which include many simple processing elements connected together in a special way to perform a considered task. ANNs are strong computational devices which can learn and generalize from training data; since there is no requirement for complicated feats of programming.

From the mathematics view, a neuron's network function $f(x)$ can be described as a forming of other functions $g_i(x)$, which can be defined as other functions forming. This can be easily defined as a network structure, with arrows representing the dependencies between variables. A commonly used kind of forming is the nonlinear weighted sum, where

$$f(x) = K\left(\sum_i \omega_i g_i(x)\right) \qquad (2)$$

where K represents a predefined function, like the hyperbolic tangent. It will be easy for the following to assign a collection of functions g_i as simply a vector $g = (g_1...g_n)$.

From different techniques, Backpropagation (BP) is a commonly used method which is employed for feedforward networks. It evaluates the error on all of the training pairs and regulates the weights to fit the desired output. This is performed in several iterations to achieve the minimum value for error of the training set. After training
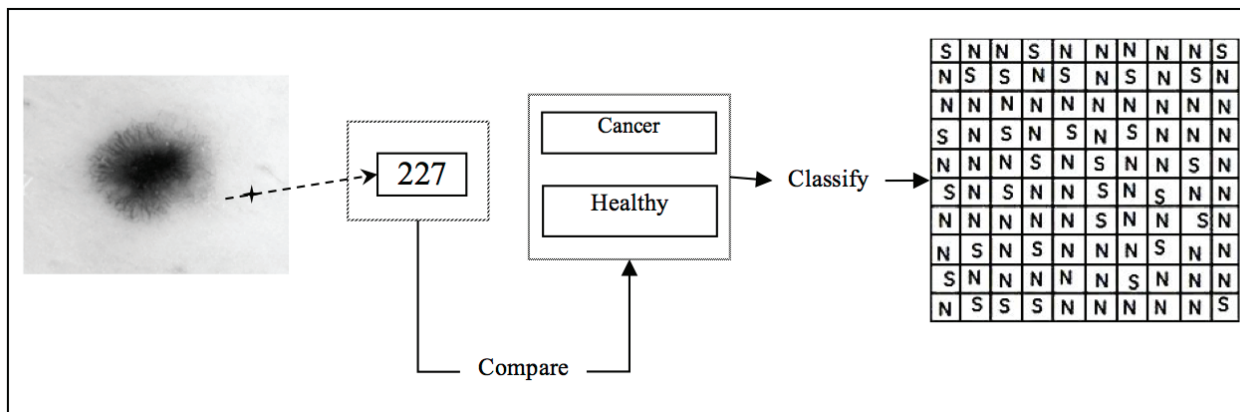

Fiure 2: Steps in supervised classification

process, the network weights are ready to use for evaluating output values for new given samples.

BP uses gradient descent algorithm to minimize error space. This algorithm has the drawback of trapping to the local minimum which is entirely dependent on initial (weight) settings. This objection can be removed by an algorithm by an exploration based algorithm, like the evolutionary algorithms.

5 World cup optimization algorithm

In the last decades, meta-heuristic algorithms have been considered as higher-level procedures to find, generate, or select a heuristic to provide a sufficiently good solution to an optimization problem, especially with incomplete or imperfect information or limited computation capacity.

There are different meta-heuristic algorithms like Genetic algorithm [24], particle swarm optimization [25, 26] and quantum invasive weed optimization [27] have been introduced to employ for solving complicated problems from different applications of science and technology.

In the recent years, a new meta-heuristic algorithm has been introduced which is inspired from the FIFA world cup competitions and shows good results in different applications; the algorithm is known as World Cup Optimization (WCO) algorithm.

The main purpose of WCO algorithm is to attention into the competition among different teams until one of them reach the best score and become the champion.

In WCO algorithm, a coefficient is introduced as rank. Rank has an important impact on every team's success. After achieving the rank scores, strong teams have been categorized as the first seed, the second seed includes the teams weaker than the first seed and the others have been categorized like the second team hierarchically. In this algorithm, in the first step, the seed one arises to the next level with no competitions. Afterwards, the challenge starts. Here, the competition starts with challenging the teams separately in their seeds to win the competition, raise their scores and upgrade their rank for the next games and cups.

After early competitions, the best two teams from each group arise to the next level and the rest has been eliminated. The third place of each competition in the seeds has a second chance to arise itself into the next level by winning the other same score teams from the other seeds (Play-Off). The final competition is held between two teams with the most scores to define the champion of the competitions. The flowchart of WCO algorithm is shown in the figure below.

6 ANN weights development using WCO (HNNWCO)

An important aspect of an ANN model is training process; because the performance of ANNs is directly dependent on the training process success. The main purpose of the training step is to minimize the mean squared error (MSE) between its actual and target outputs by adjusting weights and biases.

Selecting a proper algorithm for achieving this purpose has become a challenge for researchers. Back-propagation (BP) algorithm is one of the most popular algorithms which has been proposed by researchers as a training phase. After some time, researchers have pointed out that the BP algorithm based on gradient descends have some drawbacks. Slow convergence rates and trapping in local minima are some of the important drawbacks.

Recently, Meta-heuristic algorithms are known for their ability to produce optimal or near-optimal solutions for optimization problems [22]. In this paper, we utilized WCO algorithm to search for weight values as below:

At first, ANN is trained using WCO algorithm to find the optimal initial weights. After that, the neural network is trained by using a back-propagation algorithm which involves an optimal back-propagation network.

Check whether the network has achieved the considered error rate or the definite number of generations has been reached then to end the algorithm.

For representing the ANN, a two-layered network can be considered as follows:

$$\sum_{i=1}^{H} w_i \sigma \left(\sum_{j=1}^{d} w_j x_j + b \right) \tag{3}$$

where H illustrates the number of neurons in the hidden layer, w is the network weights, b denotes the value of the bias and σ is the activation function of each neuron which is considered as sigmoid in this case.

The network is trained by employing the WCO algorithm to achieve the value of the weights for each node interconnection and bias terms until the output layer neurons values are as close as possible to the actual outputs. The mean squared error of the network (MSE) can be defined as below:

$$MSE = \frac{1}{2}\sum_{k=1}^{g}\sum_{j=1}^{m}\left(Y_j(k) - T_j(k)\right)^2$$

(4)

Here m is the number of nodes in the output, g is the number of training samples, $Y_j(k)$ defines the desired output, and $T_j(k)$ is the real output.

The procedure for this HNNGWO algorithm can be summarized as follows:

1. Initialize the whole teams and groups randomly in the range of [0, 1].

2. Evaluate each initialized team's fitness value

3. Find the best team with the highest score based on its rank, competitions and other operators

4. If the maximal iterative generations are achieved, go to (7), else, go to (5).

5. Update and repeat the competition based on the previous ranks

6. Utilize the backpropagation algorithm to search around the best cost for some epochs; if the search result is better than the best cost, the output will be the achieved search result; otherwise, previous output will be selected.

7. End of algorithm

7 Dataset description

Different databases are employed to analyze and compare the proposed technique results with other methods for performance analysis. Major images are acquired from Australian Cancer Database (ACD) as a well-known and broadly used skin cancer database. The main purpose of this research is to diagnose cancer in the skin from skin cancer images. In the following, we will show the results of the proposed method.
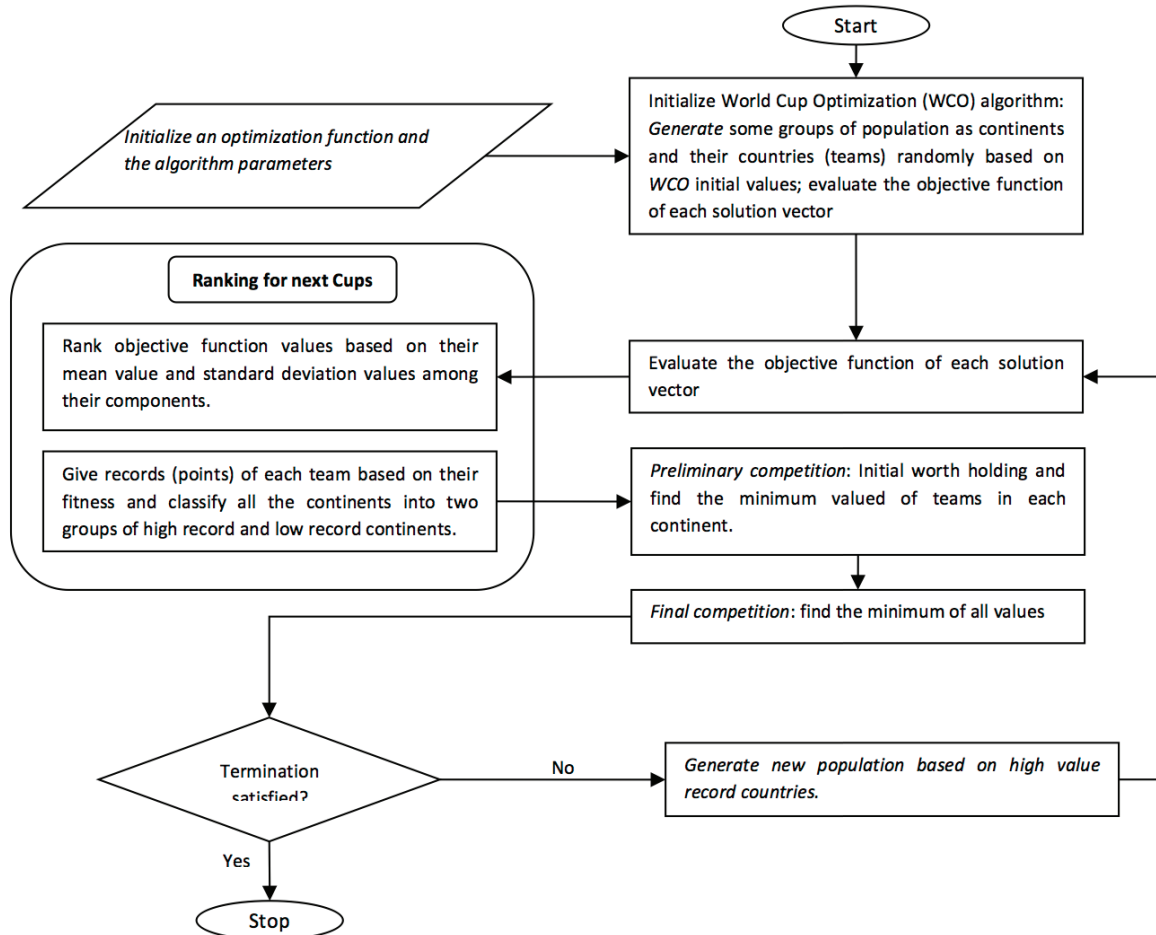

Figure 3: Flowchart of world cup optimization algorithm.

8 Simulation results

Here, we considered two area for classification (cancer and healthy). The proposed method is based on pixel classification for classifying pixels independently from the neighbors. The input layer of the network comprises 3 neurons from each image either cancer or non-cancer image. In this study, a sigmoid function is used as the activation function of the MLP network. The output is between 0 and 255 (uint8 class).

After training the neural network and entering the input images into it, a single threshold value is used to characterize cancer and non-cancer pixels. Here, to analyzing the proposed method's efficiency, three performance metrics are introduced. Correct detection rate (CDR) is the first metric which is defined in Eq. (5). False acceptance rate (FAR) illustrates the percentage of identification moments in which false acceptance happens. False rejection rate (FRR) is the percentage of identification moments in which false rejection happens. The FAR and FRR are defined in Equations (6) and (7), respectively:

$$CDR = \frac{No.of.Pixels.Correctly.Classified}{Total.Pixels.in.the.Test.Dataset} \tag{5}$$

$$FAR = \frac{No.of.non-cancer.Pixels.Classified.as.cancer.PixelsClassified}{Total.Pixels.in.the.Test.Dataset} \tag{6}$$

$$FRR = \frac{No.of.cancer.Pixels.Classified.as.non-cancer.PixelsClassified}{Total.Pixels.in.the.Test.Dataset} \tag{7}$$

Fig.6. shows some examples of the input skin image and their output as the melanoma detected regions:

Table.1 presents the efficiency of the presented segmentation algorithm inaccuracy.

Table.1: Classification comparison of performance in the proposed method

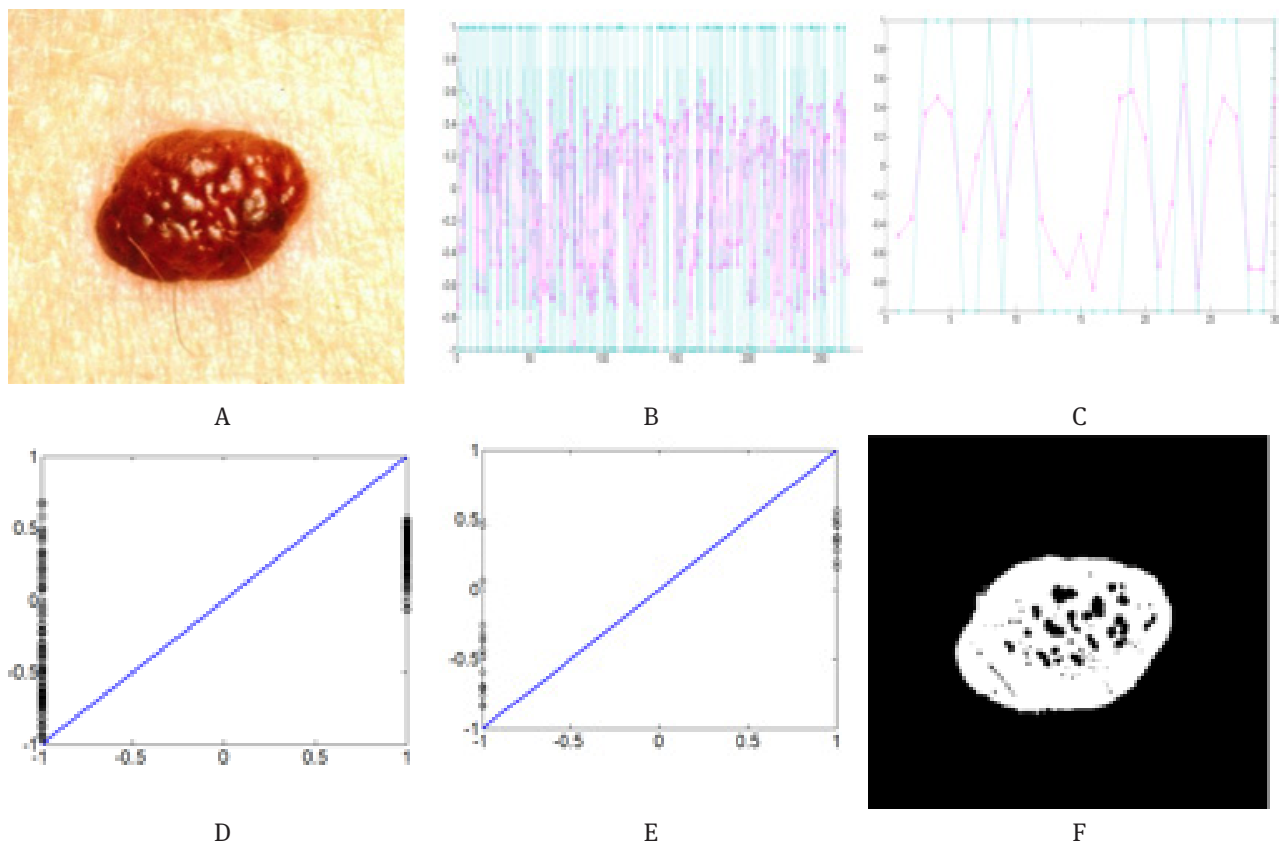| Metric	Standard MLP	MLP-WCO
CDR(%)	88	92
FAR(%)	7.5	4.5
FRR(%)	4.5	3.5

A

B

C

D

E

F

Figure 4: (A) input image, (B) and (C) train and test approximation resemblance (red: original and blue: approximation) respectively, (D) and (E) classified train and test data, (F) output melanoma segmented image
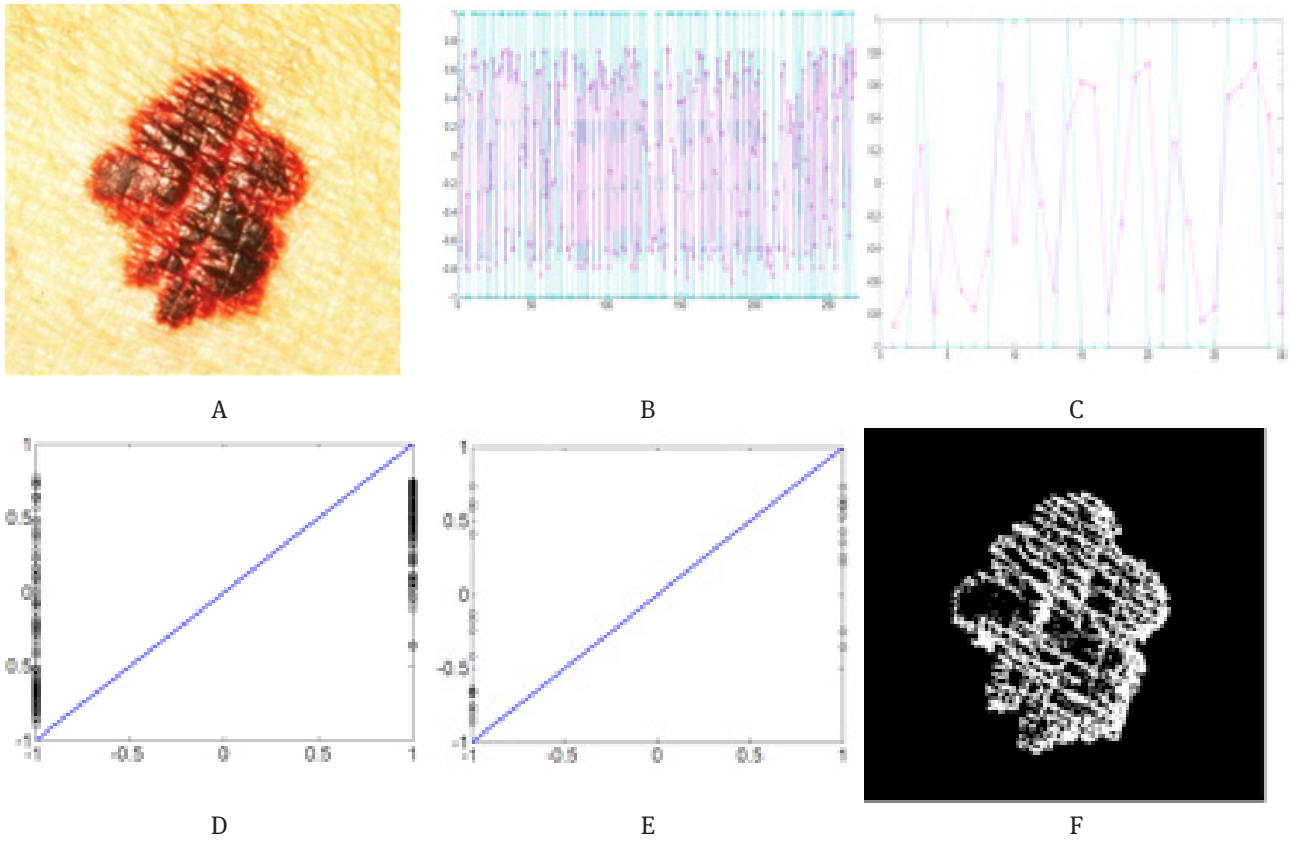
Figure 5: (A) input image, (B) and (C) train and test approximation resemblance (red: original and blue: approximation) respectively, (D) and (E) classified train and test data, (F) output melanoma segmented image
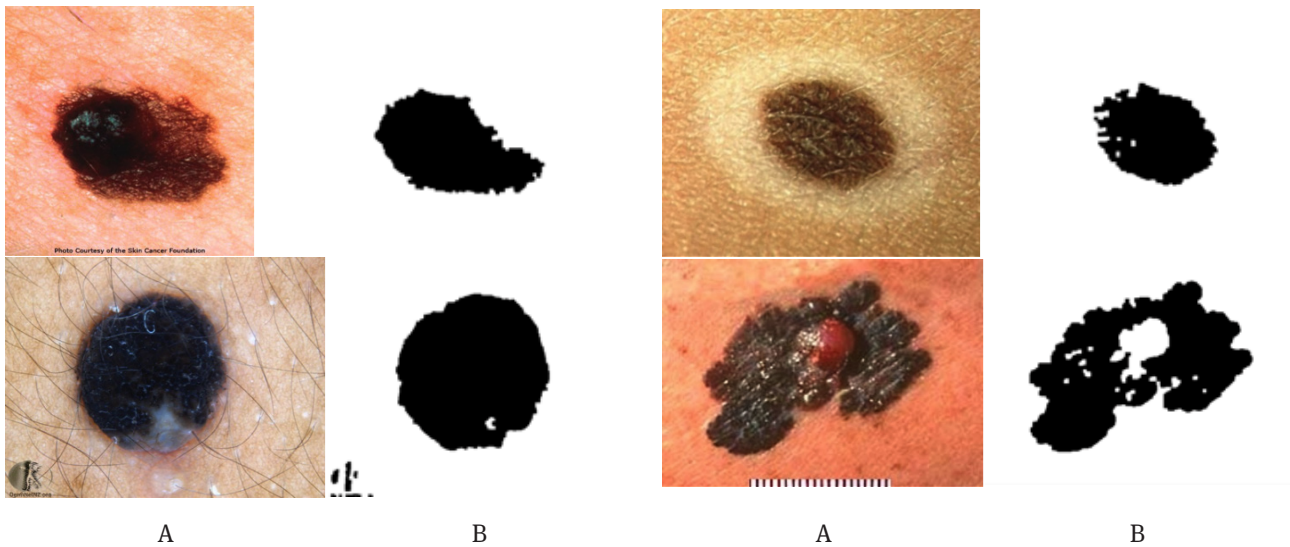

Figure 6: Some of the results of the algorithm: (A) original image and (B) segmented image.

We can see from the above results that the proposed algorithm has better efficiency in the accuracy. It is obvious from the above that MLP-WCO has better performance accuracy.

9 Conclusions

A new optimized method is proposed for diagnosing melanoma. The proposed method is a new hybrid algorithm between the artificial neural network and world cup opti-

mization for enhancing the back-propagation algorithm efficiency and for escaping from trapping in the local minima. Simulation results showed that WCO helps ANN to find the optimal initial weights and to speed up the convergence speed and reduce the RMSE error. To compare the performance of the proposed method by the ordinary ANN, three metrics (CDR, FAR and FRR) are employed and the results show good efficiency for the proposed ANN-WCO algorithm toward ordinary ANN.

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Razmjooy, 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

[2] Lie, W.-R., Lipsey, J., Warmke, T., Yan, L., and Mistry, J., Quantitative protein profiling of tumor angiogenesis and metastasis biomarkers in mouse and human models, ed: AACR, 2014

[3] Rashid Sheykhahmad, F., Razmjooy, N., and Ramezani, M., A Novel Method for Skin Lesion Segmentation, Int. J. Inf., Sec. Sys. Manage., 2015, 4(2), 458-466

[4] Parsian, A., Ramezani, M., and Ghadimi, N., A hybrid neural network-gray wolf optimization algorithm for melanoma detection, Biomed. Res., 2017, 28(8)

[5] Razmjooy, N., Ramezani, M., and Ghadimi, N., Imperialist competitive algorithm-based optimization of neuro-fuzzy system parameters for automatic red-eye removal, Int. J. Fuzzy Syst., 2017, 19(4), 1144-1156

[6] Patwardhan, S. V., Dhawan, A. P., and Relue, P. A., Classification of melanoma using tree structured wavelet transforms, Comput. Methods Programs Biomed., 2003, 72(3), 223-239.

[7] Garg, N., Sharma, V., and Kaur, P., Melanoma Skin Cancer Detection Using Image Processing, in Sens. Image Proc., ed: Springer, 2018, pp. 111-119

[8] Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., *et al.*, Segmentation of skin cancer images, Image Vis. Comput., 1999, 17(1), 65-74

[9] Ganster, H., Pinz, P., Rohrer, R., Wildling, E., Binder, M., and Kittler, H., Automated melanoma recognition, IEEE Trans. Med. Imaging, 2001, 20(3), 233-239

[10] Zagrouba, E. and Barhoumi, W.,A prelimary approach for the automated recognition of malignant melanoma, Image Analys. Stereology, 2011, 23(2), 121-135

[11] Ghadimi, N. and Ojaroudi, M., A novel design of low power rectenna for wireless sensor and RFID applications, Wirel. Pers. Commun., 2014, 78(2), 1177-1186

[12] Celebi, M. E., Aslandogan, Y. A., and Bergstresser, P. R., Unsupervised border detection of skin lesion images, in Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on, 2005, pp. 123-128

[13] Zouridakis, G., Doshi, M., and Mullani, N., Early diagnosis of skin cancer based on segmentation and measurement of vascularization and pigmentation in nevoscope images, in Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE, 2004, pp. 1593-1596

[14] Fassihi, N., Shanbehzadeh, J., Sarrafzadeh, H., and Ghasemi, E., Melanoma diagnosis by the use of wavelet analysis based on morphological operators, 2011

[15] Moallem, P., Razmjooy, N., and Ashourian, M., Computer vision-based potato defect detection using neural networks and support vector machine, Int. J. Robot. Autom., 2013, 28(2), 137-145

[16] Razmjooy, N. and Ramezani, M., Training Wavelet Neural Networks Using Hybrid Particle Swarm Optimization and Gravitational Search Algorithm for System Identification

[17] Mousavi, B. S., Soleymani, F., and Razmjooy, N., Color image segmentation using neuro-fuzzy system in a novel optimized color space, Neural Comput Appl, 2013, 23(5), 1513-1520

[18] Razmjooy, 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

[19] Moallem, P. and Razmjooy, 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

[20] Ghadimi, N., An adaptive neuro-fuzzy inference system for islanding detection in wind turbine as distributed generation, Complexity, 2015, 21(1), 10-20

[21] Razmjooy, N., Khalilpour, M., and Ramezani, M., A New Meta-Heuristic Optimization Algorithm Inspired by FIFA World Cup Competitions: Theory and Its Application in PID Designing for AVR System, J. Control Autom. Elect. Syst., 2016, 27(4), 419-440

[22] Anoraganingrum, D., Cell segmentation with median filter and mathematical morphology operation, in Image Analysis and Processing, 1999. Proceedings. International Conference on, 1999, pp. 1043-1046.

[23] Erhan, D., Szegedy, C., and Anguelov, D., Training a neural network to detect objects in images, ed: Google Patents, 2016

[24] Mousavi, B. S. and Soleymani, F., Semantic image classification by genetic algorithm using optimised fuzzy system based on Zernike moments, Signal Image Video Process., 2014, 8(5), 831-842

[25] Manafi, H., Ghadimi, N., Ojaroudi, M., and Farhadi, P., Optimal placement of distributed generations in radial distribution systems using various PSO and DE algorithms, Elekt.Elektrotech., 2013, 19(10), 53-57

[26] Moallem, P. and Razmjooy, N., Optimal threshold computing in automatic image thresholding using adaptive particle swarm optimization, J. Appl. Res. Tech., 2012, 10(5), 703-712

[27] Razmjooy, N. and Ramezani, M., An Improved Quantum Evolutionary Algorithm Based on Invasive Weed Optimization, Indian J. Sci. Res, 2014, 4(2), 413-422