IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 19 July 2023, accepted 24 August 2023, date of publication 29 August 2023, date of current version 7 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3309711

RESEARCH ARTICLE

Biomedical Image Analysis for Colon and Lung Cancer Detection Using Tuna Swarm Algorithm With Deep Learning Model

MARWA OBAYYA[1], MUNYA A. ARASI[2], NUHA ALRUWAIS[3], RAED ALSINI[4], ABDULLAH MOHAMED[5], AND ISHFAQ YASEEN[6]

[1]Department of Biomedical Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[2]Department of Computer Science, College of Science and Arts at Rijal Almaa, King Khalid University, Abha 62529, Saudi Arabia
[3]Department of Computer Science and Engineering, College of Applied Studies and Community Services, King Saud University, Riyadh 11495, Saudi Arabia
[4]Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[5]Research Centre, Future University in Egypt, New Cairo 11845, Egypt
[6]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam Bin Abdulaziz University, Al-Kharj 16278, Saudi Arabia

Corresponding authors: Munya A. Arasi (marasi@kku.edu.sa) and Ishfaq Yaseen (i.yaseen@psau.edu.sa)

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/61/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R203), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2023R608), King Saud University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

ABSTRACT The domain of Artificial Intelligence (AI) is made important strides recently, leading to developments in several domains comprising biomedical diagnostics and research. The procedure of AI-based systems in biomedical analytics takes opened up novel avenues for the progress of disease analysis, drug discovery, and treatment. Cancer is the second major reason of death worldwide; around one in every six people pass away suffering from it. Among several kinds of cancers, the colon and lung variations are the most frequent and deadliest ones. Initial detection of conditions on both fronts significantly reduces the probability of mortality. Deep learning (DL) and Machine learning (ML) systems are exploited to speed up such cancer detection, permitting researchers to analyze a huge count of patients in a lesser time count and at a minimal cost. This study develops a new Biomedical Image Analysis for Colon and Lung Cancer Detection using Tuna Swarm Algorithm with Deep Learning (BICLCD-TSADL) model. The presented BICLCD-TSADL technique examines the biomedical images for the identification and classification of colon and lung cancer. To accomplish this, the BICLCD-TSADL technique applies Gabor filtering (GF) to preprocess the input images. In addition, the BICLCD-TSADL technique employs a GhostNet feature extractor to create a collection of feature vectors. Moreover, AFAO was executed to adjust the hyperparameters of the GhostNet technique. Furthermore, the TSA with echo state network (ESN) classifier is utilized for detecting lung and colon cancer. To demonstrate the more incredible outcome of the BICLCD-TSADL system, an extensive experimental outcome is carried out. The comprehensive comparative analysis highlighted the greater efficiency of the BICLCD-TSADL technique with other approaches with maximum accuracy of 99.33%.

INDEX TERMS Cancer, biomedical imaging, artificial intelligence, colon cancer, tuna swarm algorithm, GhostNet.

I. INTRODUCTION

Cancer takes place because of the uncontrollable growth of abnormal cells from the body's tissues or organs. Cancer cells may follow in distinct tissues or organs of the body [1].

In 2020, Colon and Lung cancers were expected to rank in the top 3 most common types of cancer, as per a statistical study done in the US. Lung tumours may occur concurrently with colon cancer; almost 17% of cases of these two tumours arise at the same time [2]. Also, there is a risk of tumour cells spreading among 2 organs if an initial diagnosis is lacking. Smoking has a harmful effect on the growth of lung tumours,

The associate editor coordinating the review of this manuscript and approving it for publication was Humaira Nisar.

and it can be considered that an unaware diet leads to the growth of colon cancer [3]. Thus, lung tumour is the second cancer with colon cancer. In simple, a patient can affect by both colon and lung cancer simultaneously. Therefore, it is dynamic to examine both cancer types in patients and to detect them in advance [4].

The common symptoms are muscle pain, fatigue, cough, etc., follow by various kinds of syndromes [5]. Some of the frequently used radiographic imaging methods are ultrasound, mammography, histopathological imaging, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) for cancer detection [6]. Of these, histopathology images comprising phenotypic data are vital for the evaluation and diagnosis of cancer diseases. Manual examination of these medical images by professionals is a difficult and delicate task. Hence, it also necessitates a strong focus and time-consuming task [7]. Likewise, the recognition of cases is more difficult in the event of initial identification; the symptoms are difficult and vague to identify.

Likewise, identifying tumours consumes a lot of time and is dependent on different opinions of physicians in the initial stages [8]. A distinct domain of healthcare can sort out these difficulties. Artificial intelligence (AI) methods are utilized in the medical domain, like early identification of health disasters, biomedical image, and disease forecast [9]. DL methods are highly capable of analyzing data from anatomical representations, high-dimensional images, and videos. Likewise, DL methods extract hidden characteristics and features from medical images that are invisible to the naked eye for the initial identification of cancers and discrimination among their phases [10]. In this study, due to the same features of abnormal cells in the initial phases, numerous hybrid systems have been established with extraction features by mixed techniques.

Multi-level hyperparameters, also called hierarchical hyperparameters, can be employed for controlling the configuration and behavior of complex DL models. These hyperparameters are organized in a hierarchical manner, where higher-level hyperparameters affects the values or ranges of lower-level hyperparameters. It offers a structured way of managing and optimizing the settings for each component, making it easier to maintain, understand, and tune the overall model. When optimizing hyperparameters for a feature extractor and classifier, it is important to consider the hierarchical relationship between these components. The feature extractor is responsible for transforming the input data into a meaningful feature representation, while the classifier uses these features to make predictions. The learning rate is a critical hyperparameter in many feature extractor architectures, such as neural networks. It controls the step size at which the model updates its internal parameters during training. For classifiers like neural networks, the number of hidden units in the classifier's layers significantly impacts the model's capacity to learn complex patterns. Tuning this hyperparameter using random search or Bayesian optimization can help achieve the right balance between model complexity and generalization.

This study develops a new Biomedical Image Analysis for Colon and Lung Cancer Detection using Tuna Swarm Algorithm with Deep Learning (BICLCD-TSADL) model. The BICLCD-TSADL technique applies Gabor filtering (GF) to preprocess the input images. In addition, the BICLCD-TSADL technique employs a GhostNet feature extractor to make a collection of feature vectors. Moreover, AFAO can be executed to adjust the hyperparameters of the GhostNet technique. Furthermore, the TSA with echo state network (ESN) classifier is utilized for detecting lung and colon cancer. To demonstrate the greater efficiency of the BICLCD-TSADL algorithm, an extensive experimental outcome is carried out. In short, the key contributions of the study are listed as follows.

- Develop an automated colon and lung cancer detection model, comprising GF preprocessing, GhostNet feature extraction, AFAO-based hyperparameter tuning, ESN classification, and TSA-based parameter optimization. To the best of our knowledge, the BICLCD-TSADL technique never existed in the literature.
- Employ AFAO with the GhostNet model for the feature extraction process, which contributes to the accurate representation of colon and lung cancer image data.
- Employ the ESN model to effectively learn and classify cancerous and non-cancerous patterns in colon and lung cancer images, contributing to accurate cancer detection.
- Hyperparameter optimization process using AFAO and TSA helps to improve the cancer detection performance of the BICLCD-TSADL model.

II. RELATED WORKS

In [11], CNN methods were used to examine imaging data of colon cells. CNN with average pooling and max pooling layers and MobileNetV2 methods have been applied for colon cell image classification. Wahid et al. [12] introduced a CAD mechanism through the CNN to find colon, lung and cancer tissues. Here the author, namely ResNet18, ShuffleNet V2, and Google Net, in addition to one simple customized CNN model, uses three pre-trained CNN models. Kumar et al. [13] devise 4 CNN, namely, 3-block, baseline, 3-block, and 2-block CNNs with data augmentation, for classifying colon tissue histopathological images (HPI). To four CNN, the HPI was fed as input.

Garg and Garg [14] purposes of using and altering the present pre-training CNN-related technique for finding colon and lung cancer with the help of HPI, including better augmentation approaches. Here, 8 various pre-trained CNN methods like MobileNet, InceptionResNetV2, VGG16, InceptionV3, DenseNet169, ResNet50, NASNetMobile, and Xception were trained on LC25000 data. Adu et al. [15] present a novel DHS-CapsNet abbreviated as dual horizontal squash capsule network for categorizing the colon and lung cancers on HPI. DHS-CapsNet is a new horizontal squash (HSquash) function develop for encoder feature fusion (EFF). While a squash function, HSquash is modelled to make sure that vectors can be effectually squashed and makes sparsity for higher discriminative capsules for extracting significant data from images with different backgrounds.

In [16], chooses the gene elements with maximum correlation with tumour by Weighted Gene Co-expression Network Analysis (WGCNA) derived specific genes to differential expression outcomes utilizing the least absolute shrinkage and selection operator (Lasso) technique and effectuated survival study followed by combining the genes from the components with Lasso-extracting feature genes are joined to detect colon cancer vs healthy controls with the help of DT, RF, and SVM and colon cancer staging can be identified through differently expressed genes for all stages. Sakr et al. [17] present an innovative lightweight DL method that depends on a CNN for potential colon recognition. In this technique, the input HPI was normalised before presenting them as CNN approach and the next colon tumour identification was executed.

In [18], a classification method for colon cancer has been generated as 2 classes such as polyps and adenocarcinomas. This method was created utilizing the CNN approach with the MobileNet structure. Mohalder et al. [19] presented a DL approach to forecasting CRC tumours in HPI. CNN system employs to analyze difficult information. By CNN algorithm it is examined our difficult tumor images to recognize abnormal or suspicious tumor designs. The authors utilized the ReLU activation function in the hidden state and softmax function in the resultant layer. Bhattacharya et al. [20] examine a framework in which either DL or meta-heuristic techniques are utilized to predict colon or lung cancer, or both, in HPIs with near-perfect precision. In [21], a colon cancer classification method was established to support medical workers categorize 2 kinds of cancer colon adenocarcinomas and benign colonic tissues. The classification approach utilizes CNN with the structure of VGG19, VGG16, ResNet152, ResNet101, DenseNet201, MobileNetV2, and InceptionV3.

In spite of the ML and DL models that existed in earlier studies, it is still needed to enhance lung and colon cancer classification performance. Because of the continual deepening of the model, the number of parameters of DL models also increases quickly which results in model overfitting. At the same time, different hyperparameters have a significant impact on the efficiency of the CNN model. Particularly, hyperparameters such as epoch count, batch size, and learning rate selection are essential to attain effectual outcomes. Since the trial and error method for hyperparameter tuning is a tedious and erroneous process, metaheuristic algorithms can be applied. Therefore, in this work, we employ IAFO and TSA algorithms for the parameter selection of the GhostNet and ESN models respectively.

III. THE PROPOSED MODEL

In this work, we have introduced a novel BICLCD-TSADL system for automated colon and lung cancer detection. The purpose of the BICLCD-TSADL technique is to investigate biomedical images for the detection and classification of colon and lung cancer. To accomplish this, the BICLCD-TSADL technique comprises GF-based preprocessing, GhostNet feature extraction, AFAO, ESN classification, and TSA-based parameter tuning. Fig. 1 represents the workflow of the BICLCD-TSADL approach.

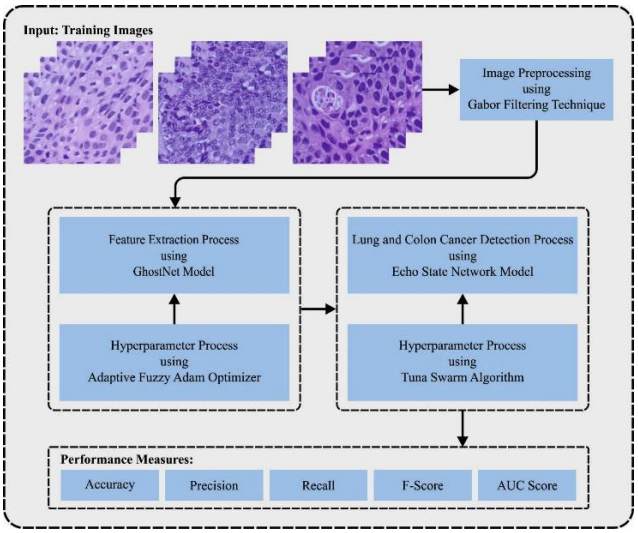

FIGURE 1. Workflow of the BICLCD-TSADL approach.

A. GF BASED PREPROCESSING

Primarily, the BICLCD-TSADL technique preprocesses the GF system to eliminate the noise. It is dependent upon the employ of Gabor functions, which are a family of difficult functions which are utilized for representing local structures from the signal/image. The GF is a kind of linear filtering which utilizes a Gabor function as its kernel. This filter can be planned to capture data on the frequency and orientation of local image features like textures and edges.

B. OPTIMAL GHOSTNET FEATURE EXTRACTOR

For the feature extraction process, the GhostNet model is used. A GhostNet-based DL method has been performed for extracting the feature from the image [22]. GhostNet recommended an innovative Ghost model that produces additional feature maps through a reasonable operation. This basic neural network unit makes image feature with fewer computations and inputs. There exist 2 features to this implementation of the module. To construct a feature map with other channels, GhostNet first implemented the typical convolution. Then, it implemented a simple operation to construct an additional mapping feature. Fig. 2 represents the structure of the GhostNet approach. Lastly, it concatenated numerous mapping features to make a novel output.

Ghost bottleneck is the essential part of GhostNet that comprised 2 ghost models. The procedure of making the M mapping feature from the ghost module was formulated by Eq. (1):

$$\gamma = X * f + b \tag{1}$$

In Eq. (1), w and h denote the width and height of the input, correspondingly; c shows the channel counts, b indicates the bias term, and $*$ denotes the convolution function; $f \in R^{c\times k\times k\times m}$, $X \in R^{h\times c\times w}$ is the convolutional kernel. $k*k$ indicates the convolution size f.

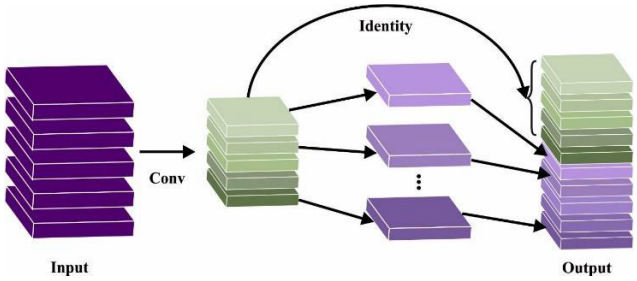

FIGURE 2. Structure of GhostNet model.

Firstly, the $W \times H \times C$ size-based input mapping feature was downscaled by the typical convolution. Then, a cheap linear operation is exploited on $W' \times H' \times C$ feature maps with a $k \times k$ small kernel convolutional function that directly produces a considerable quantity of ghost features. Eventually, the outcome of both procedures is merged to form an outcome mapping feature of dimensional $W' \times H' \times c$ that is similar to the original. The regular convolution and linear alteration used in the Ghost Module permit to brilliant preservation of novel features. Finally, the feature maps are converted as the last 1280-dimension feature vector utilizing the convolutional and global average pooling layers. The computation cost is lesser than classical convolution directly.

To adjust the hyperparameters of the GhostNet model, the AFAO is used. The AFAO algorithm is a modified variant of stochastic gradient descent (SGD) that is becoming ever more popular in DL applications [23]. AFAO algorithm is a stochastic optimizer technique used to improve the feature for the training procedure of the GhostNet model. AFAO algorithm increases the machine's efficacy and accelerates the learning process by letting the proposed method converge fast. The hyperparameters of the GhostNet technique were repeatedly updated dependent on the training dataset. The significant formula to determine the optimization algorithm using the AFAO algorithm is given below.

$$g_z = \nabla_\theta f_z\left(\theta_{z-1}\right) \tag{2}$$

$$h_z = \beta_1 . h_{z-1} + \left(1 - \beta_1\right) . g_z \tag{3}$$

$$v_z = \beta_2 . v_{z-1} + \left(1 - \beta_2\right) . g_z^2 \tag{4}$$

$$\hat{h}_z = \frac{h_z}{1 - \beta_1^z} \tag{5}$$

$$\hat{v}_z = \frac{v_z}{1 - \beta_2^z} \tag{6}$$

$$\theta_z = \theta_{z-1} - \frac{\alpha \cdot \hat{h}_z}{\sqrt{\hat{v}_z} + \varepsilon} \tag{7}$$

where $f(\theta)$ represents the stochastic objective function, α denotes the size of the step, β_1 and β_2 characterize the exponential decay rate, and z represents the time-step. θ_0 and θ_z correspondingly signify the first and last variable vectors, and h_z and v_z represent 1st and 2nd-moment vectors, correspondingly. $\hat{\theta}_{Hz}$ and $\hat{v}_z$ denote bias-corrected moment estimates, and g_z^2 shows the component-wise square.

The objective function, step size, and exponential decay rate are confirmed. The 1st and 2nd-moment vectors, time-step, and variable vectors are initialized. Next, every part of the loop was updated frequently; still, the variable θ_z converges. The biased 1st and 2nd moments are evaluated. Next, the bias-corrected 1st and 2nd-moments can be estimated. The predictive accuracy of GhostNet with the AFAO algorithm was assessed against the testing set.

C. CANCER DETECTION USING ESN MODEL

For the cancer detection and classification process, the ESN model is used. Statistical techniques can be comparatively easy to model but are insufficient forecasted power. The ML technique needs several computations and proceeds a long-time [24]. The combined learning system generates the model design further difficulty that could not conducive to updating the model. The ESN can be regarded as the most efficient approach to training RNNs that are the benefits of an easily trained procedure and short time consumption. It determines the $u(n) = [u_1(n) \cdots u_k(n)]^T$ as the input instance at time $n.y(n)$ refers to the resultant equivalent to (n). An input matrix W_{in} and the reservoir layer weighted matrix $\hat{W}$ are uniformly distributed among $[-1, 1]$ that maintained constant. In the trained procedure, with the instances input, the reservoir layer can be upgraded utilizing the subsequent equation.

$$x\left(n+1\right) = f\left(W_{in} u\left(n+1\right) + \hat{W} x\left(n\right)\right) \tag{8}$$

whereas $\hat{W}$ refers to the reservoir state connection matrix, f represents the activation function inside the reservoir state that is generally obtained strictly as hyperbolic tangent functions, and W_{in} implies the input connection matrix. Based on the above state of reservoir layers, the resultant ESN is computed utilizing the subsequent formula.

$$y\left(n+1\right) = f_{out}\left(W_{out} x\left(n+1\right)\right) \tag{9}$$

W_{out} stands for the resultant connection matrix, and f_{out} represents the resultant excitation function. In the trained method, the reservoir state can be gathered as a state matrix X. The last network resultant weighted W_{out} is computed utilizing the subsequent formula.

$$W_{out} = (X^T X)^{-1} X^T Y \tag{10}$$

whereas The superscript T implies the matrix transpose and -1 denotes the inverse of a matrix. X defines the matrix procedure of the input reservoir layer, and Y demonstrates the resultant matrix method. An optimum solution of the resultant matrix is initiated utilizing least squares or MSE alone.

D. PARAMETER TUNING USING TSA

Finally, the TSA is used to optimally select the parameters related to the ESN model. The TSA method begins the optimization process by randomly and uniformly generating the initial population in the search range [25]. The TSA can be mathematically formulated as follows:

$$X_1^{int} = rand \cdot (ub - lb) + lb, \quad i = 1, 2, \ldots, NP \tag{11}$$

In Eq. (11), X_i^{int} denotes the initial location of the $i - th$ individual, and Ub and lb show the upper as well as lower

boundaries of searching spaces correspondingly. NP indicates the amount of tuna population. Typically, this value affects the optimization rate of the TSA and is a uniformly distributed random vector. Spiral foraging is the primary foraging method of tuna schools that chases the prey by making tight spirals. Along with chasing prey, schools of tuna exchange data with all the others. Every tuna is sequenced and strongly related; hence adjacent tuna share data. The spiral foraging strategy can be mathematically expressed as:

$$X_1^{t+1} = \begin{cases} \alpha_1 \cdot \left(X_{best}^t + \beta \cdot X_{best}^t - X_i^t\right) + \alpha_2 \cdot X_i^t, \\ \qquad i = 1, \\ \alpha_1 \cdot \left(X_{best}^t + \beta \cdot X_{best}^t - X_i^t\right) + \alpha_2 \cdot X_{t-1}^i, \\ \qquad i = 2, 3, \ldots, NP, \end{cases} \tag{12}$$

$$\alpha_1 = \partial + (1 - a) \cdot \frac{t}{t_{\max}}, \tag{13}$$

$$\alpha_2 = (1 - a) - (1 - a) \cdot \frac{T}{t_{\max}}, \tag{14}$$

$$\beta = e^{bl} \cdot \cos(2\pi b) \tag{15}$$

$$l = e^{3\cos(t_{\max}+1/t)-1)\pi)} \tag{16}$$

where *best* denotes the existing better individual (food), X_i^{t+1} denotes the $i-th$ individuals of the $t+1$ iteration, a indicates the constant, defines to which extent the tuna follows the better individual and the prior individual at the initial phase, t and $t_{\max}$ indicates the existing and the maximum amount of iterations, α_1 and α_2 denotes the weight coefficient which controls the movement trends of the individual to the better and the prior individuals, and b is a uniformly distributed random integer within [0,1]. Once the best individual could not find food, blindly following the optimum individual foraging is not advantageous to group foraging. Hence, to assist every individual in having the best spatial search abilities, a reference point for spiral search should be given to produce a random coordinate in the search range, thereby allowing TSA to have the best global exploration abilities, and it can be a mathematical equation by Eq. (17):

$$X_1^{t=1'} = \begin{cases} \alpha_1 \cdot \left(X_{\text{rand}}^t + \beta \cdot \left|X_{\text{rand}}^t - X_i^t\right|\right) + \alpha_2 \cdot X_i^t, \\ \qquad i = 1, \\ \alpha_1 \cdot \left(X_{\text{rand}}^t + \beta \cdot \left|X_{\text{rand}}^t - X_i^t\right|\right) + \alpha_2 \cdot X_{i-1}^t, \\ \qquad i = 2, 3, \ldots, \text{N}P \end{cases} \tag{17}$$

where X_{rand}^t denotes the random reference point in the search ranges, TSA is typically explored extensively globally at an early stage and then slowly transitioned to accurate local exploitation. Thus, with the increasing amount of iterations, TSA slowly changes the reference point of spiral foraging from a random individual at the beginning to an optimum individual. The spiral foraging strategy can be mathematically expressed by Eq. (18), as shown at the bottom of the next page.

Tuna chooses to spiral foraging along with parabolic cooperative foraging. Tuna forms a parabola with the targeted food as a reference to the Z-point. Tuna finds the targeted food by searching around the parabola. Both foraging approaches of tuna are performed based on the probability allocation, and if the selection probability for the two foraging techniques is 1/2, then they are concurrently performed, and it can be mathematically formulated as follows.

$$X_1^{t+1} = \begin{cases} X_{best}^t + (rand \cdot X_{best}^t - X_1^t) + TF \cdot^2 \cdot X_{best}^t - X_i^t), \\ \qquad if rand < 0.5, \\ TF \cdot p^2 \cdot X_1^t, \\ \qquad if rsnd \geq 0.5, \end{cases} \tag{19}$$

$$p = \left(1 - \frac{t}{t_{\max}}\right)^{(t/t_{\max})}, \tag{20}$$

where TF denotes the random integer within $[1, -1]$.

The fitness chosen is a vital feature in the TSA approach. An encoder solution can be utilised to grow the ability of candidate solutions. Here, the accuracy value is the critical factor utilized to propose a fitness function.

$$Fitness = max\,(P) \tag{21}$$

$$P = \frac{TP}{TP + FP} \tag{22}$$

whereas FP signifies the false positive and TP denotes the true positive value.

TABLE 1. Details of database.

Class Name	Description	No. of Instances
LunAd	Lung Adenocarcinoma	5000
LunBe	Lung Benign Tissue	5000
LunSC	Lung Squamous Cell Carcinoma	5000
ColAd	Colon Adenocarcinoma	5000
ColBe	Colon Benign Tissue	5000
Total Number of Samples		25000

IV. EXPERIMENTAL VALIDATION

The proposed model is simulated using Python 3.6.5 tool on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

The experimental validation of the BICLCD-TSADL algorithm was tested on the LC25000 database [26] comprising five classes with 5000 samples under each class, as depicted in Table 1. Fig. 3 depicts the sample of Lung and colon images. The 5 classes are colon adenocarcinomas, benign colonic tissues, lung adenocarcinomas, lung squamous cell carcinomas and benign lung tissues. For experimental validation, 70:30 of training/testing data.

Fig. 4 establishes the classifier outcomes of the BICLCD-TSADL system under the test database. Figs. 4a-4b represents the confusion matrix offered by the BICLCD-TSADL model on 70:30 of TRP/TSP. The figure denoted that the BICLCD-TSADL approach has identified and classified all 5 class labels accurately. Likewise, Fig. 4c represents the PR investigation of the BICLCD-TSADL

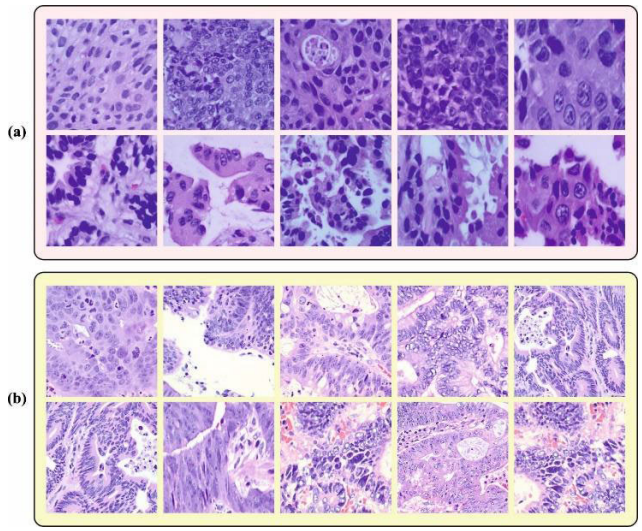

FIGURE 3. a) Lung Images b) Colon Images.

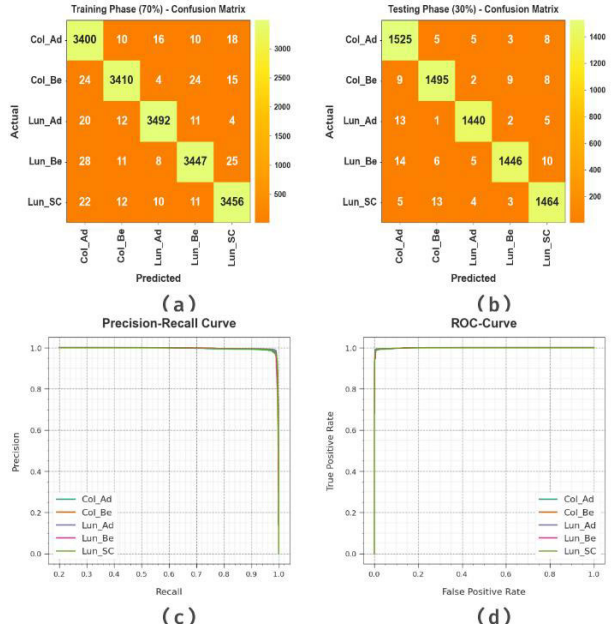

FIGURE 4. Classifier outcome of (a-b) Confusion matrices, (c) PR curve, (d) ROC curve.

approach. The figures reported that the BICLCD-TSADL approach has obtained higher PR performance under 5 classes. Eventually, Fig. 4d illustrates the ROC investigation of the BICLCD-TSADL model. The figure stated that the BICLCD-TSADL approach has led to able outcomes with maximal ROC values under 5 classes.

In Table 2 and Fig. 5, the overall outcomes of the BICLCD-TSADL system are examined on 70% of TRP.

TABLE 2. Classifier outcome of BICLCD-TSADL system on 70% of TRP.

Training Phase (70%)					
Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}	AUC_{Score}
ColAd	99.15	97.31	98.44	97.87	98.88
ColBe	99.36	98.70	98.07	98.38	98.88
LunAd	99.51	98.92	98.67	98.80	99.20
LunBe	99.27	98.40	97.95	98.18	98.78
LunSC	99.33	98.24	98.43	98.34	99.00
Average	99.33	98.31	98.31	98.31	98.95

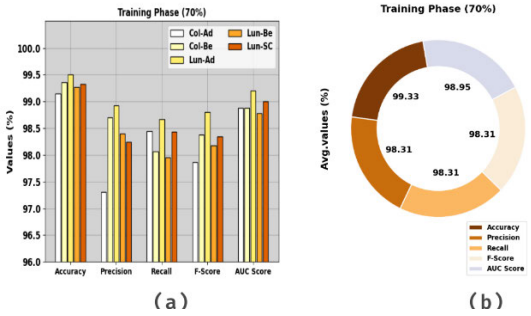

FIGURE 5. (a) Classification outcome on 70% of TRP and (b) Average outcome on 70% of TRP.

The results indicate that the BICLCD-TSADL technique reaches proficient results in each class. For instance, on ColAd, the BICLCD-TSADL technique gains $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.15%, 97.31%, 98.44%, 97.87%, and 98.88%, respectively. Followed by, on LunAd, the BICLCD-TSADL method attains $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.51%, 98.92%, 98.67%, 98.80%, and 99.20% correspondingly. Concurrently, on LunSC, the BICLCD-TSADL system reaches $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.33%, 98.24%, 98.43%, 98.34%, and 99% correspondingly. Finally, the BICLCD-TSADL technique reaches average $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.33%, 98.31%, 98.31%, 98.31%, and 98.95%, respectively.

In Table 3 and Fig. 6, the overall outcomes of the BICLCD-TSADL system are examined on 30% of TSP. The outcomes stated that the BICLCD-TSADL methodology reaches proficient outcomes under each class. For instance, on ColAd, the BICLCD-TSADL algorithm gains $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.17%, 97.38%, 98.64%, 98.01%, and 98.98% correspondingly. Afterwards, on LunAd, the BICLCD-TSADL technique reaches $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.51%, 98.90%, 98.56%, 98.73%, and 99.15% correspondingly. Simultaneously, on LunSC, the BICLCD-TSADL algorithm reaches $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.25%, 97.93%, 98.32%, 98.12%, and 98.90% correspondingly. At last, the BICLCD-TSADL approach obtains average

$$X_1^{t+1\prime} = \begin{cases} \begin{cases} \alpha_1 \cdot \left(X_{best}^t + \beta \cdot X_{best}^t - X_i^t\right) + \alpha_2 \cdot X_i^t, & i = 1, \\ \alpha_1 \cdot \left(X_{best}^t + \beta \cdot X_{best}^t - X_i^t\right) + \alpha_2 \cdot X_{i-1}^t, & i = 2, 3, \ldots, \mathrm{N}P, \end{cases} & if\ rand \geq \dfrac{t}{t_{\max}} \\ \begin{cases} \alpha_1 \cdot \left(X_{rand}^t + \beta \cdot X_{rand}^t - X_i^t\right) + \alpha_2 \cdot X_i^t, & i = 1, \\ \alpha_1 \cdot \left(X_{rand}^t + \beta \cdot X_{rand}^t - X_i^t\right) + \alpha_2 \cdot X_{i-1}^t, & i = 2, 3, \ldots, \mathrm{N}P, \end{cases} & if\ rand < \dfrac{t}{t_{\max}} \end{cases} \tag{18}$$

TABLE 3. Classifier outcome of BICLCD-TSADL algorithm on 30% of TSP.

Testing Phase (30%)					
Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}	AUC_{Score}
ColAd	99.17	97.38	98.64	98.01	98.98
ColBe	99.29	98.36	98.16	98.26	98.87
LunAd	99.51	98.90	98.56	98.73	99.15
LunBe	99.31	98.84	97.64	98.23	98.68
LunSC	99.25	97.93	98.32	98.12	98.90
Average	99.31	98.28	98.26	98.27	98.92

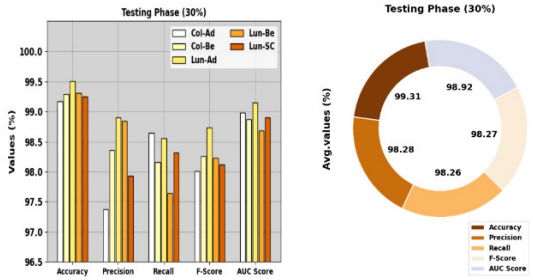

FIGURE 6. (a) Classification outcome on 30% of TSP and (b) Average outcome on 30% of TSP.

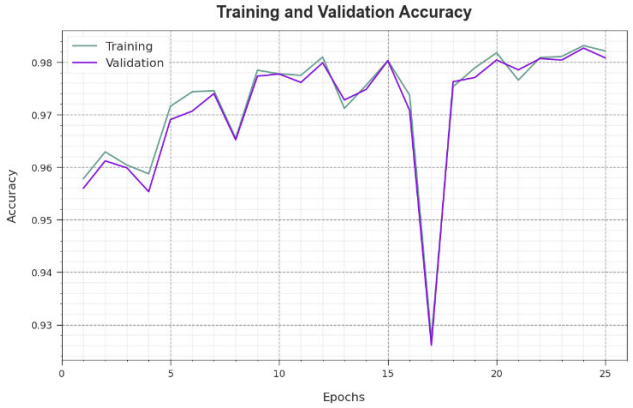

FIGURE 7. Accuracy curve of the BICLCD-TSADL approach.

$accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 99.31%, 98.28%, 98.26%, 98.27%, and 98.92% correspondingly.

Fig. 7 examines the accuracy of the BICLCD-TSADL technique under the training and validation process on the test dataset. The figure notifies that the BICLCD-TSADL method reaches maximal accuracy values over increasing epochs. In addition, the increasing validation accuracy over training accuracy exhibits that the BICLCD-TSADL technique learns efficiently on the test dataset.

The loss investigation of the BICLCD-TSADL algorithm at the time of training and validation is depicted on the test database in Fig. 8. The outcomes indicate that the BICLCD-TSADL system reaches closer values of training and validation loss. It can be clear that the BICLCD-TSADL algorithm learns efficiently on the test dataset.

A comparative classification outcome of the BICLCD-TSADL technique is made with recent approaches in Table 4 and Fig. 9 [27]. The outcomes depicted the ineffectual outcomes of the mSRC model, whereas the ResNet50, CNN, and DL models obtain slightly improved performance. At the same time, the MPADL-LC3, Faster

FIGURE 8. Loss curve of the BICLCD-TSADL approach.

TABLE 4. Comparative outcome of BICLCD-TSADL approach with recent algorithms.

Methods	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}
BICLCD-TSADL	99.33	98.31	98.31	98.31
MPADL-LC3	98.81	97.92	97.87	97.91
mSRC	88.25	85.20	91.74	86.66
Faster RCNN	98.70	96.44	97.84	97.28
DAELGNN	98.56	98.08	96.56	96.65
RESNET50	93.96	96.28	97.48	96.95
CNN Model	97.05	96.92	97.41	97.74
DL Model	96.29	96.96	96.14	98.03

RCNN, and DAELGNN models accomplish closer outcomes. However, the BICLCD-TSADL technique gains effectual outcomes with a maximum $accu_y$ of 99.33%, $prec_n$ of 98.31%, $reca_l$ of 98.31%, and F_{score} of 98.31%. These results show the improvement of the BICLCD-TSADL system in medical image analysis. The proposed IAFO and TSA models choose the optimal values for the hyperparameters of a given GhostNet and ESN models. Hyperparameters are settings that are not learned during training but must be set before training. They can have a significant impact on the performance of the model, and selecting the optimal values can lead to better accuracy. By exploiting IAFO and TSA algorithms, the BICLCD-TSADL model can achieve even better results by focusing on the selection of the optimal settings for the algorithm. These results ensured the improved performance of the BICLCD-TSADL technique over other existing techniques.

V. CONCLUSION

In this article, we have developed a novel BICLCD-TSADL algorithm for automated colon and lung cancer detection. The purpose of the BICLCD-TSADL technique is to investigate biomedical images for the identification and classification of colon and lung cancer. To accomplish this, the BICLCD-TSADL technique comprises GF-based preprocessing, GhostNet feature extraction, AFAO, ESN classification, and TSA-based parameter tuning. The design of AFAO and TSA for parameter tuning techniques helps to accomplish enhanced classification performance. To demonstrate the greater performance of the BICLCD-TSADL algorithm,

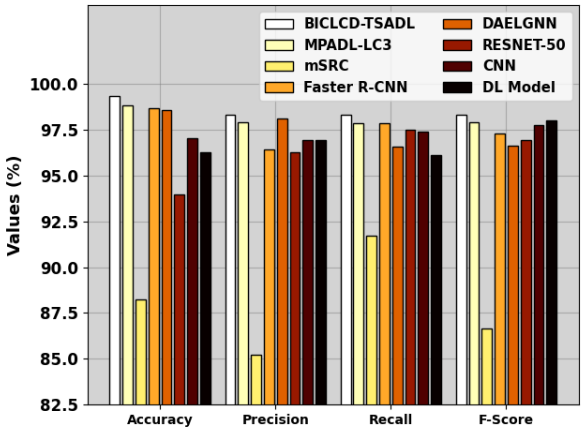

FIGURE 9. Comparative outcome of BICLCD-TSADL approach with recent algorithms.

an extensive experimental outcome is carried out. A detailed comparative outcome highlighted the greater outcome of the BICLCD-TSADL technique with state-of-the-art approaches with maximum accuracy of 99.33%. In future, the performance of the proposed model can be improved by feature fusion and ensemble learning process. In addition, computation complexity of the proposed model can be investigated in future.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/61/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R203), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2023R608), King Saud University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444).

REFERENCES

[1] M. S. N. Raju and B. S. Rao, "Lung and colon cancer classification using hybrid principle component analysis network-extreme learning machine," *Concurrency Comput., Pract. Exper.*, vol. 35, no. 1, Jan. 2023.
[2] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, K. F. Hasan, and M. A. Moni, "Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning," *Expert Syst. Appl.*, vol. 205, Nov. 2022, Art. no. 117695.
[3] O. Attallah, M. F. Aslan, and K. Sabanci, "A framework for lung and colon cancer diagnosis via lightweight deep learning models and transformation methods," *Diagnostics*, vol. 12, no. 12, p. 2926, Nov. 2022.
[4] M. A. Fahami, M. Roshanzamir, N. H. Izadi, V. Keyvani, and R. Alizadehsani, "Detection of effective genes in colon cancer: A machine learning approach," *Informat. Med. Unlocked*, vol. 24, 2021, Art. no. 100605.
[5] H. C. Reis and V. Turk, "Transfer learning approach and nucleus segmentation with MedCLNet colon cancer database," *J. Digit. Imag.*, vol. 36, no. 1, pp. 306–325, Sep. 2022.
[6] N. Kumar, M. Sharma, V. P. Singh, C. Madan, and S. Mehandia, "An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103596.
[7] 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, Jan. 2021.
[8] A. H. Chehade, N. Abdallah, J.-M. Marion, M. Oueidat, and P. Chauvet, "Lung and colon cancer classification using medical imaging: A feature engineering approach," *Phys. Eng. Sci. Med.*, vol. 45, no. 3, pp. 729–746, Sep. 2022.
[9] M. Toğaçar, "Disease type detection in lung and colon cancer images using the complement approach of inefficient sets," *Comput. Biol. Med.*, vol. 137, Oct. 2021, Art. no. 104827.
[10] I. Pacal, D. Karaboga, A. Basturk, B. Akay, and U. Nalbantoglu, "A comprehensive review of deep learning in colon cancer," *Comput. Biol. Med.*, vol. 126, Nov. 2020, Art. no. 104003.
[11] Z. Tasnim, S. Chakraborty, F. M. J. M. Shamrat, A. N. Chowdhury, H. A. Nuha, A. Karim, S. B. Zahir, and M. M. Billah, "Deep learning predictive model for colon cancer patient using CNN-based classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 8, pp. 687–696, 2021.
[12] R. R. Wahid, C. Nisa', R. P. Amaliyah, and E. Y. Puspaningrum, "Lung and colon cancer detection with convolutional neural networks on histopathological images," *AIP Conf. Proc.*, vol. 2654, Feb. 2023, Art. no. 020020.
[13] A. Kumar, A. Vishwakarma, V. Bajaj, A. Sharma, and C. Thakur, "Colon cancer classification of histopathological images using data augmentation," in *Proc. Int. Conf. Control, Autom., Power Signal Process. (CAPS)*, Dec. 2021, pp. 1–5.
[14] 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 *Proc. 3rd Artif. Intell. Cloud Comput. Conf.*, Dec. 2020, pp. 38–45.
[15] K. Adu, Y. Yu, J. Cai, K. Owusu-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," *Int. J. Imag. Syst. Technol.*, vol. 31, no. 4, pp. 2075–2092, Dec. 2021.
[16] Y. Su, X. Tian, R. Gao, W. Guo, C. Chen, C. Chen, D. Jia, H. Li, and X. Lv, "Colon cancer diagnosis and staging classification based on machine learning and bioinformatics analysis," *Comput. Biol. Med.*, vol. 145, Jun. 2022, Art. no. 105409.
[17] A. S. Sakr, N. F. Soliman, M. S. Al-Gaashani, P. Pławiak, A. A. Ateya, and M. Hammad, "An efficient deep learning approach for colon cancer detection," *Appl. Sci.*, vol. 12, no. 17, p. 8450, Aug. 2022.
[18] Y. G. C. Kelana, S. Rizal, and S. Saidah, "Classification of histopathological images of colon cancer using convolutional neural network method," in *Proc. Int. Conf. Comput. Sci., Inf. Technol. Eng. (ICCoSITE)*, Feb. 2023, pp. 821–826.
[19] R. D. Mohalder, F. Bin Ali, L. Paul, and K. Hasan Talukder, "Deep learning-based colon cancer tumor prediction using histopathological images," in *Proc. 25th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2022, pp. 629–634.
[20] A. Bhattacharya, B. Saha, S. Chattopadhyay, and R. Sarkar, "Deep feature selection using adaptive β-Hill climbing aided whale optimization algorithm for lung and colon cancer detection," *Biomed. Signal Process. Control*, vol. 83, May 2023, Art. no. 104692.
[21] I. D. Irawati, I. Andrea Larasaty, and S. Hadiyoso, "Comparison of convolution neural network architecture for colon cancer classification," *Int. J. Online Biomed. Eng. (iJOE)*, vol. 18, no. 3, pp. 164–172, Mar. 2022.
[22] S. S. R. Bairaboina and S. R. Battula, "Ghost-ResNeXt: An effective deep learning based on mature and immature WBC classification," *Appl. Sci.*, vol. 13, no. 6, p. 4054, Mar. 2023.
[23] S. Sorguli and H. Rjoub, "A novel energy accounting model using fuzzy restricted Boltzmann machine—Recurrent neural network," *Energies*, vol. 16, no. 6, p. 2844, Mar. 2023.
[24] Y. Li and Y. Li, "Predicting chaotic time series and replicating chaotic attractors based on two novel echo state network models," *Neurocomputing*, vol. 491, pp. 321–332, Jun. 2022.
[25] S.-M. Guo, J.-K. Guo, Y.-G. Gao, P.-Y. Guo, F.-J.-A. Huang, S.-C. Wang, Z.-C. Lou, and X. Zhang, "Research on engine speed control based on tuna swarm optimization," *J. Eng. Res. Rep.*, pp. 272–280, Dec. 2022.
[26] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides. *Lung and Colon Cancer Histopathological Image Dataset (LC25000)*. Accessed: Mar. 14, 2023. [Online]. Available: https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancerhistopathological-images?resource=download
[27] H. A. Mengash, M. Alamgeer, M. Maashi, M. Othman, M. A. Hamza, S. S. Ibrahim, A. S. Zamani, and I. Yaseen, "Leveraging marine predators algorithm with deep learning for lung and colon cancer diagnosis," *Cancers*, vol. 15, no. 5, p. 1591, 2023, doi: doi.org/10.3390/cancers15051591.