*Research Article*

# Skin Cancer Detection Based on Extreme Learning Machine and a Developed Version of Thermal Exchange Optimization

**Shi Wang**[1] **and Melika Hamian** 🆔 [2]

[1]*Department of Computer Engineering, Dongguan Polytechnic, Dongguan 523808, Guangdong, China*
[2]*Department of Engineering, Payame Noor University (PNU), Tehran, Iran*

Correspondence should be addressed to Melika Hamian; mhamian92@gmail.com

Melanoma is defined as a disease that has been incurable in advanced stages, which shows the vital importance of timely diagnosis and treatment. To diagnose this type of cancer early, various methods and equipment have been used, almost all of which required a visit to the doctor and were not available to the public. In this study, an automated and accurate process to differentiate between benign skin pigmented lesions and malignant melanoma is presented, so that it can be used by the general public, and it does not require special equipment and special conditions in imaging. In this study, after preprocessing of the input images, the region of interest is segmented based on the Otsu method. Then, a new feature extraction is implemented on the segmented image to mine the beneficial characteristics. The process is then finalized by using an optimized Deep Believe Network (DBN) for categorization into 2 classes of normal and melanoma cases. The optimization process in DBN has been performed by a developed version of the newly introduced Thermal Exchange Optimization (dTEO) algorithm to obtain higher efficacy in different terms. To show the method's superiority, its performance is compared with 7 different techniques from the literature.

## 1. Introduction

The skin cancer has been known as one of the deadliest types of cancers which has recently witnessed an exponential growth worldwide. With all these interpretations, it can be cured if it is detected in the early stages, and, in most cases, a simple biopsy can prevent cancer from growing [1]. This cancer has been growing dramatically during the recent years and the importance of the initial therapy has been more to be seen. The melanoma skin cancer has been ranked as the 3rd most common malignant cancer among the skin cancer cases [2]. This cancer alters the skin color as a result of the irregular effects of the cells [3]. Notwithstanding this danger, it is known as a curable cancer if it is detected in the early stages. With all these interpretations, early detection of the melanoma from the other types of the benign skin moles is a thought-provoking mission. Any variations in the size, color, or the shape of a skin mole can be considered as initial signs for this cancer.

Based on the statistical information for 2019, skin cancer with 15000 cases is ranked as the fourth most common cancer, and with 1900 deaths cases are ranked as one of the 10th common cancers in the world [4]. The most common method for melanoma detection is by the physicians and their sampling and testing. However, in some cases, even experienced specialists make some mistakes in melanoma detection. Furthermore, laboratory sampling and testing is a time-consuming and expensive experiment that bothers the patients. Therefore, designing a (computer) system that can detect malignant lesions will be very useful. According to researches, the effectiveness of the computer-aided system is more accurate than that of a specialist doctor. Recently, many research works are accomplished for automatic early detection of skin cancers. Also, the high capability of the medical imaging assists the medical doctors to lessen the diagnosis difficulty of this work and to enhance the speed of the initial diagnosis of the disease.

The application of Artificial Neural Networks (ANNs) in different parts of image processing and machine vision has been increasing. The ANNs are imitated from the human brain actions and are used in a variety of applications, from medicine to economics and engineering [5].

The ANN is made up of lots of highly unified processing components called neurons which can solve problems with their interconnectedness. Artificial Neural Networks, like humans, learn by example to solve different kinds of problems from pattern recognition to information classification throughout a learning procedure.

Deep learning is a sort of ANNs which uses mathematical techniques to give a structure like the human brain. On the other hand, advances in technology have led to algorithms for optimizing normal neural networks so that we can count the number of neural layers. Neural networks are increased from multiple layers to thousands of layers and thousands of neurons in each layer, which until a few years ago could not have created such a structure [6]. This type of neural network is called a deep learning network. Several methodologies of using different kinds of ANNs and other machine learning techniques were introduced for skin cancer diagnosis [7, 8]. For instance, Li et al. [9] proposed a new data synthesis methodology to combine individual skin lesions images and heavily expand them to create noteworthy amounts of data. The study worked on a convolutional neural network (CNN) to provide superior efficiency to the old-style detection and tracking methods. Moreover, the system was trained by humans with simple criteria.

Esteva et al. [10] presented a supervised technique for skin lesions using deep learning. The method's efficiency was verified by twenty-one medical images with two critical binary classifications: malignant and benign groups. Two cases were analyzed here. The first one recognizes the most common cancers, and the second case signifies the deadliest skin cancer identification. The final results indicated the high performance of the proposed method.

Jafari et al. [11] proposed a deep learning-based procedure to provide a precise skin cancer area segmentation. After image denoising and preprocessing, a CNN was utilized for segmentation. They combined local and global structure information and outputted a label for the pixels to create a segmentation mask that indicates the lesion area. This mask was then improved with some postprocessing operators. The empirical results indicated the superiority of the presented method compared to the other techniques.

Mohamed et al. [12] proposed a learning technique to categorize skin lesions to determine melanoma. The method was based on applying a CNN with multiple outlines. The results of the suggested technique were performed by 14 layers into the International Skin Imaging Collaboration (ISIC) dataset. Final results indicated successful ratio for the method in the segmentation of the skin cancer area.

Xu et al. [2] proposed an automatic diagnosis method for melanoma. The presented technique was based on an optimized image segmentation methodology along with a CNN optimized by satin bowerbird optimization (SBO). After segmentation, feature extraction was implemented on the processed image. The main features were then selected using the SBO algorithm to trim additional information. At last, the results were presented by performing an SVM into two groups of normal and melanoma cases. The method's results were authenticated by comparison with other techniques from the literature to indicate its performance.

As the research background shows, the application of deep learning, including CNNs in medical images processing, especially in skin lesion diagnosis, has been exponentially increasing. The present study proposes an optimized technique of skin lesion segmentation using a new design of the Thermal Exchange Optimization algorithm which has been used to reach this purpose with higher performance. The overall technique for the proposed technique has been shown in Figure 1.

Therefore, the main contributions of the present study can be highlighted as follows:

(i) A new computer-aided method for the diagnosis of malignant melanoma

(ii) Segmenting the region of interest based on the Otsu method

(iii) Feature extraction to extract the beneficial features

(iv) A new optimized Deep Believe Network (DBN) to classify the melanoma cases

(v) Optimization based on a developed version of the newly introduced Thermal Exchange Optimization (dTEO) algorithm

## 2. Skin Cancer Preprocessing

Medical imaging is the method and process used to create images of the human body to examine or advance medical science and is widely used in medicine [13]. Diagnostic images are the subject of many scans, examinations, and images used in medicine. Researches have shown that medical images taken are highly susceptible to vulnerability to noise and other factors such as low contrast. Here are two simple ways to do this purpose.

### 2.1. Noise Reduction.
Images noise removal is a significant part of all medical image processing. This process tries to reduce or remove the noise from the original image. Features that expose them to noise are available on all recording devices [14]. The noise may be white or random noise. Noise reduction recreates and restores the noisy image to its original model based on ideal models in which we analyze how this method is done. The noise has a major impact on the image in certain situations, particularly through the identification of the image edge that needs differentiation which increases the influence of high-frequency pixels that specifically contain noise [15, 16].

The noise in medical images is typically Gaussian noise. Gaussian noise is dispersed in the image and is located between the image main pixels, which makes some inconsistencies for the next operations, and because the processing must be done only on the image main pixels, it requires a proper noise elimination for this purpose. Noise pixels are usually recognized as a different pixel among a group of
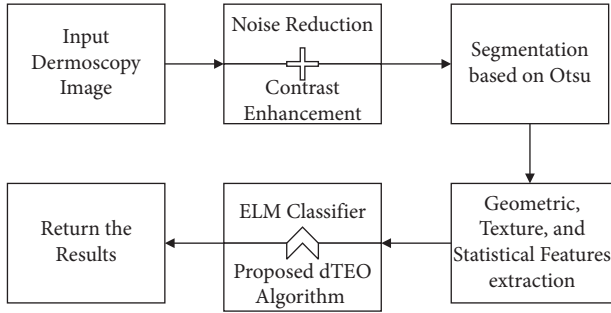
FIGURE 1: The overall procedure of the suggested method.

pixels in the original image that differs from adjacent pixels in different terms such as light intensity and transparency. Wang-Mendel algorithm is an efficient fuzzy-based noise removal tool [17]. Because of the simple conception of the fuzzy techniques, they are popular for different applications. Furthermore, because of their speediness as an early-stage process, they are very useful for producing the initial fuzzy model [18]. In this method, the input and the output datasets show the behavior of the solved state of the problem. The production of the rule database is performed as follows:

(1) Consider fuzzy isolation from the space that includes the input variables, which is achieved by the knowledge of an expert along with normalization. Then, it splits it into equivalent or unequal sections by doing a fuzzy-based method for parting the space of input variable. Then a form of the membership function is chosen and each component is given as a fuzzy package. Afterward, the membership and fuzzy set has been established to all the parts.

(2) A set of candidate language rules are produced, which are generated by choosing the most inclusive rules for the samples.

(3) Validity degree is assigned for the laws, which is achieved by multiplying the membership function values for all the components.

(4) The final rules database is obtained from the rules of the candidate language set, which is done by categorizing these rules into some groups, where the groups include the rules of candidates with the same assumption. The final rule is obtained by selecting the highest degree for all the groups.

*2.2. Image Contrast Enhancement.* Dermoscopy images typically lack high image contrast, which makes them impossible to process next stages. The concern arises from multiple circumstances, such as the poor quality of measuring instruments and cameras, the low degree of the user interface in photography, the environmental factors, and the prevalence of noise. Sometimes, any crucial details from the images disappear from the above examples, rendering the processing too complex [19]. Contrast enhancement is a method for addressing the problems of contrast consistency. In this research, image contrast enhancement is used to enhance and help demonstrate the specifics of the cancer

regions. Global enhancement of contrast using the lookup table is used in the present research. For categorizing and storing the received images on the disk, an 8-bit lookup table is used. The strategy is usually developed as follows [20]:

$$PDF_{out} = \frac{PDF_{In} - PDF_{min}}{PDF_{Max} - PDF_{min}}, \tag{1}$$

where $PDF_{In}$ and $PDF_{out}$ represent the probability density function of the input and the output corrected images, respectively, and $PDF_{min}$ and $PDF_{Max}$ represent the minimum and the maximum probability density levels, respectively. Figure 2 shows an example of the preprocessed image for more clarification.

## 3. Skin Cancer Segmentation

*3.1. Color Space.* The RGB is the most common and basic color space in images processing, which usually is used for displaying images. Due to the dependence of these three colors, which are considered as base colors in the RGB color space, as well as their dependence on the intensity of ambient light, this color space is less used for main processing. To compensate for this shortcoming, the XYZ color model has been employed. This color space is a quantifiable link between the physiologically supposed colors in human color vision and wavelengths distributions in the electromagnetic visible spectrum . In this color space, the values of $X$ and $Z$ give color information, and $Y$ defines the luminance. To convert color space from RGB to XYZ, the following equation has been used [21]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \times \begin{bmatrix} 0.49 & 0.31 & 0.2 \\ 0.17697 & 0.8124 & 0.01063 \\ 0 & 0.01 & 0.99 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{2}$$

The most significant benefit of $XYZ$ color model is that it is completely independent of the device.

*3.2. Method of Segmentation.* In dermoscopy images, the red color space is the only dimension in RGB color space that gives nearly the strength of the images. As mentioned before, both $X$ and $Z$ values give related information from the color for the XYZ color space. Therefore, the red (R) dimension and the $X$ dimension are just normalized for the segmentation; that is [21],

$$\widehat{R} = \frac{R}{\sqrt{R^2 + G^2 + B^2}},$$
$$\widehat{X} = \frac{X}{\sqrt{X^2 + Y^2 + Z^2}}. \tag{3}$$

On any pixel of the input images, these normalization values are performed. After this normalization, the Otsu thresholding is used for giving a low-cost segmentation in terms of time complexity.
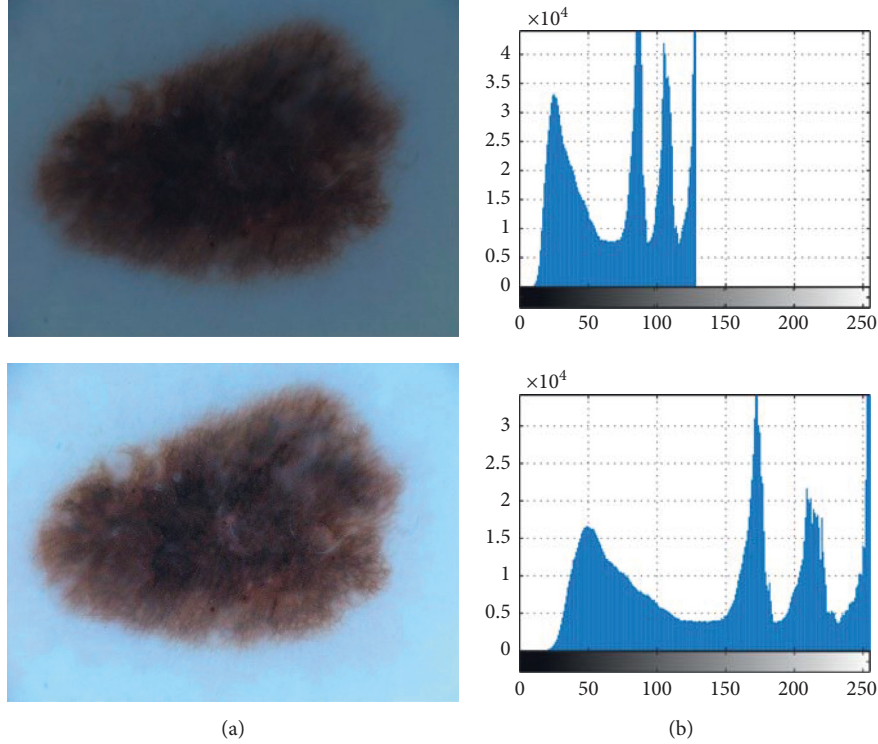
FIGURE 2: Two examples of the preprocessing of the image: (a) original image; (b) PDF of (a).

The Otsu technique is a popular method for proper thresholding of the images. The method is based on the conception of the variance of the intergroup maximization and the intragroup minimizing of the pixels. The global threshold has a problem when the image background resolution is inadequate. A local threshold can be used to eliminate the heterogeneity effect. This problem is solved by image preprocessing to eliminate inhomogeneities and apply a global threshold to the processed image. Based on Otsu technique, threshold level is comprehensively searched to provide minimum in-between variance class, which is expressed as follows:

$$v(t) = \omega_1(t)v_1^2(t) + \omega_2(t)v_2^2(t), \tag{4}$$

where $\omega_i$ denotes the probability of two separate classes with a threshold value of $t$. $\sigma_i^2$ in this case is the value of the variance of these classes. In fact, Otsu shows that minimizing the amount of variance in a class is like maximizing interclass variance:

$$v_b^2(t) = v^2 - v_\omega^2(t) = \omega_1(t)\omega_2(t)[m_1(t) - m_2(t)]^2, \tag{5}$$

where $v_i^2$ describes the variance and $m_i$ signifies the mean value that can be updated intermittently. This technique is briefly given as follows:

(i) Evaluate image histogram and its intensity level probability

(ii) Give initial values of $\omega_i(0)$ and $v_i(0)$ for all levels of possible threshold ($t = 1, 2, \ldots$)

(iii) Renew $\omega_i$ and $m_i$

(iv) Calculate $v_b^2(t)$

The optimal threshold here is the maximum of $\sigma_b^2(t)$.

Then, a morphology operation including closing, opening, and filling is used for better results. In this stage, by applying the mathematical filling operator, the extra image holes are filled. This operator is as follows [22]:

$$X_k = (X_{k-1} \oplus D) \cap C^c, \quad k = 1, 2, 3\ldots, \tag{6}$$

where $A$ and $B$ represent the area and the constructing component, respectively.

Then, the opening operation has been established on the filled image for eliminating the ignitor particulars with no changes on other gray surfaces. This operator is as follows:

$$C \circ D = (C \ominus D) \oplus D. \tag{7}$$

Finally, for connecting the narrow parts, the closing operation is performed based on the following equation:

$$C \cdot D = (C \oplus D) \ominus D. \tag{8}$$

Here, the structural element in this study is set as a $5 \times 5$ identity matrix. Algorithm 1 shows the pseudocode of the image segmentation.

Figure 3 displays some samples of image segmentation using the suggested segmentation.

## 4. Features Extraction

Feature extraction is the process of collecting more detailed information from the image. The features extraction in this part of processing is utilized for extracting the main features

Input: the preprocessed skin dermoscopy image
*Step 1.* Normalize $R$ dimension from RGB color space
*Step 2.* Convert RGB color space to XYZ
*Step 3.* Normalize $X$ dimension from XYZ color space
*Step 4.* Apply Otsu threshold algorithm
*Step 5.* Apply postprocessing method based on mathematical morphology
Output: the segmented image

ALGORITHM 1: Pseudocode of the proposed image segmentation.



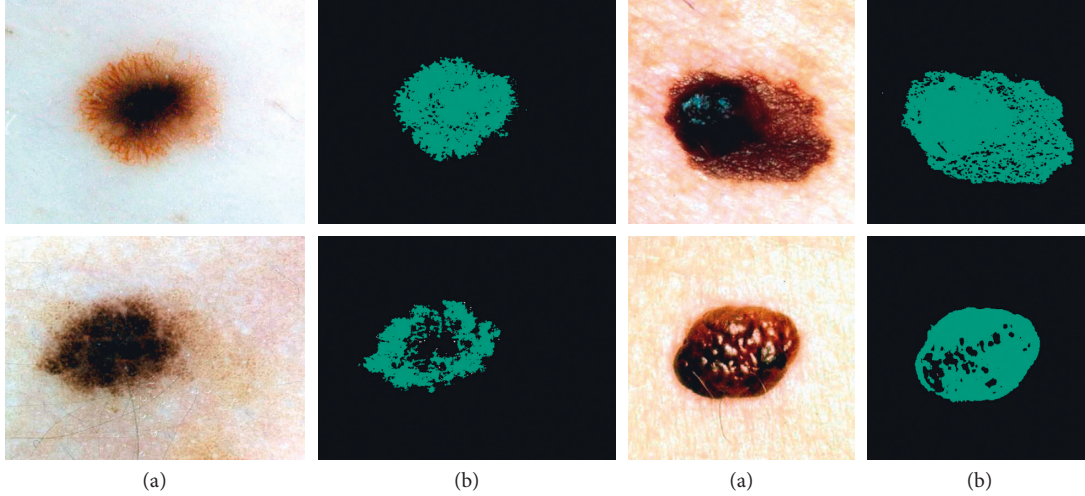|       (a)       |       (b)       |       (a)       |       (b)       |

FIGURE 3: Some samples of the skin cancer segmentation using the suggested technique: (a) the input image and (b) the segmented image.

of the segmented skin cancer area to simplify the diagnosis successful. There are numerous features for feature extraction. This research uses statistical features and texture features, and geometric features have been adopted for this purpose. Table 1 gives the features utilized here.

In the above table, MN describes the size of image, $B_p$ describes the length of external side in the pixel boundary, $p(i, j)$ represents the intensity values of the pixels at position, $\mu$ defines the mean value, $\sigma$ describes the standard deviation, and $a$ and $b$ define the major axis and the minor axis, respectively. However, some achieved features are not useful for the process and just increase the method's complexity. To prevent this issue, a feature selection procedure has been utilized. To do so, in this study, an optimized technique has been used. For giving optimal results of features selection, the following cost function has been considered:

$$\text{fitness} = \frac{(T_P \times T_N) - (F_P \times F_N)}{\sqrt{((T_N + F_P) \times (T_P + F_P) \times (T_P + F_N) \times (T_N + F_N))}},$$

$$(9)$$

where $F_P$ describes the false positive, $F_N$ describes the false negative, $T_P$ defines true positive, and $T_N$ defines the true negative. The idea is to minimize the above equation based on a developed optimization method.

## 5. Extreme Learning Machine (ELM)

One of the neural network models which has recently received many considerations is the ELM model. Some of the advantages for ELM are their high speed in learning, their simple using, and their ability to be used in numerous activation functions and nonlinear kernel functions. The ELM model can provide an integrated template with a wide variety of feature transmissions that can be used in the hidden layer, which can be utilized directly for regression and categorization. This algorithm is a learning method for Single Hidden Neural Networks (SHNN) based on initializing the input biases and weights randomly and the output weights evaluation. Thus, the network can be trained in just a few steps. Figure 4 shows the arrangement of a simple ELM network.

For a classification problem with $D$ dimension and $N$ number of training samples [23],

$$\left(x^{(n)}, t^{(n)}\right), \quad n = 1: N, \qquad (10)$$

where $x^{(n)} \in \mathbb{R}^D$ and $t^{(n)} \in \mathbb{R}^K$.

An ELM-based feed-forward neural network can be formulated as follows:

$$\sum_{m=1}^{M} \beta_m g\left(w_m^T x^{(n)} + b_m\right) = t^{(n)}, \qquad (11)$$

where $g(\cdot)$ describes the activation function, $b_m$ describes the bias of the $m^{\text{th}}$ hidden neuron, $M$ signifies the hidden neurons, $w_m = [w_{m_1}, w_{m_2}, \ldots, w_{mD}]$ determines the input weight vector which connects the input neurons to the $m^{\text{th}}$ the neuron of the hidden layer, and $\beta_m = [\beta_{m_1}, \beta_{m_2}, \ldots, \beta_{mK}]$ determines the vector of weight that connects the $m^{\text{th}}$ neuron.

This conception is given as follows:

$$H\beta = T, \tag{12}$$

where

$$H = \begin{bmatrix} g\left(w_m^T x^{(1)} + b_1\right) & \cdots & g\left(w_M^T x^{(1)} + b_1\right) \\ \vdots & \ddots & \vdots \\ g\left(w_1^T x^{(N)} + b_1\right) & \cdots & g\left(w_M^T x^{(N)} + b_M\right) \end{bmatrix}_{N \times M}, \tag{13}$$

$$H = \left[\beta_1^T, \beta_2^T, \ldots, \beta_M^T\right]_{M \times N}^T,$$

$$T = \left[t_1^T, t_2^T, \ldots, t_M^T\right]_{N \times K}^T.$$

Meanwhile, $H$ defines a nonsquare matrix. The training samples quantity is bigger than the hidden neurons quantity. Therefore, to solve this problem, the following formulation has been used:

$$\widehat{\beta} = H^\dagger T, \tag{14}$$

where $H^\dagger$ describes the generalized Moore-Penrose matrix inverse.

Thus, the ELM networks can be shortened to three main steps:

(1) Initializing the input weights and bias with random values

(2) Evaluation of the hidden neurons' outputs matrix $H$

(3) Evaluation of the hidden neurons' outputs weights matrix by equation (16)

# 6. The Developed Thermal Exchange Optimization Algorithm

Optimization is the process of using specific techniques to solve the optimization problems. However, in recent years, the performance of the classic optimization methods due to increasing the complexity and nonlinearity of them has been decreased or even failed [24]. Recently, several methods have been introduced for resolving these problems [25]. Metaheuristics are one class of these methods that have been popular due to their simple and suitable results in various problems. The metaheuristic algorithms are derived from various phenomena from nature, human society, animal hunting behavior, and so forth. Several algorithms have been proposed in this field [26–29], for example, World Cup Optimizer [30], Ant Lion Optimizer (ALO) [31], Chimp Optimization Algorithm [32], Harris Hawks Optimization [33], and mayfly optimization algorithm [34]. In the present research, a novel modified model of the Thermal Exchange Optimization (TEO) algorithm has been presented to achieve optimal results for the considered methodology [5].
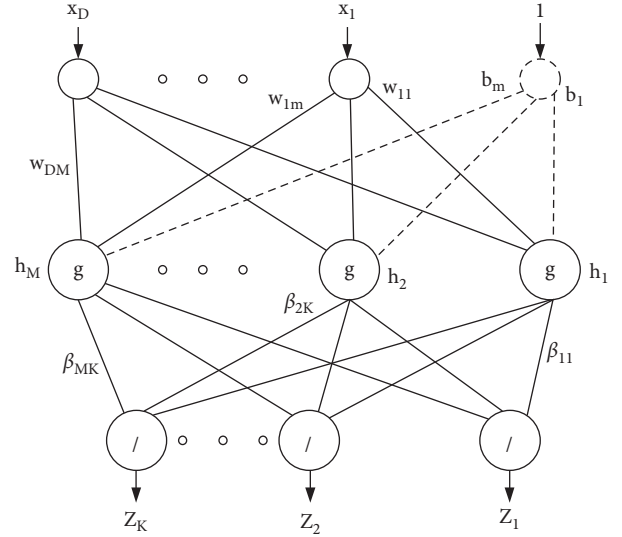


FIGURE 4: A sample configuration of the ELM network.

The TEO algorithm is a novel metaheuristic technique that is inspired by the temperature of the objects and their position that is switched between cold and warm places, indicating the updated positions. In the following, more explanation about this algorithm has been provided.

6.1. The Newton Cooling Law. The tests indicated that the rate of cooling was roughly related to the temperature difference between the environment and the heated object. This is formulated as in the following equation:

$$\frac{dQ}{dt} = \alpha \times A \times (T_s - T_b), \tag{15}$$

where $A$ describes the area of the body surface that transmits heat, $Q$ signifies the heat, $\alpha$ defines the coefficient of heat transfer that is reliant on different cases like object geometry, heat transfer mode, and surface state, and $T_b$ and $T_s$ represent the temperature of the body and the ambient temperature, respectively.

The heat loss in time $dt$ is $\alpha \times A \times (T_s - T) \, dt$, which states the variation in kept heat as the temperature drops $dT$; in other words,

$$V \times d \times c \times dT = -\alpha \times A \times (T - T_b) \, dt, \tag{16}$$

where $d$ signifies the density (kg/m³), $c$ defines the specific heat (J/kg/K), and $V$ describes the volume (m³).

So,

$$\frac{(T - T_b)}{(T_M - T_b)} = \exp\left(\frac{-\alpha \times t \times A}{V \times c \times d}\right), \tag{17}$$

where $T_M$ describes the initial high temperature. Equation (19) is valid if when $\alpha \times A \times t / V \times d \times c$ is $T$-independant,

$$\zeta = \frac{(\alpha \times A)}{(V \times c \times d)}. \tag{18}$$

So, by considering a constant $\zeta$,

$$\frac{T - T_b}{T_M - T_b} = \exp(-\zeta t). \tag{19}$$

So, consequently, by reformulating equation (19),

$$T = (T_M - T_b) \times \exp(-\gamma t) + T_b. \tag{20}$$

*6.2. The Algorithm.* In Thermal Exchange Optimization, some of the candidates have been assumed as the cooling objects and the other remaining candidates are assumed as the environment; afterward, the reverse operation has been done. In TEO algorithm, like any other metaheuristic algorithm, the candidates are first generated randomly. So, the initial temperature for each object is formulated by the following equation:

$$T_i^0 = T_{\min} + \text{rnd} \times (T_{\max} - T_{\min}), \tag{21}$$

where rnd signifies a random vector limited between 0 and 1, $T_i^0$ describes the initial solution vector of the $i^{\text{th}}$ object, and $T_{\min}$ and $T_{\max}$ represent the minimum and maximum limitations of the candidates.

The randomly generated candidates are then performed to the cost fiction to evaluate their cost value. After that, the position of some best $T$ candidate vectors is saved as *Thermal Memory* (TM) to be utilized then for improving the algorithm's efficacy by less computational cost. Afterward, some best TM candidates have been given and the equal numbers of population with the worst values have been eliminated.

The candidates are separated into two equal types that are shown in Figure 5.

For more clarification, $T_1$ determines the environment object for $T_{n/2+1}$ cooling object, and, inversely, if any object has lower value than $\zeta$, the temperature exchanging is performed slowly. Therefore, $\zeta$ will be achieved by the following formula:

$$\gamma = \frac{\text{Cos (object)}}{\text{Cos (worst object)}}. \tag{22}$$

Another important term in the algorithm is time, which is associated with the iteration quantity. This parameter is formulated in the following:

$$t = \frac{\text{iteration}}{\text{max. iteration}}. \tag{23}$$

To give an exploration term to the algorithm, the environmental temperature changing has been modeled, which is given in the following equation:

$$T_i^e = (1 - (m_1 + m_2 \times (1 - t) \times \text{rand})) \times T_i^{'e}, \tag{24}$$

where $m_1$ and $m_2$ describe the control variables and $T_i^{'e}$ signifies the object earlier temperature which has been modified to $T_i^e$.

By considering the former models, the new temperature of the objects is updated as follows:

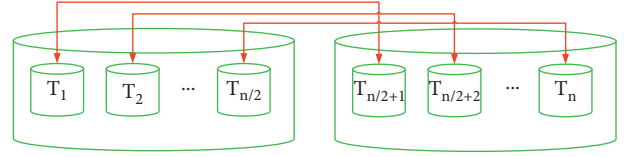$$T_i^+ = T_i^e + (T_i^{\text{old}} - T_i^e)\exp(-\zeta t). \tag{25}$$



FIGURE 5: The cooling and heating transfer and environment objects pairs.

The next term is Pr, which is used to indicate whether the cooling objects' element should be altered or not.

The Pr components are compared with $R(i)\,(i = 1, 2, \ldots, n)$, which has a randomly distributed value between 0 and 1. If $R(i) < \text{Pr}$, the $i^{\text{th}}$ candidate will be randomly chosen as 1-dimensional and the value is reformulated by the following:

$$T_{i,j} = T_j^{\min} + \text{rnd}\left(T_j^{\max} - T_j^{\min}\right)\exp(-\zeta t). \tag{26}$$

where $T_{i,j}$ determines the $j^{\text{th}}$ variable of the $i^{\text{th}}$ candidate and $T_j^{\min}$ and $T_j^{\max}$ are the minor and the major limitations of the variable number $j$, respectively. At last, the algorithm is stopped when stopping criteria are checked to terminate the algorithm.

*6.3. Developed TEO Algorithm.* Although TEO algorithm has numerous benefits in providing the best global solution, it also has some drawbacks which could be resolved. The main drawback of the algorithm could be its local minimum results and low convergence in some cases. Numerous improvements were introduced for resolving the exploration of the metaheuristics. In this research, chaotic theory has been utilized as the first improvement. The reason for using chaos in this algorithm is to improve the algorithm's efficiency to resolve the local optimization followed by an advanced speed of convergence. In the considered dTEO algorithm, the modified randk is replaced with the "rnd," i.e.,

$$\text{rand}_k = \alpha f_k^2 \sin(\pi f_k),$$
$$f_0 \in [0, 1], \, \alpha \in (0, 4], \tag{27}$$

where $k$ signifies the number of iterations.

The other improvement for the algorithm to give a proper trade-off between exploration and exploitation in the algorithm using the Gaussian mutation mechanism is formulated below [35]:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \tag{28}$$

where $\sigma^2$ signifies the variance of the Gaussian PDF and $\mu$ determines the Gaussian distribution expectation. This difference is applied to the considered location by renewing the TEO algorithm ($T_{i,j}^{\text{new}}$) as follows:

$$T_{i,j}^{\text{new}} = T_{i,j} \times (1 + \gamma \times g(0, 1)). \tag{29}$$

where $\gamma$ describes a decreasing random value between 0 and 1 and $g(0, 1)$ describes the standard Gaussian distribution.

*6.4. Algorithm Authentication.* In this section, the proposed developed Thermal Exchange Optimization (dTEO) algorithm that is stated in the previous section has been programmed by the MATLAB 2017b environment and then validated based on a 16 GB RAM Core™ i7/4720HQ 1.60 GHz processor. The proposed dTEO algorithm validation is based on applying it to some standard test cases. The results are also compared with some other approaches, that is, Multiverse Optimizer (MVO) [36], Locust Swarm Optimization (LSO) [37], Spotted Hyena Optimizer (SHO) [38], and the original TEO [39], to show the proposed method's efficacy. Table 2 shows the details of the studied functions.

In Table 3, the parameters settings of all studied algorithms for use in this study are given below.

For determining the performance of the algorithm, they performed on the benchmark functions and two measurement indicators including the minimum value and the standard values are extracted from the results to show the algorithms' accuracy and reliability. It should be noted that, for all algorithms, the population size and the number of iterations are set as 120 and 100, respectively. The simulations have been repeated 20 times to give consistent results. Table 4 illustrates the simulation results.

As can be seen from Table 4, the suggested dTEO has the least value of the minimum value for the minimization problems. This means that the suggested method provides the best accuracy toward the other state-of-the-art methods for solving the studied functions. Plus, based on the results achieved for the standard deviation, it is definite that the suggested method has the minimum value of that. This shows that using the proposed method has the highest reliability against the comparative techniques in this study. So, we can conclude that using this algorithm in this study is a good idea in terms of accuracy and reliability.

## 7. Optimized ELM Network Based on Algorithm

The classification of the final extracted features is performed based on the proposed optimized ELM network. The main task of the ELM classifier here is to determine the skin cancer dermoscopy image into two groups of normal and malignant melanoma classes.

In this study, activation functions of the network are considered for enhancement. Generally, the significant effect of activation function on the ANNs is too clear, because a network with various activation functions gives good generalization capacities. This study considers a sigmoid activation function as one of the popular functions in this area. The formulation for this function is explained as follows:

$$g(X) = \frac{1}{1 + \exp(-(as.x + bs))}. \tag{30}$$

For improving the efficiency of the ELM-based classifier, different sigmoid functions have been utilized in the hidden neurons. In other words, different values are utilized for as and bs. Figure 6 displays the impact of as and bs on the function shape.

In other words, the problem of finding the suitable parameter of the sigmoid function can be considered as an optimization problem. Here, we used a developed Thermal Exchange Optimization algorithm to refine the issue. However, both "as" and "bs" have almost equal effect on the function shape; we used the "bs" for optimization of the network. Therefore, two optimization policies have been used.

The input weights have been first initialized. Then, the developed Thermal Exchange Optimization algorithm is utilized for optimal parameters selection by assuming the initial weights. The objective function of the developed TEO is based on the square error between the output of the desired value and the network by the following:

$$E = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{k} \left( d_{ji} - y_{ji} \right)^2, \tag{31}$$

where $n$ and $k$ describe the numbers of training samples and the output layers, respectively, and $y_{ji}$ and $d_{ji}$ represent the output of the network and the desired value, respectively.

The primary target of the present study is to utilize the mentioned optimized ELM network for diagnosis of the skin cancer. The optimization is using a newly designed metaheuristic, namely, the developed Thermal Exchange Optimization algorithm.

## 8. Results and Discussion

In this paper, a new optimized methodology is used to detect the skin cancer from dermoscopy images.

*8.1. The Database.* For validating the proposed skin cancer identification system, there are many databases. This analysis uses a well-known dataset, called SIIM-ISIC Melanoma. The images are generated in DICOM format to be retrieved by popular libraries and include both image and metadata. Two formats of TFRecord and JPEG are considered for the dataset. The size of the images is uniform ($1024 \times 1024$). Some examples of the utilized images in the study are displayed in Figure 7.

*8.2. Results and Discussion.* As mentioned before, the present study is authenticated by the ACS database and the results have been verified by some different methodologies. The simulations are performed on a laptop with MATLAB 2018b environment with the following information: 16 GB RAM and Core™ i7-4720HQ-1.60 GHz processor. The pipeline methodology of the method is explained in the following:

(1) Start

(2) Apply preprocessing

(2.1) Apply noise reduction

(2.2) Apply image contrast enhancement

(3) Perform skin cancer segmentation

TABLE 1: The utilized features of this study.

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

TABLE 2: The details of the studied functions.

| # | Test function | Minimum value | Boundary |
|---|---|---|---|
| 1 | $F_1 = \sum_{n=1}^{N} x_n^2$ | 0 | $-\infty \le x \le \infty$ |
| 2 | $F_2 = \sum_{n=1}^{N-1} (100 \times [x_{n+1} - x_n^2]^2 + [1 - x_n]^2)$ | 0 | $-\infty \le x_n \le \infty$ |
| 3 | $F_3 = \sum_{n=1}^{N} |x_n| - 10 \cos(\sqrt{|10x_n|})$ | 0 | $-\infty \le x_n \le \infty$ |
| 4 | $F_4 = x \sin(4x) + 1.1y \sin(2y)$ | $-18.5547$ | $0 \le x, y \le 10$ |
| 5 | $F_5 = [\sum_{n=1}^{N} nx_n^4] + N_n(0,1)$ | Varies | $-\infty \le x \le \infty$ |
| 6 | $F_6 = 10N + \sum_{n=1}^{N} [x_n^2 - 10 \cos(2\pi x_n)]$ | 0 | $-\infty \le x_n \le \infty$ |
| 7 | $F_7 = 1 + \sum_{n=1}^{N} \frac{x_n^2}{4000} - \prod_{n=1}^{N} \cos(x_n)$ | 0 | $-\infty \le x_n \le \infty$ |
| 8 | $F_8 = 1/2 + \sin^2 \sqrt{x^2 + y^2} - 0.5/1 + 0.1(x^2 + y^2)$ | $-0.5231$ | $-\infty \le x, y \le \infty$ |

TABLE 3: The utilized parameters settings of all studied algorithms.

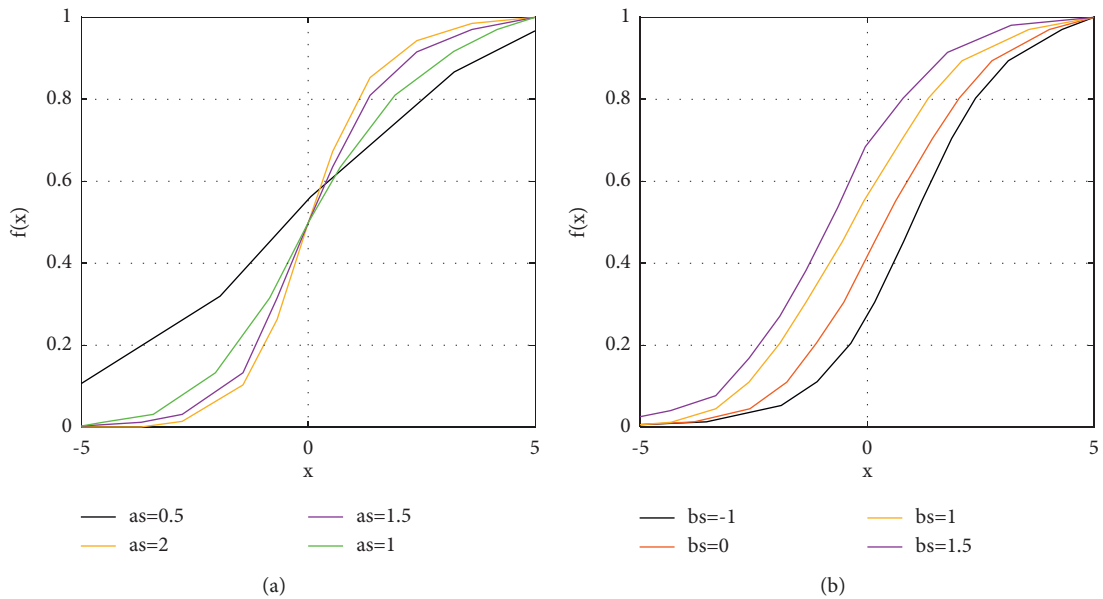| Algorithm | Parameter | Value | Algorithm | Parameter | Value |
|---|---|---|---|---|---|
| | F | 0.6 | ALO [31] | w | [2, 6] |
| LSO [37] | L | 1 | | No. of search agents | 50 |
| | $\vec{g}$ | 20 | | Traveling distance rate | [0.6, 1] |
| SHO [38] | $\vec{M}$ | [0.5, 1] | MVO [36] | Wormhole existence probability | [0.2, 1] |
| | $\vec{h}$ | [5, 0] | | | |



FIGURE 6: The impact of (a) as and (b) bs on the function shape.

FIGURE 7: Some examples of the utilized images in the study.

TABLE 4: The comparison results of the studied algorithms implemented on the test functions.

| Function | | Algorithm | | | | | |
|---|---|---|---|---|---|---|---|
| | | LSO [37] | SHO [38] | MVO [36] | ALO [31] | TEO | dTEO |
| $f_1$ | Min | $3.837e-24$ | $2.1201e-28$ | $-2.2800e-27$ | $3.3197e-28$ | $3.4511e-28$ | $8.2193e-31$ |
| | Std | $2.559e-19$ | $2.4937e-27$ | $4.0865e-28$ | $2.9938e-29$ | $2.0073e-30$ | $2.37928e-32$ |
| $f_2$ | Min | $7.2874e-5$ | $7.564e-4$ | $4.824e-4$ | $2.56384e-5$ | $4.5463e-6$ | $6.7811e-5$ |
| | Std | $3.2084e-6$ | $4.18297e-5$ | $2.1167e-5$ | $3.5918e-6$ | $5.0648e-7$ | $2.1253e-5$ |
| $f_3$ | Min | $-7.2553$ | $-8.15754$ | $-10.00$ | $-7.19376$ | $-10.00$ | $-10.00$ |
| | Std | $0.35$ | $0.46$ | $0.27$ | $0.15$ | $0.15$ | $0.10$ |
| $f_4$ | Min | $-5.2846$ | $-18.053$ | $-17.1146$ | $-16.3927$ | $-18.1106$ | $-18.1683$ |
| | Std | $3.126$ | $1.294$ | $2.391$ | $4.190$ | $1.630$ | $1.01$ |
| $f_5$ | Min | $11.46e-10$ | $2.597e-16$ | $2.18769e-7$ | $5.0913e-9$ | $2.8096e-23$ | $3.7938e-22$ |
| | Std | $8.942e-11$ | $4.0973e-17$ | $2.89430e-8$ | $6.5514e-10$ | $2.8897e-25$ | $7.1937e-24$ |
| $f_6$ | Min | $4.276e-11$ | $2.1953e-12$ | $2.1967e-21$ | $3.1957e-10$ | $2.0054e-18$ | $5.6124e-23$ |
| | Std | $7.297e-12$ | $1.5374e-14$ | $4.2849e-23$ | $6.8091e-11$ | $2.3169e-19$ | $4.6498e-24$ |
| $f_7$ | Min | $2.623e-15$ | $3.3733e-10$ | $5.0406e-8$ | $3.7628e-13$ | $3.6781e-9$ | $6.39482e-17$ |
| | Std | $1.167e-16$ | $4.1967e-12$ | $4.9364e-9$ | $3.1936e-15$ | $4.1945e-10$ | $2.22864e-19$ |
| $f_8$ | Min | $0.0067$ | $-0.1472$ | $-0.3492$ | $-0.5846$ | $-0.5790$ | $-0.5273$ |
| | Std | $0.653$ | $0.467$ | $0.385$ | $0.805$ | $0.152$ | $0.100$ |

TABLE 5: The performance validation of the suggested technique compared with other methods.

| Method | Performance metric | | | | |
|---|---|---|---|---|---|
| | Accuracy | Sensitivity | Specificity | PPV | NPV |
| Fractal analysis [40] | 88.23 | 91.18 | 88.23 | 80.88 | 92.65 |
| CNN [41] | 82.35 | 83.82 | 61.76 | 66.18 | 83.82 |
| Delaunay Triangulation [42] | 86.76 | 89.70 | 86.76 | 77.94 | 88.23 |
| Side-by-Side method [43] | 67.64 | 83.82 | 58.82 | 58.82 | 83.82 |
| Genetic Algorithm [44] | 80.88 | 79.412 | 79.41 | 75.00 | 85.29 |
| Fusion method [45] | 83.82 | 82.35 | 76.47 | 70.58 | 80.88 |
| SVM [46] | 83.82 | 83.82 | 76.47 | 75.00 | 89.70 |
| Proposed method | 92.65 | 91.18 | 89.70 | 86.76 | 94.12 |

(4) Apply features extraction based on texture features, statistical features, and the geometric features

(5) Perform dTEO-based ELM classifier for the final diagnosis

(6) Return the results

(7) End

In the presented dTEO-based ELM classifier, 80% of data are set to train data, and the remaining 20% are set to test the data. The training step of the network is assumed by 700 iterations. The training process has been iterated 15 times to achieve a confidential result. The validation has been performed using 5 measurement indexes: specificity, sensitivity, accuracy, PPV, and NPV, which are given as follows:

$$
\begin{aligned}
\text{accuracy} &= \frac{\text{correctly detected cases}}{\text{total cases}}, \\[2mm]
\text{specificity} &= \frac{\text{correctly detected healthy skin cases}}{\text{total healthy skin cases}}, \\[2mm]
\text{sensitivity} &= \frac{\text{correctly detected skin cancer cases}}{\text{total skin cancer cases}}, \\[2mm]
\text{PPV} &= \frac{\text{correctly detected skin cancer cases}}{\text{detected skin cancer cases}}, \\[2mm]
\text{NPV} &= \frac{\text{correctly detected healthy skin cases}}{\text{detected healthy skin cases}}.
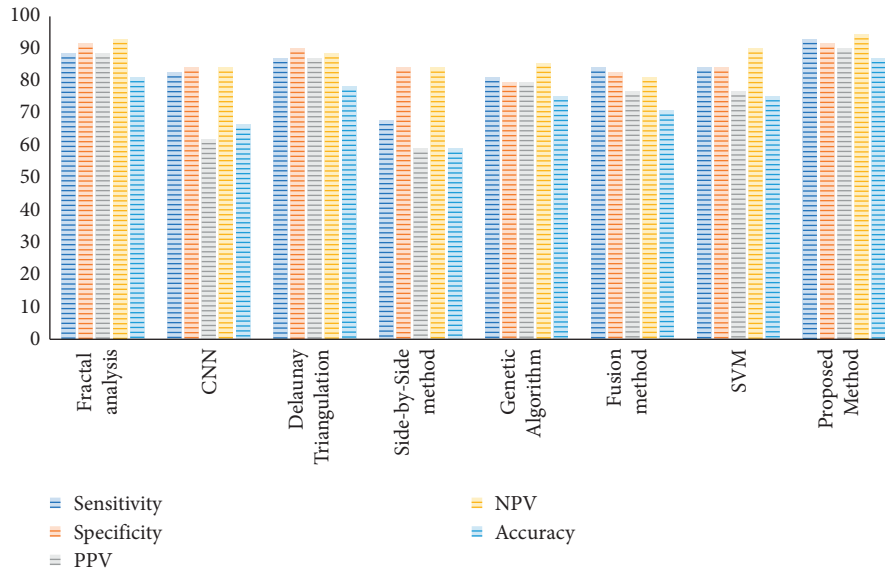\end{aligned}
\tag{32}
$$

FIGURE 8: The bar plot of the performance validation for the proposed method compared with other state-of-the-art methods.

The results are compared with 7 other methods, fractal analysis [40], CNN [41], Delaunay Triangulation [42], Side-by-Side method [43], Genetic Algorithm [44], fusion method [45], and SVM [46], to verify the method's performance in different terms.

Table 5 tabulates the performance validation for the suggested technique compared with other methods. The results show the mean value of 30 runs for the methods.

As is concluded from Table 5, the suggested method with 95% accuracy provides the uppermost precision among the other compared techniques. Also, with 95% of sensitivity, which is the highest against other methods, the reliability of the proposed method has been proved. This can be proved also for NPV, PPV, and specificity toward the others. The better results of NPV and PPV indicate higher commonness of the condition to determine the probability of a test diagnosing cancer. Also, better results of the specificity and sensitivity for the suggested technique show higher prevalence-independent results for the algorithm. For more clarification, the results of Table 5 are shown in Figure 8.

## 9. Conclusions

In recent years, skin cancer has been recognized as the most dangerous and common type of cancer in humans. There are different types of skin cancer. Melanoma is a common type of skin cancer, where early detection can be helpful in its treatment and can suggestively stop death from this deadly skin cancer. Designing an approach that facilitates the skin cancer detection in the early stages is very useful and valuable. In this study, an optimized pipeline procedure was utilized for the optimal detection of melanoma from dermoscopy images. In the proposed method, after preprocessing of the input dermoscopy images based on noise reduction and contrast enhancement, the region of interest was segmented. Afterward,

feature extraction was implemented on the segmented images to extract useful features from them. Finally, an optimized Deep Believe Network (DBN) was utilized to separate images into two classes of healthy and cancerous. The optimization of the DBN was using a new meta-heuristic method called the developed Thermal Exchange Optimization algorithm to improve the network efficiency in terms of reliability and accuracy. So, the main contribution of the proposed method is to use a newly developed version of the newly introduced Thermal Exchange Optimization for the diagnosis of malignant melanoma. The main advantage of using this technique is that the results showed that it improved the system's efficiency in both terms of accuracy and precision or, in other words, its reliability during different runs. The performance of the proposed technique was authenticated by comparing it with 7 other methods: fractal analysis, CNN, Delaunay Triangulation, Side-by-Side method, Genetic Algorithm, fusion method, and SVM. Simulation results specified that, based on some various performance indexes, the suggested procedure gave the best results against the compared techniques. The main limitation of the proposed method is that, due to different soft techniques, it needs a lot of time. In the future study, we will work on designing a method to simplify the method from a theoretical methodology to a real-time system for practical applications.

## Data Availability

The data are available at https://www.kaggle.com/c/siim-isic-melanoma-classification.

## Conflicts of Interest

The authors declare that there are no conflicts of interest in this research.

## Acknowledgments

## References

[1] S. Bao, A. Ebadi, M. Toughani et al., "A new method for optimal parameters identification of a PEMFC using an improved version of monarch butterfly optimization algorithm," *International Journal of Hydrogen Energy*, vol. 45, no. 35, pp. 17882–17892, 2020.

[2] Z. Xu, F. R. Sheykhhamad, N. Ghadimi, and N. Razmjooy, "Computer-aided diagnosis of skin cancer based on soft computing techniques," *Open Medicine*, vol. 15, no. 1, pp. 860–871, 2020.

[3] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[4] A. Costa, Y. Kieffer, A. Scholer-Dahirel et al., "Fibroblast heterogeneity and immunosuppressive environment in human breast cancer," *Cancer Cell*, vol. 33, no. 3, pp. 463.e10–479.e10, 2018.

[5] D. Yu, Y. Wang, H. Liu, K. Jermsittiparsert, and N. Razmjooy, "System identification of PEM fuel cells using an improved elman neural network and a new hybrid optimization algorithm," *Energy Report*, vol. 5, pp. 1365–1374, 2019.

[6] V. Rajinikanth and M. S. Couceiro, "RGB histogram based color image segmentation using firefly algorithm," *Procedia Computer Science*, vol. 46, pp. 1449–1457, 2015.

[7] U. R. Acharya, S. L. Fernandes, J. E. WeiKoh et al., "Automated detection of alzheimer's disease using brain MRI images-a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, p. 302, 2019.

[8] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using tsallis entropy and region growing segmentation," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2018.

[9] Y. Li, A. Esteva, B. Kuprel, R. Novoa, J. Ko, and S. Thrun, "Skin cancer detection and tracking using data synthesis and deep learning," 2016, https://arxiv.org/abs/1612.01074.

[10] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[11] M. H. Jafari, N. Karimi, E. Nasr-Esfahani et al., "Skin lesion segmentation in clinical images using deep learning," in *Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR)*, December 2016.

[12] A. Mohamed, W. A. Mohamed, and A. H. Zekry, "Deep learning can improve early skin cancer detection," *International Journal of Electronics and Telecommunications*, vol. 65, 2019.

[13] R. Pugalenthi, M. P. Rajakumar, J. Ramya, and V. Rajinikanth, "Evaluation and classification of the brain tumor MRI using machine learning technique," *Journal of Control Engineering and Applied Informatics*, vol. 21, no. 4, pp. 12–21, 2019.

[14] S. C. Satapathy, N. Sri Madhava Raja, V. Rajinikanth, A. S. Ashour, and N. Dey, "Multi-level image thresholding using otsu and chaotic bat algorithm," *Neural Computing and Applications*, vol. 29, no. 12, pp. 1285–1307, 2018.

[15] D. Anoraganingrum, "Cell segmentation with median filter and mathematical morphology operation," in *Proceedings of the 10th International Conference on Image Analysis and Processing*, September 1999.

[16] T. Loupas, W. N. McDicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Transactions on Circuits and Systems*, vol. 36, no. 1, pp. 129–135, 1989.

[17] V. Rajinikanth and S. C. Satapathy, "Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and fuzzy-tsallis entropy," *Arabian Journal for Science and Engineering*, vol. 43, no. 8, pp. 4365–4378, 2018.

[18] N. Dey, V. Rajinikanth, A. Ashour, and J. M. Tavares, "Social group optimization supported segmentation and evaluation of skin melanoma images," *Symmetry*, vol. 10, no. 2, p. 51, 2018.

[19] S. R. Fernandes, J. T. de Assis, V. V. Estrela et al., "Nondestructive diagnosis and analysis of computed microtomography images via texture descriptors," in *Advances in Multidisciplinary Medical Technologies—Engineering, Modeling and Findings*, pp. 249–261, Springer, Berlin, Germany, 2020.

[20] K. B. Singh, T. V. Mahendra, R. S. Kurmvanshi, and C. V. Rama Rao, "Image enhancement with the application of local and global enhancement methods for dark images," in *Proceedings of the 2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)*, April 2017.

[21] M. N. Carney and W. M. Johnston, "A novel regression model from RGB image data to spectroradiometric correlates optimized for tooth colored shades," *Journal of Dentistry*, vol. 51, pp. 45–48, 2016.

[22] P. Maragos, "Differential morphology and image processing," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 922–937, 1996.

[23] J. Cao, K. Zhang, M. Luo, C. Yin, and X. Lai, "Extreme learning machine and adaptive sparse representation for image classification," *Neural Networks*, vol. 81, pp. 91–102, 2016.

[24] M. Ramezani, D. Bahmanyar, and N. Razmjooy, "A new optimal energy management strategy based on improved multi-objective antlion optimization algorithm: applications in smart home," *SN Applied Sciences*, vol. 2, no. 12, pp. 1–17, 2020.

[25] M.-W. Tian, S.-R. Yan, S.-Z. Han, S. Nojavan, K. Jermsittiparsert, and N. Razmjooy, "New optimal design for a hybrid solar chimney, solid oxide electrolysis and fuel cell based on improved deer hunting optimization algorithm," *Journal of Cleaner Production*, vol. 249, Article ID 119414, 2020.

[26] L. Yanda, Z. Yuwei, and N. Razmjooy, "Optimal arrangement of a micro-CHP system in the presence of fuel cell-heat pump based on metaheuristics," *International Journal of Ambient Energy*, pp. 1–12, 2020.

[27] Z. Yuan, W. Wang, H. Wang, and N. Razmjooy, "A new technique for optimal estimation of the circuit-based PEMFCs using developed sunflower optimization algorithm," *Energy Report*, vol. 6, pp. 662–671, 2020.

[28] Y. Guo, X. Dai, K. Jermsittiparsert, and N. Razmjooy, "An optimal configuration for a battery and PEM fuel cell-based hybrid energy system using developed krill herd optimization algorithm for locomotive application," *Energy Report*, vol. 6, pp. 885–894, 2020.

[29] Z. Yang, Q. Liu, L. Zhang, J. Dai, and N. Razmjooy, "Model parameter estimation of the PEMFCs using improved barnacles mating optimization algorithm," *Energy*, vol. 212, Article ID 118738, 2020.

[30] R. Navid, M. Khalilpour, and M. Ramezani, "A new meta-heuristic optimization algorithm inspired by FIFA world cup competitions: theory and its application in PID designing for AVR system," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 419–440, 2016.

[31] M. Mani, O. Bozorg-Haddad, and X. Chu, "Ant lion optimizer (ALO) algorithm," in *Advanced Optimization by Nature-Inspired Algorithms*, pp. 105–116, Springer, Berlin, Germany, 2018.

[32] M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Systems with Applications*, vol. 149, Article ID 113338, 2020.

[33] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.

[34] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Computers & Industrial Engineering*, vol. 145, Article ID 106559, 2020.

[35] N. Higashi and H. Iba, "Particle swarm optimization with gaussian mutation," in *Proceedings of the 2003 IEEE Swarm Intelligence Symposium. SIS'03 (Cat. No. 03EX706)*, April 2003.

[36] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: a nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, no. 2, pp. 495–513, 2016.

[37] E. Cuevas, F. Fausto, and A. González, "The locust swarm optimization algorithm," in *New Advancements in Swarm Algorithms: Operators and Applications*, pp. 139–159, Springer, Berlin, Germany, 2020.

[38] G. Dhiman and V. Kumar, "Spotted hyena optimizer: a novel bio-inspired based metaheuristic technique for engineering applications," *Advances in Engineering Software*, vol. 114, pp. 48–70, 2017.

[39] A. Kaveh and A. Dadras, "A novel meta-heuristic optimization algorithm: thermal exchange optimization," *Advances in Engineering Software*, vol. 110, pp. 69–84, 2017.

[40] R. Dobrescu, M. Dobrescu, S. Mocanu, and D. Popescu, "Medical images classification for skin cancer diagnosis based on combined texture and fractal analysis," *WISEAS Transactions on Biology and Biomedicine*, vol. 7, no. 3, pp. 223–232, 2010.

[41] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, and B. Badami, "Skin cancer diagnosis based on optimized convolutional neural network," *Artificial Intelligence in Medicine*, vol. 102, Article ID 101756, 2020.

[42] A. Pennisi, D. D. Bloisi, D. Nardi, A. R. Giampetruzzi, C. Mondino, and A. Facchiano, "Skin lesion image segmentation using delaunay triangulation for melanoma detection," *Computerized Medical Imaging and Graphics*, vol. 52, pp. 89–103, 2016.

[43] G. Babino, A. Lallas, M. Agozzino et al., "Melanoma diagnosed on digital dermoscopy monitoring: a side-by-side image comparison is needed to improve early detection," *Journal of the American Academy of Dermatology*, vol. 85, no. 3, pp. 619–625, 2020.

[44] C. Salem, D. Azar, and S. Tokajian, "An image processing and genetic algorithm-based approach for the detection of melanoma in patients," *Methods of Information in Medicine*, vol. 57, no. 1, pp. 74–80, 2018.

[45] R. Das, S. De, S. Bhattacharyya, J. Platos, V. Snasel, and A. E. Hassanien, "Data augmentation and feature fusion for melanoma detection with content based image classification," in *Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications*, Cairo, Egypt, January 2020.

[46] H. Alquran, I. A. Qasmieh, A. M. Alqudah et al., "The melanoma skin cancer detection and classification using support vector machine," in *Proceedings of the 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, October 2017.