Biomedical Signal Processing and Control 84 (2023) 104965

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

New bidirectional recurrent neural network optimized by improved Ebola search optimization algorithm for lung cancer diagnosis

Mohammad Hosein Sabzalian [a], Farzam Kharajinezhadian [b], AmirReza Tajally [c], Reza Reihanisaransari [d], Hamzah Ali Alkhazaleh [e], Dmitry Bokov [f,g,*]

[a] Department of Mechanical Engineering, University of Santiago of Chile (USACH), Avenida Libertador Bernardo O'Higgins 3363, Santiago 9170022, Chile
[b] Faculty of Biomedical Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
[c] School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran
[d] Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA
[e] College of Engineering and IT, University of Dubai, Academic City, 14143 Dubai, United Arab Emirates
[f] Institute of Pharmacy, Sechenov First Moscow State Medical University, 8 Trubetskaya St., bldg. 2, Moscow 119991, Russian Federation
[g] Laboratory of Food Chemistry, Federal Research Center of Nutrition, Biotechnology and Food Safety, 2/14 Ustyinsky pr., Moscow 109240, Russian Federation

ARTICLE INFO

Keywords:
BRNN (Bidirectional Recurrent Neural Network)
Diagnosis
Lung cancer
Improved Ebola Optimization Search Algorithm

ABSTRACT

The early detection of cancerous and malignant lung cancer by medical imaging techniques, CT-scan for example, which never needs to do sampling reduces the risk of cancer growth and spreading. Accordingly, computer image processing and diagnostic system development, followed by cancer's classification into malignant and benign, is of primary importance in the early discovery of lung cancer which plays a pivotal role in the treatment improvement and saving the patient's life. This work intended to improve malignant and benign gland categorization accuracy and, as a result, detection accuracy. Here, a new methodology has been proposed to get an accurate lung cancer diagnosis system using an improved Bidirectional Recurrent neural network. The improvement of the network has been done by designing an improved form of an Ebola optimization search algorithm. Before applying the major diagnosis system, some preprocessing techniques have been done. The model is then applied to IQ-OTH/NCCD lung cancer dataset and its results are compared with some published works to indicate the eminence of the suggested method toward the comparative ones.

1. Introduction

Abnormal cells in lung tissue are the indicator of a lung malignant tumor. If the cell growth is left untreated, it can influence adjacent tissues and organs because it can move outs of the lungs in a process under the name of metastasis. Cough (with bloody sputum), weight loss, and shortness of breath are the most prevalent symptoms [1]. In 90% of cases, lung cancer is due to prolonged exposure to secondhand smoke [2]. Non-smokers are 15% more likely to develop lung cancer than smokers [3]. Airborne and workplace pollutants include asbestos, nickel, uranium, and radon (a gas without color and scent that is the production of the uranium's natural decomposition in water, soil, and ancient building materials). They are lungs. Genetic factors may also increase the risk of lung cancer [4]. The lungs are a big bodily organ. As a result, a malignant growth might grow for a long time in the lungs without any trace of symptoms, increasing the likelihood of a late diagnosis [5]. Cough and fatigue are two signs that can be created by lung cancer, but most individuals ignore them and attribute them to other reasons [6].

Based on the World Health Organization data [7], each year,6.7 million people die of cancer globally. Cancer is responsible for 11% of all fatalities worldwide. A lung cancer diagnosis is critical since the doctor will begin therapy for the same form of lung cancer. Early detection of cancer improves the effectiveness of treatment, as well as the patient's recovery and survival [8]. Various studies have been conducted to date to discover and describe lung illnesses. Diagnosing nodules from vessels, wounds, and other sources is challenging for a specialist physician because of the elimination and the vast majority of lung radiographic pictures, in addition to their complicated and irregular structure.

Image processing as a decision-making technique helps clinicians to

* Corresponding author at: Institute of Pharmacy, Sechenov First Moscow State Medical University, 8 Trubetskaya St., bldg. 2, Moscow 119991, Russian Federation.
E-mail address: bokovdmitry64@gmail.com (D. Bokov).

https://doi.org/10.1016/j.bspc.2023.104965
Received 30 June 2022; Received in revised form 24 March 2023; Accepted 10 April 2023
Available online 28 April 2023

detect cancer soon. The key factor in reducing cancer mortality is the early diagnosis of cancer by imaging screening [9]. Forecasting, screening, recognition, staging, prognosis, therapy planning, and response to treatment all benefit from medical imaging [10].

In other words, pictures are a crucial part of cancer clinical protocols because they can give functional, metabolic, structural, and morphological information, as well as assist in clinical decision-making when combined with other diagnostic techniques.

Instances of imaging systems are as followed: X-rays, radiography, fluoroscopy, Computerized Tomography (CT) image scans, nuclear remedy, MRI (Magnetic Resonance Imaging), and ultrasound. X-rays, radiography, and CT scans are common diagnostic methods for lung cancer.

Medical image processing with recognition of images helps to effectively identify tumors alongside the growth of cancer cell diagnostic techniques employing the sciences of engineering. Processing medical photos frequently necessitates the use of a software system and accompanying specialist software. Limited computed tomography (CT) has recently become popular for lung cancer screening.

It has been shown new computer technologies, and image processing mechanisms, for instance, are effective in cancer diagnosis and classification processes in studies. In the contemporary era, the demand for sophisticated diagnostic tools that are both accurate and timely has increased due to the rising prevalence of lung ailments in developed areas.

A computer diagnostic aid system helps the doctor diagnose the disease. The benefits of computational detection methods have reduced selection error and positive error as well as increased accuracy. Diagnosis of swellings from wounds, vessels, etc. is a demanding task for a specialist physician because of the elimination and the great number of lung radiographic images, as well as its intricate and unusual structure. A computer recognition system is a useful system that helps the doctor in disease diagnosis.

Lung glands are one of the signs of lung cancer. Because these tiny glands are difficult to spot with the human eye on CT scans of the lungs, machine intelligence or computer-aided detection (CAD) technologies are used. They can assist in discovering, finding, and assessing gland quality as an expert. The most significant problem facing present intelligent systems is balancing improvements in accuracy, detection, sensitivity, and the decrease in false positive error rate (FPr), as well as the complexity of these systems, reducing efficiency and speed of execution. Agility and problem optimization are also taken into account.

A method that converts an image to digital and implements some operations on it to gain an improved image or some useful information is Image processing.

Manual approaches, such as sequential floating feature selection algorithms or evolutionary algorithms, were previously employed to extract image attributes. They had less of an impact on large-scale properties like photos. Deep learning and the trend of CNN systems, which have been frequently applied in the interpretation of medical pictures in recent years, are two new subfields of artificial intelligence.

Many research works on the application of artificial intelligence based on image processing have been conducted, some of which are included below.

Zhou et al. [11] used artificial neural networks to detect lung malignant cells. An artificial neural network ensemble is a learning paradigm in which numerous artificial neural networks work together to solve a problem. This work proposes Neural Ensemble-based Detection (NED), an automatic disease detection approach. With great confidence, the ensemble is based on two-level ensemble architecture. According to the findings, NED not only reaches the highest percentage of overall recognition but also a low percentage of false negative identification.

Lin et al. [12] utilized a fuzzy neural network to detect the recovered lung nodes. A neural fuzzy model is presented to extract the diagnosis guidelines for detecting pulmonary nodules. The proposed technique has an 89.3 percent detection rate and a false positive rate of about 0.3 per image. In comparison to other approaches, these results demonstrate that our approach improves prediction accuracy and reduces false positives. It also implies that this system has clinical application potential.

Šarić et al. [13] applied the CNN-based technique to detect lung cancer. Lung cancer must be diagnosed early to increase patient survival. A standard approach for early diagnosis is a histopathological examination of tissue. Tissue analysis is normally done by a pathologist although this is a time-consuming and error-prone action. We offer a full mechanical technique for the identification of lung cancer in all slide photographs of lung clinical samples. To categorize images at the patch level, Convolutional Neural Networks (CNN) are utilized. The findings demonstrate that a CNN-based strategy could aid pathologists in diagnosing the disease.

Zang et al. [2] employed R-CNN-based PET imaging to find lung cancers. One of the most reliable ways for the recognition of many cancers, specifically lung tumors, is PET (Positron Emission Tomography). Radiologists are becoming significantly overworked with the increased use of PET scans., A novel approach, “computer-aided diagnostics”, is being considered to reduce the heavy workloads. The authorities introduced a multiscale Mask R-CNN–based (MaskRegion–Based Convolutional Neural Network) method on PET imaging basis for lung tumor detection. The proposed technique had recall, precision, and F score values of 1,0.90, and 0.95, respectively. The results of the experiments demonstrated a firm belief in the utility of the method in identifying lung tumors, as well as its capacity to identify a healthy chest pattern and significantly reduce inaccurate tumor detection. Agarwal et al. [14] diagnosed and classified lung cancer based on Alexnet CNN. One of the most significant challenges that humans have faced in recent years has been Lung tumor identification and medication. This research describes a solution that includes an AlexNet Network Model and a Convolutional Neural Network (CNN) to categorize lung cancers as malignant or benign. The suggested CNN results are highly accurate which is more efficient in comparison with CNN (Conventional Neural Network) systems.

Lung cancer is the leading cause of cancer-related deaths worldwide. Early diagnosis of lung cancer is critical for successful treatment and improved survival rates. Machine learning and deep learning techniques have emerged as promising candidates to enable accurate and timely diagnosis of the disease. Existing lung cancer diagnosis systems based on deep learning techniques still have limitations in terms of accuracy, sensitivity, and specificity. The lack of standardized and annotated datasets for training and evaluating the models poses significant challenges to the development of reliable and robust diagnosis systems.

This work aims to utilize deep learning to detect lung cancer with higher accuracy. The use of deep learning produces acceptable results for disease diagnosis, but in the presented work, an effort has been made to invent a modified deep learning, called a bidirectional recurrent neural network, for the sake of producing more satisfactory outcomes. In addition to the deep neural network, an optimization procedure is introduced in this study to promote the system’s function in disease diagnosis. Optimization procedures have been increasingly applied in BRNN optimization due to their outstanding performance in solving difficult optimization problems. Here are some advantages of using metaheuristic algorithms in BRNNs:

- Better optimization results: optimization procedures are capable of finding the global or near-global optimum of complex, nonlinear optimization problems. By using metaheuristic optimization, BRNNs can arrive at better solutions and achieve better performance.
- Reduced computation costs: optimization procedures can reduce the computational burden of BRNN optimization by avoiding the need for exhaustive search methods, such as grid search. This is particularly useful when the number of network parameters is large.
- Flexibility: optimization procedures are flexible, i.e., they can accommodate different types of optimization goals and constraints.

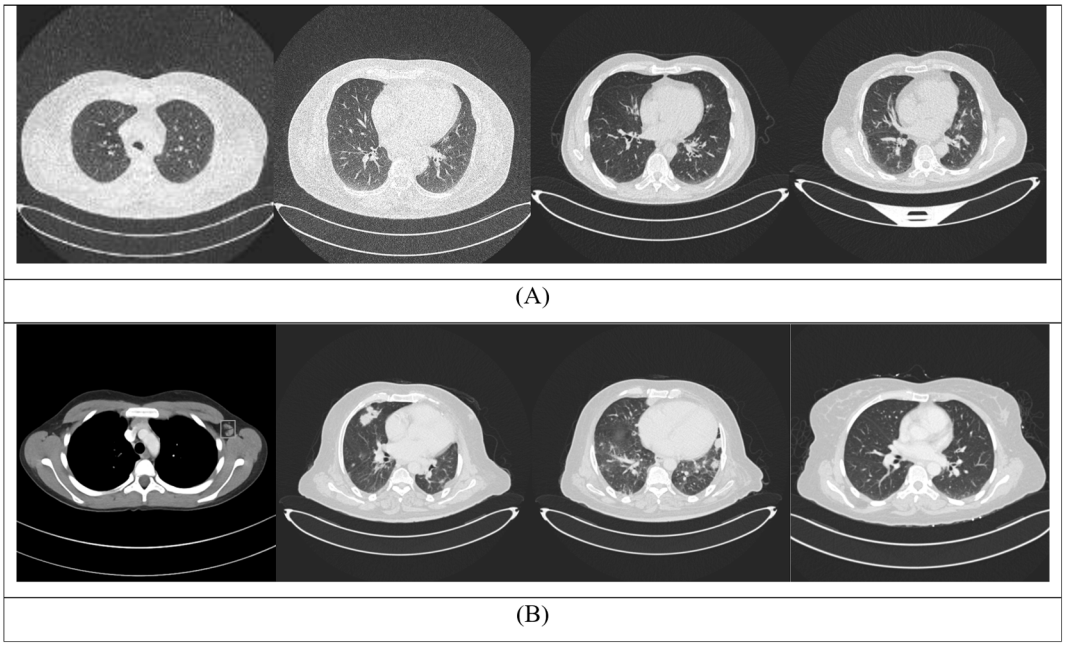

Fig. 1. Some lung CT scans of malignant and benign cases from the IQ-OH/NCCD dataset.

BRNNs with different optimization objectives, such as classification or prediction, can benefit from using metaheuristic algorithms.
- Robustness to noisy data: optimization procedures are generally robust to noise and uncertainty in data, which is important for BRNNs that operate on real-world data sets.
- Shorter training time: By converging faster to near-optimal solutions, metaheuristic algorithms can reduce the time required to train BRNNs. This is especially beneficial in real-time applications where time is critical.

Overall, a combination of optimization procedures and BRNNs can yield powerful optimization results that can enhance the performance of BRNNs beyond what standard gradient-based optimization methods can achieve. The new optimizer approach is based on improving the Ebola search optimization algorithm to provide more efficient results. The proposed Ebola search optimization algorithm is based on opposition-based learning (OBL). The OBL involves generating complementary and conflicting solutions and using them to diversify the search space, leading to better optimization results. Using OBL provides enhanced exploration capability, improved solution quality by creating complementary solutions, increased convergence speed, and higher robustness. The method of diagnosis will be presented in the following section along with a full description of the suggested strategy. Therefore, the main contributions of the present study can be highlighted as follows:

- New optimal deep method for lung cancer detection from CT Images
- A new methodology based on improved Bidirectional Recurrent neural network (BRNN).
- Designing an improved form of Ebola optimization search algorithm for optimizing the BRNN.
- Applying to IQ-OTH/NCCD lung cancer dataset and comparison with some published works.

2. Dataset description

In the current study, the dataset of lung cancer obtained from IQ-OH/NCCD (Iraq Oncology-teaching Hospital or National Center specialized for Cancer Diseases) was used to validate the represented methodology. The IQ-OH/NCCD dataset is collected from special hospitals over a 3-month interval in the fall of 2019. This data covers Computed Tomography (CT) images of cases infected with lung cancer and healthy ones in various stages. Radiologists and oncologists have also marked this dataset. The IQ-OH/NCCD dataset includes over-all 1190 CT images consisting of 110 cases. The dataset includes three classes of healthy, benign, and malignant where, fifteen cases were identified as benign, fifty-five cases were identified as normal cases, and forty cases were identified as malignant. The format for the images is in DICOM. All photos are in "*.jpg" format and are acquired from the Kaggle website [15].

We also used the Lung-PET-CT-Dx dataset for more evaluation. In this dataset, lung cancer patients' CT and PET-CT DICOM pictures are annotated with XML files that show the bounding boxes where the tumors are located. According to a tissue histopathological diagnosis, subjects were divided into groups. To create algorithms for medical diagnosis, the locations of each tumor were marked by five academic lung physicians with experience in lung cancer. Fig. 1 illustrates some Computed Tomography (CT) images of malignant and benign cases of the lung from the IQ-OH/NCCD dataset.

3. Data preprocessing

The captured CT images have sometimes inconsistencies which reduce their quality. This process increases the complexity of the main process, like the diagnosis of the disease. Therefore, we need some preprocessing stages to enhance the images' quality [16]. There are diverse kinds of preprocessing techniques for this purpose. In this study, we use two preprocessing methods, including noise cancellation and contrast enhancement to improve the images' quality.

3.1. Median filtering

The preprocessing stage usually tries to remove the noise. Undoubtedly, CT images also have noises that need to be removed. Of course, removing noise as much as possible should not damage the edges and reduce the clearness of the image [17]. The average and the median filter are the same, in a median filter, each achieved pixel is established from the average intensity values of the adjacent pixels corresponding to the input ones [18–21]. However, by median filtering, the value of one obtained pixel is controlled by the median of the adjacent ones instead of the average pixel. The sensitivity of the center filter for throwing values

Input: Image X of size $M \times N$, filter size n
Output: Image Y of the same size as X
Initialize histogram H *and* starting point k
for $i = 1$ **to M do for** $j = 1$ *to* N **do**
for k = $-n/2$ to n/2 **do**
Remove $X_{i+k,j-n/2-1}$ from H
Add $X_{i+ki+n/2}$ to H
Update k
end for
$Y_{i,j} \leftarrow$ median(H, k)
end for
End for

Fig. 2. Fast median filtering algorithm.

is low, so it has a high capability to delete these points without a reduction in the image resolution.

Based on Huang's technique, we consider an approach for determining the median value by applying a beginning point to a histogram in this study. The focus of the median filtering method is on how quickly the position of the median values is discovered in the histogram. The suggested methodology is identical to Huang's histogram update method, except it employs the previously determined beginning point to calculate the median value more rapidly [22–25]. When the histogram is updated, the location at which the histogram's value begins is also changed, which is referred to as the starting point. In contrast to the present technique, which finds the median value progressively from 0, the median value is found from the prepared beginning point position [26].

Achieving the histogram's position of the median value while utilizing the beginning point is faster than Huang's approach if the size of the filter is small. The pseudocode of the fast median filter is given in Fig. 2.

Although, the fast median technique with a large filter size is designed to eliminate noise from high-resolution images quickly. However, when used in applications, there is a serious issue with removing a little structure that is a component of the image. This is a major issue in areas where median filters are utilized, such as the diagnosis system. As a result, using a small-size median filter is appropriate. Currently, the suggested method has the advantage of being sufficiently quick, helpful, and simple. Fig. 3 shows a sample example of noise cancelation by median filtering for a CT image.

3.2. Contrast enhancement

Image enlargement aims to either give far better input for automated-processing systems or improve the analysis of the contained information, for the viewer. Contrast is one of the most crucial quality factors in the processing of diverse images.

Image contrast enhancement techniques are mathematical that assist to increase the quality of an image. The result is an improved image that shows certain features in a way that is better than the original image. It is also possible to produce multiple processed versions of the original image, each offering a more accurate appearance. One of the most attractive and simplest areas of digital medical imaging is this technique (enhancement of image contrast). This step aims to clarify details of the obscure image or to highlight specific parts of the image that are being looked for [27,28]. Image enhancement features include Brightness and Contrast. The image information will be cleared in brighter regions and consistently focused providing that image contrast is more black, brighter, or too focused. As a result, the image's contrast ought to be increased to display all information from the input picture.

One of the most common methods in contrast enhancement is General histogram equalization (GHE), this method is used in approximately all kinds of images owing to its almost great operation and simplicity. The fundamental idea of GHE is the reprogramming of the input pictures' gray surfaces based on Probability Density Function (P.D.F) uniform expansion. General histogram equalization (GHE) levels out the image P.D.F dynamic ranges, rising the input image contrast, as much as possible. This method is likely to produce an intensity saturation dilemma, create irritating artificial effects, and increase the noise in the obtained image due to over-enhancement. Numerous methods were offered to deal with these limitations.

DSHE (Dualistic sub-image histogram equalization) was one of the initially suggested methods. In this method, the main histogram is divided into two sub-histograms based on the average intensity values. Afterward, the adjustment process is separately implemented in both sub-histograms. It is imperative to note that GHE maintains the average brilliance of the original image more than this method, but, in some cases, it has a problem with intensity saturation. What is more, the histogram adjustment method was presented under the dual image. In this procedure, the major histogram is divided into two sub-histograms with the same quantity of pixels, additionally, the middle of the intensity values is utilized as the separating point. Allegedly, it works far more

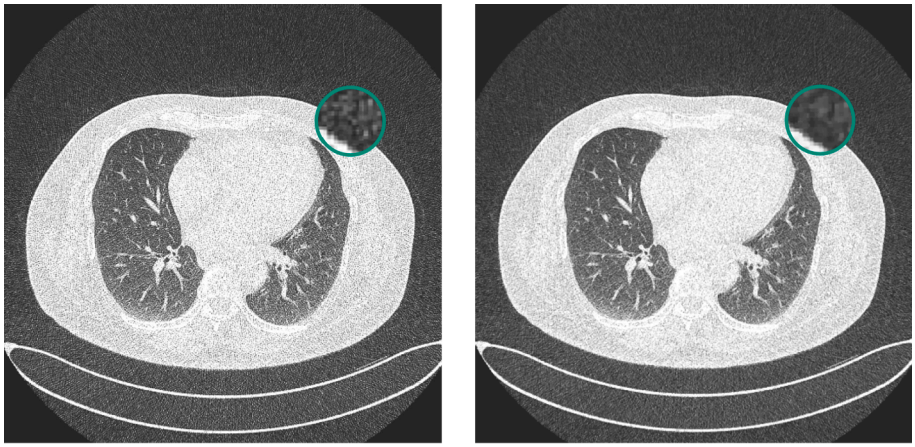

Fig. 3. Sample example of noise cancelation by median filtering for a CT image.

Input: raw image with low contrast.
Output: enhanced image with higher contrast.
Generate the histogram of the input image.
Evaluate the median value of the histogram
Divide histogram into two sections based on median value of the histogram
Equalize the parts independently using PDF and CDF

Fig. 4. DSHE algorithm pseudocode.

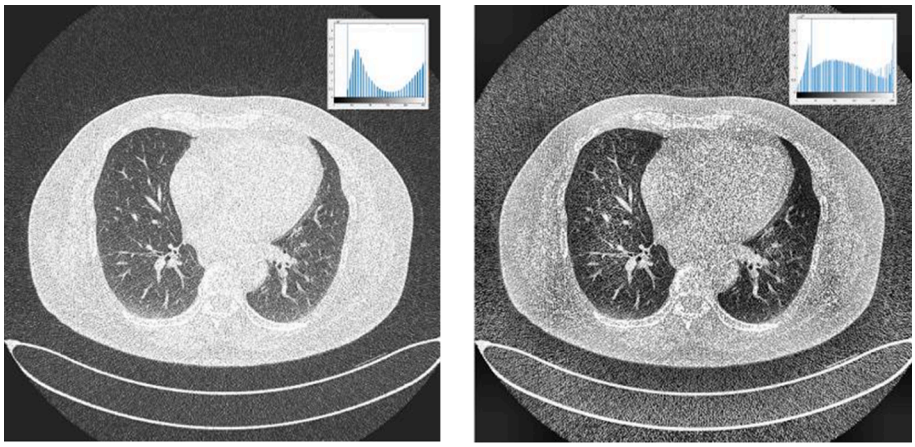

Fig. 5. A sample example of contrast enlargement based on the CT image on the DSHE basis (A) original image, and (B) contrast-enhanced image.

effectively than the previous method in information retention (entropy) and maintaining the average brightness terms. Fig. 4 illustrates the pseudocode of the DSHE algorithm.

Fig. 5 indicates an example of contrast enlargement based on the CT image on the DSHE basis.

3.3. Data normalization

Normalization of data is the act of organizing data within a database so that users can better utilize it for future queries and analysis, taking into account all of the many explanations included therein. In reality, data normalization is a way of structuring data attributes to raise the coherence of various entities in the data form.

Creating excellent data is commonly referred to as data normalization. However, a more in-depth examination of the meaning or goal of data normalization has yielded nothing. The sequence in which information seems to be the same in all documents and fields is referred to as data normalization. Improving the coherence of many forms of data input, resulting in greater quality sorting, segmentation, and data. During the data normalization, the data limit is assumed identical in the range [0, 1], with zero means. Here, the data minimum and maximum values are set to 1 and 0. The present study uses the Min-Max approach for data normalization. Based on the Min-Max approach, by considering an image as I, with d dimension, the normalized output for this image (I_n) can be achieved as follow:

$$I_n = \left(I - \underline{I}\right) \times \frac{\overline{I}^n - \underline{I}^n}{\overline{I} - \underline{I}} + \underline{I}^n \quad (1)$$

where, $\underline{I}$ and $\overline{I}$ represent the minimum and maximum values of image I.

3.4. Image augmentation

Image data augmentation is a technique that may be used to artificially enhance the amount of data without using new images and merely by deploying current images. More data can be utilized for training the convolutional neural networks (CNNs) and image data augmentation techniques may be used to produce different sorts of images. As well as the original images, the network learns additional aspects of those images, representing the understanding capacity.

The CNNs are modern deep-learning algorithms that can learn characteristics that are independent of their placement in the image. However, amplifying and adding data by changing rotation, light, and color can aid in the learning of independent features in the image; eventually, the grid learns to recognize the object correctly even if it was turned from left to right in the image, had little light and clarity, or a variety of other factors. In this study, the SdSmote technique is utilized for image augmentation.

Not all minority samples are ideal for developing new synthetic samples, as certain minority samples are relatively easy to learn. As a result, instances that are difficult to teach should be chosen. Most algorithms hunt for these samples at the decision border while generating fake data. By representing the degree of support as a criterion, that is consequent of utilizing the sample distance concept, this plan searches for pick boundary samples to generate fake samples. To evaluate the degree of support, the total distance of this sample from all negative classes is obtained as follows:

$$S_i = \sum_{i=1}^{p} \sqrt{\|x_i - x_j\|^2}$$

$$S = \sum_{j=0}^{q} x_i \tag{2}$$

where, m and n represent the positive and negative class data quantity, respectively, and x_i assume the positive class samples, and S_i describes the sum of the previous step. Eventually, the distance average value is obtained by equation (3) as followed:

$$\widehat{S} = p^{-1} \times q^{-1} \times S \tag{3}$$

where, $\widehat{S}$ specifies the distance average value.

4. Bidirectional RNN (Recurrent neural Networks)

The RNN (Recursive Neural Network) is one of the most powerful and well-constructed types of neural networks, it is also one of the most promising types of neural networks because of its internal memory. RNNs, like many other neural networks, are becoming old. Although they were originally developed in the 1980 s, it only took a few years for their entire potential to be realized [29]. It is worth mentioning that three factors namely, computer processing power development, a great number of data, and the burgeoning of long-term short-term memory (LSTM), provide fertile grounds for the progression of the RNN.

While standard deep neural networks consider that inputs and outputs are fully independent of one another, RNNs may recall critical details about the inputs they've received, allowing them to anticipate what will happen in the future [30,31]. This indicates the RNN's effectiveness with different types of data such as financial data, time series, text, voice, audio, and video to name but a few. Compared to other algorithms, the RNN can get a better comprehension of the sequence and context knowledge of the inputs.

A feed-forward neural network maps an input to an output, but a return neural network does not. The lengths of the inputs and outputs in these networks can also vary. RNN networks are utilized in a variety of applications, including music composition, emotion identification, and machine translation.

By considering the $L = [l_t]$ as the input of RNN, and $Z = z_t$ as an output vector, we aim to model the distribution $P(Z|L)$. However, the RNNs can map the input to output, they can be utilized for the succeeding input forecasting. One of the popular models of the RNNs is unidirectional RNN which can be defined as follows:

$$P\left(z_t|[l_i]_{i=1}^{t}\right) = \sigma(W_z x_t + b_z) \tag{4}$$

where $l_t \in R^N$, $z_t \in R^M$ are input and output vectors for time step t.

$$x_t = \tanh(W_x x_{t-1} + W_l l_t + b_h) \tag{5}$$

where, b_x and b_z represent the vector of hidden and output layers, respectively, W_l, W_x and W_Z denote the weight matrix to connect the input layer to hidden, hidden to hidden, and hidden to output layers, respectively.

Based on this form, the RNN calculates the output based on the distributed information over the hidden layer yet of whether it depends on the values l_i.

With adding an extra hidden layer, bidirectional RNN (BRNN) has been made, where the hidden-to-hidden layers are arranged in reverse sequential order.

Bidirectional Recurrent Neural Networks (BRNN) is a type of artificial neural network that allows for the analysis of sequential data in both forward and backward directions. This means that the network can take past and future context into account while processing the current input. Some of the advantages of BRNN include:

Improved context awareness: One of the primary advantages of BRNN is its ability to analyze the context in both forward and backward directions. This allows the network to better understand the relationships between elements in a sequence, leading to improved performance in tasks that require context awareness.

Enhanced accuracy: BRNN improves the accuracy of predictions by considering both forward and backward contexts, making it well-suited for tasks such as speech recognition, machine translation, and sentiment analysis.

Efficient use of data: BRNN is capable of effectively using all the available data by taking advantage of the entire sequence. This makes it particularly useful in applications such as handwriting or speech recognition where the input data is typically sequential.

Robustness: The bidirectional nature of the network allows it to better handle noisy inputs and variations in the data. This makes it more robust and less likely to make errors than unidirectional networks.

Overall, the use of a bidirectional recurrent neural network can lead to improved accuracy, better handling of sequential data, and more robust performance in a variety of applications.

Accordingly, this model can search both earlier and upcoming directions. Her, model output is achieved as follows:

$$P\left(z_t|[l_i]_{i\neq t}\right) = \sigma\left(W_z^g x_t^g + W_z^b x_t^b + b_z\right) \tag{6}$$

$$x_t^g = \tanh\left(W_x^g x_{t-1}^g + W_l^g l_t + b_x^g\right) \tag{7}$$

$$x_t^b = \tanh\left(W_x^b x_{t+1}^b + W_l^b l_t + b_x^b\right) \tag{8}$$

It's worth noting that in the bidirectional RNN's backward pass, there are two steps for back-propagation. This is responsible for changing the weights to reduce the MSE (Mean Square Error); however, in the represented work, a novel version of the Ebola optimization search algorithm is used.

5. Parameter optimization

The weights of a BRNN are one of the primary parameters that have a considerable impact on network efficiency. This word indicates how much of an impact hidden layers have on input–output relationships and learning rates. The output has been formed by the input layer and the weights multiplication followed by adding up. The degree of neurons' impact on each other is determined by their weights. To put it another way, the neurons' output achieved by Eq. (9) provided that the inputs and their weights are described as $X = [x_1, x_2, ..., x_m]$ and $W = [w_1, w_2, \cdots, w_m]$.

$$z = f(l) = \sum_{i=1}^{m} l_i w_i + b \tag{9}$$

where,b describes the bias for the output adjusting, and m represents the number of inputs.

Numerous deep neural networks, notably the BRNN, suffer from the unstable gradient problem. The elementary source of this issue is that as we travel backward from the buried layers, the gradient grows lesser. This means that neurons in the higher layers learn far faster than those in the lower layers. This phenomenon is called the vanishing gradient issue. Developed Ebola Optimization Search Algorithm was used in this work to overcome the BRNN problem's unstable gradient issue. This algorithm was utilized to discover the best biases and weights to minimize error, i.e.,

$$F_E = \frac{1}{M} \sum_{i=1}^{M} (d_i - z_i)^2 \tag{10}$$

where, z_i and d_i denote the output and the desired data, respectively.

Therefore, the main objective of this study is to minimize F_E to get the best accuracy.

6. Ebola optimization search algorithm

Algorithms inspired by natural phenomena and biology are among the algorithms of the *meta*-heuristic that have been introduced to figure out various optimization problems in engineering and other relevant

Table 1
Explanations and symbols for parameters and variables used in the current study (for SEIR-HDVQ).

Symbols	Data Type	Explanations
E	Vector	Exposed individuals
S		Susceptible individuals
H		Hospitalized infected individuals
I		Infected individuals
D		Diseased from infection individuals
R		Recovered infected individuals
PE		Agents capable of infecting individuals
V		Vaccinated infected individuals
π	Scalar	Recruitment rate of susceptible human individuals
α		Rate of hospitalization of infected individuals
η		Decay rate of Ebola virus in the environment
Γ		Disease-induced death rate of human individuals
β_1		Contact rate of infectious human individuals
β_2		Contact rate of pathogen Individuals/ environment
β_3		Contact rate of deceased human individuals
β_4		Contact rate of recovered human individuals
τ		Natural death rate of human individuals
γ		Recovery rate of human individuals
ϑ		Rate of vaccination of individuals
δ		Rate of burial of deseased human individuals
μ		Rate response to vaccination
ω		Rate of response to hospital treatment
ζ		Rate of quarantine of infected individuals

subjects [32]. These algorithms provide solutions to complex life problems by exploiting and converting the basic processes of nature into specific procedures or laws and then mathematically modeling them [33,34]. Generally speaking, they are algorithms on a population basis, for instance, WHO (Wildebeest Herd Optimization), EOA (Earthworm Optimisation Algorithm), SMA (Slime Mould Algorithm), IWO (Invasive Weed Colonization Optimization), Emperor Penguin Salp Swarm Algorithm (ESA), SBO (Satin Bowerbird Optimizer), COA (Coronavirus Optimization Algorithm), Virus Colony Search (VCS), BBO (Biogeography- Based Optimization). Despite the swarm-based algorithms such as WOA, ABC, and PSO, and evolutionary-based algorithms such as DE and GA contributing some features of an algorithm inspired by biology, we have opted for restricting our analysis to listed algorithms [35].

Two phenomena combination, extracted from the E.P.A (Emperor Penguin Algorithm) and the S.P.A (Salp Swarm Algorithm), provides ESA in which the conduct of these two creatures is modeled. The offered algorithm reveals far better performance in convergence, sensitivity analysis, and scalability in comparison with other analogous metaheuristic algorithms. Based on the coronavirus propagation strategy, COA is offered. Likewise, in the human safety strategy, another type of optimization called CHIO is proposed. By implementing COA in CNN (the convolutional neural network) problem design, its productiveness was assessed. CHIO displays a powerful role in practical engineering issues. The optimization algorithm of earthworms is inspired by the nature of earthworm reproduction, which includes two reproductive schemes. In the first strategy, an earthworm can generate one child, but in the second strategy, it can generate two or more. This reproducibility is managed by the C.M. (Cauchy Mutation) procedure permitting crossing operators.

6.1. EOSA (Ebola optimization search Algorithm)

Given that, we are in urgent need of understanding the performance of a model, based on SEIR, in the transmission of an ailment if we aim to devise an algorithm for optimization, this area represents an improved SEIR model on EVD. Furthermore, a representation of the EOSA proceeding flow and the related flow chart is displayed and reviewed [36]. Finally, to validate the suggested algorithm, the SEIR model is represented which utilizes a mathematical model, and the algorithm after that.

6.2. SIR model of EOSA

The current study chooses two corresponding models among the present SEIR models by adding and determining new sections that are considered elimination. Considering that the Ebola virus is only transmitted by an infected person from the reservoir, the determined containers are repositories for viruses, vaccinations, and quarantines, indicated by *PE*, *V*, and *Q*. What is more, the virus prevalence is overshadowed by quarantined infected individuals and the role of vaccination. For this purpose, the dissemination model is re-modeled to acquire SEIR-HDFVQ: I (Infected), R (Recovered), S (Susceptible), E (Exposed), H (Hospitalized), D (Death or dead), F (Funeral), V (Vaccinated), and Q (Quarantine). The proposed model assumes that a small number of improved cases can still carry the disease and keep the virus in their body fluid. Accordingly, the EVD propagation model is employed for the improvement of an optimization algorithm, notably, investigating all factors is crucial for supporting increased infection.

Table 1 illustrates the Model of the SEIR-HDVQ variables. The propagation of EVD is considered to supply an appropriate method to solve some optimization issues given that communities are affected by its aggressive infection rate. A group of susceptible people has considered the closeness of this group creates a community of other subgroups. Agent reservoirs, vulnerable individuals, and polluted environments can accidentally catch people from susceptible individuals to the infected group, which may be because of their proximity to any of the infected subgroups. The subgroups of infected people include infected person, dead person, recovered person, infected agent reservoir, and contaminated environment. We suppose that the virus can rot in its contaminated environment. Affected people can also heal without hospitalization or die without moving to the hospital. Also, any vaccinated person was considered hospitalized and it was assumed that any hospitalized person (H) or any non-hospitalized person could be in the dead group (D). Meanwhile, the ones who recuperated (R) from getting vaccinated (V) are considered susceptible (S). Table 1 summarizes the rate of parameters and variable alteration used in the current study.

6.3. Pseudocode of the EOSA

1) The structure of the EOSA algorithm is inspired by the execution of the SEIR-HDVQ model. Then, to execute the algorithm, each scalar and vector quantities (variables and people), are initialized: infected (I), Dead (D), Recovered (R), susceptible (S), Hospitalized (H), Vaccinated (V) and Quarantine (Q).
2) Create the index items (I_1) from susceptible people haphazardly.
3) Compute the amount of fitness of the index item and place it as the present best and the global best.
4) When the numeral of repetitions is not finished and there is one afflicted person at any rate:
a) Any susceptible people create and update their location founded on their relocation. It is imperative to note that the more increased the displacement of an infected item, the higher the quantity of infections, and the short movement explains the exploitation mode, otherwise exploration of the algorithm is done.
I. Create recently infected people (nI) according to(a).
II. The recently created cases are added to (I).
b) Calculate the number of people who are added to D, H, R, Q, B, and V, utilizing their linked rates about I size.
c) I and S are updated on an nI basis.
d) the present finest is selected from I in comparison with the worldwide finest.
e) If the termination conditions are not met, we return to stage 6
5) The global best solution and every solution are replaced.

The mathematical model of the algorithm as applied in the flowchart is represented below.

Table 2
Parametric values of the algorithms in the literature.

Algorithm	Parameter	Value
PIO [43]	Dimension of space	20
	Compass and map determinant	0.2
	Map and compass limitation	150
	Limitation of landmark	200
	w	1
	C_1	1.2
	C_2	1.2
COA [44]	r_1 and r_2	Random
	m	Chaotic
HBO [45]	Degree	4
	C	$T/25$
	p_1	$1-t/T$
	p_2	$p_1 + [(1-p_1)/2]$
SDO [46]	MaxIteration	200
	MarketSize	40
	FunIndex	1

6.4. EOSA: Mathematical model

For updating the locations of any exposed person, Eq. (11) operates:

$$mI_i^{t+1} = mI_i^t + \rho M(I) \tag{11}$$

where, the scale factor of relocation of a person is indicated by ρ, the updated and primary locations at time t and $t+1$ are represented by mI_i^{t+1} and mI_i^t. $M(I)$ is the motion ratio, created by people, and is defined as follows:

$$M(I) = srate^{*}rand(0,1) + M(Ind_{best}) \tag{12}$$

$$M(s) = lrate^{*}rand(0,1) + M(Ind_{best}) \tag{13}$$

Assuming that the infected person remains at a distance of zero (0) or relocations in a range that is not more than *srate*, the exploitation step is prepared in the algorithm. Where *srate* indicates short space motion. Also, considering that the infected person's motions exceed the average neighborhood *lrate*, the exploration step is performed in the algorithm. The greater the relocations, the more quantity of people in S is prone to infection. These two states are clarified in Equations (12) and (13). A neighborhood variable sets *srate* and *lrate* so that if neighborhood ≥ 0.5, a person has motioned exceeds the neighborhood causing the mega infection; if not, it stays inside the neighborhood, that harnesses disease.

- the susceptible population Initialization

In the starting, by distributing accidental numbers with zero beginning locations, a primary population is created. The candidate is created according to formula (14). The higher and lower ranges for the i^{th} the candidate is indicated by U_i and L_i, and i defines from $[1,2,3,...N]$, in the size of the population.

$$individual_i = L_i + rand(0,1)^{*}(U_i + L_i) \tag{14}$$

The following equation shows by far the Best present calculation obtained on the afflicted people at t:

$$bestS = \begin{cases} gBest, fitness(cBest) < fitness(gBest) \\ cBest, fitness(cBest) \geq fitness(gBest) \end{cases} \tag{15}$$

Where, the global finest solution, finest solution, and present finest solution at t are denoted by *gBest*, *bestS*, and *cBest*; the cost function devoted to the problem is indicated by *fitness*. cBest and gBest are determined as infected candidates who are Ebola virus Spreaders and Superspreaders.

Given that differential calculus relates to the rate of variation of one quantity involving another, integrals find different properties of derivatives and integrals. In the present study, using differential calculus, the change rates of the S, H, I, R, D, V, and Q quantities at time t are obtained, which are presented below:

$$\frac{\partial S(t)}{\partial t} = \pi - (\beta_1 I + \beta_3 D + \beta_4 R + \beta_2 (PE))S - (\tau S + \Gamma I) \tag{16}$$

$$\frac{\partial I(t)}{\partial t} = (\beta_1 I + \beta_3 D + \beta_4 R + \beta_2 (PE)\lambda)S - (\Gamma + \gamma)I - (\tau)S \tag{17}$$

$$\frac{\partial H(t)}{\partial t} = \alpha I - (\gamma + \varpi)H \tag{18}$$

	Predicted positive	Predicted negative
Positive class	True Positive (TP)	False Negative (FN)
Negative class	False Positive (FP)	True Negative (TN)

Fig. 6. Confusion matrix of TP, TN, FP, and FN.

Table 3
Comparison of the statistically gained answers of the suggested DESO algorithm and other state-of-the-art metaheuristic algorithms.

Function	Indicator	SDO [46]	ESO [36]	PIO [43]	COA [44]	HBO [45]	DESO
F1	Avg	7.58×10^2	7.65×10^2	9.5×10^2	7.00×10^9	7.97×10^6	3.55×10^7
	StD	5.00×10^2	$3.38 \times 10\text{-}^1$	5.41×10^4	3.01×10^4	4.43×10^3	3.50×10^3
F3	Avg	6.27×10^2	6.41×10^{-1}	10.63×10^2	4.91×10^2	6.55×10^2	6.91×10^1
	StD	7.33×10^2	5.85	8.07×10^2	4.79×10^1	7.72×10^2	3.85×10^1
F5	Avg	8.11×10^2	6.10×10^{-14}	8.73×10^2	3.06×10^1	8.59×10^2	3.48×10^1
	StD	9.35×10^2	7.58×10^2	10.51×10^2	4.79×10^1	9.95×10^2	4.32×10^1
F7	Avg	10.31×10^2	6.25	10.98×10^2	4.45×10^1	10.62×10^2	3.78×10^1
	StD	21.00×10^2	5.51	4.44×10^3	9.85×10^2	3.66×10^3	6.87×10^2
F9	Avg	3.77×10^3	2.22	4.98×10^3	6.45×10^2	3.61×10^3	4.71×10^2
	StD	7.58×10^2	7.65×10^2	9.5×10^2	7.00×10^9	7.97×10^6	3.55×10^7
F11	Avg	5.00×10^2	$3.38 \times 10\text{-}^1$	5.41×10^4	3.01×10^4	4.43×10^3	3.50×10^3
	StD	6.27×10^2	6.41×10^{-1}	10.63×10^2	4.91×10^2	6.55×10^2	6.91×10^1
F13	Avg	7.33×10^2	5.85	8.07×10^2	4.79×10^1	7.72×10^2	3.85×10^1
	StD	8.11×10^2	6.10×10^{-14}	8.73×10^2	3.06×10^1	8.59×10^2	3.48×10^1
F15	Avg	9.35×10^2	7.58×10^2	10.51×10^2	4.79×10^1	9.95×10^2	4.32×10^1
	StD	10.31×10^2	6.25	10.98×10^2	4.45×10^1	10.62×10^2	3.78×10^1
F17	Avg	21.00×10^2	5.51	4.44×10^3	9.85×10^2	3.66×10^3	6.87×10^2
	StD	3.77×10^3	2.22	4.98×10^3	6.45×10^2	3.61×10^3	4.71×10^2
F19	Avg	7.58×10^2	7.65×10^2	9.5×10^2	7.00×10^9	7.97×10^6	3.55×10^7
	StD	5.00×103	3.38×10^4	5.41×10^4	3.01×10^4	4.43×10^3	3.50×10^3
Friedman mean rank		6.32×10^2	6.27×10^2	6.41×10^2	10.63×10^2	4.91×10^2	6.25×10^2
Rank		2	3	5	4	6	1

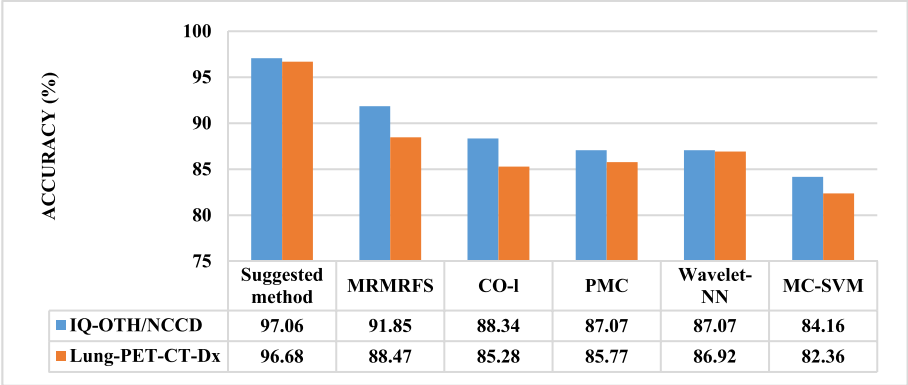

Fig. 7. Accuracy results of all methods applied to IQ-OTH/NCCD and Lung-PET-CT-Dx dataset with preprocessing.

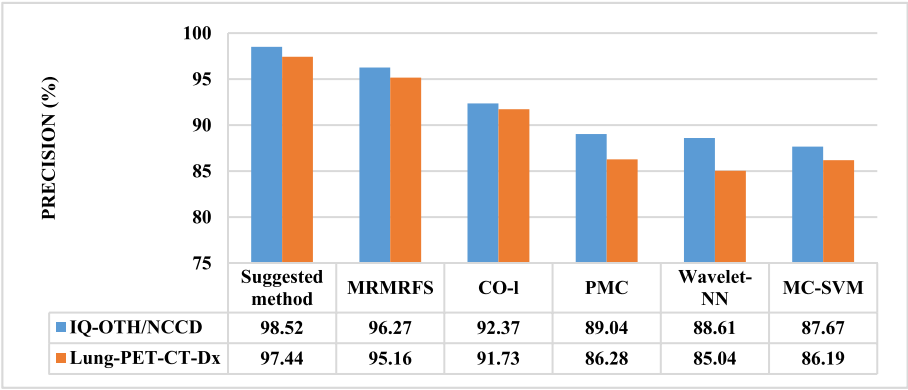

Fig. 8. Precision results of all methods applied to IQ-OTH/NCCD and Lung-PET-CT-Dx datasets with preprocessing.

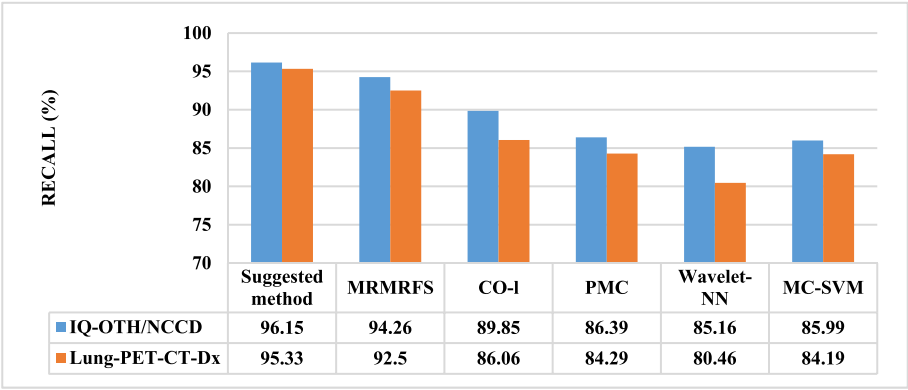

Fig. 9. Recall results of all methods applied to IQ-OTH/NCCD and Lung-PET-CT-Dx datasets with preprocessing.

$$\frac{\partial R(t)}{\partial t} = \gamma I - \Gamma R \tag{19}$$

$$\frac{\partial V(t)}{\partial t} = \gamma I - (\mu + \vartheta)V \tag{20}$$

$$\frac{\partial D(t)}{\partial t} = (\tau S + \Gamma I) - \delta D \tag{21}$$

$$\frac{\partial Q(t)}{\partial t} = (\pi I - (\gamma R + \Gamma D))) - \xi Q \tag{22}$$

Eqs. (6)–(11) are considered scalar functions, which means that each equation has one number which can be described as a float.

To acquire the number of susceptible individuals number at t-time, first of all, we define the variation rate in the susceptible population. Afterward, we use it as the sensitive vector present size. The identical method is devoted to calculating the collection of candidates in I, D, V, R, H, and Q vectors, Table 1 represented employing rates. In the presented study the primary circumstances are considered as follows: $S(0) = S0$, $I(0) = I0$, $R(0) = R0$, $D(0) = D0$, $P(0) = P0$, and $Q(0) = Q0$, where t follows the iteration, and δ is for the rate of burial in Eq.(21). The quarantine rate of Ebola cases is modeled by Eq. (22).

6.5. Developed Ebola optimization search algorithm

However, Ebola Optimization Search Algorithm is stated as one of the newest metaheuristics that shows satisfying results toward different standard benchmark functions, it has some drawbacks that can be

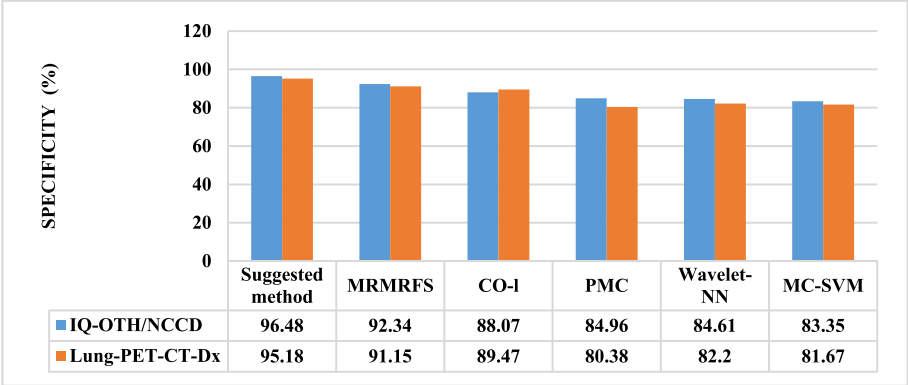

Fig. 10. Specificity results of all methods applied to IQ-OTH/NCCD and Lung-PET-CT-Dx datasets with preprocessing.

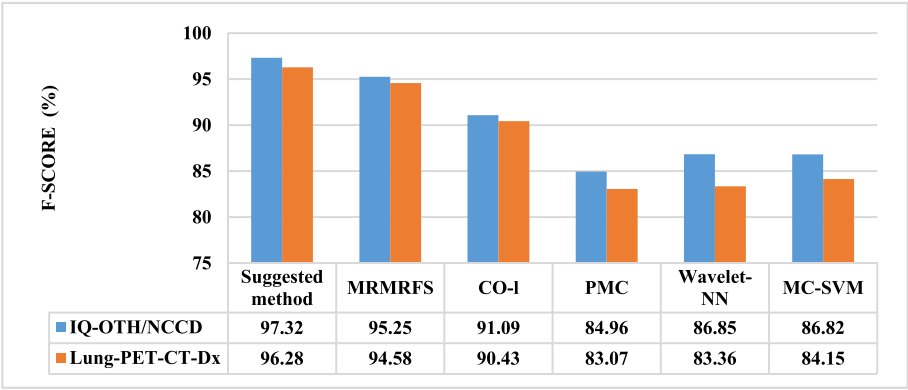

Fig. 11. F-Score results of all methods applied to IQ-OTH/NCCD and Lung-PET-CT-Dx datasets with preprocessing.

Table 4
Compared results of the procedures applied to IQ-OTH/NCCD dataset without preprocessing.

Method	Accuracy	Recall	Specificity	Precision	F-Score
DEOS/BRNN	93.46	92.65	92.18	94.15	93.26
Simple BRNN	85.26	87.04	88.33	89.71	89.22
MRMRFS [47]	88.62	89.85	90.63	91.89	91.65
CO-l [48]	80.18	85. 34	88.50	88.18	87.29
PMC [49]	77.29	82.17	80.11	84.49	85.32
Wavelet-NN [50]	76.16	80.08	75.42	82.05	84.16
MC-SVM [51]	74.37	79.96	74.63	80.68	81.20

Table 5
The proposed method's Confusion matrix.

Dataset		Cancer	Non-cancer
IQ-OH/NCCD	Cancer	395	22
	Non-cancer	13	762
PET-CT DICOM	Cancer	526	20
	Non-cancer	14	440

improved as much as possible [37,38]. One drawback is its generating. Although this term is so good in the first steps, its diversity has decreased after some steps, such that, the difference in position among Ebola candidates has gradually decreased, resulting in a low diverge searching. This issue makes local optimization for the algorithm. To resolve this problem, the Opposition-based learning (OBL) mechanism has been used.

The OBL is a technique for better selection of candidates by considering both their value and their opposing value. This may be expressed numerically as follows:

$$mI_i^{t+1} = \underline{ml} + \overline{ml} - mI_i^t \tag{23}$$

where, $\underline{ml}$ and $\overline{ml}$ express the process maximum and minimum solution.

After generating the opposite value and achieving the objective value, the best candidate is selected as the newly updated candidate of the algorithm.

We also change the scale factor (ρ) from a constant coefficient into a self-learn parameter. This is done by adding the following equation:

$$\rho = \underline{\rho} + \left(\overline{\rho} - \underline{\rho}\right) \times \left(1 - \arccos\left(\frac{\left(-2 \times \frac{d}{it} + 1\right)}{\pi}\right)\right) \tag{24}$$

where, it specifies the highest value of the iteration, d signifies the iteration number, and $\overline{\rho}$ and $\underline{\rho}$ describe the lowest and the highest ρ value.

6.6. Authentication

The efficiency of the suggested algorithm has been authenticated in this part [39,40]. The benchmark CEC 2017 is applied for the efficiency assessment of the suggested DEOS (Developed Ebola Optimization Search) algorithm [41]. Here, the benchmark 2017 test suite with ten test functions has been utilized. More information about the utilized CEC can be found in [41,42]. The analyzed benchmark functions in this research include F1-F10. To validate the suggested method, the following 5 previous studies in the literature have been used:

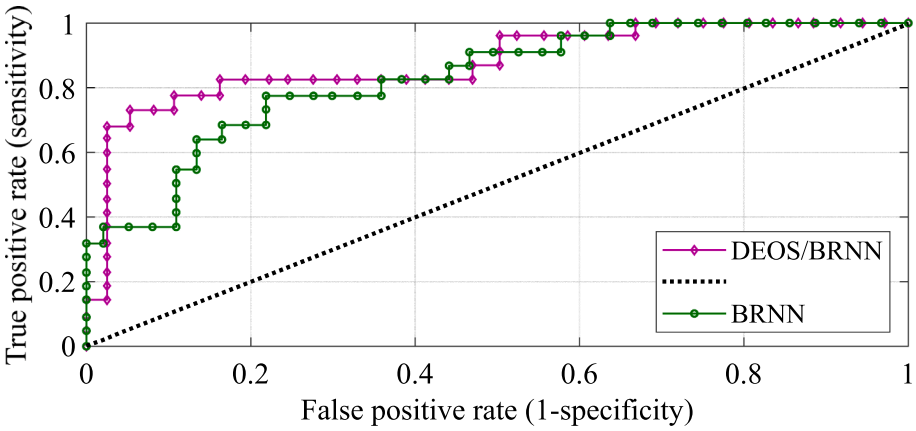

Fig. 12. The ROC curve for the suggested DEOS/BRNN compared with the traditional BRNN.

- Pigeon-Inspired Optimization Algorithm (PIO) [43]
- Chimp Optimization Algorithm (COA) [44]
- Heap-based optimizer (HBO) [45]
- Supply-Demand-Based Optimization (SDO) [46]
- Ebola Search Optimization Algorithm (ESO) [36]

The studied algorithms' parameter values are determined in Table 2.

The number of populations in all algorithms is set to 50 and the maximum iteration of them is set to 200 to provide validation with far higher proper comparison. It should be noted that for presenting reliable results, all of the algorithms have been run 25 times, independently. Table 3 depicts the mean value of algorithms and the normal alteration of the functions.

It can be observed from Table 3 that the suggested developed Ebola Search Optimization Algorithm proposes by far the best average value in accuracy terms in comparison with the other algorithms used in previous works. In a similar vein, the far higher value of the standard deviation for the proposed algorithm illustrates its prominence toward the other comparative algorithms in finding reliable results during several runs.

However, using the proposed DESO for minimizing F_E increases its complexity than the conventional BRNN model, due to the need for training for once and the impact of accuracy in medical applications, this can be neglectable.

7. Simulation results

The experimental findings of the suggested model are discussed in this section. MATLABR2019b is used to implement the model and all simulations were organized on an Intel® Core™ i7-9750H CPU rate 2.60 GHz with 16.0 GB RAM and 8 GB RTX 2070 Nvidia GPU. The proposed method was authenticated by applying it to a standard dataset, namely IQ-OTH/NCCD lung cancer dataset. For validation of the proposed method, different measures namely Specificity, Accuracy, Precision, F-score, and Recall have been employed and a comparison was done between the answers and some published state-of-the-art methods to provide an impartial validation. The employed methods contained minimum redundancy maximum relevance feature selection (MRMRFS) method [47], Co-learning feature fusion maps (CO-l) [48], probability of malignancy calculation (PMC) [49], FNN classifier, and wavelet feature descriptor (Wavelet-NN) [50], and MC-SVM classifier [51]. The mathematical formulation of the measures is given in the following:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{25}$$

$$Precision = \frac{TP}{TP + FP} \tag{26}$$

$$Recall = \frac{TP}{TP + FN} \tag{27}$$

$$Specificity = \frac{TN}{\mathrm{TN} + \mathrm{FP}} \tag{29}$$

$$\mathrm{F-Score} = \frac{2 \times precision \times recall}{precision + recall} \tag{30}$$

where, the concept of TP, TN, FP, and FN are shown in Fig. 6 as a 2×2 confusion matrix.

This study authenticated the method based on two different situations. The first one is to get the results of the suggested procedure with preprocessing (Scenario 1), and the second one is based on the proposed method in the absence of preprocessing (Scenario 2).

7.1. Scenario 1

As explained before, scenario 1 considers the proposed method with preprocessing step. By applying this model to the lung cancer dataset of the IQ-OTH/NCCD, and comparing the conclusions with the aforementioned methods, the comparison results are achieved and are recorded in the following.

- Accuracy

Based on the formulation, accuracy does not have to be the only model statistic that data scientists pursue; straightforward explanations of other metrics should be included as well. Fig. 7 demonstrates the comparison results of the accuracy measure among the suggested procedure and the other condition of the art methods.

Regarding Fig. 7, it is clear that the suggested procedure provided by far the highest accuracy (97.06% for IQ-OTH/NCCD and 96.68% for Lung-PET-CT-Dx) which clarifies that it has the lowest rate of error and the best rate of success for lung cancer diagnosis.

- Precision

Precision is a measure to show the positive predictive value. This measure is the percentage of relevant examples achieved among the retrieved instances. The comparison results of the Precision measure between the suggested procedure and the other condition of the art methods are demonstrated. Fig. 8

Concerning Fig. 8, the suggested method is endowed with 98.52% precision in IQ-OTH/NCCD and 97.44% in Lung-PET-CT-Dx, providing the highest quantity for this measure. A higher value of precision by the proposed method indicates its prominence in real positive rate diagnosis for suitable identification of lung cancer.

- Recall

The recall measure, which is also acknowledged as sensitivity, is the quantity of the achieved relevant instances. Fig. 9 compares the results of the Recall measure between the suggested procedure and the other condition of the art methods.

According to Fig. 8, it is obvious that the suggested procedure, with 96.15% in IQ-OTH/NCCD and 95.33% Lung-PET-CT-Dx dataset precision, offers more satisfying results in terms of this measure. This higher value of sensitivity by the proposed method toward the other comparative methods shows its dominance in diagnosing the unaffected positive rate, which indicates appropriate recognition of the cancer pixels.

- Specificity

Specificity, which is also acknowledged as a True Negative Rate, is the negative test likelihood if it is negative. This measure refers to the capacity of the test to appropriately reject healthy people who do not have a disease. This measure defines the relation of people who do not have the disease who test negative for it. Fig. 10 depicts the comparison answers of the Specificity measure between the suggested procedure and the other ones in the literature.

Fig. 10 shows that the highest value of the specificity 96.48% for IQ-OTH/NCCD and 95.18% Lung-PET-CT-Dx is obtained from the designed method. This shows that the proposed technique provides the highest true negative rate for all comparative techniques, which describes its capability in the correct identification of the normal pixels.

- F-Score

A measure of a test's precision is defined as F-score. It is derived from the test's accuracy and recall. Fig. 11 illustrates the comparison results of the F-score measure between the suggested methodology and other state-of-the-art methods.

Regarding the results, the suggested procedure, with the highest F-score i.e., 97.32% for IQ-OTH/NCCD and 96.28% for Lung-PET-CT-Dx, provides a good harmonic mean between the recall and precision.

7.2. Scenario 2

As aforementioned, scenario 2 assumes the situation in which the proposed method is applied without preprocessing. By performing this model on the IQ-OTH/NCCD lung cancer dataset, and its validation with before mentioned techniques, the comparison results were achieved and recorded below (in Table 4).

As can be seen, in Scenario 1, the prominence of the proposed method toward the others is observable. Also, the lower efficiency of the proposed method and the others when not using the preprocessing stage are observable. We also compared the method by its simple network without DEOS to show the impact of using the proposed method on increasing the accuracy of the method than its conditional method. The proposed classification technique includes a confusion matrix as an additional measurement for assessing the classification outcomes. Table 5 indicates the proposed method's confusion matrix for IQ-OH/NCCD and PET-CT DICOM datasets.

The expected number of true positives (Cancer) for IQ-OH/NCCD and PET-CT DICOM were 395 and 526, respectively, and the expected number of false negatives (Non-cancer) were 762 and 440, respectively.

while the expected number of false positives was 22 and 20, respectively and the expected number of false positives was 13 and 14, respectively. This shows the high accuracy of the proposed method for lung cancer diagnosis.

Additional statistical tests comparing the suggested model to the traditional BRNN are conducted to assess the results in addition to the results mentioned above. Therefore, the accuracy can be assessed using the area under the curve or the area under the ROC curve. The Receiver Operating Characteristic (ROC) of the proposed DEOS/BRNN compared with the traditional BRNN is shown in Fig. 12.

The area under the ROC curve is a measure of accuracy. For instance, a test that scores a 0 or a 1 is considered to be perfect. In comparison to the traditional BRNN, the DEOS/BRNN produced the best area under the ROC curve with an AUC of 0.92 and the BRNN produced the lowest area under the ROC curve with an AUC of 0.72. Generally, although, based on some literature works, the convolutional neural networks don't need any preprocessing, it is concluded that this stage can provide far more efficient results for the diagnosis system.

8. Conclusions

This paper proposed a lung cancer diagnosis system based on an improved Bidirectional Recurrent neural network and metaheuristic technique. The network was altered on a reformed form of the Ebola search optimization algorithm basis for providing higher accuracy for the suggested algorithm. Before applying the main diagnosis system, some preprocessing techniques, including Median filtering for eliminating the noise, a contrast enhancement for enhancing the image quality, and data normalization for simplifying the process were performed. The system was applied to the lung cancer dataset of the IQ-OTH/NCCD for proper authentication of the algorithm. To increase the number of dataset images artificially, Image augmentation was also utilized. The proposed procedure results were then approved by comparing some published works, including the minimum redundancy maximum relevance feature selection (MRMRFS) method, Co-learning feature fusion maps (CO-l), probability of malignancy calculation (PMC), wavelet feature descriptor (Wavelet-NN) and FNN classifier, and MC-SVM classifier to show the prominence of the suggested procedure. When compared to previous approaches, the experimental findings reveal improved performance. Precision, recall, specificity, F-score, and accuracy are the metrics that are being compared.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

References

[1] Z. Guo, et al., Novel computer-aided lung cancer detection based on convolutional neural network-based and feature-based classifiers using metaheuristics, Int. J. Imaging Syst. Technol. (2021).
[2] R. Zhang, et al., Multiscale mask R-CNN–based lung tumor detection using PET Imaging. Mol. Imaging, 18 (2019) p. 1536012119863531.
[3] F.R.S. Navid Razmjooy, N. Ghadimi, A Hybrid Neural Network – World Cup Optimization Algorithm for Melanoma Detection, Open Med. 13 (2018) 9–16.
[4] Q. Tian, et al., A New optimized sequential method for lung tumor diagnosis based on deep learning and converged search and rescue algorithm, Biomed. Signal Process. Control 68 (2021) 102761.
[5] N. Razmjooy, et al., Imperialist competitive algorithm-based optimization of neuro-fuzzy system parameters for automatic red-eye removal, Int. J. Fuzzy Syst. 19 (4) (2017) 1144–1156.
[6] Y. Zhou, Y. Lu, Z. Pei, Accurate diagnosis of early lung cancer based on the convolutional neural network model of the embedded medical system, Microprocess. Microsyst. 81 (2021) 103754.
[7] X. Cai, et al., Breast Cancer Diagnosis by Convolutional Neural Network and Advanced Thermal Exchange Optimization Algorithm, Comput. Math. Methods Med. 2021 (2021).
[8] Y. Xu, et al., Lung cancer diagnosis in CT images based on Alexnet optimized by modified Bowerbird optimization algorithm, Biomed. Signal Process. Control 77 (2022) 103791.
[9] Z. Xu, et al., Computer-aided diagnosis of skin cancer based on soft computing techniques, Open Med. 15 (1) (2020) 860–871.

[10] B. Bahat, P. Görgel, Lung Cancer Diagnosis via Gabor Filters and Convolutional Neural Networks, in: 2021 Innovations in Intelligent Systems and Applications Conference (ASYU). 2021. IEEE.
[11] Z.-H. Zhou, et al., Lung cancer cell identification based on artificial neural network ensembles, Artif. Intell. Med. 24 (1) (2002) 25–36.
[12] D.-T. Lin, C.-R. Yan, Lung nodules identification rules extraction with neural fuzzy network, in: Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP'02. 2002. IEEE.
[13] M. Šarić, et al. CNN-based method for lung cancer detection in whole slide histopathology images, in: 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech). 2019. IEEE.
[14] A. Agarwal, K. Patni, D. Rajeswari, Lung cancer detection and classification based on alexnet CNN, in: 2021 6th International Conference on Communication and Electronics Systems (ICCES). 2021. IEEE.
[15] alyasriy, h. *IQ-OTH/NCCD - Lung Cancer Dataset*. 2020; Available from: https://www.kaggle.com/datasets/adityamahimkar/iqothnccd-lung-cancer-dataset.
[16] Z. Yang, et al., Robust multi-objective optimal design of islanded hybrid system with renewable and diesel sources/stationary and mobile energy storage systems, Renew. Sustain. Energy Rev. 148 (2021) 111295.
[17] R. Ranjbarzadeh, et al., Nerve optic segmentation in CT images using a deep learning model and a texture descriptor, Complex Intell. Syst. (2022) 1–15.
[18] G. Aghajani, N. Ghadimi, Multi-objective energy management in a micro-grid, Energy Rep. 4 (2018) 218–225.
[19] P. Akbary, et al., Extracting appropriate nodal marginal prices for all types of committed reserve, Comput. Econ. 53 (1) (2019) 1–26.
[20] M. Bagheri, et al. A novel wind power forecasting based feature selection and hybrid forecast engine bundled with honey bee mating optimization, in: 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe). 2018. IEEE.
[21] W. Cai, et al., Optimal bidding and offering strategies of compressed air energy storage: A hybrid robust-stochastic approach, Renew. Energy 143 (2019) 1–8.
[22] M. Dehghani, et al., Blockchain-based securing of data exchange in a power transmission system considering congestion management and social welfare, Sustainability 13 (1) (2021) 90.
[23] H. Ebrahimian, et al., The price prediction for the energy market based on a new method, Econ. Res.-Ekonomska istraživanja 31 (1) (2018) 313–337.
[24] M. Eslami, et al., A New Formulation to Reduce the Number of Variables and Constraints to Expedite SCUC in Bulky Power Systems. Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, 2018: p. 1–11.
[25] X. Fan, et al., High voltage gain DC/DC converter using coupled inductor and VM techniques, IEEE Access 8 (2020) 131975–131987.
[26] J. Bae, H. Yoo, Fast Median Filtering by Use of Fast Localization of Median Value, Int. J. Appl. Eng. Res. 13 (12) (2018) 10882–10885.
[27] M. Ghiasi, N. Ghadimi, E. Ahmadinia, An analytical methodology for reliability assessment and failure analysis in distributed power system, SN Appl. Sci. 1 (1) (2019) 44.
[28] A.R. Gollou, N. Ghadimi, A new feature selection and hybrid forecast engine for day-ahead price forecasting of electricity markets, J. Intell. Fuzzy Syst. 32 (6) (2017) 4031–4045.
[29] A. Alferaidi, et al., Distributed Deep CNN-LSTM Model for Intrusion Detection Method in IoT-Based Vehicles, Math. Probl. Eng. 2022 (2022).
[30] D. Yu, et al., Energy management of wind-PV-storage-grid based large electricity consumer using robust optimization technique, J. Storage Mater. 27 (2020) 101054.
[31] Z. Yuan, et al., Probabilistic decomposition-based security constrained transmission expansion planning incorporating distributed series reactor, IET Gener. Transm. Distrib. 14 (17) (2020) 3478–3487.
[32] W. Jiang, et al., Optimal economic scheduling of microgrids considering renewable energy sources based on energy hub model using demand response and improved water wave optimization algorithm, J. Storage Mater. 55 (2022) 105311.
[33] G. Bo, et al., Optimum structure of a combined wind/photovoltaic/fuel cell-based on amended Dragon Fly optimization algorithm: a case study, Energy Sources Part A 44 (3) (2022) 7109–7131.
[34] M. Ramezani, et al., A New Improved Model of Marine Predator Algorithm for Optimization Problems, Arab. J. Sci. Eng. (2021) 1–24.
[35] M. Ghiasi, et al., A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future, Electr. Pow. Syst. Res. 215 (2023) 108975.
[36] O.N. Oyelade, et al., Ebola optimization search algorithm: a new nature-inspired metaheuristic optimization algorithm, IEEE Access 10 (2022) 16150–16177.
[37] M. Ghiasi, et al., Evolution of smart grids towards the Internet of energy: Concept and essential components for deep decarbonisation, IET Smart Grid 6 (1) (2023) 86–102.
[38] N. Ghadimi, et al., An innovative technique for optimization and sensitivity analysis of a PV/DG/BESS based on converged Henry gas solubility optimizer: A case study, IET Gener. Transm. Distrib. (2023).
[39] Y. Guo, et al., An optimal configuration for a battery and PEM fuel cell-based hybrid energy system using developed Krill herd optimization algorithm for locomotive application, Energy Rep. 6 (2020) 885–894.
[40] E. Han, N. Ghadimi, Model identification of proton-exchange membrane fuel cells based on a hybrid convolutional neural network and extreme learning machine optimized by improved honey badger algorithm, Sustain. Energy Technol. Assess. 52 (2022) 102005.
[41] G. Wu, R. Mallipeddi, P.N. Suganthan, Problem definitions and evaluation criteria for the CEC 2017 competition on constrained real-parameter optimization. National University of Defense Technology, Changsha, Hunan, PR China and Kyungpook National University, Daegu, South Korea and Nanyang Technological University, Singapore, Technical Report, 2017.
[42] V. Stanovov, S. Akhmedova, E. Semenkin, LSHADE algorithm with rank-based selective pressure strategy for solving CEC 2017 benchmark problems, IEEE, 2018.
[43] Z. Cui, et al., A pigeon-inspired optimization algorithm for many-objective optimization problems, Sci. China Inf. Sci. 62 (7) (2019) 1–3.
[44] M. Khishe, M.R. Mosavi, Chimp optimization algorithm, Expert Syst. Appl. (2020) 113338.
[45] Q. Askari, M. Saeed, I. Younas, Heap-based optimizer inspired by corporate rank hierarchy for global optimization, Expert Syst. Appl. 161 (2020) 113702.
[46] W. Zhao, L. Wang, Z. Zhang, Supply-demand-based optimization: a novel economics-inspired algorithm for global optimization, IEEE Access 7 (2019) 73182–73206.
[47] M. Toğaçar, B. Ergen, Z. Cömert, Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks, Biocybernet. Biomed. Eng. 40 (1) (2020) 23–39.
[48] A. Kumar, et al., Co-learning feature fusion maps from PET-CT images of lung cancer, IEEE Trans. Med. Imaging 39 (1) (2019) 204–217.
[49] S. Manoharan, Early diagnosis of lung cancer with probability of malignancy calculation and automatic segmentation of lung CT scan images, J. Innovative Image Processing (JIIP) 2 (04) (2020) 175–186.
[50] R. Arulmurugan, H. Anandakumar, Early detection of lung cancer using wavelet feature descriptor and feed forward back propagation neural networks classifier, in: Computational vision and bio inspired computing, Springer, 2018, pp. 103–110.
[51] J. Alam, S. Alam, A. Hossan, Multi-stage lung cancer detection and prediction using multi-class svm classifie, in: 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2). 2018. IEEE.