Received: May 14, 2023. Revised: July 9, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

LSTM based Modified Remora Optimization Algorithm for Lung Cancer Prediction

Manaswini Pradhan[1]* Ioana L. Coman[2] Subhankar Mishra[3] Thanh Thieu[4] Alauddin Bhuiyan[5]

[1]*P.G. Department of Computer Science, Fakir Mohan University, Balasore, Odisha, India*
[2]*Department of Computer Science, State University of New York, Oswego, New York, USA*
[3]*School of Computer Sciences, National Institute of Science Education and Research, Bhubaneswar, Odisha, India*
[4]*Department of Computer Science, Oklahoma State University, Stillwater, Oklahoma, USA*
[5]*Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, NY, USA*
* Corresponding author's Email: mpradhan.fmu@gmail.com

Abstract: Early detection of lung cancer in patients can decrease the mortality rate and helps in early diagnosis. An efficient lung cancer prediction system is proposed in this paper known as a long short-term memory (LSTM) based modified remora optimization algorithm (MROA) with multi-objective criteria. Although the MRO is proposed by several researchers, it was based on single-objective with traditional objective functions which tried to maximize the accuracy. This led to the biased classifier with high accuracy and sacrificed sensitivity which resulted in insufficiency of one class and prevalence of other class. To overcome this limitation, a Multi-objective MROA (MMROA) for hyperparameter tuning is proposed, where both accuracy and sensitivity are equally considered as a fitness function. Initially the histopathology images of Lung and Colon 25000 (LC25000) dataset is normalised using colour normalization and segmented with saliency driven edge dependent top-down level set (SDREL) method. The features are extracted using Grey Level Cooccurrence-Matrix (GLCM) and GoogleNet followed by the feature selection using enhanced grasshopper optimization algorithm (EGOA). The selected features are optimized (hyperparameter tuning) using MMROA and fed to LSTM classifier. The proposed model is compared with the existing models such as convolutional neural network (CNN) and enhanced grasshopper optimization algorithm based random forest (EGOA-RF) and obtained remarkable results with accuracy, precision, recall, and f-measure values of 99.02%, 99.17%, 99.03%, and 99.24% respectively.

Keywords: GoogleNet, Gray level co-occurrence matrix, Histopathology images, Long-short term memory, Lung cancer classification, Multi-objective modified remora optimization algorithm.

1. Introduction

One of the most prevalent causes of cancer related deaths is due to lung cancer which is highly dangerous compared to colon, prostate, and breast cancers. It is claimed that more than 6.1 million people die from direct tobacco use and 900,000 people die from being exposed to that smoke. As more than 62% of diagnosed cases are in people over 65, the life expectancy of lung cancer patients has been relatively low, at 10% [1-4]. To increase the life expectancy of people with lung cancer, accurate detection, early diagnosis and prompt treatment are the best ways. Since the 1970s, pulmonary nodules (the medical term for cancerous tissues in the lung) have been found using computerized tomography (CT) and chest radiographs (CXR) to diagnose lung cancer [5, 6]. Smokers typically have one or more pulmonary nodules. More than 25% of smokers have multiple pulmonary nodules. There are different types of lung cancer images used for clinical imaging, but CT scans were frequently preferred due to their reduced noise. The high-intensity and resolution images are generated by well-known digital pathology systems compared to the currently used

imaging methods, which demonstrates the true source of the patient's illness. It has been identified that Deep Learning models are the best way for clinical imaging, highlight extraction, and item characterization from various existing research [7-9]. Decision trees, support vector machines, and neural networks are the frequently used classifiers for classifying types of lung cancer. Neural networks such as auto encoders, recurrent neural networks, and feedforward neural networks have made constant progress among the classifiers [10, 11]. Computer aided diagnosis (CAD) is the most used system in the medical field, for fast and effective diagnosis and detection of cancer with better classification accuracy. As determined by a CAD system using texture analysis and feature extraction, tumour heterogeneity is the key factor in determining how aggressive a tumour is. [12, 13]. Many deep learning models such as artificial neural network (ANN), CNN, deep convolutional neural network (DCNN) has been proposed for predicting lung cancer in its initial stages [14, 15]. Other existing methods and their drawbacks are also discussed in this manuscript, to validate the proposed model's performance. The major contributions of this work are:

- The attributes obtained from GLCM and GoogleNet are used for extracting active attribute vectors which has the benefit of characterizing image texture by analyzing statistical measures of the image matrix. A GoogleNet model with deep layers is used as it trains the deeper layer networks faster with less computational time.
- The obtained output is given to the LSTM classifier to classify lung cancer tissues based on time series data. If there exists a vanishing gradient problem in the network, it is addressed by using the LSTM model
- The hyperparameter optimization is carried out using MMROA to enhance the performance of the LSTM model in classification and prediction tasks. The MMROA deals with the multi-level threshold segmentation of complex grey-scale images.
- The implementation of MMROA-LSTM results in better accuracy in predicting lung cancer compared to the existing models by considering accuracy and sensitivity as fitness function.

The structure of the remaining paper is as follows: section 2 explains the literature review, section 3 describes the proposed methodology, section 4 illustrates the experimental results, and section 5 summarizes the entire manuscript.

2. Literature review

This section provides the existing research on lung cancer prediction and classification performed on LC25000 dataset.

K. Adu [16] proposed a new lung cancer classification model known as a dual horizontal squash modeled capsule network (DHS-CapsNets). It involves two functions namely horizontal squash (H-Squash) model and an encoder based feature fusion (EFF) for classifying the colon from lung cancer. To extract the values of the active features from the various histopathological background images, the H-Squash function was used. Divergent spherocylinder is also generated by the H-squash function. The extracted feature vectors from the convolutional layers were integrated by the EFF function which generates valid information about features. The performance evaluation of this model was evaluated on the LC25000 dataset in terms of precision and recall. The DHS-CapsNets model increased the depth because of the fading gradient that occurred in the network, which was a drawback of this work.

M. Masud [17] proposed a CNN classification model to detect an overall five different types of colon and lung cancers which include 2 cancerous and 3 non-cancerous tissues using the LC25000 dataset. Digital image processing techniques like 2-dimensional-discrete fourier transform (2D-DFT) and 2D Wavelet Transform was used for feature extraction. The proposed CNN model has the advantage of being less prone toward a specific class while decision-making. However, it was computationally expensive due to its large requirement of dataset images for training the network, which was a drawback of this work.

Manaswini Pradhan [18] proposed a novel EGOA based lung cancer classification model with a random forest classifier for lung cancer from histopathological images of the LC25000 dataset. An SDREL method was applied for segmenting images followed by GLCM features for feature extraction. The extracted features were selected using Alexnet. At last, the enhanced grasshopper optimization algorithm was combined with a known RF classifier, for selecting optimal features and performing classification. The proposed model has achieved the highest accuracy with the EGOA, but the accuracy of histopathological images was not enhanced, which was a drawback of this work.

Bijaya Kumar Hatuwal [19] proposed a CNN model for classifying an image of three different categories, namely, adenocarcinoma, benign, and squamous cell carcinoma, from the LC25000 dataset. Image acquisition techniques like horizontal and

vertical flipping and zooming were applied to the images to observe variations in the data pattern in this work. An adaptive moment estimation (Adam) optimizer was used in measuring the learning rates of different parameters. Further, the model was trained and tested on the GPU: 0 devices known as the Google Collaboratory GPU. However, the accuracy of this classification model is not evaluated which represents the robustness of the model.

Satvik Garg [20] proposed an approach of using eight pre-trained CNN-based models, such as VGG-16, ResNet 50, DenseNet169, NASNetMobile, MobileNet, Xception, InceptionV3, and InceptionResNetV2 for lung and colon cancer detection. This approach was modified and validated on LC25000 dataset. To classify the malignant and benign images and to visualize the attention images, two methods, namely: GradCAM and SmoothGrad, were used in this work. This approach was not suitable for classifying all cancer types, which was a limitation of this work.

Nora Yahia Ibrahim [21] proposed a deep-learning model for diagnosing colon and lung cancer from the LC25000 dataset. Inverse discrete wavelet transform (IDWT) and double-contrast limited adaptive histogram equalization (Double-CLAHE) techniques were used to increase the poor boundary edges for better image contrast. The EfficientNetB7 was also used in this research which has an advantage for this work by employing a compound scaling strategy to increase network capacity, but its training time is relatively more. The results are calculated using accuracy, recall, precision, and f-measure.

Mizuho Nishio [22] proposed a CAD model for automatically classifying three types of lung cancer tissues, namely: benign, cell carcinoma, adenocarcinoma, and lung squamous from LC25000 dataset. The obtained results have shown that homology-based image processing acquired better accuracy results than conventional texture analysis in the automatic classification of lung cancer tissues. The limitations of this work state that, there was no external validation done for the CAD model as it may cause an overfitting issue.

Purba Daru Kusuma [23] presented an improved version of pelican optimization algorithm (POA) known as guided pelican algorithm (GPA) to optimize the portfolio issues and achieved better results. It improved GPA in three ways such as: replacing the randomized target with the global best solution as the deterministic target, replaced the present location with the search space size, implemented multiple candidates in first two cases instead of individual candidates as it was employed in the original POA. However, the real-world optimization problems are still challenging for GPA.

Purba Daru Kusuma [24] presented a stochastic komodo algorithm (SKA) which focused on conducting diversification and intensification without redundancy. This algorithm is a modified version of Komodo mlipir algorithm (KMA) inspired by Komodo's behaviour during the foraging and mating. The effectiveness of SKA was evaluated on the five unimodal and five multimodal functions. This approach has achieved better results through female dominant formation. Although, the algorithm still has real world optimization problems.

Purba Daru Kusuma [25] introduced a quad tournament optimizer (QTO), which performed four searches in every iteration by agent. The four searches include: searching for the best solution, searching between the global solution and random solution, searching relative to a randomly selected solution, and the neighbourhood search around the corresponding solution and the global best solution. The mechanism has carried out a tournament to find a best candidate. The QTO was challenged to find the optimal solution out of 23 classic functions and achieved better output results for 9 functions when compared to exiting metaheuristic algorithms. The sorting mechanism at the beginning of the iterations consumes more computation, which is a drawback of QTO.

SHAHID MEHMOOD [26] suggested a transfer learning based class selecting image processing (CSIP-TL) for the accurate prediction of lung and colon cancer using LC25000. A pre-trained AlexNet network was tuned before classifying the dataset images. This approach has achieved better classification accuracy and efficient computation along with reduced time and computational cost due to its ability to select all classes of images. However, the results are largely incomparable, due to its selection of various classes.

To overcome aforesaid limitations such as high computational time, low network capacity, overfitting issues, and insufficient data size in the existing algorithms a deep learning lung cancer prediction model is proposed. The proposed model is composed of deep layer networks which perform histopathological image classification and accurate prediction of lung cancer.

3. Methodology

The proposed histopathological lung cancer prediction model includes six stages as shown in

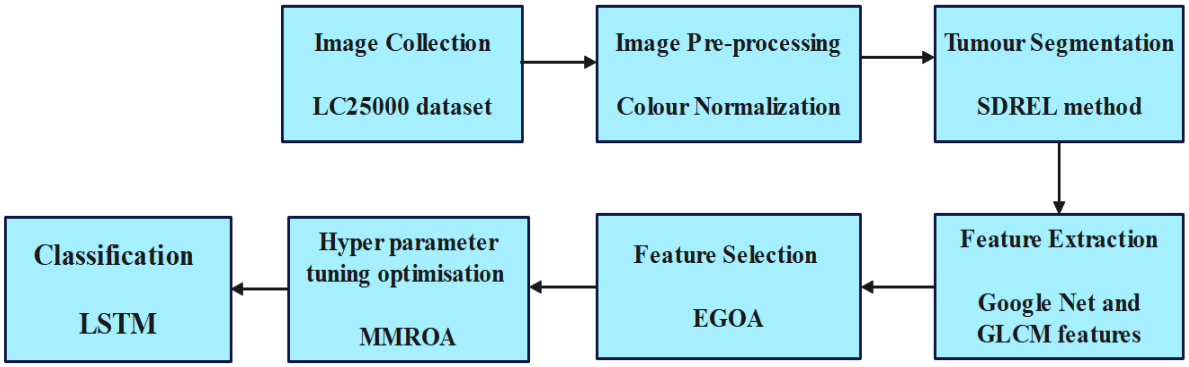
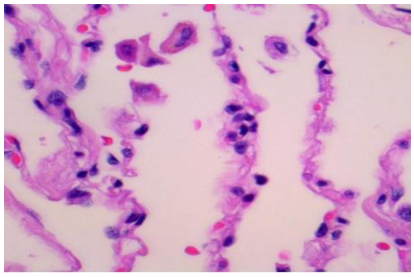
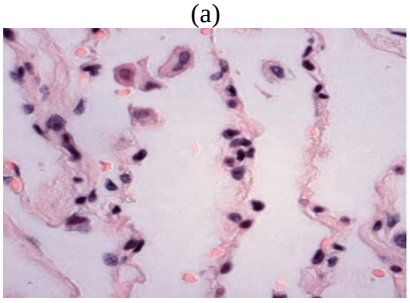


Figure. 1 Block diagram for the proposed lung cancer prediction model



(a)



(b)

Figure. 2: (a) Sample image input of collected histopathology, and (b) Sample image input of normalized histopathology

Fig. 1. The six stages are namely: 1. Image collection from LC25000 dataset, 2. Image pre-processing with color normalization technique, 3. Image segmentation using SDREL set approach, 4. Obtain GLCM feature as the process of feature extraction and GoogleNet model, 5. Selection of useful features through EGOA mechanism and 6. Classification using MMROA-LSTM.

3.1 Image collection

The suggested prediction model for lung cancer experiments on the publicly available standard dataset LC25000 which consists of 25,000 histopathology images with 768×768 resolution. The images are divided into 5 classes in a dataset with 5000 of each type like benign colon tissue images, benign lung tissue images, carcinoma of lung squamous cell, and adenocarcinomas of colon and lung. [18]. The proposed experimental investigation uses 15,000 lung-related tissue images from this dataset like benign lung tissue, carcinoma of lung squamous cell and adenocarcinoma of the lung. The sample input and its normalized form are shown in Fig. 2 for the set of input data gathered for lung cancer with histology.

Dataset link:
https://academictorrents.com/details/7a638ed187a6180fd6e464b3666a6ea0499af4af

3.2 Pre-processing

Pre-processing digital pathology images to find deformation and imaging variations to enable clear vision is the first step in analyzing them. Color normalization is the general pre-processing technique that is applied to the images to enable a clear vision of them. The mathematical equation for calculating color normalization is represented in Eq. (1).

$$IN_r = (I - Min) + \frac{newMax - newMin}{Max - Min} + newMin \quad (1)$$

Where, IN_r is the image normalized, I is the input histopathology image sample from collected data. The value of maximum and minimum in the pixel range is denoted by Max and Min and $newMax$ and $newMin$ is the pixel range of normalized images.

3.3 Segmentation

The cells with nuclei and without nuclei (non-nuclei) are segmented from the pre-processed images such as image normalized, information of gradient and information based upon the region by using an SDREL technique. Here, the edge-based information is given by maps of gradient and the information dependent on the region is given by maps of saliency and intensity of colour. The input images of the sensitive salience knowledge are integrated within

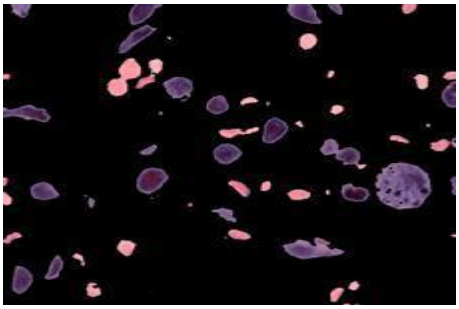


Figure. 3 Output of SDREL method

the suggested model that cannot visible to the naked eye. The SDREL method has two phases in the image segmentation in which the first phase includes generating external energy from both gradient and region-based information, and in the second phase, from the internal energy, the evolution is dependent completely. The segmentation of the pre-processed images by using the SDREL mechanism has important advantages such as:

- Many levels set-based techniques have fundamental problems, including their sensitivity to the initiation condition. The SDREL method addresses this issue by utilizing a two-stage evolution process. The suggested model is insensitive to the initialization condition according to the top-down evolution strategy, which leads to quick and reliable level-set evolution.
- The level set energy function and saliency maps are combined in a new energy term, which significantly enhances level set-based image segmentation. As a result, Fig. 3 graphically represents the segmented image.

3.4 Feature extraction

The images segmented based upon metrics such as entropy measures of images normalized, normalized with an inverse difference, contrast, sum, prominence of the cluster, energy, variance of sum, the difference among entropies, correlation, shades of the different cluster, the inverse of the difference values, the probability at maximum, variance in differences, correlation measure, an average of all summated entropy values, moment normalized with an inverse difference, dissimilarity, the sum of squares, variance, autocorrelation and homogeneity are obtained as active 21 feature vectors by GoogleNet model and GLCM features. The occurrence of a smaller number of false positives is the main advantage of GLCM features. The GoogleNet model is a variant of the inception network which consists of 22-layer Deep CNN. This model detects images with 5 to 22 layers and also has high accuracy. To decrease the size of the inception module layer, first, a 1×1 matrix convolution is performed before performing convolution with 3×3 and 5 ×5 metrics.

To demonstrate the spatial information of image pixels, features such as energy, entropy, contrast, correlation, and homogeneity are used.

Energy: The sum of square elements in the GLCM matrix is given by energy, which is said to be high when the frequency of repeated image pixels is high. The mathematical formula to measure energy in GLCM is represented by Eq. (2).

$$Energy = \sum_{i,j=0}^{N-1}(P_{i,j})^2 \tag{2}$$

Entropy: The measure of randomness in intensity of neighborhood pixels is known as entropy and it is measured using Eq. (3).

$$Entropy = \sum_{i,j=0}^{N-1} -\ln(P_{ij})P_{ij} \tag{3}$$

Contrast: The measure of the spatial frequency of an image which gives the difference between the highest and lowest values of pixels that are adjacent to each other. It is measured using Eq. (4).

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2 \tag{4}$$

Correlation: correlation specifies how a specific pixel is related to its neighbor pixel, and measures the gray tone linear dependencies. The correlation between pixels is calculated using Eq. (5).

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij}\frac{(i-\mu)(j-\mu)}{\sigma^2} \tag{5}$$

Where μ is the mean of GLCM matrix, σ is the variance of all pixel's intensities which is calculated using Eqs. (6) and (7).

$$\mu = \sum_{i,j=0}^{N-1} iP_{ij} \tag{6}$$

$$\sigma^2 = \sum_{i,j=0}^{N-1} P_{ij}(i-\mu)^2 \tag{7}$$

Homogeneity: The similarity of pixels, which also measures the closeness of distributed elements in the GLCM is known as the homogeneity of the pixels. The mathematical expression for the homogeneity of pixels is represented by Eq. (8).

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} \tag{8}$$

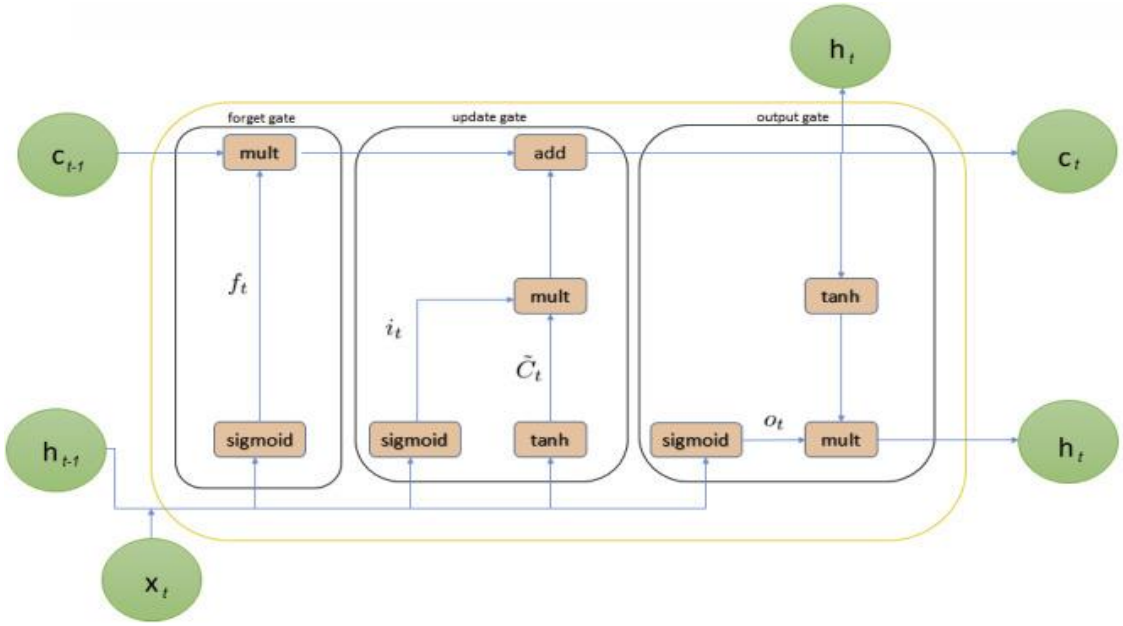


Figure. 4 LSTM architecture

Where, $P_{i,j}$ is the ith and jth element of the normalized image and N is the number of gray levels in the image.

3.5 Feature selection

While choosing the best feature vectors for image classification, the EGOA is provided with the extracted features as input. The traditional GOA imitates grasshopper swarming behavior. The Eq. (9) represents the position of the i^{th} grasshopper mathematically and it is indicated by Xi.

$$Xi = Si + Gi + Ai \tag{9}$$

Where, Si stands for social interaction, Gi for gravitational force, and Ai for the i^{th} grasshopper's exposure to wind. The mathematical representation of these three components is given in Eqs. (10), (11), and (12) which describes the motion of grasshoppers in the GOA.

$$Si = \sum_{j=1, j\neq i}^{N} s(d_{ij})\grave{d}_{ij} \tag{10}$$

$$Gi = -geg \tag{11}$$

$$Ai = uew \tag{12}$$

Where, $\grave{d}_{ij} = \frac{xj-xi}{d_{ij}}$ and $d_{ij} = |xj - xi|$ indicates unit vector, $s(r) = \frac{fe-r}{l-e-r}$ stands for social forces, f for attraction intensity, l for the scale of attractive length, g for a constant of gravitational attraction, u for drift constant. The ew represents the vector unity that uses the aligned direction of the wind, and eg represents the vector unity aimed towards the spatial center of the globe. Eq. (13) is used to calculate the movement of nymph grasshopper.

$$Xi = j = 1, jiN\, s\big(d_{ij}\big)(d_{ij}) - geg + uew \tag{13}$$

Where N represents several grasshoppers. The mechanism of exploration GOA completed and exploitation of the same to find the optimal value of approximation globally to solve the optimization problem. Hence Eq. (13) is revised as shown in Eq. (14).

$$Xid = cj = 1, jiNcubd - lbd2s(|xjd - xid|) \\ xj - xid_{ij} + Td \tag{14}$$

Where, ubd and lbd stand for upper and lower bound values, and Td for target value.

A decreasing coefficient in direct proportion to the number of iterations is denoted by C and represented in Eq. (15)

$$C = cmax - itrcmax - cmin \times Maxitr \tag{15}$$

Where, itr stands for iteration number, maximum iterations are denoted by $Maxitr$ which is equal to 100, minimum c value is denoted by $cmin$, and maximum iterations with c value is given by $itrcmax$

3.6 Classification

After extracting and selecting active feature vectors, the images undergo a tumor classification process by using the LSTM classifier whose architecture is represented in Fig 4. For hyperparameter optimization of selected features, a MMROA is used in this research. The hyper parameters are optimized for effective classification of lung tissue images into lung tissue of benign, carcinomas of the lung squamous cell, and the lung adenocarcinomas.

The principle of the input gate of LSTM is shown using the Eqs. (16), (17), and (18).

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{16}$$

$$\widetilde{C_t} = (W_i.[h_{t-1}, x_t] + b_i) \tag{17}$$

$$C_t = f_t C_{t-1} + i_t \widetilde{C_t} \tag{18}$$

Where, x_t refers to the current input, h_t and h_{t-1} refers to the current and previous outputs, C_t is the $tanh$ output, $\widetilde{C_t}$ is the current moment information, C_t and C_{t-1} are the new and previous cells. W_i refers to weight matrices, b_i denotes the input gate bias of LSTM, and σ is the sigmoid function.

The LSTM's output gate determines the continuation of h_{t-1} and x_t using the Eqs. (19) and (20).

$$O_t = \sigma(W_o.[h_{t-1}, x_t] + b_0) \tag{19}$$

$$h_t = O_t \tan(C_t) \tag{20}$$

Where, W_o and b_0 are the output gates' weighted matrices and LSTM bias.

3.6.1. Optimizing LSTM classifier Hyper parameters with MMROA

Although the MRO is proposed by several researchers, it was based on single-objective criteria with traditional objective functions which tried to maximize the accuracy. This led to the biased classifier with high accuracy and sacrificed sensitivity. To overcome this limitation, a Multi-objective MROA (MMROA) for hyperparameter tuning is proposed, where both accuracy and sensitivity are equally considered as a fitness function. The hyper parameters are initialized first which are randomly generated in an attempt to improve the classification accuracy. This phenomenon is repeated until the stopping criteria are met. For evaluating and returning the classification accuracy, the LSTM networks are responsible which is considered to be a fitness function. The Remora optimization algorithm (ROA) replicates moments of remora in the ocean based on parasitic attachment, direct attack, and feeding behavior of the host. To update the remora moments and positions, ROA uses some formulas of sailed fish optimizer (SFO) and whale optimization algorithm (WOA) as shown in the following steps:

Step 1: Calculating the initialization of the remora population using Eq. (21)

$$X_{i,j} = lb_j + rand \times (ub_j - lb_j) \tag{21}$$

where, $X_{i,j}$ is the position of ith remora in jth dimension, lb_j and ub_j are the lower and upper boundaries of the search space, and rand is a random number from 0 to 1.

The initialization of lower bound and upper bound values is based on the minimum and maximum values of the hyperparameters.

Step 2: Exploration

It has been demonstrated that remora will follow the sailfish once it has been absorbed by it. The SFO algorithm's formula is enhanced based on its elite strategy, and the resulting formula shown is Eq. (22)

$$X_i^{t+1} = X_{Best}^t - (rand \times \left(\frac{X_{Best}^t + X_{rand}^t}{2}\right) - X_{rand}^t) \tag{22}$$

Where, the current iteration number, current best position, and current random position of remora are given by t, X_{Best}^t, and X_{rand}^t. If there is a replacement of the host, the remora will move in a circular motion of the host by considering the previous positions of the host. The mathematical expression of this phenomenon is given in Eq. (23)

$$X_{att} = X_i^t + (X_i^t + X_{pre}) \times randn \tag{23}$$

Where, X_i^t is the position of ith remora in the current iteration, $randn$ is the random number between 0 to 1. Remoras' tentative moment and its previous generation's position is represented by X_{att} and X_{pre}. The parameters assumed in the MMROA are listed as follows: population size is 30, maximum iteration is 100, and dimension is 4. The parameter settings for the LSTM hyper parameters are as follows: max-epoch between 5, 10, 15, and 20; dropout range between 0.1 and 0.4; learning rate range between 0.003 and 0.1; and L2 regularisation range between 0.003 and 0.1.

- **Fitness function**

The classifier output of LSTM networks acts as fitness which is responsible for the evaluation and returning the classification accuracy. The fitness values of X_i^t and X_{att} determines whether the remora will switch hosts after a brief range of movement as shown in Eq. (24).

$$f(X_i^t) > f(X_{att}) \tag{24}$$

$f(X_i^t)$ and $f(X_{att})$ is the fitness function, which represents the values of the current position (X_i^t) and previous position (X_{att}) of remora. The mathematical expression for selecting a host can be given by Eq. (25)

$$H(i) = round(rand) \tag{25}$$

Where, $round$ is the rounded function, $H(i)$ is the host absorbed by remora with starting value 0 or 1.

The host will be sailfish if the value of $H(i)$ is 0. Similarly, the host will be whale if the value is equal to 1.

Step 3: Exploitation

If the host is the whale, the remora will move in a synchronous direction along with the whale and it is represented as shown in Eq. (26).

$$X_i^{t+1} = D \times e^l \times cos\ cos\ (2\pi a) + X_i^t \tag{26}$$

Where, e is constant value with 2.7182, D is the distance between current best position and optimal position and it is given by Eq. (27).

$$D = \left|X_{Best}^t - X_i^t\right| \tag{27}$$

Where, l from Eq. (26) is a random number within [-1, 1] and represented mathematically as shown in Eq. (28).

$$l = rand \times (a - 1) + 1 \tag{28}$$

Where, a is a linear decrease during iteration between [-2, 1] and it is represented by Eq. (29).

$$a = -(1 + \frac{t}{T}) \tag{29}$$

Where, T in the above Eq. (30) represents the number of maximum iterations.

The search range is narrowed in the exploitation stage, as remora seeks food near the host and its calculated as shown in Eq. (30).

$$X_i^{t+1} = X_i^t + A \tag{30}$$

1. Where A is the distance of remora's movement that is proportional to the remora's volume and host which is measured as shown in Eq. (31).

$$A = B \times (X_i^t - C \times X_{Best}) \tag{31}$$

Where B determines the host volume simulation which is calculated using Eq. (13) and C is the remora factor used in limiting the remora's position by using Eq. (32).

$$B = 2 \times V \times rand - V \tag{32}$$

Where, remora's volume V is given by Eq. (33)

$$V = 2 \times (1 - \frac{t}{T}) \tag{33}$$

However, the performance of RAO is not great due to its weak functionality when considering the feeding habits of remora, which depends on the host to find food. Hence MMROA is used that includes a host-switching mechanism, joint opposite selection, and restart strategy, to solve the imbalance issue between exploration and exploitation.

3.6.1.1. Host-switching mechanism

The new switching mechanism judges whether the remora needs to switch the host or not by reducing the self-feeding ability of remora. The mathematical expression for the new host-switching mechanism is given by Eq. (34).

$$X_{new} = X_i^t + k \times step\ if\ rand < P \tag{34}$$

Where, X_{new} is the new solution, P is the decreasing factor between 1 to 0.5 used for frequency control of new solutions, and k is the random factor which is measured using Eq. (35).

$$k = \beta \times (1 - rand) + rand \tag{35}$$

Where, β = 0.2 which is a constant value. $step$ in Eq. 15 is calculated using Eq. (36).

$$step = X_{r1}^t - X_{r2}^t \tag{36}$$

$step$ is between the distance between two random functions X_{r1}^t and X_{r2}^t. The $f(X_{new})$ is the fitness value of X_{new} and it is given by Eq. (37).

$$f(X_i^t) < f(X_{new}) \quad (37)$$

3.6.1.2. Joint opposite selection

Union of selective leading opposition (SLO) and dynamic opposition (DO) is chosen in the joint opposition selection phase that balances the expedition and utilization of the algorithm.

The variation among the current best position and previous best position is measured by SLO in every dimension using Eq. (38).

$$\overline{X_{\iota,d_c}} = lb + ub - X_{i,d_c} \; if \; src < 0 \text{ and size } (d_c) > \text{size } (d_f) \quad (38)$$

Where, d_c and d_f are the dimension of close distance and far distance, $\overline{X_{\iota,d_c}}$ is the contrary ith solution in closed dimension.

opposite ith solution of close dimension.

The DO is used to attain the advantage of allowing the algorithm to escape the local optimum by searching space in dynamic nature and moving asymmetrically in the search space. A mathematical representation of this phenomenon is given in Eq. (39).

$$\overline{X_{do}} = X_i + rand \times (X_r - X_i) \text{ if } rand < jr \quad (39)$$

Where, $\overline{X_{do}}$ is the opposite dynamic solution, jr is the jump rate which is the probability of carrying out dynamic opposition, X_r is calculated using Eq. (40).

$$X_r = rand \times \overline{X}_\iota \quad (40)$$

Where, X_r is the random opposite solution, $X_i$is calculated in Eq. (41).

$$\overline{X}_\iota = lb + ub - X_i \quad (41)$$

Where, $\overline{X}_\iota$ is the opposite ith solution of X_i.

3.6.1.3. Restart strategy

A trail vector $trail(i)$ is used to maintain a record of individual movements when stooped at a local optimum. Next, to assist the individual, a restart strategy is initialized in moving the range outside of the local optimum and escaping the state of stagnation. The value of $trail(i)$ is increased by 1, if there is no improvement in the solution for the long period, or else it will be reset to zero. When the limitation of $trail(i)$ exceeds a predefined limitation, it will produce two new solutions as shown in Eqs. (42) and (43).

$$T_1 = (ub - lb) \times rand + lb \quad (42)$$

$$T_2 = (ub + lb) \times rand - X_i \quad (43)$$

Where, T_1 and T_2 are the new solutions to replace the current solution with the best-chosen one. If the T_2 crosses the boundary, it is pulled back using Eq. (44).

$$T_2 = (ub - lb) \times rand + lb \quad \text{if } T_2 > ub || T_2 < lb \quad (44)$$

The Pseudo code of MMROA

Input: Population Size, Number of Iterations, Hyperparameters of Classifier
Output: Optimal Hyperparameters of classifier
Process:
Initialize the size of population = 30, Number of iterations at maximum = 100
The initial position of the X_i(i= 1,2,…30) is randomly declared
Declare Iteration t=0
While(t<Number of iterations at maximum)
If the search agent reaches beyond the search space
Amend the Search agent again to random position in the search space
For all search agents in X_i:
Compute the fitness of each remora
If (Fitness (X_i) > Fitness (X_{Best})):
$X_{Best} = X_i$
For each operation using Eq. (38) by SLO:
For each remora indexed as i:
If H(i) ==0
Update position of Sailfish using Eq. (22)
Else If H(i) ==1
Update position of Whale using Eqs. (26)-(28)
End if
Do experience attack by use of Eq. (23)
If P > Rand:
Using the Eq. (34) compute a prediction around the host
If $f(X_i^t) > f(X_{new})$:

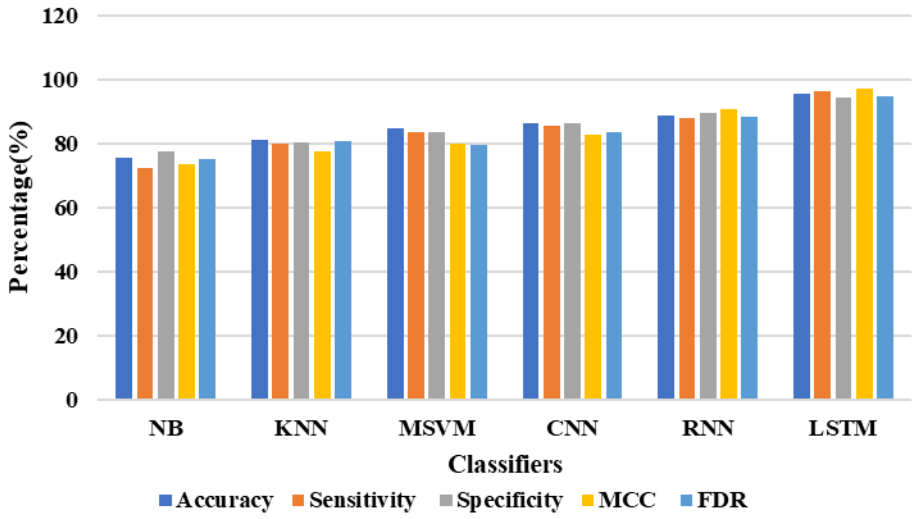


Figure. 5 Analysis of Hyper parameter optimizations for various deep learning classifiers

Switch the hosts using Eq. (25)
End if
End if
Perform host feeding mode for remora using Eq. (30)
For each position of agent perform DO using Eq. (39)
For each remora update the trail(i)
If trail(i) > limit:
Position generation for remora using Eqs. (43) and (44)
Select the best fitness value's corresponding position
End if
End For
End While

Return X_{Best}

4. Results and discussions

The performance evaluation of the proposed MMROA based LSTM hyperparameter models on the LC25000 dataset. In a system formation over Windows 10, the experimental evaluation is performed on MATLAB 2020, Intel core i9 processor, 16 GB RAM, and 4 TB hard disk. The performance metrics like Matthew's correlation coefficient (MCC), accuracy, specificity, sensitivity, and false discovery rate (FDR) are used in evaluating the performance of the proposed model. The value of MCC ranges from 0 to 1, where 1 is the best agreement and 0 is a no agreement. The mathematical expression of MCC is given in Eq. (45).

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \times 100 \quad (45)$$

Accuracy is the prediction of true classes o the total number of predictions and its mathematical formula is given by Eq. (46).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (46)$$

Specificity is the proportion of predicted true negative class to the total observations and is represented mathematically as shown in Eq. (47).

$$Specificity = \frac{TN}{TN+FP} \times 100 \quad (47)$$

Sensitivity is also known as true positive rate (TPR) which measures the true observation of positive classes and it is represented mathematically as shown in Eq. (48).

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (48)$$

The false Discovery Rate is determined as the proportion of false positives to the total number of total positive observations and is represented as shown in Eq. (49).

$$FDR = \frac{FP}{FP+TP} \times 100 \quad (49)$$

Where, TP, and TN are true observations of positive and negative outcomes; FP and FN are the False observations of positive and negative outcomes in the above equations.

Table 1. Performance analysis of Hyper parameter optimization for various deep learning classifiers

Hyper parameter optimization					
Classifiers	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**MCC (%)**	**FDR (%)**
NB	75.57	72.39	77.39	73.55	74.93
KNN	80.98	79.77	80.38	77.49	80.65
MSVM	84.79	83.34	83.66	80.05	79.67
CNN	86.32	85.44	86.23	82.54	83.66
RNN	88.53	87.76	89.41	90.65	88.34
LSTM	**95.67**	**96.45**	**94.33**	**96.95**	**94.54**

Table 2. Performance analysis of LSTM classifier with different hyperparameter optimization

LSTM Hyper parameter classifier					
Optimizers	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**MCC (%)**	**FDR (%)**
SFO	84.45	83.29	80.78	83.88	84.32
WOA	86.67	85.33	85.67	85.46	86.45
ROA	93.98	92.87	93.45	92.77	93.24
GPA	94.26	93.43	94.23	93.52	94.57
SKA	94.93	94.62	94.56	95.67	95.63
QTO	95.27	95.44	95.66	95.95	96.34
MMROA	**98.29**	**97.79**	**96.74**	**97.56**	**97.56**

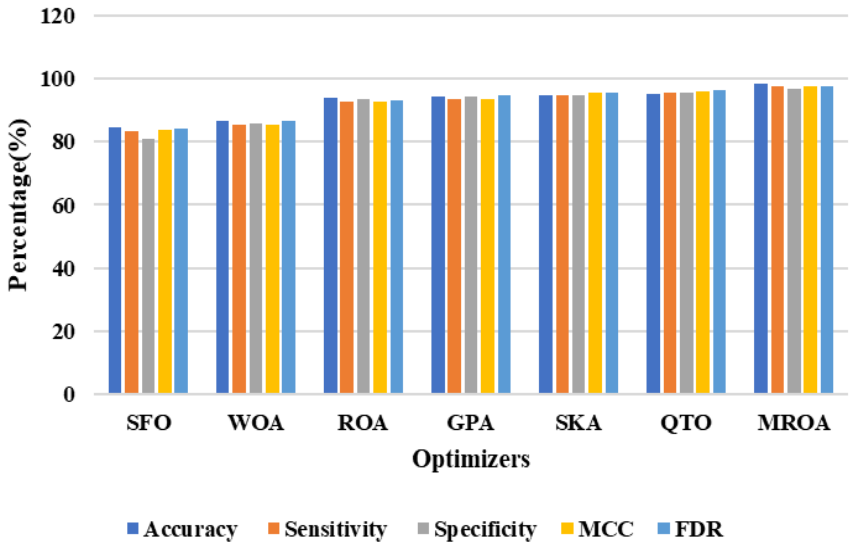


Figure. 6 Performance analysis of LSTM classifier with hyperparameter optimization

4.1 Quantitative evaluation

The potential of the proposed MMROA-LSTM is validated on the LC25000 dataset which consists of 250000 histopathological images. Table 1 represents the effectiveness hyperparameter optimization for LSTM compared to different classifiers like Naïve Bayes, K-nearest neighbor (KNN), and modified support vector machine (MSVM). The performance metrics are calculated in terms of accuracy, sensitivity, specificity, MCC, and FDR. From the observations, the hyper parameter optimization for the LSTM classifier has shown better results with an accuracy of 95.67%, sensitivity of 96.45%, specificity of 94.33%, MCC of 96.95%, and 94.54%. The graphical representation of these results is demonstrated in Fig. 5. The attained simulation result of the proposed MMROA-LSTM network is better compared to existing machine learning models like NB, KNN, and MSVM. In the lung cancer classification, the LSTM network provides an enormous range of parameters like input biases, learning rates, and output biases, where it does not require any fine adjustments.

Table 2 represents the effectiveness LSTM classifier with the proposed hyper parameter optimization (MMROA) and compared to different optimization techniques like sail fish optimization (SFO), whale optimization algorithm (WOA), Remora optimization algorithm (ROA), guided Pelican algorithm (GPA), stochastic Komodo

Table 3. Performance analysis of existing lung cancer prediction models to the proposed MMROA-LSTM

Models	Dataset	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)
CNN [17]	LC25000	96.33	96.38	96.39	96.37
EGOA-RF [18]		98.50	99.08	98.92	97.98
CSIP-TL [26]		98.40	-	-	-
MMROA-LSTM		99.02	99.24	99.17	99.03

algorithm (SKA), and Quad tournament optimizer (QTO). From the observations, the LSTM hyper parameter optimization for the MMROA has shown better results with an accuracy of 98.29%, sensitivity of 97.79%, specificity of 96.74%, MCC of 97.56%, and 97.56%. The graphical representation of these results is demonstrated in Fig. 6. The ability of both exploration and exploitation was balanced with the help of joint selection in MMROA. Along with this process, the global nature of the algorithm is improved to know the optimal solution.

4.2 Comparative evaluation

The existing lung cancer prediction models are compared with the proposed deep learning-based lung cancer prediction model in this section. The comparative results are evaluated in terms of accuracy, f-measure, precision, and recall on LC25000 histopathological image dataset as shown in Table 3. The proposed model is compared to existing models like convolutional neural network (CNN) [17], Enhanced grasshopper optimization algorithm based random forest (EGOA-RF) [18], and transfer learning based class selecting image processing CSIP-TL [26]. The results of the proposed model achieved the highest values with an accuracy of 99.02%, f-measure of 99.24%, precision of 99.17%, and recall of 99.03%.

The proposed model overcomes the following limitations: The given dataset images were insufficient to train the CNN model in classifying lung cancer tissues. As proposed by M. Masud [17], the CNN model requires a large number of histopathological images to train the model and it is computationally expensive. To overcome the limitation of limited data for training in existing CNN, an LSTM model with hyper parameter tuning optimization is proposed, that can train the network with large amount of input data. In this publication, the future work proposal of Ref [18] to improve the histopathology lung cancer diagnosis is implemented by adding hyper parameter optimisation to the LSTM. The extensive experimental results show that the proposed MMROA-LSTM network effectively minimized the information loss, and analyzed the recognition accuracy in lung cancer classification.

4.2.1. Discussion

It is observed that the proposed method has achieved remarkable results with an accuracy of 99.02%, f-measure of 99.24%, precision of 99.17%, recall of 99.03% as given in Table 3. The CNN [17] has achieved better results but this approach was computationally expensive due to its large requirement of dataset images for training the network. The EGOA-RF [18] has also achieved better classification accuracy with 98.50%, however, the input images are not enhanced in this method which resulted in low illumination images. The CSIP-TL [26] has also achieved 98.40% but the results are largely incomparable due to its selection of various classes. To overcome these limitations, the proposed method has introduced hyperparameter tuning using MMROA with multi-objective criteria and obtained remarkable classification with LSTM of 99.02% accuracy. The proposed method is effective in classifying histopathology images with multi-objective criteria by considering accuracy and sensitivity as a fitness function which is an advantage of finding best optimum solution.

5. Conclusion

A Deep Learning based LSTM hyperparameter tuning with MMROA is proposed in this manuscript for effective lung cancer detection. Initially, the LC25000 dataset is considered for histopathological images, after which a colour normalization method is applied to those images for clear vision. After pre-processing, the features from the images are extracted by applying GLCM features and GoogleNet, where the combination allows for extracting active feature vectors within the deep layers of the network. The chosen features are then fed to LSTM to classify the lung tissues. The proposed model's effectiveness has achieved greater results compared to the existing models on LC25000 dataset in terms of accuracy, f-measure, precision, and recall. The simulation result has shown that the proposed hyper parameter optimization model achieved greater results compared to existing lung cancer prediction models such as CNN, EGOA-RF, and CSIP-TL on LC25000 dataset. The proposed model achieved 99.02% accuracy, 99.17% precision, 99.03% recall, and 99.24% f-measure. An over-fitting issue is decreased with the accurate feature selection in this work.

Along with this, the training time is less and LSTM addresses the over-fitting issues if there exists any. Further, the proposed model's effectiveness should be validated on other datasets to know its ability in classification tasks. In future work, the proposed method should be implemented with other classifiers to compare LSTM classification in lung cancer.

Nomenclature

Terms	Representation
I	Sample input image
IN_r	Normalized image
N	Gray levels of image
$P_{i,j}$	ith and jth element of the normalized image
μ and σ	Mean of GLCM and variance of all pixel intensities
Xi	Position of ith grasshopper
X_i^t	Current position of ith remora
X_i^{t+1}	Improved location update of remora
X_{att}	Remoras previous position and tentative moment

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st and 3rd author. The supervision and project administration, have been done by 2nd, 4th and 5th author.

References

[1] G. Bansal, V. Chamola, P. Narang, S. Kumar, and S. Raman, "Deep3DSCan: deep residual network and morphological descriptor based framework forlung cancer classification and 3D segmentation", *IET Image Processing*, Vol. 14, No. 7, pp. 1240-1247, 2020.

[2] W. Wang and G. Charkborty, "Automatic prognosis of lung cancer using heterogeneous deep learning models for nodule detection and eliciting its morphological features", *Applied Intelligence*, Vol. 51, No. 4, pp. 2471-2484, 2021.

[3] A. Srinivasulu, K. Ramanjaneyulu, R. Neelaveni, S. R. Karanam, S. Majji, M. Jothilingam, and T. R. Patnala, "RETRACTED ARTICLE: Advanced lung cancer prediction based on blockchain material using extended CNN", *Applied Nanoscience*, Vol. 13, No. 2, p. 985, 2023.

[4] M. A. Heuvelmans, P. M. A. V. Ooijen, S. Ather, C. F. Silva, D. Han, C. P. Heussel, W. Hickes, H. U. Kauczor, P. Novotny, H. Peschl, M. Rook, R. Rubtsov, O. V. Stackelberg, M. T. Tsakok, C. Arteta, J. Declerck, T. Kadir, L. Pickup, F. Gleeson, and M. Oudkerk, "Lung cancer prediction by Deep Learning to identify benign lung nodules", *Lung Cancer*, Vol. 154, pp. 1-4, 2021.

[5] D. Yu, Z. Liu, C. Su, Y. Han, X. Duan, R. Zhang, X. Liu, Y. Yang, and S. Xu, "Copy number variation in plasma as a tool for lung cancer prediction using Extreme Gradient Boosting (XGBoost) classifier", *Thoracic cancer*, Vol. 11, No. 1, pp. 95-102, 2020.

[6] R. Selvanambi, J. Natarajan, M. Karuppiah, S. K. H. Islam, M. M. Hassan, and G. Fortino, "RETRACTED ARTICLE: Lung cancer prediction using higher-order recurrent neural network based on glowworm swarm optimization", *Neural Computing and Applications*, Vol. 32, No. 9, pp. 4373-4386, 2020.

[7] A.M. Khalil, S. G. Li, Y. Lin, H. X. Li, and S. G. Ma, "A new expert system in prediction of lung cancer disease based on fuzzy soft sets", *Soft Computing*, Vol. 24, No. 18, pp. 14179-14207, 2020.

[8] G. Carioli, M. Malvezzi, P. Bertuccio, P. Boffetta, F. Levi, C. L. Vecchia, and E. Negri, "European cancer mortality predictions for the year 2021 with focus on pancreatic and female lung cancer", *Annals of Oncology*, Vol. 32, No. 4, pp. 478-487, 2021.

[9] S. Suresh and S. Mohan, "ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis," *Neural Computing and Applications*, Vol. 32, No. 20, pp. 15989-16009, 2020.

[10] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier", *Neural Computing and Applications*, Vol. 34, No. 12, pp. 9579-9592, 2020.

[11] P. Murchan, C. Ó. Brien, S. O. Connell, C. S. McNevin, A. M. Baird, O. Sheils, P. Ó. Broin,

and S. P. Finn, "Deep learning of histopathological features for the prediction of tumour molecular genetics", *Diagnostics*, Vol. 11, No. 8, p. 1406, 2021.

[12] Y. Su, K. He, D. Wang, and T. Peng, "An Improved Level Set Method on the Multiscale Edges", *Symmetry*, Vol. 12, No. 10, p. 1650, 2020.

[13] M. M. Yamunadevi and S. S. Ranjani, "RETRACTED ARTICLE: Efficient segmentation of the lung carcinoma by adaptive fuzzy–GLCM (AF-GLCM) with deep learning based classification", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 12, No. 5, pp. 4715-4725, 2021.

[14] P. Nagabushanam, S. T. George, and S. Radha, "EEG signal classification using LSTM and improved neural network algorithms", *Soft Computing*, Vol. 24, No. 13, pp. 9981-10003, 2020.

[15] P. M. Shakeel, A. Tolba, Z. A. Makhadmeh, and M. M. Jaber, "Automatic detection of lung cancer from biomedical data set using discrete AdaBoost optimized ensemble learning generalized neural networks", *Neural Computing and Applications*, Vol. 32, No. 3, pp. 777-790, 2020.

[16] K. Adu, Y. Yu, J. Cai, K. O. Agyemang, B. A. Twumasi, and X. Wang, "DHS-CapsNet: Dual horizontal squash capsule networks for lung and colon cancer classification from whole slide histopathological images", *International Journal of Imaging Systems and Technology*, Vol. 31, No. 4, pp. 2075-2092, 2021.

[17] M. Masud, N. Sikder, A. A. Nahid, A. K. Bairagi, and M. A. AlZain, "A Machine Learning Approach to Diagnosing Lung and Colon Cancer Using a Deep Learning-Based Classification Framework", *Sensors*, Vol. 21, No. 3, p. 748, 2021.

[18] M. Pradhan, A. Bhuiyan, S. Mishra, T. Thieu, and I. L. Coman, "Histopathological Lung Cancer Detection Using Enhanced Grasshopper Optimization Algorithm with Random Forest", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 6, pp. 11-20, 2022, doi: 10.22266/ijies2022.1231.02.

[19] B. K. Hatuwal and H. C. Thapa, "Lung cancer detection using convolutional neural network on histopathological images", *International Journal of Computer Trends and Technology*, Vol. 68, No. 10, pp. 21-24, 2020.

[20] S. Garg and S. Garg, "Prediction of lung and colon cancer through analysis of histopathological images by utilizing Pre-trained CNN models with visualization of class activation and saliency maps", In: *2020 3rd Artificial Intelligence and Cloud Computing Conference*, Kyoto Japan, pp. 38-45, 2020.

[21] N. Y. Ibrahim and A. S. Talaat, "An Enhancement Technique to Diagnose Colon and Lung Cancer by using Double CLAHE and Deep Learning", *International Journal of Advanced Computer Science and Applications*, Vol. 13, No. 8, pp. 276-282, 2022.

[22] M. Nishio, M. Nishio, N. Jimbo, and K. Nakane, "Homology-based image processing for automatic classification of histopathological images of lung tissue", *Cancers*, Vol. 13, No. 6, p. 1192, 2021.

[23] P. D. Kusuma and A. L. Prasasti, "Guided Pelican Algorithm", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 6, pp. 179-190, 2022, doi: 10.22266/ijies2022.1231.18.

[24] P. D. Kusuma, and M. Kallista, "Stochastic Komodo Algorithm", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 4, pp. 156-166, 2022, doi: 10.22266/ijies2022.0831.15.

[25] P. D. Kusuma and M. Kallista, "Quad Tournament Optimizer: A Novel Metaheuristic Based on Tournament Among Four Strategies", *International Journal of Intelligent Engineering & Systems*, Vol. 16, No. 2, 2023, doi: 10.22266/ijies2023.0430.22.

[26] S. Nazir and M. Kaleem, "Federated Learning for Medical Image Analysis with Deep Neural Networks", *Diagnostics*, Vol. 13, No. 9, p. 1532, 2023.