Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 15s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

A Hybrid Graph Convolutional Bidirectional Lstm Model Based Classifier And Segmentation Using Modified Birch Algorithm For Early Detection Of Lung Cancer

Ms. M. Vaishnava Priya[1], Dr. R. Tamilselvi[2]

[1]Research Scholar, Department Of Artificial Intelligence And Data Science, VET Institute Of Arts And Science (Co- Ed) College, Thindal, Erode.

Email ID: mvaishnavapriya.tup@gmail.com

[2]Assistant professor & Head, Department of Artificial intelligence and Data Science, VET institute of Arts and science co education College, Thindal, Erode

Cite this paper as: Ms. M. Vaishnava Priya, Dr. R. Tamilselvi, (2025) A Hybrid Graph Convolutional Bidirectional Lstm Model Based Classifier And Segmentation Using Modified Birch Algorithm For Early Detection Of Lung Cancer. *Journal of Neonatal Surgery*, 14 (15s), 1607-1619.

ABSTRACT

Lung cancer (LC) has an unbelievable annual incidence of over five million deaths, making it one of the leading causes of mortality worldwide for both men and women. Detecting malignant lung nodules (LN) on the provided input lung image and classifying the LC along with its severity are the primary objectives of this study. Using cutting-edge Hybrid Deep learning (HDL) techniques, this study detects the malignant LNs. This research proposes an intelligent framework for lung cancer detection in PET images, integrating advanced techniques for noise removal, segmentation, classification, and hyperparameter tuning. First, aAdaptive Weiner Filter (AWF) is applied to PET images to effectively remove noise and enhance image clarity, ensuring more accurate analysis. Subsequently, the Modified BIRCH algorithm is utilized for segmentation, enabling the delineation of regions of interest within the images. For LC classification, a Hybrid Graph Convolutional Bidirectional Long Short-Term Memory (Bi-LSTM)framework is developed. This novel architecture combines the effectiveness of Graph Convolutional Networks (GCN) and Bi-LSTM units, facilitating comprehensive analysis of spatial and temporal features in PET images. To enhance model performance, Hyperparameter (HP) tuning is performed using the Enhanced HBO (Honey bee optimization) Algorithm, optimizing model parameters for improved accuracy and robustness. The suggested framework is evaluated using the PET Dataset, demonstrating its effectiveness in accurately detecting LC in PET images.

Keywords: *Lung Cancer (LC),Adaptive Weiner Filter (AWF), Modified BIRCH algorithm, hybrid GC (Graph Convolutional)Bi-LSTM (Bidirectional LSTM) model, Enhanced HBO (Honey Bee Optimization) Algorithm.*

1. INTRODUCTION

One of the most prevalent malignancies, LC causes around 225,000 cases, 150,000 deaths, and $12 billion in medical expenses per year in the United States [1]. LC is among the worst malignancies; in the United States, just 17% of patients receive a diagnosis and live for 5 years afterward; in developing nations, the survival rate is even lower. The extent to which a cancer has metastasized is indicated by its stage [2]. Localized malignancies of the lungs are referred to as stages 1 and 2, whereas tumors that spread to other organs are referred to as later stages. Biopsies and imaging techniques like CT scans are currently used for diagnosis [3]. The chances of survival are greatly increased by the detection of LC; at present, there are less symptoms in the early stages of the disease, it is also more challenging to diagnose it.

Figure 1: 2D CT scan slice containing a small (5mm) early stage LC nodule

To identify any signs of LC in patient CT scans of the lungs with and without early stage LC, one must solve Binary Classification (BC) problem [4]. The objective is to create an accurate classifier by utilizing techniques from DL and CV (Computer Vision), namely 2D and 3D CNN (Convolutional Neural Networks). More early detection and better survival could be made possible with a more rapid and cost-effective LC classifier. A Computer-Aided Diagnosis (CAD) system that can determine whether a patient has LC based on chest CT scans as input is the objective [5]-[6]. Fig. 1 represents an instance of an early stage LC nodule inside a 2D CT scan segment. Moreover, a CT scan is noisy due to surrounding tissues, bone, and air; therefore, preprocessing of this noise is necessary before the CAD system can search efficiently. One of the key issues in DM (Data Mining) and ML challenges is to identify an appropriate representation of the data from all features [7]. Regression or classification problems may not always benefit from making use of all original features [8]. In order to optimize the value of a criterion function over all subsets of size M, FS involves choosing a subset of M features from a collection of N features, where M<N.Determining the optimal subset while maintaining classification accuracy is the primary objective of FS techniques, which ideally search among feature subsets. The tool utilized for classifying patterns or assess the extent to which all subset predicts the class output or pattern is called a classification function.

Human cognitive processes are to be emulated by Artificial Intelligence (AI). Thanks to the rapid growth of analytics tools and the growing availability of healthcare data, it is revolutionizing the field of healthcare. CNNs are composed of neurons that train to maximize themselves, just like regular ANNs [10]. The foundation of ANNs is the concept that each neuron will keep receiving input and performing an action (e.g. producing a scalar product that complies with a Non-Linear Function (NLF)). Managing the data directly makes classification more challenging [11]. Introducing a smart framework that can execute Medical Imaging (MI) more accurately is the primary objective of the study that has been suggested. The primary objectives of this study is to classify the LC and its severity as well as identify any cancerous LN from the provided input lung image. Using cutting-edge HDL techniques, this study detects the malignant LN. This study integrates cutting-edge methods for HP tuning, segmentation, (NR) Noise Reduction, and classification to present a sophisticated framework for the diagnosis of LC in PET scans.

The remaining research has been arranged as follows: Section 2 analyzes some of the most recent methods for classifying and detecting LC utilizing DLT. The suggested methodology's approach is presented in section 3. The results and the discussion are given in section 4. Future work and the conclusion are covered in section 5.

2. LITERATURE REVIEW

This section analyzes some of the most recent methods for classifying LC utilizing DL algorithms.

Histogram of Oriented Gradients (HoG), Wavelet Transform (WT)-based features, Local Binary Pattern (LBP), Scale Invariant Feature Transform (SIFT), and Zernike Moment are a few of the FE approaches that Asuntha et al. [12] developed. The optimal feature is chosen using the Fuzzy Particle Swarm Optimization (FPSO) technique following the extraction of textural, geometric, volumetric, and features with intensity. At last, DL is employed to classify these features. An Optimal Deep NN (ODNN) and Linear Discriminate Analysis (LDA) was suggested by Lakshmanaprabu et al. [13]. To classify LN as benign or malignant, deep features from CT lung images are retrieved, and the dimensionality of the features is subsequently reduced via LDR. To determine the classification of LC, the ODNN is applied to CT scans and then optimized with the Modified Gravitational Search Algorithm (MGSA). Based on comparative analysis, the suggested classifier achieves 96.2% sensitivity, 94.2% specificity, and 94.56% accuracy.DL techniques and structures were presented by Riquelme et al. [14] as potential CAD systems for the detection of LC. They fall into 2 categories: (1) False positive (FP) reduction frameworks, that classify a collection of given candidate nodules into benign or malignant tumors; and (2) Nodule detection systems identify candidate nodules from the original CT scan. Utilizing CT scans from SPIE-AAPM-LungX data, Cengil et al. [15] demonstrated a tensorFlow and 3D CNN for the classification of LN. In recent years, DL has become a popular option for the classification process. It is specifically utilized in the DL frameworks' TensorFlow and 3D CNN structure implementation.

In order to reliably and efficiently classify whole-slide images from The Cancer Genome Atlas into LUAD, LUSC, or normal lung tissue, Coudray et al. [16] trained a Deep CNN (DCNN) (Inception v3) on the images. Researchers found that the technique's Average Area Under the Curve (AUC) of 0.97 is equivalent to that of clinicians. Utilizing 3D cubes taken from the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), LUng Nodule Analysis 2016 (LUNA16), and Kaggle Data Science Bowl 2017 datasets, Tekade et al. [17] developed a 3D Multipath (MP) VGG-like network that is tested. The final outcomes are a combination of 3D MP VGG-like network and U-Net prediction. Considering 95.60% accuracy and 0.387732 logloss, the framework is employed to classify the LN and determine the malignancy level.Utilizing a dataset from the 2017 Kaggle Data Science Bowl, Alakwaa et al. [18] created a Computer-Aided Diagnosis (CAD) method for classifying CT scans with unidentified nodules as LC. The first method, which aimed to classify the segmented CT images directly into 3D CNNs, was ineffective. zThis was done because the U-Net nodule detection produced a lot of FPs. A 86.6% test set accuracy was achieved by the 3D CNNs.

DeepLung, a fully automated approach for lung CT cancer diagnosis, was introduced by Zhu et al. [19]. Based on image technique, the nodule classification subnetwork outperformed medical professionals and SOTA approaches in the validation

process conducted on a publicly available dataset from LIDC-IDRI.The challenge of classifying lung nodule malignancy was investigated by Lyu et al. [20] using a Multi-Level CNN (ML-CNN). 3 CNNs make up ML-CNN, which is employed to extract (MS) Multi-Scale features from CT scans of lung nodules. Using only one extra hand-crafted preprocessing procedure, the experimental results demonstrate that the ML-CNN achieves 84.81% accuracy. In ternary classification, it is also shown that the model performs effectively.Nodule segmentation and FE in CAD systems can be minimized by utilizing CNNs to accomplish an End-to-End (E2E) classification from raw 3D nodule CT patches, as developed by Zhang et al. [21]. In particular, the most advanced CNN models like VGG16, VGG19, ResNet50, DenseNet121, MobileNet, Xception, NASNetMobile, and NASNetLarge are converted into 3-D CNN models and tested on the LIDC-IDRI, a public CT lung dataset. Experimental outcomes indicate that DenseNet121 and Xception perform superior for LN diagnosis byaccuracy, sensitivity, specificity, precision, and AUC. A Deep Transfer Learning (DTL) method for classifying LNs as malignant has been offered by Da Nóbrega et al. [22]. Accuracy (ACC), Area Under the Curve (AUC), True Positive Rate (TPR), Precision (PPV), and F1-Score were also computed as evaluation metrics to compare the classifiers' performance with each other and with other ones in the existing research. Cancer Cell Detection utilizing Hybrid NN (CCDCHNN) is a unique approach for an early and accurate detection put forth by ShaliniWankhade et al. [23]. DNNs are employed for FE from the CT scan images. To protect the patient from this deadly disease, early CCD depends critically on the accuracy of FE. The outcomes validate the applicability of the suggested HDL strategy for the early diagnosis of LC when they are assessed using accepted statistical methods.

3. PROPOSED METHODOLOGY

This research work proposes an intelligent framework for lung cancer detection in PET images, integrating advanced techniques for noise removal, segmentation, classification, and HP tuning. First, anAdaptive Weiner Filter (AWF) is applied to PET images to effectively remove noise and enhance image clarity, ensuring more accurate analysis. Subsequently, the Modified BIRCH algorithm is utilized for segmentation, enabling the delineation of regions of interest within the images. For LC classification, a hybrid Graph Convolutional Bidirectional LSTM model is developed. This novel architecture combines the robustnesss of GCN and Bi-LSTM units, facilitating comprehensive analysis of spatial and temporal features in PET images. To enhance model performance, hyperparameter tuning is performed using the Enhanced HBO Algorithm, optimizing model parameters for improved accuracy and robustness. The figure 2. shows the overall process of the suggestedprocedure.

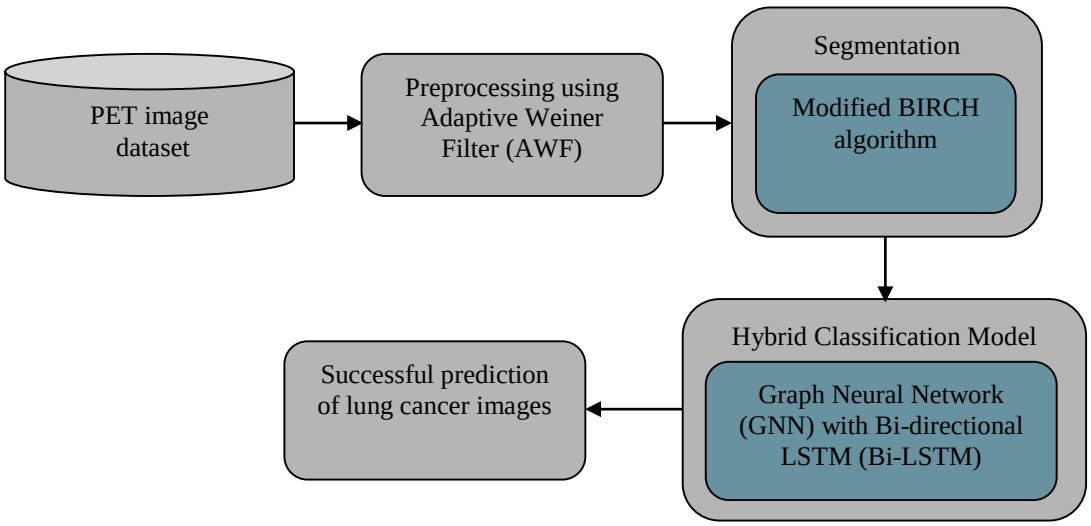

Figure 2. The overall process of the suggestedprocedure

3.1. Preprocessing using Adaptive Weiner Filter (HAF)

By adjusting the filter's output based on the image's local variance, AWF produces better results. Minimizing the MSE (Mean Square Error) among the original image and the restored image is the primary Objective [24]. However, the center sample in the moving windows should be appropriately employed in rough areas and neglected in smooth regions to decrease anomalies that are subjectively difficult.

$$y(i,j) = x(i,j) + n(i,j) \qquad (1)$$

Here,the additive noiseisrepresented as n(i,j), the imae with noise-free is symbolizedasx(i,j), and the noisy measurement can be denoted as y(i,j). Eliminating noise or "denoise",i.e., y(i,j) is the objective. The variance and mean of each pixel in the image are compared for each (WS) Window Size, such as $(3+2i)^2$, i = 0, 1, 2, 3, and so on. The window with the lowest

average value is then selected as the processing window that will be utilized in the final stage. Adaptively choosing the filter template based on several regions is possible. To enhance efficiency and preserve edges and texture, the tiny Window Filter (WF) is utilized in the detail section and the wide WF in the smooth area. To deal with the pixels and obtain the desired output, the formula below is utilized.

$$r(i,j) = \mu + (1 - q + \Delta) * (s(i,j) - \mu) \qquad (2)$$

$$q = \frac{\sigma_{avg}}{\sigma_{var}+1} \qquad (3)$$

$$\Delta = \frac{\sigma_{var}}{\sigma_{avg}+\sigma_{max}+1} \qquad (4)$$

The source pixel is denoted by $s(i, j)$, the output pixel is denoted by $s(i, j)$, the variance of the current pixel is denoted by σ_{var}, and the maximum variance of all the pixels in the image is represented by σ_{max}. The outcome of the inner window can be applied to a bigger outer window in the computer, which results in a significant improvement.

3.2. Segmentation using Modified BIRCH procedure

Combination of HCA (Hierarchical Clustering Algorithm) is the BIRCH technique. For the generic cluster description, it utilizes the concepts of the Clustering Features (CF) and CF Tree (CFT)concepts. The CFT facilitates the clustering of important data, and its storage space in memory is significantly less than that of the meta-data collection (MDC). This enhances the technique's scalability and performance when clustering big datasets. And it's effective for solving clustering problems involving discrete and continuous features. The dataset's objects are clustered into a CF form for subclustering [25]. Then, using the conventional HCA, k-groups formed from this CF clustering. The data triple CF is made up of CF = (N, LS, SS), where N is the number of data points, LS is the value of the attribute X added, and SS is the value of the squared features. Theorem: In the event that two CFs merge (N1 + N2, 1 + 2, SS1 + SS2) equals CF12. BIRCH computes the CF sub cluster analysis separately. Vector CF is employed to represent clusters, and CF Vector is kept in storage. This CF value is sufficient to compute Sub-Cluster (SC) parameters like centroid, radius, and diameter. It serves as an effective storage solution through compressing SC data relatively maintaining every point. Using the distance formula, one could locate a cluster feature that is acceptable for merging. So, D2 is employed.

$$D2 = \frac{\sqrt{(N_1 SS_2)+(N_2 SS_1)-2LS_1 LS_2}}{N_1 N_2} \qquad (5)$$

Then, applying the following calculation to get the radius for CF-leaf:

$$R = \frac{\sqrt{SS-(LS)^2/n}}{n} \qquad (6)$$

Less resilience to outliers, which could emerge as new clusters or even force other clusters to merge, is a primary drawback of BIRCH and most other HCA.

Modified BIRCH algorithm

Some researchers reported that BIRCH could be modified at the (TV) Threshold Value to increase cluster quality, based on earlier studies. The CF-entry threshold value will change for this investigation. The author attempts to change the TV to dynamic while analyzing this research in an effort to enhance the cluster's quality. CF = (N, LS, and SS) is the CF formula employed in conventional BIRCH. The CF-Leaf value will be altered for this investigation. CF-Leaf (modif) = (N, LS, SS, T) is the CF-leaf modif formula that needs to be applied. The most recent changes from the threshold are stored by adding the T parameter to CF-Leaf (modif).The CF-Node continues to employ the formulationCF = (N, LS, and SS), whereas the T Additional parameter serves only for CF-Leaf data. After determining the nearby distance, the data point in a conventional BIRCH determines the CF-node. T surpasses the radius of the leaf; the data point will proceed to the CF-leaf. However, if it over the TV, the splitting of the leaf will result, and if it beyond the leaf limit, a new leaf will be created. Alternatively, in BIRCH-modif, the TV will be adjusted for any novel data point that exceeds it. Specifically, it can lessen split leaf in BIRCH by increasing the size on the leaf radius. The (ABF) Adaptive Benefit Factor is utilized in the computation of this TV.

Adaptive benefit factor (ABF)

The ABF is then utilized to update each novel solution in turn. Only the ABF has been used in the suggested strategy and it is expressed as follows. The term "collaboration process" refers to a symbiotic relationin which two diverse species benefit individually from the synergy. Let X_j be a randomly selected organism in the ecosystem, and let X_i and X_j stand for the ith and jth organisms, respectively. Eqs. (7) and (8) below describe the manner in which X_i and X_j interact to produce new candidate solutions:

$$X_{inew} = X_i + rand[0,1] \times (X_{best} - Mutual_{vector} \times BF1) \qquad (7)$$

$$X_{jnew} = X_j + rand[0,1] \times (X_{best} - Mutual_{vector} \times BF2) \qquad (8)$$

where $Mutual_{vector} = (X_i + X_j)/2$

Here, the best organism in the ecosystem can be denoted as X_{best}; the benefit factors (BF1 and BF2) that are randomly generated as either 1 or 2, and rand (0, 1) is a random number that is evenly distributed in the interval [0,1]. Mutual Vector represents the distinctive relationship among 2 organisms, X_i and X_j, while these components signify the degree of benefit for each organism. By comparing X_{inew} and X_{jnew} with X_i and X_j, the fittest organism in every pair is then determined. Based on theX_{best}, new organisms are created during this stage. This contributes in improving the capacity for local search or exploitation.

The following constitutes the improved BIRCH technique:

1. Applying the formulation CF = (N, LS, SS), all data points are transformed into CF procedure.

2. Following the transformation of all the data into CFs, CFT begins to function in order to merge several CFs that have been produced. The number B (Brancing) needs to be entered in this section.

3. The initial CFT threshold needs to be initialized before any data points from the database are analyzed. That T will be employed as the starting point TV for every new CF entry that won't be altered throughout the clustering procedure.

4. As L will be added when the cluster needed it, they does not need it for enhancing BIRCH.

5. Let's add two parameters, m and b, to assist in the computations. The amount of branches on the CF-non leaf is counted using parameter b, while the amount of leaf branches on the CF-leaf is counted using parameter m.

6. Utilizing the (ED) Euclidean Distance, BIRCH will compare the locations of all CFs that have formed at the RN (Root Node) for every CF SC that enters. In BIRCH, the CF SC is extended to the next CF RN.

7. After selecting the CF node, the SC falls to the non-leaf child node. BIRCH correlates every non-leaf CF's position with that of the sub cluster. The SC that entered the non-leaf CF node, that was nearest to the incoming SC, continued using BIRCH.

8. After the non-leaf CF node selected closest to it, the SC descends to the child leaf node. The sub cluster's position is compared to every leaf's sub cluster using BIRCH. SCarrived the leaf close to the arriving CF SC was temporarily bypassed via BIRCH.

9. The closest leaf will be identified by the SC, which will then determine whether to enter the leaf (leaf-CF (modif)) if the leaf radius that encompasses the novel SCfail to surpass T.

10. The system will increase the cluster scale if the leaf radius chosen, which includes the CF sub cluster, is greater than T. Next, double-check. The SCmayacquire the leaf (leaf-CF (modif)) and the TV of T will be adjusted if the radius is less than the novel TV.

11. However, if the radius value continues to rise beyond the T threshold due to changes in the T, a leaf (leaf-CF (modif)) made up entirely by incoming sub clusters is generated. The ABF is employed to update the TV. For new data points, the CF parent was updated to root.

12. Global clustering, or stage 3, will subsequently get carried out by BIRCH.

3.3. Hybrid Graph Convolutional Bidirectional LSTM model

This novel architecture combines the robustness of GCN and Bi-LSTM units, facilitating comprehensive analysis of spatial and temporal features in PET images. To enhance model performance, hyperparameter tuning is performed using the Enhanced HBO Algorithm, optimizing model parameters for improved accuracy and robustness.

3.3.1. Graph Neural Network (GNN)

In recent times, GNNs have become more and more common in a variety of fields. The major development in the field of graph analysis is made possible by GNN's ability to model the dependencies among nodes in a network [26]. One kind of NN that directly manipulates the graph structure is the GNN. Node classification is a common application for GNN. The objective is to predict the labels of the nodes in the graph without using ground truth, since each node in the network is simply assigned a label. The neighborhood of every node is a discrete computation graph Utilize data propagation throughout the graph to calculate embeddings for nodes. All of the data transferred from a node's neighbors, who have been notified by their own neighbors , and it is is effectively aggregated.

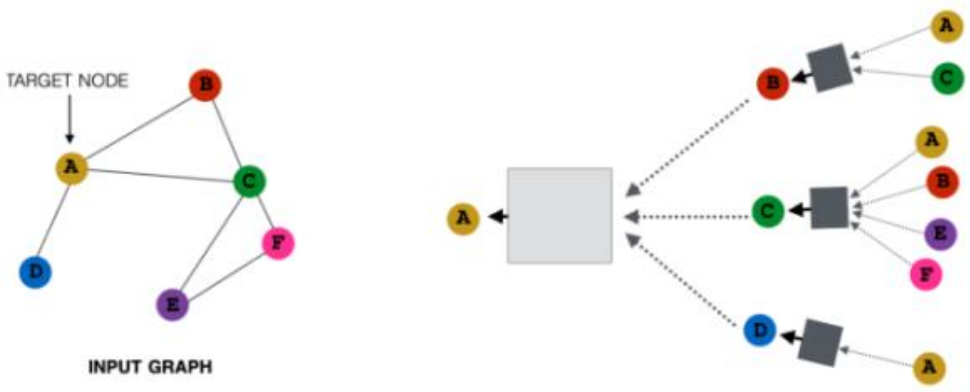

Figure 3. Structure of input graph and target node in GNN

The computation graph of node A depending on its neighborhood is presented on the right side of the graph in Figure 3, which depicts the input graph and the target node A. It is logical, A receives all of the messages from the nodes [B, C, D] in its neighborhood and handles them in some way. The nodes [B, C, D] then handle data from their neighbors. To understand the computational graph, look at the edge directions. The original node features of nodes [A,C] are transmitted to node BBB in this computational graph of depth 2, where they are altered before being passed back to A node. Every node possesses its own CG (Computational Graph). As, dynamic recursive programming was employed, node embeddings of every node at every tier of the CG are simultaneously computed and sent into the subsequent layer. Using this concept, the disease classification is detected in this study.

The probability of diagnosing leaf l with disease *m* can then be estimated by calculating the degree of correspondence among any diseases and leaf image's embeddings. This section begins with a description of (GE) Graph Encoder framework, that's in the position of assigning every node in an arbitrary graph a low-dimensional embedding, z ∈ Rd [30]. In addition to generating embeddings for nodes that were not previously observed during training such as a recently connected leaf image in the dataset, and this configuration optimizes the application of data shared across graph regions.

The first step is to randomly characterize a disease, sign, or node for a graph $G = \{C,P\}$as $v \in G$ in order to make the representation simple. Afterward, the embedding **h***lv* of node v is computed at the *l*-thinformationTL (Transmission Layer) as follows:

$$h^l_{N(v)} = AGGREGATE\left(\left\{h^{l-1}_{v'}, \forall v' \in N(v)\right\}\right) \quad (9)$$

$$h^l_v = \sigma\left(W^l.\left[h^{l-1}_v;\ h^l_{N(v)}\right]\right) \quad (10)$$

Here, the embedding of node v at the previous layer can be demoted as h^{l-1}_v, the weight matrix to be acquired at the l-th layer can be denoted as W^l, and *L* is the overall layer size. The combination of 2 vectors is represented by [.;.], while the collection of uniformly sampled neighbor nodes of *v* is represented by *N*(*v*). It should be noted that, when *l* = 0, the node embedding h0*v*∈R*d* is initialized using, depending on availability, either side of details from the data or randmonized values. A leaf node, for example, can be initialized as a real-valued dense (FV) Feature Vector h0v, where each digit in h0*v* indicates the observed value of FV (e.g., age) given the accessible image demographics and features in the data.When the aggregation function is used to aggregate the embeddings of node *v*'s neighbors at the (*l*−1)-th layer, the synergic representation that results is represented as h*lN*(*v*). If *σ* is a non-linear AF (Activation Function) (like tanh), then mean, max pooling, or another option might be selected as the aggregator. Use *mean*(⋅) by default to aggregate the data in the model created. Then proceed to the final embedding for every node at final layer *L* after a normalization stage:

$$h_v = \frac{h^L_v}{\left\|h^L_v\right\|_2}, \qquad \forall_v \in \mathcal{g} \quad (11)$$

It is noticeable that the models learned for the symptoms nodes *p* in the concept network *C* and leaf graph *P* share a similar embedding space is notable in this study. In other words, every symptom *p* has similar embeddings in both networks, which effectively acts as a link among the disease and leaf nodes from graph that is distinct. As various node types: diseases, leaves, and symptoms, for example learn from 3distinct embedding spaces, projected all node embeddings into one space, and then perform a non-linear activation to further align their contexts:

$$z_v = \sigma(W h_v), \forall_v \in \mathcal{g} \quad (12)$$

Here,*zv* is the node *v* final embedding and *W* is the learnable projection weight. Though distinct edges in a knowledge graph

may represent various node relationships andit is unable to handle edge data. Furthermore, fixed point may not be the best option for learning to represent nodes since it can impede the diversification of node distribution. Therefore, in order to solve the abovementioned problem, this research introduced a Bi-LSTM.

3.3.2. Bidirectional LSTM

A fundamental RNN takes an input sequence, x = (x1, . . . ,xT), and iteratively solves the following equations from t = 1 to T to calculate the Hidden Vector (HV) sequence, h = (h1, . . . , hT) and the (OV) Output Vector sequence, y = (y1, . . . , yT).

$$h_t = \mathcal{H}(W_{xh}h_t + W_{hh}h_{t-1} + b_h) \quad (13)$$

$$y_t = W_{hy}h_t + b_y \quad (14)$$

Here, the (HL) Hidden Layer function can be denoted as $\mathcal{H}$, the b terms stand for bias vectors (e.g., hidden bias vector is b_h), and the W terms represent weight matrices (e.g., the input-hidden weight matrix is represented as W_{xh}).

The Sigmoid Function (SF) is often applied elementwise to H. The LSTM framework, on the other hand, has proven to be more effective in identifying and utilizing long range context [27]. It performs this by storing data in particular memory cells. one memory cell of an LSTM. The following Composite Function (CF) implements H for the version of LSTM utilized in this study:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (15)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (16)$$

$$c_t = f_t c_{t-1} + i_t tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (17)$$

$$O_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o) \quad (18)$$

$$h_t = O_t \tanh(c_t) \quad (19)$$

The forget vector can be denoted as f, output vector can be denoted as o, cell activation vector can be denoted as c, and the input vector can be denoted as i, these are all the same size as the h (HV). Here, σ represents the logistic SF. Only element m of every Gate Vector (GV) receives input from element m of the cell vector because the weight matrices from the cell to GV (such as W_{si}) are diagonal. The fact that traditional RNNs can only utilize prior context constitutes one of their limitations. Future context should also be utilized in SR (Speech Recognition), since entire statements are translated at once. Through the utilization of 2 separate HL for processing the data in both ways and then input the results back to the same (OL) Output Layer, bidirectional RNNs, or BRNNs, accomplish this. A BRNN iterates the Forward Layer (FL) from t = 1 to T, the Backward Layer (BL) from t = 1 to T, and then updates the OL to calculate the forward hidden sequence ($\overleftarrow{h}$), the backward hidden sequence ($\overleftarrow{h}$), and the output sequence (y):

$$\vec{h}_t = \mathcal{H}(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t-1} + b_{\vec{h}}) \quad (20)$$

$$\overleftarrow{h}_t = \mathcal{H}(W_{x\overleftarrow{h}}x_t + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}) \quad (21)$$

$$y_t = W_{\vec{h}y}\vec{h}_t + W_{\overleftarrow{h}y}\overleftarrow{h}_t + b_y \quad (22)$$

Bi-LSTM has access to long-range information in both input directions, is created by combining BRNNs with LSTM. Using deep structures to create increasingly complex models for acoustic information is a key component of the recent achievement of hybrid structures. By piling several RNN HL on top of one another, deep RNNs can be produced, with the input sequence for every layer being formed by the previous layer with a output series. The HV sequences h n are successively calculated from n = 1 to N and t = 1 to T, assuming that the same HL function is applied to all N layers in the stack:

$$h_t^n = \mathcal{H}(W_{h^{n-1}h^n}h_t^{n-1} + W_{h^n h^n}h_{t-1}^n + b_h^n) \quad (23)$$

Here $h^0 = x$ is defined. y_t is the network output. Then,

$$y_t = W_{h^N y}h_t^N + b_y \quad (24)$$

By substituting the forward and backward sequences, $\vec{h}^n$and $\overleftarrow{h}^n$, for each hidden sequence h n, and making sure that each HL accepts input from the FL and BL at the lowerlevel, Deep Bi RNNs (DBi-RNNs) can be created. In Fig. 4, Get Bi- LSTM if LSTM is applied to the HLs.

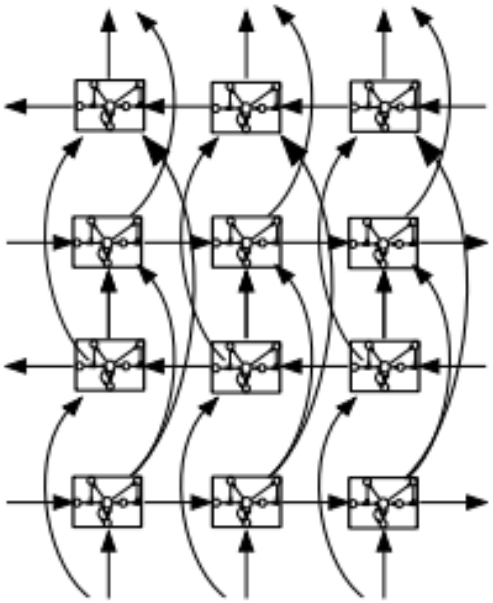

Figure 4. Structure of Bi-LSTM

3.3.3. Hyper Parameter (HP) optimization using Enhanced HBO Algorithm

Some adjustments should be made to the HBO technique to handle MO (Multi Objective) problems. Selecting the "best solutions" (queens) should no longer be done only on the basis of comparing the values of individual (OF) Objective Functions once the initial population has been generated and the corresponding OFs have been evaluated. The efficiency of the algorithm is enhanced by the application of clustering method; it additionally promotes greater distribution of the solutions along the front. The HBO technique integrates several distinct processes. Each of these relates to a distinct stage of the Honey-Bee's mating cycle, as shown in Fig. 5. A probabilistic-based annealing function can be utilized for expressing a drone mate with a queen in the following way:

$$Prob(D) = \exp[-\Delta(f)/S(t)] \quad (25)$$

Here,D(f) is the total variance among D's and the queen's fitness, $S(t)$ is the queen's speed at time t, and $Prob(D)$is the probability of adding drone D's sperm to the queen's spermatheca. When the drone is as fit as the queen is, or when the queen is moving at a rapid speed, there is a greater chance of mating. The following equation describes how the queen's speed drops with every shift in space:

$$S(t+1) = \alpha \times S(t) \quad (26)$$

Here, the measure of speed and energy loss following every phase and transition is represented by the symbol an e (0, 1). To improve the genotype of the brood, workers could indicate a variety of heuristics. The heuristic fitness value is determined by the rate of change in the genotype of the brood as a consequence of applying a heuristic to that brood. The suggested EHBO method makes use of several queens. They are kept in an external memory (repository) and regarded as a set of NDS. The non-dominated individuals included in the initial population initiate the repository. A fuzzy-based approach is employed to extract the optimum solution as the OF are not the same.

If any individual in the repository is dominated by the best solution during the search procedure, that individual is promptly eliminated from the repository. A large repository is preferable. However, because of memory limitations, the size of the repository must be limited. Large repositories also result in delays in processing. The probability that a drone will mate with a queen can be calculated using the EHBO technique as follows:

$$Prob(D) = \exp\left(-\frac{|F_{queen} - F_{drone}|}{S(t)}\right) \quad (27)$$

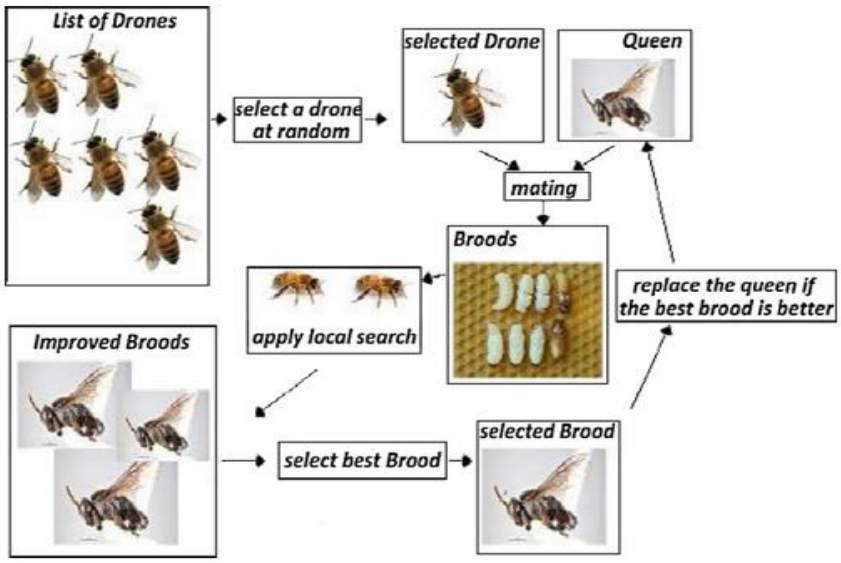

Figure 5. Procedure of HBO

$$\left|F_{queen} - F_{drone}\right| = \sqrt{(f_1^{queen} - f_1^{drone})^2 + (f_2^{queen} - f_2^{drone})^2 + (f_3^{queen} - f_3^{drone})^2 + (f_4^{queen} - f_4^{drone})^2} \quad (28)$$

Thus, the process described can be summed up as follows: for single and MO problems, respectively, the probability of mating may be computed from Eqs. (26) and (28) when one drone is selected at random from the initial population. The selected drone is then kept in the queen's spermatheca if the chance of mating is higher than another random value within [0, 1]. As a result, the selected queen can generate more NDS in the following iteration.The Decision-Maker (DM) must select one option as the best solution among the Pareto set of NDS after obtaining the Pareto-Optimal Set (POS). The Membership Function (MF) $\mu_{fi}(X)$ defines the ith OF, as a result of the DM's ineffective evaluation.

$$\mu_{fi}(X) = \begin{cases} 1 & for\ f_i(X) \le f_i^{min} \\ 0 & for\ f_i(X) \le f_i^{max} \\ \frac{f_i^{max} - f_i(X)}{f_i^{max} - f_i^{min}} & for\ f_i^{min} \le f_i(X) \le f_i^{max} \end{cases} \quad (29)$$

The lower and upper bounds as well as a strictly monotonically declining and continuous function constitute the MF. To obtain MF $\mu_{fi}(X)$ for every OF, $f_i(X)$, the lower and upper bounds (f_i^{min} and f_i^{max}) of every OF according to the stated limitations are determined. The normalized MF μ_k for each non-dominated solution k can be computed by

$$\mu_k = \frac{\sum_{i=1}^{N_{obj}} \mu_i^k}{\sum_{k=1}^{M} \sum_{i=1}^{N_{obj}} \mu_i^k} \quad (30)$$

Here, the quantity of solutions that are not dominated can be denoted as M. The selection with the highest value of μ_k is the optimal solution.The jth brood in the initial HBO is created by the following method: The mating relationship among the drones kept in the queen's spermatheca and the queen produces a population of broods. The following procedure is employed to generate the jth brood:

$$X_{best} = [X_{best}^1, X_{best}^2, \dots\dots X_{best}^n] \quad (31)$$

$$SP_i = [s_i^1, s_i^2, \dots.. s_i^n] \quad (32)$$

$$Brood_j = X_{best} + \beta \times (X_{best} - SP_i), \qquad j = 1,2, \dots.. N_{brood} \quad (33)$$

Here,$Brood_j$ is the jth brood and b is a random value among 0 and 1. The queen place is X_{best}. The queen's spermatheca contains one drone, called SP_i. 2 parent structures' traits are combined in the initial mating procedure to create an offspring that is similar to both. This research suggests a novel approach for generating the broods as follows so as to enhance the broods generated by the queen's mating flight:

3 drones ($SP_{z1}, SP_{z2}, \dots.. SP_{z3}$) are first chosen at random from the queen's spermatheca such that $z1 \neq z2 \neq z3$,Then, the following formula is employed to get an improved drone PV (Position Vector):

Through Eq.34, Every novel offspring is attained

$$X_i^{N_{New}} = pX_{b_D}^N + qX_{b_M}^N \quad (34)$$

Here,$X_{b_D}^N$ and $X_{b_M}^N$ stand for the parameters of the individual Dad and Mom, accordingly, and P is a pseudo-random distributed number from zero to one, q = (1-p). When the choice of barnacle to mate with exceeds the starting value of pl, sperm casting takes place as the technique's exploration period. The procedure is formulated in Eq. 35.

$$X_i^{N_{New}} = rand * X_{b_M}^n \quad (35)$$

In the range [0, 1], where rand denotes a random number. The mother individual creates the new progeny for the investigation process. In order to handle the solution matrix expansion from the individual dimension, the offsprings were assessed and merged with their parents. As a result, the undesirable outcomes have been eliminated and the organizing procedure is carried out to select the 50% best solutions that meet the population size.

4. RESULTS AND DISCUSSION

The suggested hybrid GNN using Bi-LSTM (HGNN-BiLSTM) classification algorithm for lung image recognition has been evaluated in this research using images acquired from the PET dataset. The metrics accuracy, precision, and recall f-measure are employed for evaluating the effectiveness of the recommendedtechnique. The current 3D-CNN-RNN and CNN are contrasted with the suggested HGNN-BiLSTM technique.The performance metrics according to this confusion matrix are calculated as follows.

Precision is defined as the ratio of accurately detected +ve observations to all predicted +ve observations.

$$Precision = TP/(TP + FP) \quad (36)$$

The ratio of correctly identifying +ve observations to the overall amount of observations is referred to as sensitivity or recall.

$$Recall = TP/(TP + FN) \quad (37)$$

The weighted average of both recall and precision is known as the F-measure. That requires both FP (False Positives) and FN (False Negatives) as an outcome.

$$F - measure = 2 * (Recall * Precision)/(Recall + Precision) \quad (38)$$

In terms of +ves and -ves, accuracy is computed using the following formula:

$$Accuracy = (TP + FP)/(TP + TN + FP + FN) \quad (39)$$

Here, TP- True Positive, FP-False Positive, TN-True Negative, FN- False Negative.

Table 1 tabulates the performance comparison results of the proposed and existing methods

Performance metrics	CNN	3D-CNN-RNN	HGNN-BiLSTM
Accuracy	85.12	93	95.69
Precision	81	90.09	96.984
Recall	84	92.2	92.499
F-measure	83	91.12	94.688

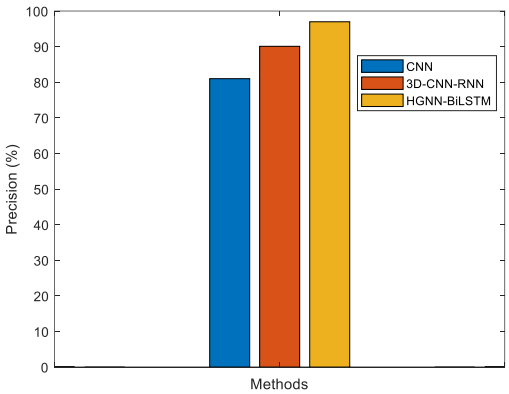

Fig.6. Comparison Outcomes of precision among the suggested and currenttechnique for the disease classification in lung images

The precision outcomes of comparison among suggested HGNN-BiLSTM and the current approach for classifying the disease in lung images are shown in Fig. 6. For increasing classifier accuracy, the suggested M-BRICH based segmentation technique offers the most effective outcomes. According to the results, the suggested HGNN-BiLSTM strategy overtakes the current classification approachesby precision.

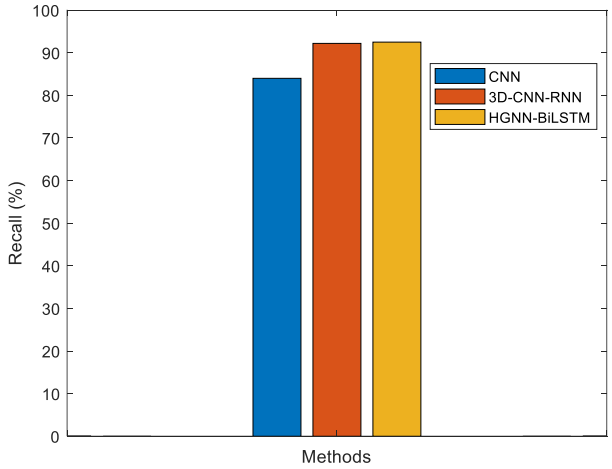

Fig.7. Comparison of recalloutcomesamong the suggested and currenttechnique for the disease classification in lung image

The recall outcomes of the comparison of the suggested and current methods for identifying the disorders in lung images are presented in fig. 7. The optimal FE result is obtained from the suggested segmentation framework. As an outcome, the suggested HGNN-BiLSTM algorithms were employed for the assigned work, and the outcomes were assessed and examined.

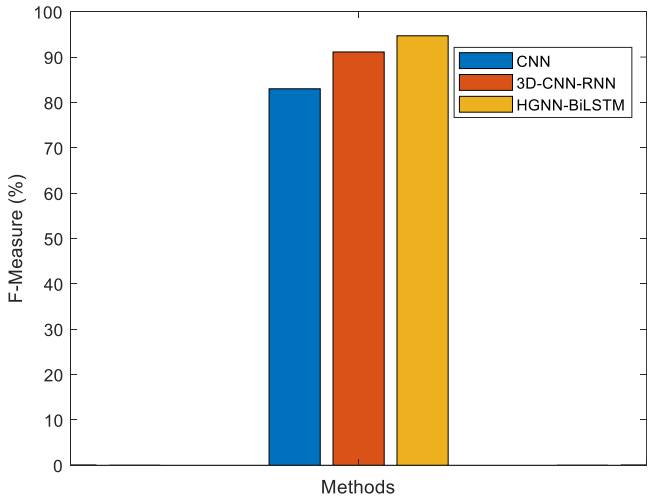

Fig.8. Comparison of F-measureoutcomesamong the suggested and currenttechnique for the disease classification in lung images

The F-measure results for the comparison among the suggested and current technique for classifying the disorders in lung images are presented in Fig. 8. Furthermore, utilizing the current classification algorithms results in a majordrop in the amount of time needed to complete a prediction. When compared to the other DLT, the suggested model has the greatest f-measure rate measurement in the database, as can be observed in the comparison above. When compared to other techniques, the database that was employed offered the best f-measure outcomes.

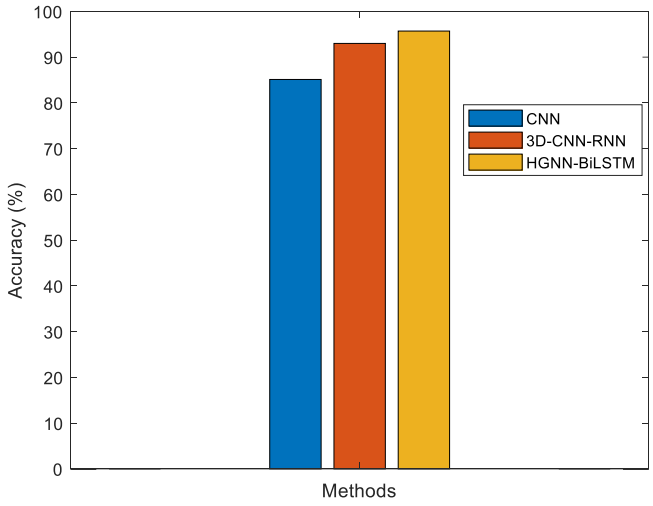

Fig.9. Comparison of Accuracyoutcomesamong the suggested and currenttechnique for the disease classification in lung images

The accuracy comparison among the suggested and current methods for classifying the disorders in lung imaging is shown in fig. 9. The ability to precisely predict the target and generalize predictions to new cases is a sign of an effective DL framework. The suggested HGNN-BiLSTM approach has a high accuracy of 95.67%, according to the simulation outcomes, compared to 93.9% for the current 3D-CNN-RNN approach and 71.41% for the CNN approach.

5. CONCLUSION

An intelligent framework for LC detection in PET images, integrating advanced techniques for noise removal, segmentation, classification, and hyperparameter tuning. First, aAdaptive Weiner Filter (AWF) is applied to PET images to effectively remove noise and enhance image clarity, ensuring more accurate analysis. Subsequently, the Modified BIRCH algorithm is utilized for segmentation, enabling the delineation of regions of interest within the images. For lung cancer classification, a hybrid Graph Convolutional Bidirectional LSTM model is developed. This novel architecture combines the robustness of GCN and Bi-LSTM units, facilitating comprehensive analysis of spatial and temporal features in PET images. To enhance model performance, hyperparameter tuning is performed using the Enhanced Honey bee optimization Algorithm, optimizing model parameters for improved accuracy and robustness. The proposed framework is evaluated using the PET Dataset, demonstrating its efficiency in detecting LC accurately in PET images.From the simulation outcomes it is identified that the suggested HGNN-BiLSTM model has improved accuracy as 95.67%. As the future work the ensemblemodelis introduced for improving the classifier accuracy.

REFERENCES

[1] Taher, F., Prakash, N., Shaffie, A., Soliman, A., & El-Baz, A. (2021). An overview of lung cancer classification algorithms and their performances. *IAENG International Journal of Computer Science*, *48*(4).

[2] Bach, P. B., Kelley, M. J., Tate, R. C., &McCrory, D. C. (2003). Screening for lung cancer: a review of the current literature. *Chest*, *123*(1), 72S-82S.

[3] Travis, W. D. (2011). Pathology of lung cancer. *Clinics in chest medicine*, *32*(4), 669-692.

[4] Schabath, M. B., & Cote, M. L. (2019). Cancer progress and priorities: lung cancer. *Cancer epidemiology, biomarkers & prevention*, *28*(10), 1563-1579.

[5] Thakur, S. K., Singh, D. P., &Choudhary, J. (2020). Lung cancer identification: a review on detection and classification. *Cancer and Metastasis Reviews*, *39*(3), 989-998.

[6] Shames, D. S., &Wistuba, I. I. (2014). The evolving genomic classification of lung cancer. *The Journal of pathology*, *232*(2), 121-133.

[7] Travis, W. D. (2011, July). Classification of lung cancer. In *Seminars in roentgenology* (Vol. 46, No. 3, pp. 178-186).

[8] Cersosimo, R. J. (2002). Lung cancer: a review. *American journal of health-system pharmacy*, *59*(7), 611-642.

[9] Mountain, C. F. (2002). Staging classification of lung cancer: a critical evaluation. *Clinics in chest*

medicine, *23*(1), 103-121.

[10] Rodriguez-Canales, J., Parra-Cuentas, E., &Wistuba, I. I. (2016). Diagnosis and molecular classification of lung cancer. *Lung Cancer: Treatment and Research*, 25-46.

[11] Kuruvilla, J., &Gunavathi, K. (2014). Lung cancer classification using neural networks for CT images. Computer methods and programs in biomedicine, 113(1), 202-209.

[12] Asuntha, A., & Srinivasan, A. (2020). Deep learning for lung Cancer detection and classification. *Multimedia Tools and Applications*, *79*(11), 7731-7762.

[13] Lakshmanaprabu, S. K., Mohanty, S. N., Shankar, K., Arunkumar, N., & Ramirez, G. (2019). Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*, *92*, 374-382.

[14] Riquelme, D., &Akhloufi, M. A. (2020). Deep learning for lung cancer nodules detection and classification in CT scans. *Ai*, *1*(1), 28-67.

[15] Cengil, E., &Cinar, A. (2018, September). A deep learning based approach to lung cancer identification. In *2018 International conference on artificial intelligence and data processing (IDAP)* (pp. 1-5). Ieee.

[16] Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., ... &Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature medicine*, *24*(10), 1559-1567.

[17] Tekade, R., &Rajeswari, K. (2018, August). Lung cancer detection and classification using deep learning. In *2018 fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-5). IEEE.

[18] Alakwaa, W., Nassef, M., &Badr, A. (2017). Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). *International Journal of Advanced Computer Science and Applications*, *8*(8).

[19] Zhu, W., Liu, C., Fan, W., &Xie, X. (2018, March). Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. In *2018 IEEE winter conference on applications of computer vision (WACV)* (pp. 673-681). IEEE.

[20] Lyu, J., & Ling, S. H. (2018, July). Using multi-level convolutional neural network for classification of lung nodules on CT images. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 686-689). IEEE.

[21] Zhang, Q., Wang, H., Yoon, S. W., Won, D., & Srihari, K. (2019). Lung nodule diagnosis on 3D computed tomography images using deep convolutional neural networks. *Procedia Manufacturing*, *39*, 363-370.

[22] Da Nóbrega, R. V. M., Peixoto, S. A., da Silva, S. P. P., &RebouçasFilho, P. P. (2018, June). Lung nodule classification via deep transfer learning in CT lung images. In *2018 IEEE 31st international symposium on computer-based medical systems (CBMS)* (pp. 244-249). IEEE.

[23] Wankhade, S., &Vigneshwari, S. (2023). A novel hybrid deep learning method for early detection of lung cancer using neural networks. *Healthcare Analytics*, *3*, 100195.

[24] Jin, F., Fieguth, P., Winger, L., & Jernigan, E. (2003, September). Adaptive Wiener filtering of noisy images and image sequences. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)* (Vol. 3, pp. III-349). IEEE.

[25] Nirmala, G., &Thyagharajan, K. K. (2019, April). A modern approach for image forgery detection using BRICH clustering based on normalised mean and standard deviation. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0441-0444). IEEE.

[26] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., &Monfardini, G. (2008). The graph neural network model. *IEEE transactions on neural networks*, *20*(1), 61-80.

[27] Graves, A., Fernández, S., &Schmidhuber, J. (2005, September). Bidirectional LSTM networks for improved phoneme classification and recognition. In *International conference on artificial neural networks* (pp. 799-804). Berlin, Heidelberg: Springer Berlin Heidelberg.

[28] Yuce, B., Packianather, M. S., Mastrocinque, E., Pham, D. T., &Lambiase, A. (2013). Honey bees inspired optimization method: the bees algorithm. *Insects*, *4*(4), 646-662.