International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Detection of Lung Cancer Disease Using Optimized Long Short-Term Memory Based on Improved Gray Wolf Optimization

V.Deepa[1] and Dr. S. KB. Sangeetha[2]

Submitted: 28/01/2024 **Revised**: 06/03/2024 **Accepted**: 14/03/2024

Abstract: Early diagnosis and reduced mortality rates are two benefits of early lung cancer detection in patients. This research proposes an optimized long short-term memory (LSTM) based on improved grey wolf optimization (IGWO) with opposition-based learning (OBL) and local search algorithm (LSA) as an effective lung cancer detection system. The OBL is utilized with GWO to boost its population diversity, and LSA is used with GWO to address its local optimum problem and improve the existing best solution. Then, the hyperparameters of the LSTM are optimized using IGWO, and the optimized LSTM is then used to detect lung cancer disease. The median filter is used to remove the unwanted noises from CT lung images, fuzzy c-means (FCM) is used to segment the affected regions, affected area is passed on to the feature extractions stage that extracts the various spectral features for effectively detecting lung disease. Finally, extracted features are considered as inputs to the optimized LSTM for detecting lung cancer diseases. High performance was achieved by the developed optimized method: 94.65% accuracy, 96.90% precision, 95.05% recall, and 94.94% f-measure. The experimental results revealed that the developed IGWO-LSTM achieved good detection accuracy and a quick convergence rate.

Keywords: *long short-term memory; grey wolf optimization; opposition-based learning; local search algorithm; lung cancer; hyperparameter optimization*

1. Introduction

Significant cancers that cause a high death rate include lung cancer. Today's rising cigarette smoking rates are one of the major factors contributing to lung cancer in developing nations. Because the number of individuals suffering from illness is increasing at a rapid pace, the healthcare sector offers physicians high-quality services for precise medical data analysis. To address issues with patient care, early disease prediction, and community services, healthcare services are growing in popularity. An important step in giving patients the right care is accurate lung cancer disease prediction [1, 2]. Deep learning (DL) is a computational model consisting of several processing layers that can acquire representations of data with various levels of abstraction[3]. The class of DL known as LSTM is incredibly potent; it can instantiate practically any dynamics. LSTM has gained popularity as a model in recent years for a wide range of issues[4]. Hyperparameters play a crucial role in the performance of several deep learning algorithms. Nevertheless, there is no set standard for choosing hyperparameters since there is no exact mathematical correlation between results and hyperparameters. In reality, grid search, random search, and Bayesian optimization are the most often utilized techniques for finding hyperparameters. However, these approaches have certain drawbacks. For instance, the hyperparameters significantly increase the computing cost of the grid search, and the random search approach runs the risk of reaching a local optimum. The task of identifying hyperparameters has turned into a bottleneck that limits deep learning's accuracy[5].

Over the past 20 years, metaheuristic optimization strategies have gained a lot of traction. Notwithstanding the variations amongst metaheuristics, the majority of them are nature-inspired; metaheuristic optimization techniques are demonstrated to possess exceptional superiority in ergodicity and convergence, as well as the capacity to eschew local optima. In the context of hyperparameter settings, metaheuristics have gained increasing. The social structure and hunting habits of grey wolves served as the model for the biological algorithm known as the GWO[6]. Four different kinds of grey wolves are used in GWO to simulate the leadership structure. Additionally, three basic processes are used in hunting: finding prey, encircling prey, and attacking prey. When compared to alternative algorithms inspired by nature, the GWO exhibits a notable advantage [7, 8]. As such, it can be effective for numerous engineering and optimization challenges. Even though the fundamental GWO has remarkable shortcomings such as population diversity, low convergence, and local optima [9, 10].

OBL has been used in many studies in the literature to improve convergence speed. However, the literature study uses OBL for population initialization and position updating. However, position updating using OBL will take more iterations and it does not guarantee finding the optimal

[1,2] *Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai-26, India*
Corresponding Mail: dv1019@srmist.edu.in,sangeets8@srmist.edu.in

Algorithm 1: Median filter
Step 1: Assume that there are M rows and N columns in the input matrix "A."
Step 2: Add zeros to the sides of the input matrix to create a matrix with M + 2 rows and N + 2 columns.
Step 3: Take a three-by-three mask.
Step 4: Apply the mask to the first element in matrix "A," that is, the element on the first row and first column.
Step 5: Choose every element that the mask lists, then arrange them in ascending order.
Step 6: Replace element A (1, 1) with the median value (the middle element) that you took from the sorted array.
Step 7: Move the mask to the subsequent component.
Step 8: Continue following steps 4 through 7 until all matrix "A" components have been swapped out for the matching median value.

position. Therefore, this will increase the complexity of the search algorithm because in each iteration they determine the opposite solutions for all solutions and take elite solutions [11]. However, in our work, we use OBL at the GWO initialization phase only. Additionally, the GWO approach uses LSA developed for position updating to avoid local optima problems. Consequently, the goal is to provide an improved GWO (IGWO) that uses OBL and LSA to increase the

GWO's precision and convergence rate. This paper's innovations are concentrated in two areas. To improve the global convergence ability and convergence rate, the IGWO first employs the OBL technique. To increase the likelihood of jumping out of the local optimum, an LSA is then introduced to search the neighborhood space of the local optimal value. This IGWO aids in improving the trade-off between the GWO's capacity for exploration and exploitation. IGWO may be parallelized to concurrently explore several hyperparameter space areas. The efficiency of hyperparameter tuning can be increased by this parallelization, particularly for intricate LSTM architectures with several hyperparameters. The contributions of the paper are as follows,

- The new optimized LSTM is used to detect lung cancer diseases.
- The upgraded GWO is used to tune the hyperparameters of LSTM to improve detection accuracy.
- The IGWO uses OBL to enhance its population diversity in the search process and LSA for solving its local optimization problem and improving the present best solution.

This study is structured as follows: section 2 covers relevant works, section 3 examines research methodology, section 4 discusses experimental outcomes, and section 5 presents research findings.

2. Related works

Relevant research is essential to the growth and improvement of DL methods for lung cancer detection. Because DL can evaluate vast amounts of medical imaging data efficiently and accurately, it is important in the identification of lung cancer. The following paragraphs discuss some related works regarding lung cancer detection using DL methods. P. M. Shakeel et al. (2022) new ML approach is presented to predict lung cancer. Images from the non-small cell lung cancer CT scan dataset are collected to detect lung cancer. The multilayer brightness preserving technique is used to analyze the acquired images, effectively examining every pixel, removing noise, and improving the lung image quality. The afflicted region is separated from the noise-removed lung image using an enhanced deep neural network (DNN), which uses network layers to segment the region and extract numerous features [12]. D. Mhaske et al. (2019) [13] created a sophisticated computer-aided diagnosis (CAD) system that uses DL algorithms to rapidly extract data from CT scan pictures and give exact and timely analysis of lung cancer. The CT scan pictures are segmented with (OTSU) Thresholding. The developed method focuses on using DL techniques, notably Convolutional Neural Networks (CNN) for feature extraction and LSTM for lung cancer classification, and achieves high accuracy.

H. Yu et al. (2020) established a new technique based on an adaptive hierarchical heuristic mathematical model (AHHMM). The lung CT picture is segmented to extract any relevant features, and a special feature extraction approach is used. The test evaluation revealed that the proposed model could detect the absence or presence of lung cancer with 96.67% accuracy [14]. P. M. Shakeel et al. (2019) [15] focus on improving the quality of lung images and diagnosing lung cancer by reducing misclassification. The lung CT images are obtained from the Cancer Imaging Archive (CIA) dataset; noise is removed using a weighted mean histogram equalization approach, which effectively removes noise while improving image quality; and the affected region is segmented using an improved profuse clustering technique (IPCT). The damaged region generates a wide range of spectral properties. These are investigated using DL to predict lung cancer. R. R. Subramanian et al. (2020) [16] suggested a new model based on CNN. Pretrained ImageNet models like LeNet, AlexNet, and VGG-16 are used to identify lung cancer. The suggested model employs the AlexNet model, with the features obtained from the network's last fully connected layer serving as distinct input to the softmax classifier. The suggested model provides a reliable and long-term diagnosis paradigm for lung cancer detection.

S. Lakshmanaprabu et al. (2019) [17] proposed using the Optimal DNN (ODNN) and Linear Discriminant Analysis (LDA) using CT scans of lung pictures. Deep characteristics are collected from CT lung images and then dimensionally reduced using LDR to categorize lung nodules as cancerous

or benign. The ODNN is applied to CT images and then optimized using the Modified Gravitational Search Algorithm (MGSA) to find lung cancer classification. A. Asuntha et al. (2020) [18] developed a new method for detecting cancerous lung nodules in the supplied input lung image and classify the lung cancer and its severity. This study employs revolutionary DL techniques to determine the location of malignant lung nodules. After collecting textural, geometric, volumetric, and intensity features, the Fuzzy Particle Swarm Optimization (FPSO) is used to determine the best feature. Finally, the characteristics are categorized using DL. A revolutionary FPSOCNN lowers the computational complexity of CNN. The experimental findings demonstrate that the new FPSOCNN outperforms other approaches.

D. Moitra et al. (2020) [19] created a more accurate DL model by combining convolutional and bidirectional RNNs. The study employed the NSCLC Radiogenomics dataset, which included 211 patients. The proposed model may contribute greatly to the automated prognosis of NSCLC and other forms of malignancies. M. Pradhan et al. (2023) [20] introduced a new Multi-objective MROA (MMROA) for hyperparameter tuning that treats accuracy and sensitivity equally as fitness functions. Initially, the histopathology images from the Lung and Colon 25000 (LC25000) dataset are adjusted using color normalization and segmented using the saliency-driven edge-dependent top-down level set (SDREL) technique. The attributes are extracted using the Grey Level Cooccurrence Matrix (GLCM) and GoogleNet, followed by selection using the improved grasshopper optimization algorithm (EGOA). MMROA optimizes the selected features, which are subsequently fed into the LSTM classifier.

3. Problem statement

The possibility of false positives and the difficulties of detecting early-stage malignancies make lung cancer detection both accurate and efficient a key challenge in healthcare. The current strategy lacks precision despite developments in screening techniques and imaging technology, which results in inferior patient outcomes, unnecessary invasive procedures, and delays in diagnosis. To enable early detection and individualized treatment approaches, a holistic solution that solves the shortcomings of current screening tools, strengthens risk assessment strategies, and increases patient engagement is urgently needed.

Let I be an illustration of a lung image, and let Y be the binary outcome variable that indicates whether lung cancer is present or not in the image; $Y = 1$ indicates that lung cancer is present, while $Y = 0$ indicates that it is not. The goal is to develop a prediction model $f(I)$ that can correctly forecast the likelihood of lung cancer in the image I given a dataset D of N images, each represented as $((I_i, Y_i))$.

4. Research methods

4.1 Pre-processing using a median filter

Median filters are more effective than linear filters at smoothing images with spiky noise distributions due to the reject extremes (outliers). The image's sharpness is preserved while noise is eliminated using the median filter. As implied by the name, the median value of the surrounding pixels is used to replace each pixel. This filter makes use of a 3 x 3 window. Among the standard filters that eliminate speckle noise, +is is one of the best. An important function of spatial processing is to maintain edge detail and remove non-impulsive noise using an adaptive median filter. The median filter keeps the edges and any minor structure in the image. The window size of the median filter is different for every pixel [21]. Algorithm 1 shows the description of the median filter.

4.2 Image segmentation using FCM

Unsupervised fuzzy logic-based FCM is an unsupervised classification technique first proposed by Dunn in 1973 and then refined by Bezdek in 1981 which is widely applied to image segmentation. In contrast to traditional partitioning techniques like k-means, where each data point is assigned to a single partition, a point in FCM is assigned to every class with a degree of membership that ranges from 0 to 1. The centers of the clusters are what represent them [22, 23]. This is how the centers are computed:

$$C_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} x_i}{\sum_{i=1}^{N} \mu_{ij}^{m}} \tag{1}$$

N denotes the total number of data points, x_i denotes the j^{th}data point, x_i m $(m > 1)$ denotes the fuzzifier, and μ_{ij} which is determined as follows, is the membership degree of the j^{th} pixel that is a member of cluster i.

$$\mu_{ij} = \frac{1}{\sum_{l}^{NC} \frac{d_{ij}^{\frac{2}{m-1}}}{d_{ij}}} \tag{2}$$

Where d_{lj} is the Euclidian distance between the j^{th} pixel and cluster the l, the c is the number of clusters. The objective function is minimized to determine the membership degrees and cluster centers that are ideal.

$$J_m = \sum_{j=1}^{N} \sum_{i=1}^{nc} \mu_{ij}^{m} d_{ij} \tag{3}$$

Constraints as follows,

$$\begin{cases} \forall j \in [1, N], \forall i \in [1, nc]: \mu_{ij} \in [0,1] \\ \forall j \in [1, N] : \sum_{i=1}^{nc} \mu_{ij} = 1 \end{cases} \tag{4}$$

4.3 Feature extraction

The segmented picture is sent to the feature extraction phase, which produces several spectral features such as

mean, third-moment skewness, standard deviation, and fourth movement kurtosis because they successfully detect lung cancer-related features [15]. The predictable spectral features are shown in Table 1.

4.4 Long short-term memory

In theory, recurrent networks can employ feedback networks to store representations of recent input events in the form of activations. However, erroneous signals that travel backward in time tend to explode or vanish. The cell state, which functions as a conveyor belt, is crucial to LSTMs. The cell state flows directly down the entire chain, with just modest linear interactions. LSTMs, which are controlled by structures known as gates, can delete or add information to the cell state. Thus, LSTM was purposefully developed to prevent the long-term reliance issue. Each cell in the LSTM has three gates. These are the forget, input, and output gates. The $x(t)$ and $h(t-1)$ are the inputs, and a series of equations can be used to define the computation for each cell

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (6)$$

$$g_t = \tanh(W_g \cdot [h_{t-1}, X_t] + b_g) \quad (7)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (8)$$

$$c_t = f_t \cdot C_{t-1} + g_t \cdot i_f \quad (9)$$

$$g_t = \tanh(f_t \cdot C_{t-1} + g_t \cdot i_t) \cdot o_t \quad (10)$$

Where the timestamp t, cell state C_t, and the output y_t which also serves as the input for the next timestamp, $'\cdot'$ denotes the hadamard product i_t, f_t, g_t and o_t are the output of the gates. W_i, W_f , W_o , W_g , b_i , b_f , b_o , b_g are coefficient matrixes. As distinct cells' input and output interact with one another, errors spread with the data.

4.5 GWO algorithm

The grey wolf optimization method replicates wolf leadership hierarchy and predatory performance before employing grey wolf abilities such as search, encirclement, hunting, and other predation activities to reach the optimization goal[24]. The number of wolves N with search area d, the wolf position can be expressed as: $X_i = (X_{i1}, X_{i2}, X_{i3}, \ldots., X_{id})$. The fittest solution, sometimes known as the alpha α, is used to replicate wolves' social hierarchy.

To simulate the social hierarchy of wolves, the fittest solution is referred to as the alpha (α). The second and third-best solutions are called beta (β) and delta (δ) wolves, respectively. The remaining possible solutions are considered to be omega (ω) wolves. The algorithm's location, the prey's location, matches to the alpha wolf's position. Grey wolves' encircling behavior can be analytically modeled as follows:

$$D = \left|C \times X_p(t) - X(t)\right| \quad (11)$$

$$X(t+1) = X_P(t) - A \times D \quad (12)$$

Where the set t is the current iteration. $X_p(t)$ and $X(t)$ are the position vectors of the prey and a grey wolf respectively. The control coefficient C is determined by the following formula:

$$C = 2r_1 \quad (13)$$

The set A is convergence factor considered as follows:

$$A = 2ar_2 - a \quad (14)$$

$$a = 2\left(1 - \frac{t}{T\,max}\right) \quad (15)$$

Where the set r_1 and r_2 are the random variable. The set a is the control coefficient linearly decreases from 2 to 0, that is, $a_{max} = 2, a_{min}$=0. When grey wolves catch prey, the leader wolf α directs the outside wolves to encircle it. The α wolf guides the β and δ wolves to catch the prey and closest to the prey, allowing their positions to be used to determine its location. The mathematical model is as follows.

$$D_\alpha = |C_1 \times (X_\alpha(t) - X(t))| \quad (16)$$

$$D_\beta = \left|C_2 \times (X_\beta(t) - X(t))\right| \quad (17)$$

$$D_\delta = |C_3 \times (X_\delta(t) - X(t))| \quad (18)$$

$$X_1 = (X_\alpha - A_1) \times D_\alpha \quad (19)$$

$$X_2 = (X_\beta - A_2) \times D_\beta \quad (20)$$

$$X_3 = (X_\delta - A_3) \times D_\delta \quad (21)$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \quad (22)$$

4.6 Opposition based learning (OBL)

The OBL was developed by H R Tizhoosh to obtain the opposite estimate from the current estimate and enhance the capability of the provided response [25]. Population-based optimization techniques often start by generating a set of solutions. Either prior information or random selection can be used to create the population. The optimization procedure is then applied to update the populations. On the other hand, if the answer is unknown beforehand, the provided solution cannot converge to a global solution. Additionally, it takes longer for the global solution to converge. Numerous research has been conducted using the advantages of the OBL method for initializing and updating the population to overcome these drawbacks.

The OBL is supposed to increase population variety and boost global search capabilities. For example, the OBL method is used to accelerate the convergence rate of numerous nature-inspired optimizations [26-31] and neural networks [32, 33]. The OBL algorithm offers a way to create optimization strategies that travel in the opposite direction of the existing solution. The current solution is selected as the best option after being compared to the opposite

alternative. This OBL technique produces a solution that approaches the ideal solution quickly. The ensuing subsections explain the OBL methodology.

a) Opposite number

Assume, x is the real number that ranges between m and n: $x \in [m, n]$. The opposite value is assumed by $\bar{x}$ determined as follows,

$$\bar{x} = m + n - x \tag{23}$$

The opposite number specified for multidimensional space, let $x = (x_1, x_2, \ldots, x_D)$ samples with D dimensional space. $x_1, x_2, \ldots, x_D \in R$ and $x_i[m_i, n_i] \forall i \in \{1,2, \ldots . D\}$. Hence, the opposite estimation $\bar{x}$ is well-defined as follows,

$$\bar{x}_i = m_i + n_i - x_i, i = 1,2, \ldots, D \tag{24}$$

b) Opposition-based optimization

Let $f(x)$ is a function and $g(.)$ is a proper evaluation function. Guess the value of $f(x)$ and $f(\tilde{x})$ in every iteration. The learning procedure continues with x if $g\,(f(x) \geq g\,(f(\tilde{x})$, otherwise with $\tilde{x}$ and it could be weight, a fitness function, or an error function.

4.7 Local search algorithm (LSA)

The best current location of any SI algorithm can be found with the freshly created LSA algorithm. At the end of each EHO iteration, LSA will be called to refine the current best solution. LSA first records the value of the best solution GWO in the local variable. LSA iterates several times in an attempt to enhance local values. Every time it runs, LSA randomly chooses three features from a saved list. LSA uses its parameters to set or reset the specified features. The fitness value of the new solution will likewise be ascertained by LSA. If it surpasses the fitness value of the previous solution, it is set to the local value; otherwise, it stays

Proposed IGWO .The Original GWO method as challenges such as local optima traps and search bias, particularly towards the coordinate system's origin. Additionally, search variety is limited. The two modifications are incorporated with the standard GWO algorithm which is discussed in this present section. First, the OBL method is used to initialize the initial population during GWO's initialization phase to increase population diversity.

Second, LSA utilizing with GWO to enhance its exploitation and prevent it from being trapped in local optima. Specifically, IGWO first uses OBL to generate the GWO population. Subsequently, the fittest wolves are selected to occupy the starting and opposing wolf positions. Furthermore, the starting population will consist of the best wolf among these n fittest wolves. In addition, these n wolves will have their positions updated by applying the main loop. Now, LSA will be applied to the current to check and find a better solution than the best one currently found at the end of the ISSA main loop. GWO will ultimately provide the finest answer.

Population initialization

The evolutionary algorithm is mainly based on the concept of population search and it enhances the algorithmic solving capabilities. The evolutionary algorithm is combined with the Global convergence for the purpose of the Population initialisation.The initial population is computed based on the fitness function.The fitness values are used to choose fronm the original population.OBL Strategy is used to generate the possible values of the present and the reverse values. The fitness values are used to choose the new population, sometimes referred to as the original population.The type of algorithm which makes use of the population set is called as the Genetic algorithm.The genetic algorithm also makes use of the cross valuation and validation for the generation of the population set,.

Algorithm 3 : LSA algorithm
Step 1: Temp= F and Lt=1
Step 2: While (Lt < maximum number of inner iterations)
Step 3: Choose three features at random from Temp
Step 4: if selected-feature == 1
Step 5: chosen-feature = 0
Step 6: Else
Step 7: chosen-feature = 1 end if
Step 8: Determine Temp's fitness value.
Step 9: If f(F) > f(Temp)
Step 10:F=Temp
Step 11: End if
Step 12: Lt= Lt+1
Step 13: End while
Step 14: return F

BL-based population initialization

OBL can effectively broaden the population under search and enhance algorithmic solving capabilities. It can be combined with evolutionary algorithms. Global convergence and optimization results are significantly impacted by population initialization. Based on the section analysis, this research presents the OBL strategy initialization approach to create initial GWO populations. The OBL expands the given solution by generating a population. An initial population X by guessing the opposite solution x̃_i for x_i. Compute the fitness values for both present and reverse values, and then compare them to get the best initial values for a given solution's beginning population. The fitness values are used to choose the new population, sometimes referred to as the original population.The new population helps in finding the possible set of values for the fitness function and the strategy for the set of Operational values.

LSA learning strategy

Throughout iterations, the optimal solution is crucial to GWO as it directs a population toward the global best value. However, the population search will quickly stagnate if the ideal solution becomes stuck at a local optimal value. The LSA contracts to the absolute minimum after adjusting to the local environment.

Therefore, throughout iterations, the optimal solution will be disturbed by the LSA approach to explore the neighborhood space of the local optimal value and strengthen the likelihood of jumping out of the local optimum. By employing this disruption on the local optimal solution, it can more effectively utilize search space and find superior persons with greater ease.

4.8 Optimized IGWO-LSTM

The LSTM network consists of memory cells and feedback connections. Much research has shown that an LSTM network can extract useful characteristics from low-volume input. The parameters of an LSTM network are often established by user experience. The detecting capabilities of the model will be affected by the parameters chosen. Therefore, for convenience, autonomous algorithmic approaches with the potential to converge faster and acquire an optimal/near-optimal solution within an acceptable period can be used to boost the performance of the LSTM. Hence, the research work uses the IGWO method for optimizing the LSTM for enhancing lung cancer detection accuracy with fast convergence. This work defines wolves as dimensional variables that reflect the parameters of LSTM. The LSTM is often used for the generation of the low volume inputs for the generation of the autonomous capability of the memory cells for the set of the parameters chosen.A random population is chosen for the determination of the fitness function.

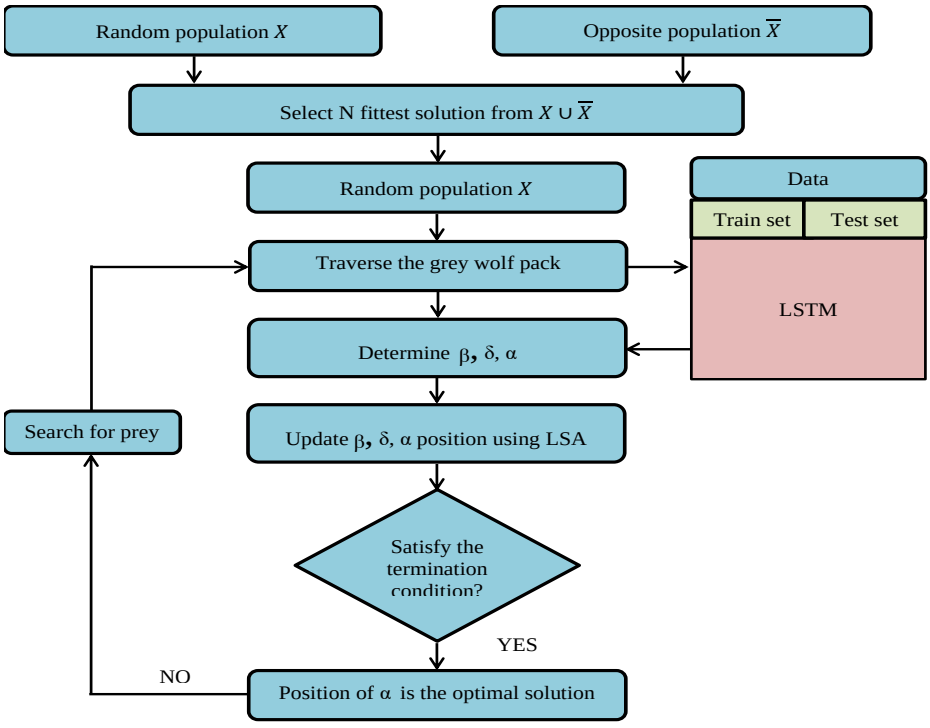

Figure 2: Flowchart of IGWO-LSTM detection method

Algorithm 4: IWGO algorithm
Step 1: Initialize the population x_i and its opposite $\bar{x}_i$
Step 2: Initialize $a, A, and\ C$
Step 3: Compute the fitness
Step 4: Select X_i fittest individuals form $\{X_i \cup OX_i\}$ as the next population X_i
Step 5: search the best agents for X_α, X_β, and X_δ
Step 6: **While** $(t < max\ number\ of\ iterations)$
Step 7: Update the dynamic interval boundaries $[da_j, db_j]$ in X_i
Step 8: For $i = 1$ to X_i
Step 9: Double $k = rand$
Step 10: For $j = 1\ to\ D$
Step 11: $\check{x}_{i,j} = k.(da_j + db_j) - x_j^e$
Step 12: *end for*
Step 13: If $\check{x}_{i,j} < a_j\ or\ \check{x}_{i,j} > b_j$
Step **14:** $\check{x}_{i,j} = rand(da_j, db_j)$
Step 15: *end if*
Step 16: Calculate the fitness value of $\check{x}_{i,}$
Step 17: *end for*
Step 18: For each search agent
Step 19: Update the current search agent's position.
Step 20: End for
Step 21: Update $a, A, and\ C$
Step 22: Evaluate the fitness of all search agents.
Step 23: Update $X_\alpha, X_\beta, X_\delta$
Step 24: Find the current best value using the LSA method
Step 25: $t = t + 1$
Step 26: End while
Step 27: Return X_α

Assign m hyperparameters to the grey wolf pack as prey, and use the real impact of LSTM's data prediction as a criterion to determine each gray wolf's position. Next, simulate an iterative search for the prey. In Figure 2, the flowchart of the proposed IGWO-LSTM is displayed. Algorithm 4 includes pseudo-code for the proposed developing LSTM model. The best hyperparameter search for LSTM is as follows. First, the population of the proposed IGWO version is randomly initialized. Each wolf represents one alternative configuration of the optimal LSTM model.

The training data is used to train the appropriate LSTM with specific structure and parameter settings for each wolf. The top three fitness solutions are identified as dominant wolves, guiding the entire wolf population toward global optimality using the suggested IGWO algorithm. The final LSTM model is based on the optimal configuration achieved by the best wolf leader. The model is re-trained with a combination of training sets before being tested on new test data. The test set is used to calculate performance metrics.

5. Experimental results analysis

Following the division of the images, the previously mentioned algorithm was applied in MATLAB 2019R to create the pre-processing, segmentation, feature extraction, and classification processes. The accuracy of segmented lung area images was compared with multiple detection algorithms based on the discussion above. The accuracy, recall, precision, and f-score of the suggested optimized LSTM approach were evaluated to determine its superiority. The performance of LSTM+IGWO is compared with some variants of LSTM such as LSTM+GWO, LSTM+IPSO, LSTM+PSO, LSTM+ GA, and LSTM.

Table 1: Details of features

Features	Formula
Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} S_i$
Standard deviations	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - \mu)^2}$
Third movement skewness	$sk = \sqrt{\frac{1}{N\times\sigma^3} \times \sum_{i=1}^{N}(S_i - \mu^3)^{1/3}}$
Fourth-moment kurtosis	$ku = \sqrt{\frac{1}{N\times\sigma^4} \times \sum_{i=1}^{N}(S_i - \mu^4)^{1/4}}$

A bad selection of network parameters usually results in subpar LSTM performance. As a result, research usually uses GA and PSO to select the best parameters. Nevertheless, several drawbacks exist, including premature convergence and local optima. Nonetheless, the PSO's weaknesses were resolved by the IPSO. Nevertheless, the MPSO's convergence rate is poor. Later, a SI optimization technique with a high solution accuracy and a quicker convergence rate was created by GWO recently [34]. The population diversity, local optima, and convergence rate are just a few of the conventional GWO's many drawbacks. To improve population diversity, convergence rate, and prevent local optima, the current paper suggests doing two adjustments using GWO. OBL was used to increase population variation, and the LSA algorithm was developed to get around the problem of local optima.

5.1 Datasets

Lung CT images were collected from the 5043 pictures in the Cancer Imaging Archive (CIA) dataset for analysis; 3000 images were utilized for training and 2043 images for testing. This includes the photos taken from the lung cohort of the National Cancer Institute Tumor Analysis Consortium. Along with clinical data, the collected patient details were connected with the proteome and genome. For experimental results analysis, we show the ten images only for performance results comparison of detection algorithms which is named as "*Image1, Image2, Image3, Image4, Image5, Image6, Image7, Image8, Image9, and Image10".*

5.2 Parameter settings

Hyperparameter tuning is critical for successfully training LSTM networks that are optimal for sequence modeling problems because of their ability to capture long-term dependencies in sequential data. To obtain optimal performance, LSTMs are trained by tweaking multiple

Table 2: Parameters details

LSTM	
Parameter	Value
Activation	Signomoid, Tanh
Loss function	MSE
Learning rate	0.057
epochs	1000
Expected error	0.0005
Weight range	-0.5 and 0.5
Dropout	0.3
Batch size	5
Time stap	30

hyperparameters. Hyperparameter optimization aids in selecting the best optimizer and tweaking its parameters to enhance training success. Hence, the present research work focused on hyperparameter optimization using IGWO to obtain optimal parameters. The IGWO's optimization objectives are weight and bias, learning rate, and time window size. The model randomly initializes the position information of each wolf based on the hyperparameter range. The algorithm creates an LSTM based on wolf position hyperparameters and trains it with training data. Finally, optimal parameters are shown in Table 2.

5.3 Performance indicators

This study used four performance metrics are used to evaluate the detection algorithms' performance. The accuracy of a detection algorithm determines its overall performance, which may be thought of as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (25)$$

The ability to foresee positive cases, such as the true positive rate, which can be considered as follows, is known as recall.

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (26)$$

Precision, also known as a positive predictive value, is well-defined as the number of relevant samples among the retrieved examples as follows:

$$Pr\,e\,cision = \frac{TP}{TP+FP} \times 100\% \quad (27)$$

The F-score is calculated using the harmonic mean of sensitivity and precision, with each given about equal weight. This enables prototype comparison, performance description, and the combination of precision and sensitivity into a single score.

$$F - score = 2 * \frac{Pr\,ecision * Re\,call}{Pr\,ecision + Re\,call} \times 100\% \quad (28)$$

False positive (FP) signifies erroneously expected normal, whereas false negative (FN) denotes incorrectly expected sickness. True positive (TP) indicates effectively predicted disease, whereas true negative (TN) indicates accurately predicted normal.

5.4 Results analysis

The following sections are critical in presenting the findings, analyzing the results, debating their implications, and evaluating the ability of the optimized LSTM. The section displays the results of the optimized LSTM lung cancer prediction model. This covers the accuracy, precision, recall, and f-measures performance indicators gathered during the model evaluation. Tables 3-6 and Figures 3-6 show the results comparisons of various lung cancer classification methods such as LSTM+IGWO, LSTM+GWO, LSTM+ IPSO, LSTM+ PSO, LSTM+ GA, and LSTM. Table 3 and Figure 3 show the performance results based on accuracy which produced the highest accuracy with 95.9%, 94.2%, 93.5%, 96%, 94.8 %, 95%, 94.8%, 94.5%, 94.4%, and 93.4%. for images 1-10 respectively.

Table 4 and Figure 4 show the performance results based on accuracy which produced the best performance with 97.9 %, 97.2 %, 96.7 %, 97.2 %, 96.5 %, 97.4 %, 96.9 %, 96.8 %, 97.4 %, and 95.7 %, for images 1-10 respectively. Table 5 and Figure 5 show the performance results based on accuracy which produced the best performance with 95.9%, 94.2%, 93.5 %, 96 %, 94.8 %, 95 %, 94.8%, 94.5%, 94.4%, and 93.4% for images 1-10 respectively. Table 6 and Figure 6 show the performance results based on accuracy which produced the best performance with 96.5 %, 95.9%, 96.2%, 96.2 %, 95.6%, 96.5%, 95.9%, 95.6%, 96.0 %, and 95.0 % for images 1-10 respectively.

The efficacy of the proposed LSTM-IGWO has been established through a variety of methods, as detailed in the previous analysis of experimental results. The optimized LSTM's structure has proven helpful in resolving real-time detection issues. The IGWO is offered as a way to improve LSTM's performance. It outperforms in terms of classification accuracy and convergence rate. The latest edition of GWO has produced encouraging results in accomplishing this goal. Overall, the LSTM-IGWO suggested method outperformed previously used detection algorithms and benchmark methodologies. The proposed model is intended to produce consistent and unbiased clinical predictions, making it appropriate for real-time diagnosis applications and another clinical diagnosis.

The overall experimental results indicate that IGWO outperforms other optimization algorithms in terms of exploration capability. The ability of IGWO to search through areas of the search space that other algorithms were unable to access lends credence to this. This proved that, in contrast to other algorithms, IGWO retains solution variety quite effectively. Furthermore, IGWO routinely achieves better fitness values than other algorithms, indicating its ability to avoid local optima. Through the selection of optimal LSTM hyperparameters, IGWO proved to be a more capable exploration algorithm than the others. The ability of IGWO to select the optimal hyperparametric and its

superiority over alternative optimization algorithms based on the accuracy of the outcomes verified its superiority in exploration.

Table 3 : Performance results based on accuracy

Detection methods	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image10
LSTM+IGWO	0.959	0.942	0.93.5	0.960	0.948	0.950	0.948	0.945	0.944	0.934
LSTM +GWO	0.936	0.934	0.924	0.927	0.928	0.939	0.938	0.924	0.934	0.924
LSTM+ IPSO	0.923	0.914	0.892	0.888	0.916	0.918	0.929	0.892	0.912	0.914
LSTM+ PSO	0.894	0.889	0.884	0.862	0.899	0.885	0.884	0.879	0.883	0.895
LSTM+ GA	0.877	0.873	0.865	0.857	0.860	0.879	0.870	0.867	0.873	0.883
LSTM	0.852	0.869	0.849	0.850	0.844	0.858	0.853	0.858	0.867	0.873

Table 4 : Performance results based on precision

Detection methods	Image 1	Image 2	Imag 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image10
LSTM+IGWO	0.979	0.972	0.967	0.972	0.965	0.974	0.969	0.968	0.974	0.957
LSTM +GWO	0.957	0.953	0.936	0.949	0.946	0.957	0.948	0.948	0.959	0.935
LSTM+ IPSO	0.936	0.928	0.913	0.924	0.925	0.936	0.939	0.925	0.936	0.925
LSTM+ PSO	0.897	0.893	0.893	0.894	0.893	0.915	0.916	0.915	0.914	0.916
LSTM+ GA	0.885	0.885	0.884	0.883	0.880	0.893	0.895	0.895	0.893	0.894
LSTM	0.873	0.873	0.878	0.877	0.869	0.870	0.883	0.884	0.884	0. 883

Table 5 : Performance results based on recall

Detection methods	Image 1	Image 2	Image 3	Image 4		Image 5	Image 6	Image 7	Image 8	Image 9	Image10
LSTM+IGWO	0.952	0.948	0.958	0.953		0.949	0.957	0.950	0.945	0.948	0.945
LSTM +GWO	0.946	0.938	0.947	0.934		0.938	0.946	0.938	0.937	0.937	0.935
LSTM+ IPSO	0.938	0.929	0.930	0.928		0.927	0.924	0.915	0.918	0.915	0.924
LSTM+ PSO	0.916	0.918	0.914	0.915		0.919	0.913	0.909	0.904	0.908	0.918
LSTM+ GA	0.895	0.898	0.885	0.896		0.890	0.891	0.896	0.897	0.891	0.884
LSTM	0.872	0.876	0.874	0.876		0.880	0.875	0.879	0.880	0.875	0.879

Table 6 : Performance results based on f-score

Detection Methods	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image10
LSTM+IGWO	0.965	0.959	0.962	0.962	0.956	0.965	0.959	0.956	0.960	0.950

LSTM +GWO	0.951	0.945	0.941	0.941	0.941	0.951	0.942	0.942	0.947	0.935
LSTM+ IPSO	0.936	0.928	0.921	0.925	0.925	0.929	0.926	0.921	0.925	0.924
LSTM+ PSO	0.906	0.905	0.903	0.904	0.905	0.913	0.912	0.909	0.910	0.916
LSTM+ GA	0.889	0.891	0.884	0.889	0.884	0.891	0.895	0.895	0.891	0.888
LSTM	0.872	0.874	0.875	0.876	0.874	0.872	0.880	0.881	0.879	0.880

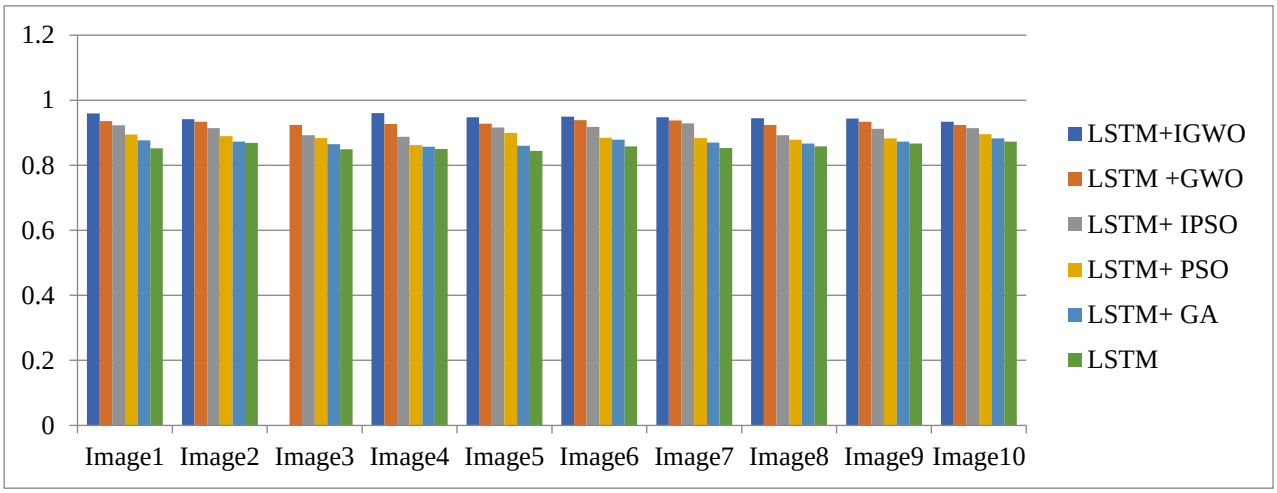

Figure 3 : Results analysis based on accuracy

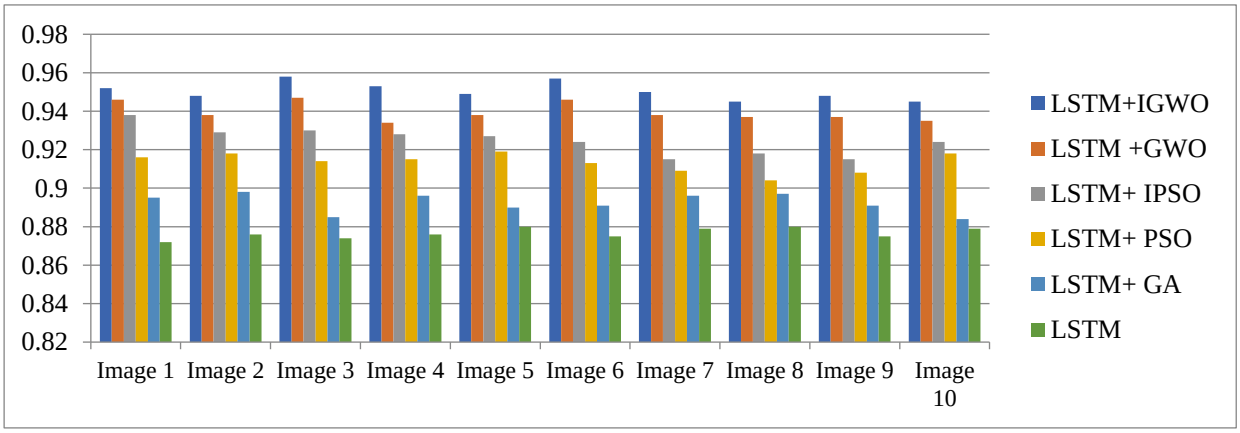

Figure 4 : Results analysis based on precision

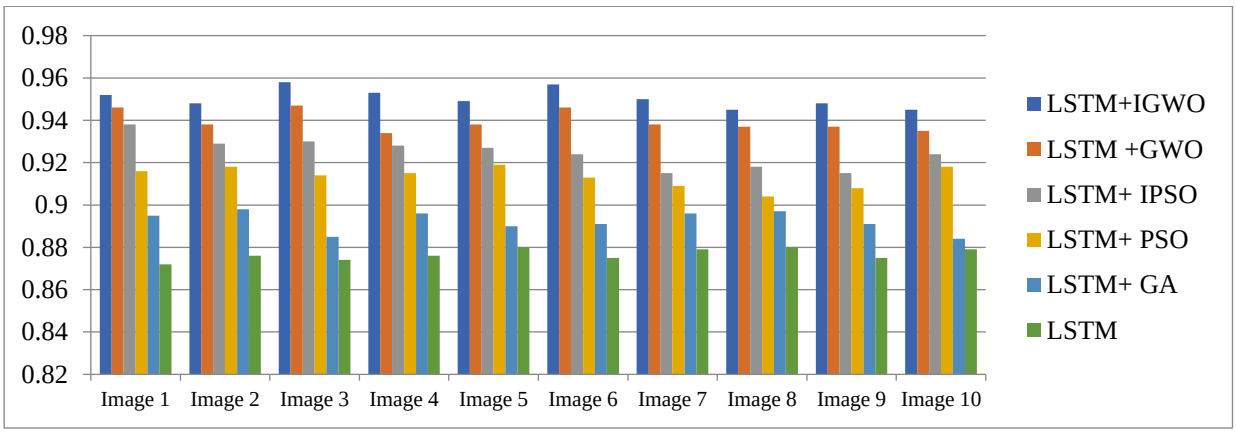

Figure 5 : Results analysis based on recall

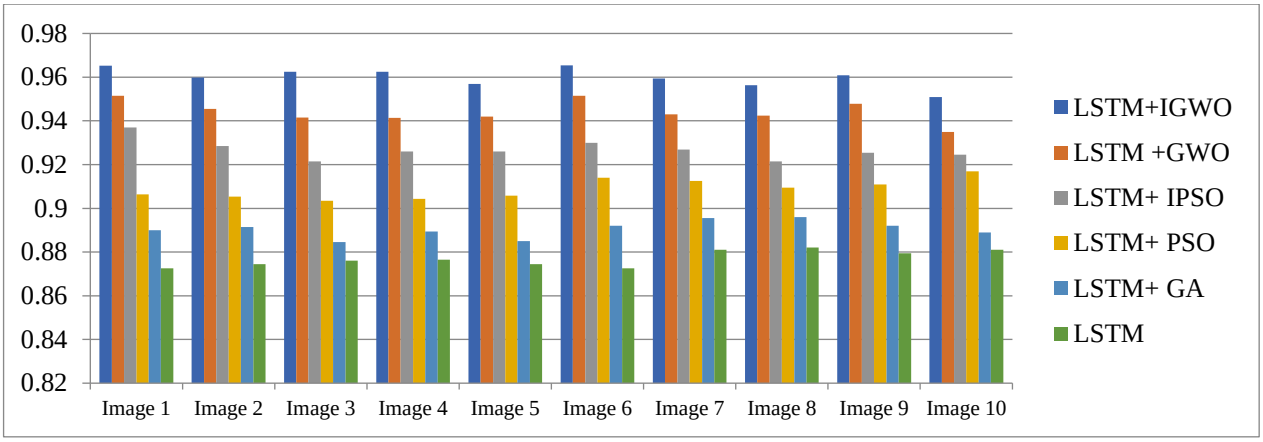

Figure 6 : Results analysis based on F-measures

6. Conclusions

The current study suggests an improved LSTM method for lung cancer diagnosis that is based on IGWO. The CIA provides CT lung imaging. After applying a median filter to denoise the gathered images, the FCM algorithm is used to segment the denoised images. Spectral features are then retrieved from the segmented images and fed into an LSTM to detect lung cancer diseases. In terms of accuracy, f-measure, precision, and recall, the benefits of the suggested model have outperformed those of the current models on the CT lung images dataset. Comparing the suggested hyperparameter optimization model to other lung cancer prediction models, the simulation's outcome demonstrates that the latter produced inferior predictions. The system's efficiency is assessed using experimental results, and the system accurately detects cancer with 94.65% of accuracy, 96.90% of precision, 95.05% of recall, and 94.94% of f-measure. The experimental results confirmed that the developed optimized LSTM method produced high performance when compared with other detection approaches.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] N. S. Nadkarni and S. Borkar, "Detection of lung cancer in CT images using image processing," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019: IEEE, pp. 863-866.

[2] W. J. Sori, J. Feng, A. W. Godana, S. Liu, and D. J. Gelmecha, "DFD-Net: lung cancer detection from denoised CT scan image using deep learning," *Frontiers of Computer Science,* vol. 15, pp. 1-13, 2021.

[3] R. Kavitha, K. Saravanan, S. A. Jebakumari, and K. Velusamy, "Machine learning algorithms for IoT applications," *Artificial Intelligence for Internet of Things: Design Principle, Modernization, and Techniques,* p. 185, 2022.

[4] N. M. Thoppil, V. Vasu, and C. Rao, "Bayesian optimization LSTM/bi-LSTM network with self-optimized structure and hyperparameters for remaining useful life estimation of lathe spindle unit," *Journal of Computing and Information Science in Engineering,* vol. 22, no. 2, p. 021012, 2022.

[5] J. Hong and W. Tian, "Prediction in Catalytic Cracking Process Based on Swarm Intelligence Algorithm Optimization of LSTM," *Processes,* vol. 11, no. 5, p. 1454, 2023.

[6] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: a review of recent variants and applications," *Neural computing and applications,* vol. 30, pp. 413-435, 2018.

[7] X. Yu, W. Xu, and C. Li, "Opposition-based learning grey wolf optimizer for global optimization," *Knowledge-Based Systems,* vol. 226, p. 107139, 2021.

[8] Q. Al-Tashi, H. Md Rais, S. J. Abdulkadir, S. Mirjalili, and H. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," *Evolutionary Machine Learning Techniques: Algorithms and Applications,* pp. 273-286, 2020.

[9] S. Zhang, Q. Luo, and Y. Zhou, "Hybrid grey wolf optimizer using elite opposition-based learning strategy and simplex method," *International journal of computational intelligence and applications,* vol. 16, no. 02, p. 1750012, 2017.

[10] S. Gupta and K. Deep, "An opposition-based chaotic grey wolf optimizer for global optimisation tasks," *Journal of Experimental & Theoretical Artificial Intelligence,* vol. 31, no. 5, pp. 751-779, 2019.

[11] [11] M. Tubishat, N. Idris, L. Shuib, M. A. Abushariah, and S. Mirjalili, "Improved Salp Swarm Algorithm based on opposition based learning and

novel local search algorithm for feature selection," *Expert Systems with Applications,* vol. 145, p. 113122, 2020.

[12] P. M. Shakeel, M. Burhanuddin, and M. I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier," *Neural Computing and Applications,* pp. 1-14, 2022.

[13] D. Mhaske, K. Rajeswari, and R. Tekade, "Deep learning algorithm for classification and prediction of lung cancer using CT scan images," in *2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, 2019: IEEE, pp. 1-5.

[14] H. Yu, Z. Zhou, and Q. Wang, "Deep learning assisted predict of lung cancer on computed tomography images using the adaptive hierarchical heuristic mathematical model," *IEEE Access,* vol. 8, pp. 86400-86410, 2020.

[15] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks," *Measurement,* vol. 145, pp. 702-712, 2019.

[16] R. R. Subramanian, R. N. Mourya, V. P. T. Reddy, B. N. Reddy, and S. Amara, "Lung cancer prediction using deep learning framework," *International Journal of Control and Automation,* vol. 13, no. 3, pp. 154-160, 2020.

[17] S. Lakshmanaprabu, S. N. Mohanty, K. Shankar, N. Arunkumar, and G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images," *Future Generation Computer Systems,* vol. 92, pp. 374-382, 2019.

[18] A. Asuntha and A. Srinivasan, "Deep learning for lung Cancer detection and classification," *Multimedia Tools and Applications,* vol. 79, pp. 7731-7762, 2020.

[19] D. Moitra and R. K. Mandal, "Prediction of non-small cell lung cancer histology by a deep ensemble of convolutional and bidirectional recurrent neural network," *Journal of Digital Imaging,* vol. 33, pp. 895-902, 2020.

[20] M. Pradhan, I. L. Coman, S. Mishra, T. Thieu, and A. Bhuiyan, "LSTM based Modified Remora Optimization Algorithm for Lung Cancer Prediction," *International Journal of Intelligent Engineering & Systems,* vol. 16, no. 6, 2023.

[21] 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, 2019.

[22] K. Velusamy and R. Manavalan, "Performance analysis of unsupervised classification based on optimization," *Int J Comput Appl,* vol. 42, pp. 22-7, 2012.

[23] H. Abdellahoum, N. Mokhtari, A. Brahimi, and A. Boukra, "CSFCM: An improved fuzzy C-Means image segmentation algorithm using a cooperative approach," *Expert Systems with Applications,* vol. 166, p. 114063, 2021.

[24] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in engineering software,* vol. 69, pp. 46-61, 2014.

[25] H. R. Tizhoosh, "Opposition-based learning: a new scheme for machine intelligence," in *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, 2005, vol. 1: IEEE, pp. 695-701.

[26] S. Rahnamayan, H. R. Tizhoosh, and M. M. Salama, "Opposition-based differential evolution," *IEEE Transactions on Evolutionary computation,* vol. 12, no. 1, pp. 64-79, 2008.

[27] O. P. Verma, D. Aggarwal, and T. Patodi, "Opposition and dimensional based modified firefly algorithm," *Expert Systems with Applications,* vol. 44, pp. 168-176, 2016.

[28] G. Xiong, J. Zhang, D. Shi, and Y. He, "Oppositional Brain Storm Optimization for Fault Section Location in Distribution Networks," in *Brain Storm Optimization Algorithms*: Springer, 2019, pp. 61-77.

[29] A. A. Ewees, M. Abd Elaziz, and D. Oliva, "A new multi-objective optimization algorithm combined with opposition-based learning," *Expert Systems with Applications,* vol. 165, p. 113844, 2021.

[30] H. Muthusamy, S. Ravindran, S. Yaacob, and K. Polat, "An improved elephant herding optimization using sine–cosine mechanism and opposition based learning for global optimization problems," *Expert Systems with Applications,* vol. 172, p. 114607, 2021.

[31] Z. Zhang, Z. Xu, S. Luan, and X. Li, "A Hybrid Max–Min Ant System by Levy Flight and Opposition-Based Learning," *International Journal of Pattern Recognition and Artificial Intelligence,* p. 2151013, 2021.

[32] K. Kalaiselvi, K. Velusamy, and C. Gomathi, "Financial prediction using back propagation neural networks with opposition based learning," in *Journal*

of Physics: Conference Series, 2018, vol. 1142, no. 1: IOP Publishing, p. 012008.

[33] D. Bairathi and D. Gopalani, "Opposition-based sine cosine algorithm (OSCA) for training feed-forward neural networks," in *2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 2017: IEEE, pp. 438-444.

[34] A. Mahmoodzadeh, H. R. Nejati, M. Mohammadi, H. H. Ibrahim, S. Rashidi, and T. A. Rashid, "Forecasting tunnel boring machine penetration rate using LSTM deep neural network optimized by grey wolf optimization algorithm," *Expert Systems with Applications,* vol. 209, p. 118303, 2022.