

## Research Article

# Lung Cancer Diagnosis Based on an ANN Optimized by Improved TEO Algorithm

Rong Shan<sup>1</sup> and Tahereh Rezaei<sup>2</sup> 

<sup>1</sup>College of Computer, Weinan Normal University, Weinan, Shaanxi, China

<sup>2</sup>Neuroscience Research Center, Shiraz, Iran

Correspondence should be addressed to Tahereh Rezaei; [dr.tahereh.rz@gmail.com](mailto:dr.tahereh.rz@gmail.com)

Received 24 May 2021; Revised 22 June 2021; Accepted 9 July 2021; Published 16 July 2021

Academic Editor: Navid Razmjooy

Copyright © 2021 Rong Shan and Tahereh Rezaei. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A quarter of all cancer deaths are due to lung cancer. Studies show that early diagnosis and treatment of this disease are the most effective way to increase patient life expectancy. In this paper, automatic and optimized computer-aided detection is proposed for lung cancer. The method first applies a preprocessing step for normalizing and denoising the input images. Afterward, Kapur entropy maximization is performed along with mathematical morphology to lung area segmentation. Afterward, 19 GLCM features are extracted from the segmented images for the final evaluations. The higher priority images are then selected for decreasing the system complexity. The feature selection is based on a new optimization design, called Improved Thermal Exchange Optimization (ITEO), which is designed to improve the accuracy and convergence abilities. The images are finally classified into healthy or cancerous cases based on an optimized artificial neural network by ITEO. Simulation is compared with some well-known approaches and the results showed the superiority of the suggested method. The results showed that the proposed method with 92.27% accuracy provides the highest value among the compared methods.

## 1. Introduction

The proliferation of lung diseases in today's industrialized societies doubles the need for modern methods of accurate and early diagnosis. Among lung diseases, lung cancer is still recognized as one of the most dangerous cancers. Cancer means the abnormal growth, and sometimes proliferation of cells in the body. All cancers have an uncontrolled growth pattern and a tendency to detach from the source and metastasize [1]. A normal lung cell may become a lung cancer cell for no apparent reason, but in most cases, the transformation is the result of repeated exposure to carcinogens such as alcohol and tobacco. The appearance and function of cancer cells are different from normal cells. A mutation or change in the DNA or genetic material of a cell occurs [2]. DNA is responsible for controlling the appearance and function of cells. When a cell's DNA changes, that cell differentiates from the healthy cells next to it and no longer does the body's normal cells. This altered cell

separates from its neighboring cells and does not know when it should stop growing and die [3]. In other words, the altered cell does not follow the internal commands and signals that other cells are in control of and acts arbitrarily instead of coordinating with other cells [4]. One-third of all cancer deaths are due to lung cancer. About 80% of patients have five years left in their best condition after being diagnosed with this type of cancer [5]. Based on a survey by the American Cancer Society (ACS), lung cancer in both men and women is the second most prevalent cancer in the United States [6]. Approximately 228,820 new cases and approximately 135,720 lung cancer deaths, based on ACS figures, will occur [5]. Figure 1 shows the statistical information for the lead cancers and their death number in 2019 based on the American Cancer Society of lung cancer screening [7].

Air pollution due to the industrialization of cities, tobacco use, and genetic factors are the main causes of these diseases [8]. Early diagnosis of lung disease will have a major

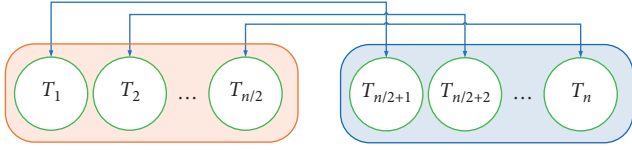


FIGURE 1: The heat transfer groups and the environment and cooling objects pairs.

impact on the possibility of definitive treatment of the disease. Major diagnostic methods for lung cancer include radiographic imaging and computed tomography, biopsy, bronchoscopy, and examination of cells in the sputum. Meanwhile, the CT scan imaging method is widely used as a superior diagnostic method. In this diagnostic method, the doctor examines possible nodules on the images. A pulmonary nodule is a small, round, opaque mass that forms inside the lung tissue [9, 10]. In other words, nodules are spherical radiographic opacities of less than three centimeters in diameter.

Formerly, lung diseases were diagnosed based on the help of experts' eye ability with no use for computer science. However, recently based on the different imaging techniques based on computer science and artificial intelligence, the diagnosis can be more precise. In most of these methods, after capturing the images from the patient, different image processing methods have been performed for tumor diagnosis [11].

## 2. Image Preprocessing

The analyzed images for validation in this study are collected from the *Lung CT-Diagnosis database* which has been provided by the Cancer Imaging Archive [12]. This database contains a collection of publicly available medical images for

different cancers [13] with contrast-enhanced CT scan images stored in the Dicom format. This database is collected from 61 patients such that 4682 different images have been acquired from them.

After image acquisition, the min-max normalization method is established for them to scale the acquired images between 0 and 1. This study uses  $250 \times 250$  scale normalization for this purpose. By considering a grayscale image with  $n$  dimension which has the following limitation:  $I_n = [X \subseteq R^n] \rightarrow [a, \dots, b]$ , the normalized image,  $I^*$ , can be achieved by the following formulation:

$$I^* = a_{\text{new}} + \frac{b_{\text{new}} - a_{\text{new}}}{b - a} \times (I - a), \quad (1)$$

where  $I^* = [X \subseteq R^n] \rightarrow [a_{\text{new}}, \dots, b_{\text{new}}]$ ,  $a$  and  $b$  describe the intensity values in the grayscale image, and  $a_{\text{new}}$  and  $b_{\text{new}}$  describe the intensity values for the normalized image [14].

Also, since the reason that all CT scan images have a kind of visual noise, they need a denoising filter to resolve the problem. A CT scan image faces different factors and includes different types, such as Gaussian noise, Shot Noise, Poisson Noise, Speckle Noise, and Salt and Pepper Noise [15]. Noise hides slight details of the CT scan image. This shows that we need a tool for noise removal before starting the CT scan images.

One of the popular noise reduction techniques for CT scan images is mean filtering. The definition of the mean (average) filter works on averaging any aspect of the picture to the neighbors [16]. The average filter calculates and divides the sum of all the pixels in the filter window by the total number of pixels [4]. It then replaces the value of the center pixel with the calculated average. The result value for each indexed pixel  $(i, j)$  is determined as follows:

$$R(i, j) = \frac{I(i-1, j-1) + I(i-1, j) + I(i-1, j+1) + I(i, j-1) + I(i, j) + I(i, j+1) + I(i+1, j-1) + I(i+1, j) + I(i+1, j+1)}{9} \quad (2)$$

## 3. Image Segmentation

One of the most important issues in image processing is the identification and separation of the image into its main components. Image segmentation determines the success or eventual failure of image analysis methods. And yet, due to its wide application, it has suitable research fields. The accuracy of this study in fields such as medicine is very important to preserve and protect human life. Thresholding is one of the most convenient methods for image segmentation [17]. By applying a thresholding method to a grayscale image, a binary image is obtained that delineates the boundaries of the objects in the image with appropriate accuracy.

The lower the threshold is, the more errors are detected and the more sensitive the results are to noise and unrelated image features [18]. On the other hand, a high threshold may

miss weak errors or parts of errors. Since the main purpose of the fabric fault detection system is to find all possible faults, the threshold must be chosen to achieve this goal [19]. There are several methods for this practice in image processing science, and some separation methods are used for specific images [20]. One way to select the appropriate threshold is to use the trial-and-error method, in which different values of the threshold are selected and the image resulting from the application of this threshold is judged by the viewer. The simplest type of separation is called general separation, which is based on the image histogram. The input of this function is a gray image or a color image [21]. Its output is also a black and white image (binary). In this method, the threshold value at any point in the image is defined based on the local properties of the image in the neighborhood of that pixel. In this paper, Kapur thresholding has been used. Assume that the gray levels in an

image with  $N$  pixels and  $L$  gray levels are in the range  $[0, 1, \dots, L - 1]$ . Kapur method obtains the thresholds based on Kapur entropy maximization and based on the information obtained from the histogram of gray surfaces, which is defined as follows in the case of two-level segmentation (thresholding):

$$\text{Maximize } f_1(t) = H(0, t) + H(t, L), \quad (3)$$

such that

$$\begin{aligned} H(0, t) &= - \sum_{i=0}^{t-1} \frac{p_i}{w_o} \ln \frac{p_i}{w_o}, & w_o &= \sum_{i=0}^{t-1} p_i, \\ H(t, L) &= - \sum_{i=0}^{t-1} \frac{p_i}{w_i} \ln \frac{p_i}{w_i}, & w_i &= \sum_{i=0}^{t-1} p_i. \end{aligned} \quad (4)$$

The value of the optimal threshold point is the amount of gray area  $t$  that maximizes the function. For correct diagnosis of the tumors, the cancerous region should be detected with high precision. As mentioned before, image thresholding has been used for identifying the cancerous region. However, mathematical morphology is needed to better filter the detected areas based on image thresholding [22]. Mathematical morphology algorithm is a new technique for processing and analyzing signals and images. The basic idea of this technique is based on the analysis of geometric information by exploring an image with a small geometric pattern called a structuring element. This study uses three popular techniques including opening, closing, and filling holes. The first operator in the region filling is based on complementation, intersections, and dilation operators and is achieved by the following equation:

$$X_k = (X_{k-1} \oplus B) \cap A^c, \quad k = 1, 2, 3, \dots, \quad (5)$$

where  $A$  indicates a group of boundaries and  $B$  defines the organizing element. This operator will end if  $X_k = X_{k-1}$ .

The mathematical opening is the second utilized operator. The opening of element  $A$  with element  $B$  has been obtained with  $A$  erosion by structure element  $B$ , followed by a dilation of the resulted image by structure element  $B$ , that is,

$$A \circ B = (A \ominus B) \oplus B. \quad (6)$$

The key goal of the mathematical opening is to remove the minor blemishes in the area that can be missed during the diagnosis of lung cancer.

Finally, the mathematical closing operator is used to smooth the counters, fuses narrow breaks, eliminates small holes, and fills gaps in the contour and long thin gulfs. This operation is formulated as follows:

$$A \cdot B = (A \oplus B) \ominus B. \quad (7)$$

## 4. Features Extraction

The purpose of feature extraction is to make raw data more usable for future statistical processing. Feature extraction is a very common process in different types of data processing

such as image processing and audio processing. Feature extraction means selecting a feature that can describe the image with little information. These features must have properties so that a set of these features is described uniquely in each image. If a set of these attributes are the same for two samples, then you will not be able to distinguish two samples with any classifier in the classification section.

The main reasons for extracting features from images are image simplification, reduced processing time and memory, and increased accuracy and efficiency. So, feature extraction is a process in which data is mapped in a high-dimension space to a lower dimension space. This mapping can be linear (such as principal component analysis) or nonlinear. How to select these features requires data properties to be examined, and to extract it, preprocessing operations and various filters must be applied to the image to turn the image into the desired information. In this study, GLCM features have been used for extracting the lung cancer images information.

The GLCM method is one of the most efficient techniques for extracting tissue from medical images. This matrix is a square matrix with dimensions  $N \times N$  where  $N$  is the number of degrees of gray in the image. Each element of this matrix represents the number of pairs of pixels that have degrees of gray on the surface of the image and are spaced in a certain direction from each other and to a certain pixel distance. After calculating the matrix, different parameters of the image texture can be extracted from it. In this study, the mentioned technique was used to extract tissue in lung tumor images. In the following, the utilized features have been explained.

**4.1. Contrast.** The contrast regulates the intensity value of the pixel and its neighbor in the image. This feature is achieved by the following equation:

$$\text{Contrast} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i-j)^2 f(i, j). \quad (8)$$

**4.2. Correlation.** The correlation feature describes the dependency on spatial features among the pixels. This feature can be mathematically given as follows:

$$\text{Cor} = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i, j) f(i, j) - \mu_i \mu_j}{\sigma_i \sigma_j}. \quad (9)$$

**4.3. Homogeneity.** Homogeneity is a local uniformity feature that makes single/multiple intervals govern the textured and nontextured characteristics. This feature is achieved by the following:

$$H = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \frac{1}{1 + (i-j)^2} f(i, j). \quad (10)$$

**4.4. Energy.** The energy feature regulates the number of repetitive pixel pairs. This feature is mathematically obtained by the following equation:

$$\text{Energy} = \sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f^2(i, j)}. \quad (11)$$

4.5. *Entropy.* The entropy is a feature that indicates the image selected interference based on the following equation:

$$\text{Ent} = - \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \log_2 f(i, j). \quad (12)$$

## 5. Features Selection

Some of the different extraction characteristics in the images are so important and crucial for classification. In the meantime, those features that contain information which is not notable can have high potential when combined with other features. Any of these attributes could also have no useful data at all. This shortcoming can be resolved by different works. In this paper, an optimization-based methodology has been proposed for this purpose. The optimization method for the features selection can be achieved by the following equation:

$$\text{FF} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{((\text{TN} + \text{FP}) \times (\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FN}))^{(1/2)}}, \quad (13)$$

where TP signifies the true positive, TN describes the true negative, FN represents the false negative, and FP defines the false positive.

## 6. Improved Thermal Exchange Optimization Algorithm

In Thermal Exchange Optimization (TEO), the temperature of the objects indicates the individual position and with objects grouping, it is started to be exchanged. Therefore, new temperatures indicate their updated positions [23].

6.1. *Newton's Law of Cooling.* In the seventeenth century, the English scientist Isaac Newton studied the cooling of objects. The experiments showed that the cooling rate was approximately proportional to the temperature difference between the heated object and the environment. This fact is written as a differential relation:

$$\frac{dQ}{dt} = \alpha \times A \times (T_s - T_b), \quad (14)$$

where  $Q$  describes the heat,  $A$  signifies the area of the body surface that transmits heat,  $T_b$  defines the body temperature,  $T_s$  represents the ambient temperature, and  $\alpha$  determines the heat transfer coefficient which is dependent on the geometry of the object, surface state, heat transfer mode, and other factors.

The heat loss in time  $dt$  is  $\alpha \times A \times (T_s - T)dt$  which defines the alteration in stored heat as the temperature falls  $dT$ , i.e.,

$$V \times \rho \times c \times dT = -\alpha \times A \times (T - T_b)dt, \quad (15)$$

where  $V$  defines the volume ( $\text{m}^3$ ),  $\rho$  describes the density ( $\text{kg}/\text{m}^3$ ), and  $c$  signifies the specific heat ( $\text{J}/\text{kg}/\text{K}$ ).

Therefore,

$$\frac{T - T_b}{T_M - T_b} = \exp\left(\frac{-\alpha \times A \times t}{V \times \rho \times c}\right), \quad (16)$$

where  $T_M$  is the starting high temperature.

The above equation is valid when  $(\alpha \times A \times t)/(V \times \rho \times c)$  is not a function of  $T$ , i.e.,

$$\gamma = \frac{\alpha \times A}{V \times \rho \times c}, \quad (17)$$

where  $\gamma$  is constant. Therefore,

$$\frac{T - T_b}{T_M - T_b} = \exp(-\gamma t). \quad (18)$$

And finally, the equation can be rewritten as follows:

$$T = (T_M - T_b) \times \exp(-\gamma t) + T_b. \quad (19)$$

6.2. *Inspiration.* During the TEO algorithm, some individuals have been considered as the cooling objects and the residual individuals have been considered as the environment; then it has been in reverse. The method of simulation of the TEO algorithm is given in the following.

The first step is initialization. The initial temperature for all of the objects has been defined in an  $m$ -dimensional solution space as follows:

$$T_i^0 = T_{\min} + \text{rand}(T_{\max} - T_{\min}), \quad (20)$$

where  $T_i^0$  signifies the initial solution vector of the object number  $i$ ,  $\text{rand}$  describes a random vector with components in the range  $[0, 1]$ ,  $T_{\min}$  and  $T_{\max}$  describe the minimum and the maximum limitations for the decision variables.

After initializing the objects, the value of the objective function for all of the individuals is evaluated. During the process, some historically best  $T$  vectors have been stored in a memory called Thermal Memory (TM) to use their position to develop the algorithm efficiency with no extra computational cost. After selecting some best values in TM, they have been added to the population and for the same numbers of them, worst individuals have been eliminated. Individuals have been divided into two equal groups. Figure 1 shows this division. For example,  $T_1$  defines an environment object for  $T_{(n/2)+1}$  cooling object and contrariwise.

During the process, if an object has lower  $\gamma$ , the temperature exchanging has been established slowly. Therefore, the value of  $\gamma$  for the objects has been established based on the following equation:

$$\gamma = \frac{\text{Cos}(\text{object})}{\text{Cos}(\text{worst object})}. \quad (21)$$

Time is another parameter that is important in this algorithm. This parameter is related to the iteration number. This parameter can be formulated by the following equation:

$$t = \frac{\text{iteration}}{\text{Max} \cdot \text{iteration}}. \quad (22)$$

Generally, an important ability of the metaheuristics is to escape from the local optimum. During this process, the environmental temperature has been altered by the following equation:

$$T_i^e = (1 - (c_1 + c_2 \times (1 - t) \times \text{rand})) \times T_i'^e, \quad (23)$$

where  $c_1$  and  $c_2$  represent the control variables.  $T_i'^e$  describes the object earlier temperature that has been modified to  $T_i^e$ .

Based on the previous models, the new temperatures of the objects have been updated by the following equation:

$$T_i^+ = T_i^e + (T_i^{\text{old}} - T_i^e) \exp(-\gamma t). \quad (24)$$

Another parameter in the algorithm is Pr with (0, 1) that stated whether a component in the cooling objects should be changed or not.

All of the Pr agents have been compared with  $R(i)$  ( $i = 1, 2, \dots, n$ ) which is a randomly distributed value in the range [0, 1]. If  $R(i) < \text{Pr}$ , one dimension of the agent number  $i$  has been randomly chosen and the value is redeveloped by the following:

$$T_{i,j} = T_j^{\text{min}} + \text{rnd} (T_j^{\text{max}} - T_j^{\text{min}}) \exp(-\gamma t), \quad (25)$$

where  $T_{i,j}$  signifies the  $j^{\text{th}}$  variable of the  $i^{\text{th}}$  agent and  $T_j^{\text{min}}$  and  $T_j^{\text{max}}$  represent the lower and the upper bounds of the  $j^{\text{th}}$  variable, respectively. To keep the structures of the agents unchanged, just one dimension has been altered.

Finally, the stopping criteria are checked to terminate the algorithm in the considered criteria.

### 6.3. Improved Thermal Exchange Optimization (ITEO).

The basic Thermal Exchange Optimization algorithm suffers some disadvantages like stability and premature convergence problems. This case leads us to design a modified version of the TEO algorithm to refine these drawbacks as possible.

The first modification is to use Lévy flight (LF) as a proper mechanism. This mechanism has been commonly employed in metaheuristic algorithms to solve premature convergence shortcomings [24]. During this mechanism, a random walk policy has been utilized for proper adjusting of the local search that is mathematically represented as follows:

$$\text{Le}(x) \approx \frac{1}{x^{1+\tau}},$$

$$x = \frac{A}{|B|^{(1/\tau)}}, \quad (26)$$

$$\sigma^2 = \left\{ \frac{\Gamma(1+\tau)}{\tau \Gamma((1+\tau)/2)} \times \frac{\sin(\pi\tau/2)}{2^{(1+\tau)/2}} \right\}^{(2/\tau)},$$

where  $A \sim N(0, \sigma^2)$ ,  $B \sim N(0, \sigma^2)$ ,  $\tau$  signifies the Lévy index which is located in the range [0, 2] (here,  $\tau = 1.5$  [25]),  $\Gamma(\cdot)$  represents Gamma function, and  $w$  defines the step size.

By assuming the above equations, the updating formulation for the TEO algorithm is as follows:

$$T_{i,j}^{\text{new}} = T_j^{\text{min}} + \text{Le}(T_j^{\text{max}} - T_j^{\text{min}}) \exp(-\gamma t). \quad (27)$$

The second modification is to use the chaos mechanism for improving the system convergence speed. Here, we used Singer function for chaos modification [26, 27]. By considering this mechanism, rnd can be updated as follows:

$$\text{rnd}_{i+1} = 1.07(7.9 \times \text{rnd}_i - 23.3 \times \text{rnd}_i^2 + 28.7 \times \text{rnd}_i^3 - 13.3 \times \text{rnd}_i^4), \quad (28)$$

where  $\text{rnd}_0 \in [0, 1]$ .

**6.4. Algorithm Authentication.** The simulations are applied to a Core (TM) i7-4720HQ 1.60 GHz with 8 GB RAM under and simulations are applied to the MATLAB 2017b environment. To prove the effectiveness of the suggested ITEO algorithm, it is performed on some different benchmark functions, i.e., Ackley, Rastrigin, Sphere, and Rosenbrock, and the results have been compared with some well-known and new optimization techniques, i.e. Locust Swarm Optimization Algorithm (LSO) [28], Crow Search Algorithm (CSA) [29], Multi-Verse Optimizer (MVO) [30], search and rescue (SAR) algorithm [31], and the basic Thermal Exchange Optimization (TEO) [10]. Table 1 tabulates the parameter settings of the studied algorithms.

To clarify the studied benchmark functions, the equations and the boundaries have been given in Table 2. During the simulation, all of the algorithms run 40 times independently for all of the benchmark functions and the maximum iteration for all of the algorithms is considered 500.

Table 3 tabulates the results of the analysis of the algorithms on the benchmark functions based on four measurement indicators including minimum, maximum, mean, and standard deviation (std).

As can be observed from the results, the results of all four indicators based on the suggested ITEO have the minimum value. This indicates that the suggested algorithm has the minimum error ratio for all of the metrics. The minimum value of the indicator ‘‘minimum’’ shows that the proposed ITEO algorithm has the highest precision among the others. Plus, the minimum value of the ‘‘std’’ for the ITEO algorithm shows its higher consistency toward the other methods.

TABLE 1: The parameter settings of the studied algorithms.

Algorithm	Parameter	Value	Algorithm	Parameter	Value
SAR [31]	SE	0.5	LSO [28]	F	0.6
	MU	20		L	1
	$c_1$	0		g	20
TEO [10]	$c_2$	1	MVO [30]	Traveling distance rate	[0.6, 1]
	pro	0.3		Wormhole existence prob.	[0.2, 1]
	TM	28.5	CSA [29]	fl	2

TABLE 2: The equations and the boundaries of the studied benchmarks in the analysis.

Function	Equation	Constraint
Rastrigin	$f_1(x) = 10D + \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i))$	$[30, 50]^D$
Rosenbrock	$f_2(x) = \sum_{i=1}^{D-1} (100(x_i^2 - x_{i+1}) + (x_i - 1)^2)$	$[-2.045, 2.045]^D$
Ackley	$f_3(x) = -20 \exp(-0.2 \sqrt{(1/D) \sum_{i=1}^D (x_i^2)}) - \exp((1/D) \sum_{i=1}^D (\cos(2\pi x_i))) + 20 + e$	$[-10, 10]^D$
Sphere	$f_4(x) = \sum_{i=1}^D x_i^2$	$[-512, 512]^D$

TABLE 3: The performance analysis of the studied algorithm.

Algorithm		$f_1$	$f_2$	$f_3$	$f_4$
LSO [28]	Min	11.428	0.035	0.022	1.327
	Max	4.284e3	6.534e2	0.679e5	6.192
	Mean	2.168e3	1.295e2	2.195e4	3.084
	Std	5.153e2	21.931	3.645e4	4.339
CSA [29]	Min	8.3161	0.028	7.65e-1	1.446
	Max	492.57	2.193	4.15e-1	11.28
	Mean	120.75	1.137	5.17e-1	8.294
	Std	69.253	0.066	1.01e-1	4.495
MVO [30]	Min	9.192	4.168e-3	8.195e-5	0.957
	Max	1.15 e2	4.279e-2	7.086e-4	1.492
	Mean	22.49	9.153e-3	1.517e-4	1.137
	Std	38.76	1.097e-3	3.349e-5	0.846
SAR [31]	Min	0.716	4.82e-16	8.16e-7	1.651
	Max	11.05	5.61e-14	2.94e-6	3.287
	Mean	4.13	5.29e-15	1.19e-6	0.201
	Std	3.62	2.46e-14	4.36e-7	0.146
TEO [32]	Min	0.356	9.208e-21	4.192e-9	3.062e-9
	Max	15.349	11.193e-19	5.930e-8	9.836e-8
	Mean	9.3462	34.255e-19	3.591e-8	3.491e-8
	Std	2.896	3.673e-19	5.086e-9	2.038e-9
ITEO	Min	0.00	8.376e-22	6.956e-10	6.763e-13
	Max	4.207	3.205e-21	3.907e-9	6.358e-12
	Mean	3.164	1.343e-21	1.2084e-9	1.928e-12
	Std	1.237	2.934e-21	3.666e-10	1.117e-13

## 7. Classification

For the final diagnosis of the features, a classifier is needed. In this study, a new optimized version of Artificial Neural Network (ANN) has been used [33]. The ANN is an almost new methodology with high efficiency for different classification applications. The ANN makes a proper relationship between input and output data. The sensitivity of the ANN is so low toward errors [34]. This method can be used for the accurate classification of medical images by correct training with no mathematical modeling [35, 36]. A popular and simple

technique among different types of ANN is the multilayer perceptron (MLP) neural network. An MLP is a mathematical model of a natural brain [34]. An MLP is a model with several numbers of weights and biases that are connected for brain performance modeling. The popular method for error minimization in MLPs is backpropagation (BP). In BP, the values of the weights and biases have been adjusted for minimizing the error between the output value and the desired value [37]. This method uses Gradient Descent (GD) algorithm for minimization. One significant problem of the GD algorithm is that they have been easily trapped into the local minimum.

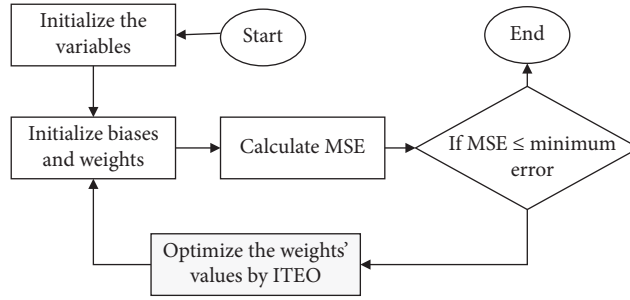


FIGURE 2: The flowchart of classification based on optimized ANN/ITEO.

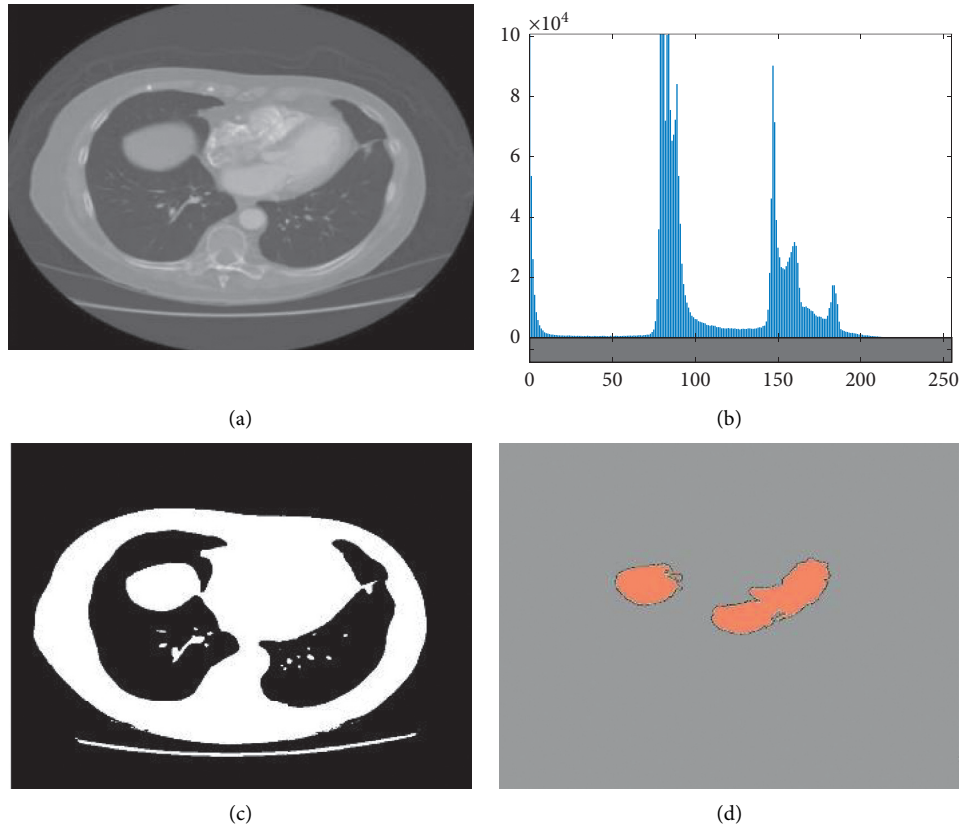


FIGURE 3: An example of the image segmentation: (a) Original image, (b) histogram of the image, (c) image segmentation based on Kapur technique, and (d) image (c) after mathematical morphology.

The output of each layer in the network is as follows:

$$z_i = \sum_{j=1}^n w_{ij} \alpha_j + b_i, \quad (29)$$

where  $\alpha_j$  indicates the  $j^{\text{th}}$  input variable,  $b_i$  signifies the  $i^{\text{th}}$  bias for the neuron, and  $w_{ij}$  describes the relation weight between  $\alpha_j$  and the  $i^{\text{th}}$  invisible neuron.

After that, the activation function has been performed to trigger the output of the neurons. This research employs sigmoid function for this purpose:

$$f_j(x) = \frac{1}{1 + e^{-z_j}}. \quad (30)$$

And the output layer gives

$$y_i = \sum_{k=1}^m w_{ki} \alpha_k + \beta_i. \quad (31)$$

The ANN calculates mean square error (MSE) between the desired and the observed output, i.e.,

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n y_i - d_i, \quad (32)$$

where  $n$  describes the number of the steps in the training data collection and  $y_i$  and  $d_i$  represent the observed value and the desired value, respectively.

TABLE 4: The GLCM data results of 20 first images from the lung CT-diagnosis database.

Image #	Homogeneity	IMC 1	IMC 2	Inverse difference	Highest probability	Sum average	Sum entropy	Sum of square variance	Sum variance
1	-0.916	-0.914	0.654	0.997	0.717	2.372	0.407	0.553	0.662
2	-0.900	-0.879	0.212	0.992	0.531	2.357	0.552	0.506	0.711
3	-0.884	-0.88	0.625	0.991	0.891	2.68	0.69	0.147	0.374
4	-0.891	-0.929	0.476	0.992	0.529	2.472	0.601	0.147	0.21
5	-0.911	-0.903	0.476	0.992	0.692	2.517	0.558	0.211	0.091
6	-0.895	-0.915	0.274	0.993	0.865	2.532	0.23	0.123	0.075
7	-0.915	-0.875	0.713	0.993	0.674	2.308	0.384	0.397	0.43
8	-0.882	-0.914	0.405	0.991	0.771	2.41	0.495	0.659	0.554
9	-0.881	-0.859	0.355	0.995	0.861	2.434	0.423	0.517	0.078
10	-0.888	-0.866	0.416	0.992	0.891	2.194	0.198	0.453	0.483
11	-0.917	-0.905	0.437	0.992	0.911	2.437	0.296	0.68	0.382
12	-0.913	-0.895	0.544	0.995	0.584	2.517	0.629	0.174	0.14
13	-0.925	-0.877	0.554	0.997	0.614	2.356	0.179	0.709	0.677
14	-0.893	-0.865	0.449	0.998	0.896	2.602	0.645	0.181	0.433
15	-0.887	-0.885	0.451	0.995	0.762	2.539	0.213	0.451	0.328
16	-0.885	-0.888	0.525	0.998	0.722	2.316	0.538	0.204	0.679
17	-0.907	-0.911	0.614	0.995	0.77	2.373	0.445	0.479	0.479
18	-0.904	-0.899	0.799	0.995	0.861	2.332	0.29	0.113	0.336
19	-0.931	-0.876	0.655	0.993	0.735	2.565	0.484	0.616	0.068
20	-0.886	-0.86	0.452	0.994	0.589	2.541	0.237	0.67	0.393

TABLE 5: The GLCM data results of 20 first images from the lung CT-diagnosis database.

Image #	Autocorrelation	Cluster prominence	Cluster shade	Contrast	Correlation	Difference entropy	Difference variance	Dissimilarity	Energy	Entropy
1	1.5292	0.9281	0.2484	0.0031	0.951	0.0101	0.001	0.001	0.8911	0.2308
2	1.7732	1.4425	0.5614	0.0049	0.9617	0.017	0.003	0.0029	0.749	0.424
3	1.5332	0.9471	0.2584	0.0024	0.9637	0.0023	0.0005	0.0004	0.8903	0.2274
4	2.0352	1.5474	0.6374	0.0086	0.957	0.0357	0.0063	0.0066	0.6421	0.5764
5	1.4202	0.5069	0.0204	0.0031	0.9247	0.0045	0.001	0.002	0.9595	0.101
6	1.3862	0.3493	0.1646	0.0015	0.9515	0.0133	0.0014	0.0014	0.984	0.0444
7	0.7428	0.7435	0.3716	0.0035	0.9413	0.029	0.0033	0.0035	0.9226	0.1792
8	0.7528	0.8409	0.4256	0.0024	0.975	0.0352	0.0046	0.0044	0.9064	0.2137
9	0.7668	0.8365	0.4216	0.0007	0.994	0.0257	0.003	0.0029	0.9105	0.2015
10	1.4782	0.7423	0.3716	0.0029	0.9603	0.0381	0.0051	0.005	0.921	0.1889
11	1.5302	0.9288	0.4746	0.0024	0.9826	0.0353	0.0045	0.0046	0.891	0.2391
12	1.5432	0.9597	0.2674	0.0039	0.9666	0.0466	0.006	0.005	0.881	0.2612
13	1.7702	1.4266	0.5524	0.0057	0.9891	0.0522	0.0077	0.0058	0.7585	0.4381
14	1.7072	0.8725	0.5024	0.0071	0.9858	0.0315	0.005	0.004	0.784	0.3995
15	1.0348	0.9531	0.7806	0.0072	0.9916	0.0294	0.0052	0.0051	0.756	0.4349
16	0.8828	0.6969	0.6166	0.0038	0.9594	0.0109	0.0018	0.0017	0.8405	0.3082
17	0.8948	0.7077	0.6256	0.0091	0.9283	0.0383	0.009	0.0069	0.8264	0.3453
18	2.0578	1.2224	0.1826	0.0204	0.94	0.105	0.0199	0.02	0.4803	0.7899
19	0.6988	0.5715	0.2796	0.002	0.955	0.0184	0.002	0.0019	0.9541	0.1186
20	1.7448	1.3354	0.5706	0.0141	0.9508	0.0811	0.0141	0.014	0.5192	0.7325

As mentioned earlier, GD algorithm for minimizing the MSE has some problems. Therefore, here for minimizing the MSE and optimizing the classifier, the suggested Improved Thermal Exchange Optimization has been used. Figure 2 shows the method of this idea.

## 8. Results and Discussion

As mentioned before, the main idea of this study is to propose a pipeline methodology for lung cancer diagnosis. The method starts with a preprocessing method for

enhancing the quality of the original image. After image preprocessing, a simple image thresholding has been done based on Kapur for lung areas. Afterward, optimal features have been selected from GLCM features based on an improved version of the Thermal Exchange Optimization algorithm. Finally, an optimized MLP system was based on the introduced Improved Thermal Exchange Optimization algorithm. The method has been validated based on the Lung CT-Diagnosis dataset collected by the Cancer Imaging Archive [12]. Figure 3 shows an example of the image segmentation.



TABLE 6: The optimum features using the proposed ITEO applied to GLCM.

Image #	Autocorrelation	Correlation	Energy	Homogeneity	Inverse difference	Highest probability	Sum average	Class
1	1.163	0.97	0.8911	0.9972	0.9971	0.9417	2.1095	Yes
2	1.407	0.9807	0.749	0.9966	0.9965	0.8594	2.2728	Yes
3	1.167	0.9827	0.8903	0.9979	0.9978	0.9411	2.1119	Yes
4	1.669	0.976	0.6421	0.9948	0.9947	0.769	2.4488	Yes
5	1.054	0.9437	0.9595	0.997	0.997	0.969	2.036	Yes
6	1.02	0.9705	0.986	0.9985	0.9984	0.9931	2.0131	Yes
7	1.111	0.9603	0.9246	0.9993	0.9972	0.9611	2.0748	Yes
8	1.121	0.954	0.9064	0.9988	0.9987	0.9512	2.0955	Yes
9	1.135	0.975	0.9105	0.9996	0.9995	0.9534	2.0928	Yes
10	1.112	0.9413	0.921	0.9985	0.9984	0.959	2.0782	Yes
11	1.164	0.9636	0.891	0.997	0.998	0.94	2.1127	Yes
12	1.177	0.9476	0.879	0.996	0.997	0.9351	2.121	Yes
13	1.404	0.9701	0.7565	0.9971	0.995	0.859	2.2727	Yes
14	1.341	0.9648	0.784	0.9952	0.9951	0.8774	2.2339	No
15	1.403	0.9706	0.756	0.9954	0.9953	0.8585	2.2723	No
16	1.251	0.9784	0.8405	0.9971	0.997	0.9142	2.1682	No
17	1.263	0.9473	0.8284	0.9945	0.9944	0.9062	2.1787	No
18	2.426	0.961	0.4823	0.9907	0.9907	0.5114	2.9592	No
19	1.067	0.976	0.9541	1.0001	1	0.9767	2.0473	No
20	2.113	0.9718	0.5192	0.994	0.9939	0.6199	2.7485	No

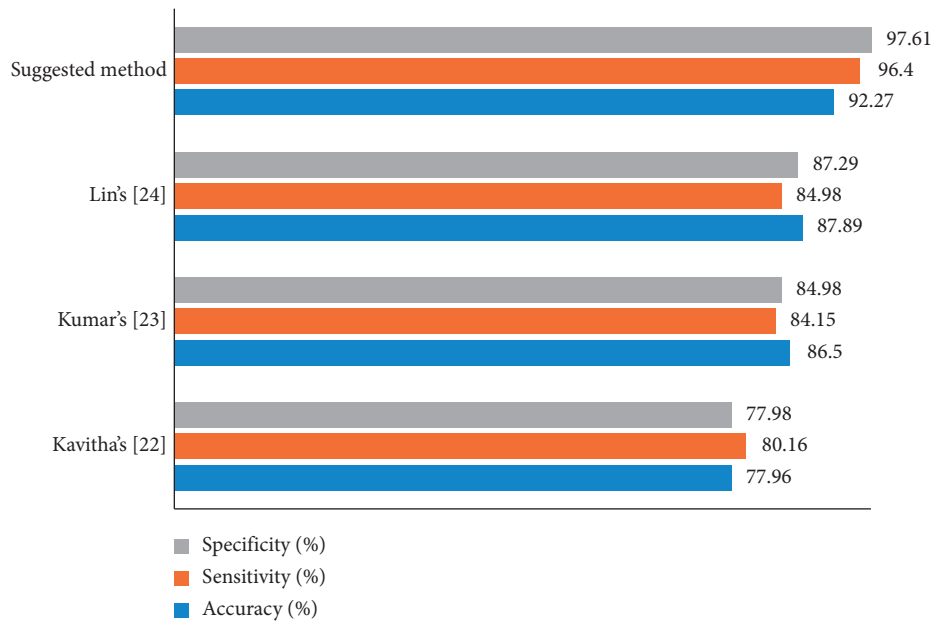


FIGURE 4: The efficiency validation of the suggested method compared with some well-known methods based on three mentioned indicators.

**8.1. Dataset Description.** In this study, the Lung CT-Diagnosis database has been utilized. This database contains several images in Dicom format, employed from capturing 61 patients. The number of the total images is 4682. The method has been simulated based on MATLAB 2017b environment and performed to the database based on the following configuration: Corei7 laptop with 16 GB RAM and CPU@2.6 GHz processor.

**8.2. Simulation Results.** Table 4 indicates the GLCM data results of 20 first images from the Lung CT-Diagnosis database. As can be observed from Table 4, nineteen numbers

of GLCM features are employed for the feature extraction (Table 5).

After feature selection based on the Improved Thermal Exchange Optimization, the optimum features based on the suggested ITEO methodology are evaluated and shown in Table 6.

The optimum threshold achieved by ITEO has been used to select the features. For the cost function, the best optimal value earned is 0.75. To test the final efficacy of the suggested technique, three measurement indicators including sensitivity, precision, and accuracy have been analyzed. The mathematical formulations of the indicators are given below:

$$\begin{aligned} \text{Accuracy (\%)} &= \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})}, \\ \text{Specificity (\%)} &= \frac{\text{TN}}{(\text{FP} + \text{TN})}, \\ \text{Sensitivity (\%)} &= \frac{\text{TP}}{(\text{TP} + \text{FN})}, \end{aligned} \quad (33)$$

where TN is truly negative, TP is truly positive, FN is false negative, and FP is false positive.

To verify the higher efficiency of the proposed method, a comparison analysis of the method has been applied toward some state-of-the-art algorithms, including Kavitha's [38], Kumar's [39], and Lin's [40], applied to the Lung CT-Diagnosis database. The comparison results are illustrated in a bar chart in Figure 4.

As can be observed from Figure 4, the proposed method has the best precision for the Lung CT-Diagnosis database and Lin's, Kumar's, and Kavitha's are placed in the next ranks. The results show also the proposed method.

## 9. Conclusions

The main purpose of this study is to propose an optimal pipeline for precise lung cancer diagnosis based on different approaches. The method started with applying a pre-processing process based on a min-max normalization for the input data and an average filter for denoising the input image. Afterward, Kapur entropy maximization along with mathematical morphology was used for segmentation of the lung area. Then, 19 numbers of the GLCM features have been extracted from the segmented images and the features with higher priority were selected based on a new optimization design. The new design, called Improved Thermal Exchange Optimization (ITEO) algorithm, was designed and employed to optimize the feature selection step by considering more accuracy and convergence ability as was shown in the validation stage. Finally, the images were classified into healthy or cancerous cases by using an artificial neural network optimized by ITEO. The simulation was compared with some different state-of-the-art methods including Lin's, Kumar's, and Kavitha's, and the results showed that the proposed method with 92.27% accuracy, 96.4% sensitivity, and 97.61% specificity has the highest efficiency toward the other state-of-the-art methods. In future work, we will work on using convolutional features of the lung cancer images to provide a method with higher accuracy in the system.

## Data Availability

The data that support the findings will be available in the Lung CT-Diagnosis database of lung cancer images at <https://doi.org/10.7937/K9/TCIA.2015.A6V7JIWX>.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

- [1] R. Navid, M. Ashourian, M. Karimifard et al., "Computer-aided diagnosis of skin cancer: a review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020.
- [2] 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.
- [3] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjoooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.
- [4] A. Hu and R. Navid, "Brain tumor diagnosis based on metaheuristics and deep learning," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 657–669, 2020.
- [5] R. Wender, E. T. H. Fontham, E. Barrera et al., "American Cancer Society lung cancer screening guidelines," *CA: A Cancer Journal for Clinicians*, vol. 63, no. 2, pp. 106–117, 2013.
- [6] N. Ghadimi, "An adaptive neuro-fuzzy inference system for islanding detection in wind turbine as distributed generation," *Complexity*, vol. 21, no. 1, pp. 10–20, 2015.
- [7] 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. 463–479, 2018.
- [8] N. Razmjoooy, V. V. Estrela, and H. J. Loschi, "Entropy-based breast cancer detection in digital mammograms using world cup optimization algorithm," *International Journal of Swarm Intelligence Research*, vol. 11, no. 3, pp. 1–18, 2020.
- [9] 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.
- [10] 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.
- [11] WHO, *Cancer: Early Diagnosis and Screening*, WHO, Geneva, Switzerland, 2020, <https://www.who.int/cancer/prevention/diagnosis-screening/en>.
- [12] O. Grove, A. E. Berglund, M. B. Schabath et al., "Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma," *PLoS One*, vol. 10, no. 3, Article ID e0118261, 2015.
- [13] M. Ghiasi, N. Ghadimi, and E. Ahmadiania, "An analytical methodology for reliability assessment and failure analysis in distributed power system," *SN Applied Sciences*, vol. 1, no. 1, p. 44, 2019.
- [14] G. Guo and N. Razmjoooy, "A new interval differential equation for edge detection and determining breast cancer regions in mammography images," *Systems Science & Control Engineering*, vol. 7, no. 1, pp. 346–356, 2019.
- [15] A. Arshaghi, N. Razmjoooy, V. V. Estrela et al., *Image Transmission in UAV MIMO UWB-OSTBC System over Rayleigh Channel Using Multiple Description Coding (MDC)*, IET, London, UK, 2020.
- [16] 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.

- [17] R. Navid, V. V. Estrela, and H. J. Loschi, "Chapter 1: a survey of potatoes image segmentation based on machine vision," in *Applications of Image Processing and Soft Computing Systems in Agriculture*, pp. 1–38, IGI Global, Hershey, PA, USA, 2019.
- [18] V. V. Estrela, A. C. B. Monteiro, R. P. França, Y. Iano, A. Khelassi, and N. Razmjoooy, "Health 4.0 as an application of industry 4.0 in healthcare services and management," *Medical Technologies Journal*, vol. 2, no. 4, pp. 262–276, 2018.
- [19] R. Sheykhahmad, N. Razmjoooy, and M. Ramezani, "A novel method for skin lesion segmentation," *International Journal of Information, Security and Systems Management*, vol. 4, no. 2, pp. 458–466, 2015.
- [20] C. E. V. Marinho, V. V. Estrela, C. Marinho et al., "A model for medical staff idleness minimization," in *Proceedings of the Brazilian Technology Symposium*, Springer, Cham, Switzerland, 2018.
- [21] R. Navid, M. Ramezani, and N. Ghadimi, "Imperialist competitive algorithm-based optimization of neuro-fuzzy system parameters for automatic red-eye removal," *International Journal of Fuzzy Systems*, vol. 19, no. 4, pp. 1144–1156, 2017.
- [22] 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.
- [23] R. Navid, V. V. Estrela, H. J. Loschi, and W. S. Farfan, "A comprehensive survey of new meta-heuristic algorithms," in *Recent Advances in Hybrid Metaheuristics for Data Clustering*, Wiley Publishing, Hoboken, NJ, USA, 2019.
- [24] C. Choi and J.-J. Lee, "Chaotic local search algorithm," *Artificial Life and Robotics*, vol. 2, no. 1, pp. 41–47, 1998.
- [25] X. Li, P. Niu, and J. Liu, "Combustion optimization of a boiler based on the chaos and Lévy flight vortex search algorithm," *Applied Mathematical Modelling*, vol. 58, pp. 3–18, 2018.
- [26] D. Yang, G. Li, and G. Cheng, "On the efficiency of chaos optimization algorithms for global optimization," *Chaos, Solitons & Fractals*, vol. 34, no. 4, pp. 1366–1375, 2007.
- [27] C. Rim, S. Piao, G. Li, and U. Pak, "A niching chaos optimization algorithm for multimodal optimization," *Soft Computing*, vol. 22, no. 2, pp. 621–633, 2018.
- [28] 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, Cham, Switzerland, 2020.
- [29] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm," *Computers & Structures*, vol. 169, pp. 1–12, 2016.
- [30] 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.
- [31] A. Shabani, B. Asgarian, M. Salido, and S. Asil Gharebaghi, "Search and rescue optimization algorithm: a new optimization method for solving constrained engineering optimization problems," *Expert Systems with Applications*, vol. 161, Article ID 113698, 2020.
- [32] A. Kaveh and A. Dadras, "A novel meta-heuristic optimization algorithm: thermal exchange optimization," *Advances in Engineering Software*, vol. 110, pp. 69–84, 2017.
- [33] R. Navid, V. V. Estrela, and H. J. Loschi, "A study on metaheuristic-based neural networks for image segmentation purposes," in *Data Science*, pp. 25–49, CRC Press, Boca Raton, FL, USA, 2019.
- [34] D. Yu, Y. Wang, H. Liu, K. Jermsittiparsert, and N. Razmjoooy, "System identification of PEM fuel cells using an improved Elman neural network and a new hybrid optimization algorithm," *Energy Reports*, vol. 5, pp. 1365–1374, 2019.
- [35] Y. Cao, Y. Li, G. Zhang, K. Jermsittiparsert, and N. Razmjoooy, "Experimental modeling of PEM fuel cells using a new improved seagull optimization algorithm," *Energy Reports*, vol. 5, pp. 1616–1625, 2019.
- [36] Y. Cao, Y. Wu, L. Fu, K. Jermsittiparsert, and N. Razmjoooy, "Multi-objective optimization of a PEMFC based CCHP system by meta-heuristics," *Energy Reports*, vol. 5, pp. 1551–1559, 2019.
- [37] R. Navid, F. R. Sheykhahmad, and N. Ghadimi, "A hybrid neural network–world cup optimization algorithm for melanoma detection," *Open Medicine*, vol. 13, no. 1, pp. 9–16, 2018.
- [38] M. S. Kavitha, J. Shanthini, and R. Sabitha, "ECM-CSD: an efficient classification model for cancer stage diagnosis in CT lung images using FCM and SVM techniques," *Journal of Medical Systems*, vol. 43, no. 3, p. 73, 2019.
- [39] K. Senthil Kumar, K. Venkatalakshmi, and K. Karthikeyan, "Lung cancer detection using image segmentation by means of various evolutionary algorithms," *Computational and Mathematical Methods in Medicine*, vol. 2019, Article ID 4909846, 16 pages, 2019.
- [40] C.-J. Lin, S.-Y. Jeng, and M.-K. Chen, "Using 2D CNN with taguchi parametric optimization for lung cancer recognition from CT images," *Applied Sciences*, vol. 10, no. 7, p. 2591, 2020.