RESEARCH ARTICLE

Lung-RetinaNet: Lung Cancer Detection Using a RetinaNet With Multi-Scale Feature Fusion and Context Module

RABBIA MAHUM[1] **AND ABDULMALIK S. AL-SALMAN**[2]

[1]Department of Computer Science, University of Engineering and Technology (UET) at Taxila, Taxila 47050, Pakistan
[2]Department of Computer Science, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding authors: Rabbia Mahum (rabbia.mahum@uettaxila.edu.pk) and Abdulmalik S. Al-Salman (salman@ksu.edu.sa)

ABSTRACT Lung cancer is one of the terrible diseases in various countries around the globe, and timely detection of the illness is still a challenging process. The oncologists consider the blood test results and CT scans to assess the tumor, which is time-consuming and involves extra human effort. Therefore, an automated system should be developed to efficiently recognize lung tumors and assess their severity to reduce mortality. Although various researchers have proposed lung disease detection systems, the existing techniques still lack significant detection accuracy for early-stage tumors. Thus, this study proposes a novel and efficient lung tumor detector based on a RetinaNet, namely Lung-RetinaNet. A multi-scale feature fusion-based module is introduced to aggregate various network layers, simultaneously increasing the semantic information from the shallow prediction layer. Moreover, a dilated and lightweight algorithm is employed for the context module to combine contextual information with each network stage layer to improve features and effectively localize the tiny tumors. The proposed methodology attained 99.8% accuracy, 99.3% recall, 99.4% precision, 99.5% F1-score, and 0.989 Auc. We evaluated our suggested model and matched the performance with state-of-the-art DL-based methods. The outcomes show that our technique provides more substantial results than existing methods.

INDEX TERMS Early detection, lungs cancer, artificial intelligence, RetinaNet.

I. INTRODUCTION

Lung cancer is considered one of the most terrible illnesses in various countries, and its death rate is 19.35% [1]. The multiple ways used by radiologists to detect cancer include sputum cytology, CT scans, X-rays, and other magnetic resonance imaging techniques. During the detection process, tumors are categorized into two categories: malignant and benign. Malignant tumors are cancerous and proliferate, having irregular shapes and sizes. It is also analyzed that the endurance rate for patients of the advanced stage is much less than for cancer diagnosed early. It has also been assessed that the timely analysis of the scans and imaging can be improved

by employing various image processing methods [2]. Various research works have been proposed for detecting early-stage cancers using image processing methods. Two main challenges may reduce the hit ratio of lung cancer detection manually. First is the technical and human accessibility, as radiology resources may be inadequate to meet the demand [3]. Second, a significant number of false positive cases are due to the first shortcoming. Therefore, high-quality training should be provided to the radiologists interpreting the images. Thus, the precision of detection and categorization for existing techniques still requires improvements.

Recent progress in machine learning (ML) and deep learning (DL) techniques has resulted in a significant shift towards computer-aided detection (CAD) systems for lung cancer detection. There exist some of the traditional ML-based

---

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

techniques in the literature aiding lung cancer detection and classification, for example, Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN) [4]. These techniques perform manual feature extraction, and then the classifier is trained using extracted features. Moreover, dealing with the various features is tiring and requires extra time. Besides this, the ML-based model gets training over a small number of samples which causes a generalization problem. Some techniques use segmentation methods for lung cancer diagnosis [5]. The region of interest (ROI) is selected based on texture, color, or grayscale from the original image as a segmented region. The standard techniques for segmentation include Region growing, Atlas, and Thresholding. The segmentation-based techniques' performance highly depends upon the segmented area and its extracted features. Many results have been attained using segmentation-based methods for lung cancer detection. However, these techniques still failover unseen samples and require modification to reduce the false ratio.

With the development in the domain of deep learning, DL-based techniques have shown superior results for the recognition of numerous diseases such as knee [6], eyes [6], potato leaves [7], brain [8], and computer vision techniques [9]. The primary benefit of DL-based models is the automatic features extraction phase, whether the model is segmentation-based or classification-based. Furthermore, DL-based models extract the most representative features due to the continuously increasing depth of the layers. The DL models comprise pooling, batch normalization (BN), convolutional, and fully connected layers. The pooling layers also minimize the feature maps' size while reducing the model's complexity. Various DL-based models have been proposed for lung tumor detection. However, most methods are based on simple classification [10]. These classification techniques consider the whole image for feature extraction, which may increase the misdiagnosis of early-stage tumors. Therefore, to solve the issues mentioned above and enhance early lung tumor detection performance, we propose a novel and improved deep learning-based model based on RetinaNet [11]. We offer a feature fusion block instead of a feature pyramid network (FPN) in RetinaNet to mine the most representative feature maps minimizing the loss of exhaustive information from input. Moreover, a dilated convolution utilizes the most powerful features of tiny tumors at shallow layers. The features from upper layers posing high classification accuracy are integrated with the bottom ones exhibiting high localization results. Utilizing contextual information during feature fusion adds more features from lower layers using a contextual block. Furthermore, the default anchors were not performing well for tiny lung tumors due to their irregular shapes and sizes. Therefore, a k-means clustering method is adopted, as used in YOLO-v3 [12], to generate precise anchors along with contextual feature fusion blocks. The outcomes show that our suggested model efficiently detects tiny lung tumors and classifies them. The main offerings of the suggested work are as follows:

- This research aims to detect and segment lung tumor automatically at an early stage. Therefore, we propose a robust multi-scale feature fusion-based module to aggregate various layers of the network, simultaneously increasing the semantic information from the shallow prediction layer.
- We propose a dilated and lightweight algorithm for the context module to combine contextual information with each network stage layer to improve features and effectively localize the tiny tumors.
- RetinaNet relies on local features for detection and lacks contextual information. Therefore, an improved Lung-RetinaNet combines the dilated context module having lateral connections at each network level and features fusion block, enriching the features at each stage. Additionally, the usage of adaptive anchors was able to detect lung tumors better.
- The proposed Lung-RetinaNet is evaluated using well-known benchmarks, and experimental outcomes indicate that our proposed model significantly outperforms existing lung tumor detectors.

The remaining paper is prepared as follows: Section II shows the related work, Section III refers to the projected model, Section IV evaluates the proposed model using various experiments, and Section V concludes the work.

II. LITERATURE REVIEW

The deep learning models have been used vastly for the cancer detection in last few decades [13], [14], [15], [16].The key attribute for lung cancer detection is the pulmonary nodules and solid clumps made up of tissues surrounding the lungs [17]. The nodules present in the lungs can be malignant or benign depending upon the nature of the cells and viewable on CT scans [18]. In 2015, Hua et al. [19] employed methods of CNN and DBN to classify lung nodules using CT scan imaging. They exhibited that deep learning techniques help extract lung swellings features and categorize them into malignant or benign without considering texture feature. Kurniawan et al. [20] and Rao et al. () [21] utilized CNN employing simple and straightforward classification to distinguish lung tumors using CT scans. Song et al. [18] performed a comparative analysis of the deep learning model, CNN and auto-encoder for lung tumor detection. They mentioned that the CNN based on classification performs better than DL and stacked auto-encoder models. Ciompi et al. [22] employed CNNs based on a multi-scale to classify lung cancer into six types: calcified, solid, part solid, non-solid, speculated, and perifissural nodules. They specifically developed a multi-stream artificial neural network that can handle multiple triplets of 2-dimensional views of lung nodules at different scales and then computes the likelihood of the presence of any tumor among six classes. Yu et al. [23] developed a system performing bone exclusion and lung segmentation before CNN's training. Shakeel et al. [24] proposed a system based on segmentation using pre-processed images. In pre-processing, the authors utilized denoising methods and

enhanced the image's appearance. In the end, the artificial neural network has been trained to categorize lung cancer.

Ardila et al. [25] proposed a system based on four modules: segmentation, cancer detection model, cancer risk estimation model, and full-volume model. After the lung segmentation, the ROI identification model estimates the maximum nodules region, whereas the full-volume module was trained to estimate cancer likelihood. The combined output is formed from these two modules as the final output for the prediction. Chen et al. [26] established a system for nodule recognition based on nodule segmentation and enhancement. Hosny et al. [27] and Xu et al. [28] employed CNN for lung cancer detection after data augmentation. They utilized various methods for data augmentation, such as flipping, rotation, and translation. A DL-based technique using temporal attention, namely visual simple temporal attention (ViSTA) in CT images [29]. 351 nodes were utilized in the work. The proposed model attained 86.4% area under the curve (AUC). The authors in [30] utilized the LUNA16 dataset for training the cancer detection model and then modified the model using the KDSB17 dataset for global feature extraction. After that, they combined local features attained from another separate classifier and achieved higher accuracy for lung nodule detection. In [31], authors employed transfer learning to train the model multiple times. The classifier was previously trained on the ImageNet dataset and then trained using the ChestX-ray14 dataset.

In the end, the JSRT dataset was used to test lung cancer detection. In [32], only a survey was performed for lung cancer detection using histopathology imaging. Squamous cell carcinoma (LUSC) and Adenocarcinoma are common types of cancers, and for their detection, the pathologist should be a very experienced person. In the proposed study, a CNN was trained on the images of slides during histopathology to accurately classify LUSC, LUAD, and normal tissues. Xu et al. [33] utilized a CNN based on LSTM for lesion detection in X-ray imaging. LSTM is a modified network of RNN and improves the architecture to reduce the issues of vanishing gradient. The proposed network exhibited a significant relationship among lesions to predict cancer precisely.

Abd El-Wahab et al. [34] developed a system using several versions of EfficientNet (B0, B1, B2) to identify the various lungs disease. The authors fused the features of pre-trained models and then they passed the features from stacked ensemble learning. The final classifier consisted of SVM and Random Forest (RF) at first phase and logistic regression at second stage. The maximum accuracy attained by the system was 99% for detection of TB in lungs. Aswathy et al [35] employed a system for the lung malignancy detection using nano-image as input and processed through a Gabor filter and color-based histogram equalization for enhancement. The segmented image of lung cancer was then obtained using the Guaranteed Convergence Particle Swarm Optimization (GCPSO) algorithm. To classify the tumor region, a graphical user interface nano-measuring tool was developed. Additionally, the Bag of Visual Words (BoVW) method and a Convolutional Recurrent Neural Network (CRNN) were utilized for feature extraction and image classification. The authors attained 99.35% accuracy for cancer detection. In [36], two benchmark datasets are downloaded, containing attribute information from multiple patients' health records. Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) are employed as standard techniques for feature extraction. In addition, deep features are obtained from the pooling layer of a Convolutional Neural Network (CNN). To select the most relevant features, the Best Fitness-based Squirrel Search Algorithm (BF-SSA) is utilized, which is considered an optimal feature selection method. They attained an accuracy of 93.15%.

III. METHODOLOGY

In this section, we demonstrate the proposed methodology. There exist two types of network-based detectors: single and two-stage. The two-stage detectors, including RCNN, CNN, Mask RCNN etc., are based on a region proposals network that finds ROI. Further, these regions are utilized for classification and regression. These models maintain high precision and accuracy; however, they become very slow due to their architecture. Moreover, some single-stage detectors are fast, including a single-shot multi-box detector. However, they may face the issue of class imbalance during the training
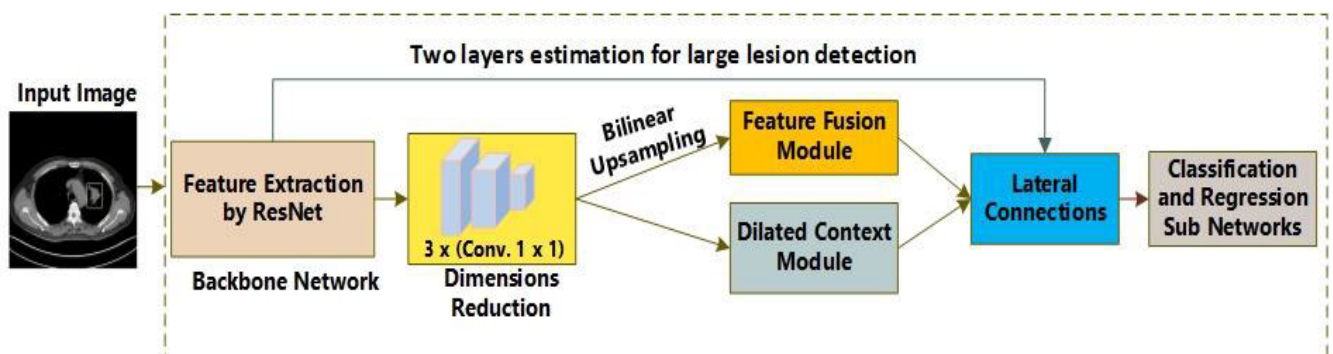

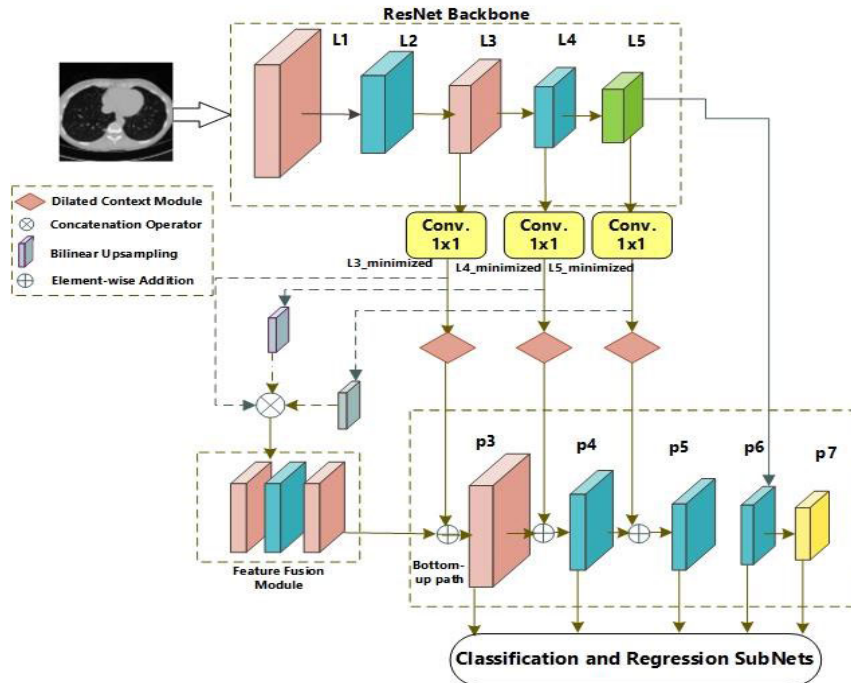**FIGURE 1.** Fused Lung-RetinaNet's architecture.

FIGURE 2. Fused Lung-RetinaNet's architecture.

phase. Therefore, our work employs a single-stage detector for lung tumor detection. The focal loss also helps the network to overcome the problem of class imbalance. The basic flow diagram of the proposed system is shown in Figure 1.

An improved RetinaNet's architecture is shown in Figure 2. We have introduced a context aggregation module instead of FPN in the original RetinaNet. Our proposed model comprises a ResNe101 as a backbone network for extracting features from input, along with dilated contextual block-based fusion, categorization and regression sub-networks. The main aim of these sub-nets is to localize and classify lesions of various sizes at varying scales. In Figure 2, the FPN is exchanged with our proposed context module. First, the final three layers are based on 1 × 1 convolution to reduce dimensions. Further, L4 and L5 are un-sampled using the bilinear approach for converting them to the same size *as* L3. A ReLU activation is employed, followed by BN in the fusion module. Then, our dilated context block is applied to L3_minimized, L4_minimized, and L5_minimized. In the end, our proposed feature fusion and dilated context block are merged for each backbone instant using a bottom-up way. The layers p6 and p7 are used for the large-size tumors.

RetinaNet is a model used for object detection that specifically addresses the challenge of class imbalance in object detection tasks. This issue arises because the majority of regions in an image do not contain objects, resulting in a substantial disparity between the positive and negative classes. To counter this imbalance, RetinaNet utilizes a focal loss function that diminishes the contribution of simple examples and concentrates more on challenging ones. By assigning a higher weight to the loss incurred by hard examples that are

incorrectly classified, RetinaNet better learns from the difficult examples and is less influenced by the easy examples.

The Huber loss is used to compute the localization loss of the model, which is the loss incurred when the model predicts the bounding boxes of the objects in the image. Hence, the Huber loss is more resistant to outliers and reduces the model's sensitivity to small differences between the predicted and ground-truth bounding boxes.

A. FEATURE FUSION BLOCK
The feature fusion block enhances the semantic knowledge of bottom-layer feature vectors [37]. Hence, it improves the results of the localization of tiny tumors. Moreover, the ReLU activation function works for non-linear functions among various layers. Then, batch normalization (BN) is employed to avoid a gradient vanishing problem. Additionally, it improves performance. As in the RetinaNet detector, the implication of ResNet-101 as a base network further helps the model extract the most powerful features [38]. In Figure 2, layer modules L1, L2, L3, L4, and L5 represent numerous backbone network scales. FPN layers ranged from L3 to L5; L3 refers to the shallowest layer. The shallow layer plays a vital role for small lesions detection. Nonetheless, it cannot extract semantic features presented at the deepest layer, L5. Following Feature Fusion Single Shot Multibox Detector (FSSD) based aggregation block, a conv. layer of 1 × 1 is utilized on each layer from L3 to L5 to generate L3_minimized, L4_minimized, and L5_minimized. The aim of 1 × 1 conv. layer is to minimize the dimensions of these layers. Then, minimized layers L4_minimized and L5_minimized are engorged through a bilinear up-sampling to convert them to the same dimension

as the L3_minimized layer. Then, these minimized layers are combined at once, employing a concatenation method [39]. The concatenated layer is transformed using a dilated context block that is enlightened later. In the end, the context-concatenated layer is employed to BN and ReLU function to form a p3 bottom estimation layer.

B. DILATED CONTEXT BLOCK
In this section, dilated blocks, developed by DetNet, and an aggregation CN [40] are employed to improve the information of surrounding tissues necessary for tiny lesion pixels. A block is proposed to extract more features, enrich more data from surrounding pixels, and improve detection, as shown in Figure 3. The block is fixed before lateral connections. It comprised a 1×1 convolutional layer for dimension minimization of a 3×3 convolutional layer to simultaneously have a dilation rate of 2. These two branches are combined using a feature's map concatenation operator [41]. Then 3×3 convolutional layer is employed with the concatenated feature map.

C. LATERAL CONNECTION
These connections ensure improvements of features and compensate the loss of information due to down-sampling. Additionally, the dense architecture enhances stability during network training. Furthermore, it also improves the lesion detection precision. Figure 3 shows these connections form 2 estimation layers, p4 and p5. Decreased feature map L4 is applied to dilated context block. Then, it was connected laterally with the down samples conc. layer through a fusion technique to show estimation layer p4. A similar process is employed on reduced layer L5 that generates estimation layer p5. Thus, estimation layers p6 and p7 are kept without any modification, and their estimations are utilized to improve large lesions detection. Each estimation layer in the contextual lower to upper lateral connections fusion block has 256 channels of features.
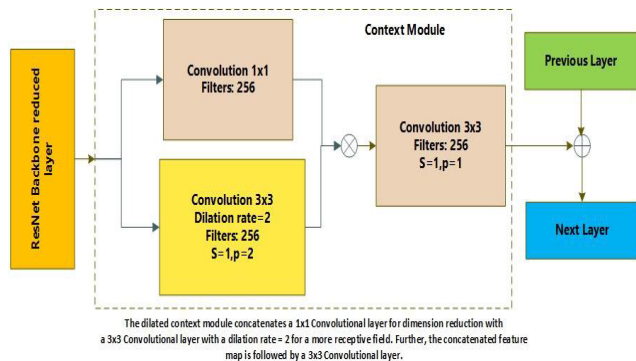

The dilated context module concatenates a 1x1 Convolutional layer for dimension reduction with a 3x3 Convolutional layer with a dilation rate = 2 for a more receptive field. Further, the concatenated feature map is followed by a 3x3 Convolutional layer.

FIGURE 3. Architecture of context module.

D. ANCHORS AND SUB-NETWORKS
In RetinaNet, anchors are bounding boxes of different scales and aspect ratios that are placed at various positions on the image. These pre-defined anchors serve as reference points during training to predict the sizes and locations of objects in the image. The use of anchors allows the network to handle objects of different sizes, shapes, and partial occlusions.

In Lung-RetinaNet, anchors are employed without any changes. Numerous anchors may have different sizes, such as 32×32 to 512×512, which are fed to multi-scale estimation layers from p3 to p7, respectively. In total, nine anchors are included in each layer, with box ratios: 1:1, 1:2, and 2:1, and varying dimensions for every box include $2^{1/3}$, $2^{2/3}$, and 2^0. The anchors with a Jaccard overlap number < 0.5 are not considered as tumors; they are considered as normal lung tissue. RetinaNet has shown that increasing the number of anchors above 9 does not improve the performance. Therefore, we utilized 6-9 k-clusters in the K-means clustering technique for anchors generation on the Lung tumor dataset.

Lung-RetinaNet is based on an improved RetinaNet's structure that is an object-detecting technique using context aggregation module, Huber loss and focal loss for data training. Two sub-networks exist along with a backbone for networks help the model in feature extraction. The sub-networks in RetinaNet are responsible for extracting features from different levels of the feature pyramid, which is a hierarchical representation of the image capturing features at various scales.

One of them is the classification sub-network that recognizes the class of image, whereas the other sub-network is known as regression which generates the bounding box (bbox). These sub-networks are comprised of four 3×3 convolutional layers having 256 channels at each layer. Then, the RELU activation function is employed to activate each layer's output. Further, classification is performed using an additional 3×3 conv. layer, which is activated using the sigmoid activation function for categorizing various objects k having specific anchors A per spatial locality. In the end, regression is performed by a conv. layer of 3×3 having four elements exhibiting offset among an estimated bbox and target bbox per spatial locality. We performed various experiments using different neural networks models with RetinaNet, such as VGG16, DenseNet201, EfficientNet82, and MobileNetv2. The results are reported in experimental section.

1) FOCAL LOSS
For the purpose of classification, RetinaNet utilized focal loss and computed, as shown in Equation 1. The focal loss is a cross-entropy having weights to overcome the issue of class imbalance. It drops out easy training examples during the training phase and considers the difficult ones.

$$CE\,(i, j) = \begin{cases} -\log(i)\,, & if\ j = 1 \\ -\log(1 - p)\,, & otherwise, \end{cases} \tag{1}$$

Here, CE refers to the cross entropy, and $j\epsilon\pm 1$ represents the target, whereas $i\epsilon[0, 1]$ exhibits the likelihood of the

estimated class having label $j = 1$.

$$FL(i_T) = -e(1 - i_T)^\gamma \log(i_T), \qquad (2)$$

where, i_T is the predicted probability of the correct class label that is equal to i for $j = 1$ and for $j = -1$ equal to $1 - i$. $-e(1 - i_T)^Y$ is the modulation element to the *CE* having $j\epsilon[0, 5]$, which can be adjusted to minimize the *CE*, and e is a hyper-parameter. e and γ represent 0.25 and 2, correspondingly for the significant performance [42].

2) HUBER LOSS
RetinaNet utilized the Huber loss function, computed as shown in Equation 3. Where e is a hyper-parameter that is adjustable and set as 1. d presents the distance between two vectors.

$$H(d) = \begin{cases} 0.5d^2, & if\ d < e \\ |d| - 0.5, & else\ if\ d \geq e, \end{cases} \qquad (3)$$

IV. EXPERIMENTAL EVALUATION
In this section, the suggested method is evaluated through various experiments, with a focus on metrics and environmental setup for performance evaluation.

A. TRAINING PARAMETERS
For training our proposed Lung-RetinaNet, we used both RetinaNet and Fused models, which employ similar parameters. The training parameters are almost same for original RetinaNet and Lung-RetinaNet as 41M parameters for ResNet50 and 64M parameters for ResNet101. The comparative graph for both models is shown in Figure 4. For each estimation level, the top 1000 estimations are considered for inference. Further, all the estimations are combined, and NMS is utilized, having a threshold of 0.5. Initially, the existing pre-trained models, such as ResNet-101 and ResNet-50 biases for ImageNet, are utilized. Weights of additional layers are assigned as $\sigma = 0.01$, $\mu = 0$, and bias b = 0. Additionally, the features are not shared among additional layers. For optimization, we utilized the Adam optimizer technique with a learning rate of 0.00001 and a clip-normalization of 0.01. The total epochs are 20, and iterations no. per epoch is 10k having batch size = 8. Total loss of regression and classification is computed as combined Huber and Focal losses. The training time required for both models was 19 to 70 hours. The time is high due to a single GPU-based system. The various configurations that are necessary for our proposed model are shown in Table 1. The model achieves 98.3% mAP on the dataset.

B. DATASET
In the proposed study, we used two datasets: Lung Image Database Consortium (LIDC-IDRI) [43] and 50 recorded CT scans of lungs from the Simba lung database [44]. We trained and tested our model using the LIDC-IDRI dataset. For cross-validation, we employed CT scan samples from the Simba database.
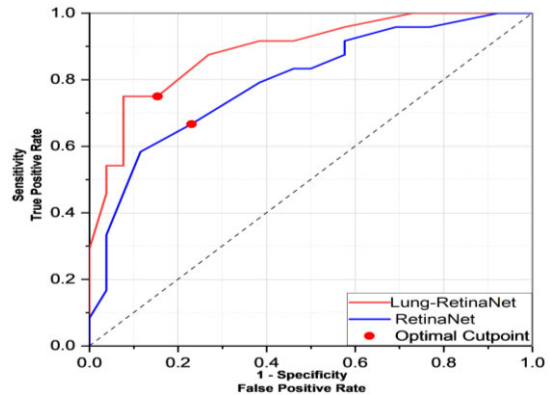

FIGURE 4. Comparative analysis of training among RetinaNet and Lung-RetinaNet.

TABLE 1. Parameters for lung-retinanet.

Parameters	Value	Description
Input Image		
Backbone	ResNet (50,101)	
Batch size	8	Mini- btach during training
GPU	Single	Nvidia GEFORCE GTX, 4 GB
Bias Initialization	Pre-trained	The models are trained on ImageNet
Focal loss	γ =2, e=0.25	Similar to loss function hyper-parameters
Epochs	20	LIDC-IDRI
IOU threshold	0.5	
optimizer	Adam	
NMS threshold	0.3, 0.5	LIDC-IDRI
Loss function	Focal and Huber loss	The aim is to minimize loss
Anchors	K-means and reference anchors	LIDC-IDRI
Total classes	2	LIDC-IDRI

In the LIDC-IDRI challenge, all CT scans were attained from the archive of The University of Chicago. All the images were from unique patients and taken from the Philips Brilliance Scanners having 1.15-mm thickness of the slice and "D" extra enhancing conv. kernel. The dataset included 1024 patients' CT images of lung nodules along with an XML file presenting results. The nodules identified by the radiologists were categorized as benign and malignant in the dataset. The total scans were 1018, among from we used 750 samples to train our proposed lung cancer detection system. The 450 samples belonged to the malignant class having various sizes of tumors, and the remaining 300 belonged to the benign class.

Second lung database comprises 50 CT images for the recognition of lung tumor. The CT scans were taken with a thickness of 1.25mm slice during a single breadth. The radiologist gave the location of nodules in the dataset. Various test sample images are shown in Figure 5.
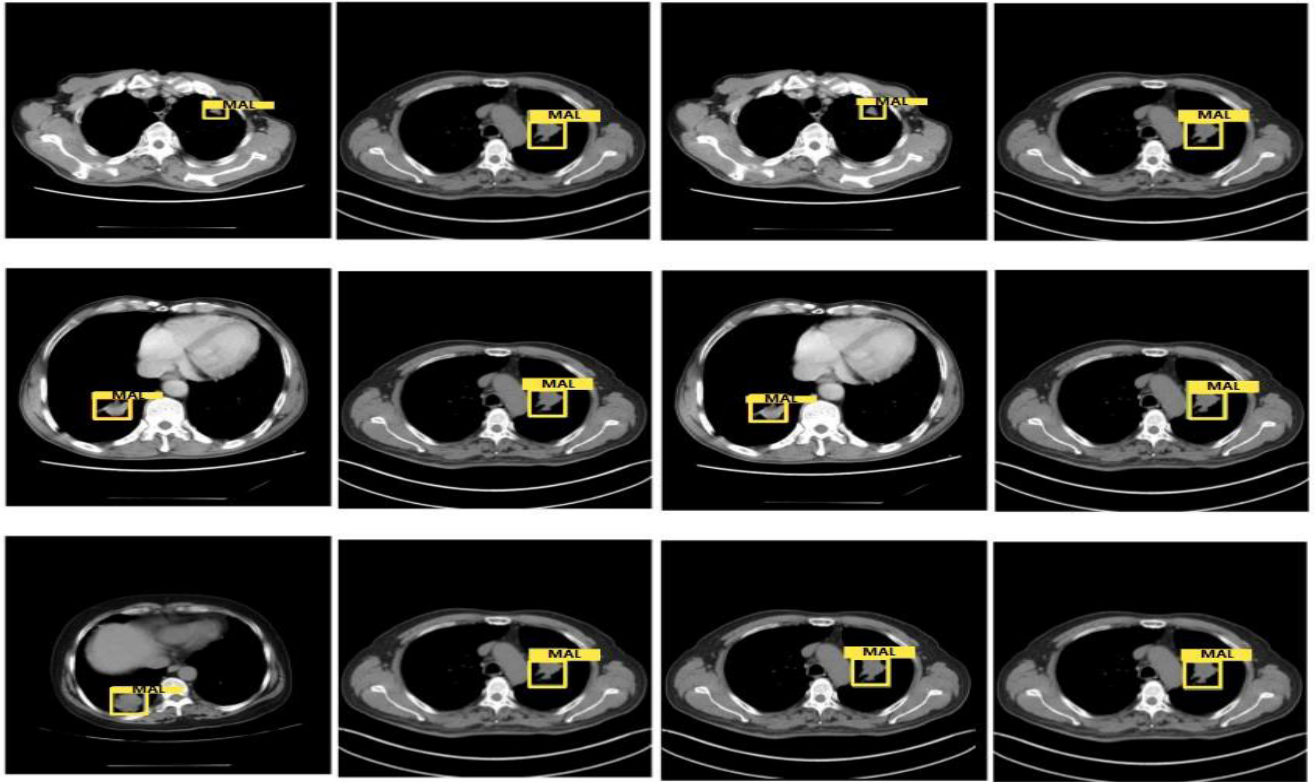
FIGURE 5. Some test samples from the simba dataset.

C. METRICS

For assessing the suggested system, we employed various parameters such as Precision, Accuracy, F1 Score, Recall, and Area under the curve (Auc). The equations are presented below.

$$Precision = \frac{TP}{TP + FP}, \tag{4}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$F1score = 2 * \frac{Precision * Recall}{Precision + Recall}, \tag{7}$$

The area under the curve (Auc) is computed as below:

$$Auc = \int_{i}^{j} f(x)\, dx, \tag{8}$$

D. ENVIRONMENTAL SETUP

Table 2 displays the details of the system used in conducting the experiments, which involved a geforce GTX GPU card with 4GB memory.

E. LOCALIZATION OF LUNG TUMOR

In this unit, the performance of our fusion-based RetinaNet is assessed utilizing four metrics: *DOI*, *TC*, no. of pixels, and area.

TABLE 2. Hardware specifications of the suggested method.

Hardware	Specifications
Graphic Processing Unit	NVIDIA GEFORCE GTX x 4
Random Access Memory	16 GB
Computer	GPU Server
Central Processing Unit	Intel Core i5

Two metrics frequently utilized in image segmentation are Texture Complexity (TC) and Degree of Interest (DOI). Texture Complexity measures the variety and intricacy of patterns within an image, while Degree of Interest evaluates the significance of different areas in the image with regards to the intended purpose. The equations of the used metrics are presented below:

$$DOI = {}^{\omega} \Big/ \frac{1 + \omega}{2} \tag{9}$$

$$TC = \sum_{x=1}^{m} \sum_{y=1}^{\sum_{y=1}^{m}(I' \cap l)} \Big/ \sum_{x=1}^{m} \sum_{y=1}^{n} (l' \cup l)' \tag{10}$$

$$area = \sum_{i=1}^{m} \sum_{j=1}^{n} I(i, j), \tag{11}$$

Our lung cancer detector's performance was evaluated by comparing the results of 50 images from each dataset, with

the TC varying between 0 and 1. The metrics were computed using the respective ground truth images, which were assessed by an expert pathologist for the LIDC-IDRI dataset. Table 3 reports the results of 10 images from the LIDC-IDRI dataset, showing that our proposed model achieved excellent results in terms of tumor localization.

TABLE 3. The localization results over LIDC_IDRI dataset.

Image	DOI	TC	Pixels	Area (nm2)
I1	0.93	0.91	2322	6.1 x 1012
I2	0.94	0.93	5311	6.1 x 1014
I3	0.97	0.93	3123	7.2 x 1013
I4	0.98	0.97	3233	3.9 x 1012
I5	0.98	0.94	3423	7.2 x 1013
I6	0.94	0.95	2932	4.3 x 1013
I7	0.95	0.94	3271	5.9 x 1012
I8	0.97	0.99	4113	5.3 x 1013
I9	0.98	0.98	3837	5.1 x 1014
I10	0.98	0.98	4344	8.2 x 1012

We have also performed a second experiment to assess the performance of our proposed Lung-RetinaNet using the Simba dataset. We compared our proposed model with various convolutional neural networks as the backbone network of RetinaNet using *DOI* and *TC*. We used DenseNet201, VGG16, EfficientNet82, ResNet101, and MobileNetV3. The results are shown in Table 4. We noticed that Efficient-Net82 and MobileNetv2 did not perform significantly. The lousy performance could be due to the varying feature maps among blocks. The VGG16 and DenseNet201 performed better than the MobileNetv2 and EfficientNet82. Although VGG16 performed better, the inference time was relatively high. The DenseNet201 performed better due to the direct connections among layers for feature extraction. Moreover, the best results were attained from our fused model, and the second highest results were achieved employing ResNet101. The reason could be that ResNet101 utilized residual blocks and feature maps transferred among layers, solving the vanishing gradient issue. Therefore, we have selected our feature fusion-based RetinaNet for lung tumor detection.

TABLE 4. The DOI and TC of various pre-trained backbone networks for the comparison with lung-retinanet using simba dataset.

Model	DOI	TC
DenseNet201	0.93	0.94
VGG16	0.92	0.94
EfficientNet82	0.87	0.86
ResNet101	0.96	0.95
MobileNetV2	0.90	0.88
Our proposed model	0.99	0.98

F. DETECTION AND CLASSIFICATION OF LUNG TUMOR

This section presents the results of our proposed model's classification performance using two datasets: LIDC-IDRI and Simba. The proposed model was trained on a total of 750 images and tested on 300 samples, with 150 images from each class: malignant and benign. Table 5 shows that the proposed model achieved significant results in classification. Additionally, we trained three object detectors: Faster RCNN, Mask RCNN, and Lung-RetinaNet, with the latter achieving the best results, including 99.8% accuracy, 99.3% recall, 99.4% precision, 99.5% F1 score, and 0.989 AUC. The next best classification results were obtained using Faster RCNN with 98.3% accuracy, followed by Mask RCNN with 97.3% accuracy. Moreover, we also compared the inference time for the techniques mentioned above, and it is clear from the outcomes that our suggested system is faster than others. The reason could be its one-stage detection mechanism and simple architecture of layers. Therefore, we believe that our proposed Lung-RetinaNet is an efficient tumor detector.

TABLE 5. The classification results of LIDC-IDRI dataset.

Algorithm	Accuracy(%)	Precision(%)	Recall (%)	F1 score (%)	AUC	Inference Time(sec)
Faster RCNN	98.3	98.9	98.1	98.5	0.912	88
Mask RCNN	97.3	97.2	96.8	97.2	0.900	45
Lung-RetinaNet	99.8	99.4	99.3	99.5	0.989	15

For the second experiment, we used the Simba dataset to cross-validate our proposed model. We used 50 images from the malignant class and achieved the best results with our proposed Lung-RetinaNet, including 99.3% accuracy, 99.1% recall, 99.2% precision, 99.3% F1 score, and 0.977 AUC. The next best classification results were obtained using Mask-RCNN with 99.0% accuracy. The lowest accuracy for cross-validation was achieved using the Faster RCNN classifier, with 98.1%. Table 6 provides more details on the cross-validation results. The better performance of Mask RCNN over Faster RCNN could be because it is easy to generalize and train. Moreover, it only adds up a little overhead to Faster RCNN, making it more efficient. Furthermore, our proposed Lung-RetinaNet is way better than Mask RCNN and Faster RCNN in efficiency as it takes minimum time for inference for both datasets.

G. ABLATION EXPERIMENTS

In this section, various methods for detecting small tumors in lung images will be presented. The Simba dataset was used for ablation study experiments. Initially, the effects of incorporating a feature fusion module, dilated context module, and lateral connections are examined. Afterwards, the impact of

TABLE 6. The classification results on simba.

Algo.	Accuracy(%)	Precision(%)	Recall(%)	F1 score (%)	AUC	Inference Time(sec)
Faster RCNN	98.1	97.9	97.4	97.2	0.901	81
Mask RCNN	99.0	98.9	98.2	98.0	0.921	39
Lung-RetinaNet	99.3	99.2	99.1	99.3	0.977	9

adjusting hyper-parameters of the focal loss on accuracy is discussed.

In the first ablation study, we performed experiment to analyze the performance due to fusion module. In original RetinaNet, fusion module and dilated context module is not present. We added the lateral connections and performed the Focal loss adjustments similar in RetinaNet, however, the mAP attained as 85.23% as shown In Table 7. It is clearly visible that when fusion module is used alone without dilated context module and lateral connections, the detection performance degrades due to information loss and down-sampling and model is unable to identify tiny tumors accurately. However, when lateral connections are added along with fusion module, the detection performance improves. Moreover, for our proposed Lung-RetinaNet, when we added the lateral connections, fusion module, dilated context block, and focal loss adjustments, we achieved remarkable results for tiny tumor detection.

TABLE 7. The performance evaluation of several modules on lung-retinanet's accuracy.

Algorithm	FPN	Lateral Connections	Fusion Module	Dilated Context Module	Focal Loss Adjustments	Accuracy
Original RetinaNet	√	√	-	-	√	85.23%
Lung-RetinaNet	-	-	√	-	-	84.33%
	-	√	√	-	-	88.23%
	-	√	√	√	-	90.7%
	-	√	√	√	√	99.3

The RetinaNet detector utilizes focal loss as its primary approach to address the class imbalance issue, with the

TABLE 8. The performance evaluation for several architectures.

Case	Accuracy(%)
1	87.34
2	99.2
3	85.5

hyper-parameters α and γ weighting factors set to 0.25 and 2, respectively, resulting in optimal RetinaNet performance. However, for Lung-RetinaNet, this setting is not the most effective, and modifying the focal loss hyper-parameters could enhance Lung-RetinaNet accuracy. Table 9 demonstrates that adjusting α and γ to 0.25 and 2.5, respectively, yields the highest Lung-RetinaNet performance, resulting in a 1.5 point increase in accuracy.

TABLE 9. The performance evaluation for adjusting hyper-parameters of focal loss on lung-retinanet.

Focusing Factor α	Weighting Factor γ	Accuracy(%)
2	0.3	82.6
3	0.3	84.3
3.5	0.25	92.4
2.5	0.25	99.3
2	0.25	91.6
3	0.25	90.3

H. COMPARISON WITH EXISTING TECHNIQUES

We compare our proposed detector for lung tumors with existing techniques. The outcomes are reported in Table 10. It is clearly seen that our suggested detector achieved 99.8% detection accuracy, and the results have been validated by two radiologists. Xie et al. [45] utilized 888 samples from the LUNA16 dataset and developed a system based on two modules. Firstly, they detected the locations in images using an improved Faster RCNN. Secondly, they trained three models for false positive reduction. The system became somehow complex and achieved 88.17% detection accuracy, which is lower than others. Chao et al. [46] developed a CNN-based technique for detecting lung tumor nodules after pre-processing images from LUNA16 and Kaggle datasets, attaining 92% accuracy. Moreover, Eali et al. [47] provided a multi-view network approach along with weighted gradient activation for the binary classification of lung tumors. The proposed model was lightweight; however, a tremendous amount of information was lost due to a class invariant problem in the max-pooling layers of the proposed CNN. Nevertheless, they attained a considerable detection accuracy of 97.17%. Comparatively, our proposed Lung-RetinaNet attained 99.8% detection accuracy while retaining the key

TABLE 10. The comparative analysis for detection with existing techniques using lidc-idri dataset.

Method	No. of samples	Training(%)	Testing(%)	Year	Accuracy(%)
Xie et al.,[45]	888	90	10	2019	88.17
Chao et al.,[46]	757	90	10	2019	92
Eali et al.,[47]	2000	-	-	2021	97.17
Lung-RetinaNet	1100	70	30	2022	99.8

information of lung CT images. Moreover, our proposed detector is based on one stage that is easy to use and modify. Therefore, our proposed Lung tumors detector and classifier achieved better results than existing techniques in terms of complexity and accuracy.

TABLE 11. The comparative analysis of detection with existing techniques on the basis of accuracy, sensitivity, and specificity.

Method	Scheme	Accuracy(%)	Sensitivity(%)	Specificity(%)
Al-shabi et al.[48]	3-D CNN	95.28	94.33	-
Abid et al.[49]	3-D CNN	97.10	97.50	96.60
Nibal et al. [50]	2-D CNN	80.90	91.07	88.64
Lung-RetinaNet	Object Detector	99.3	99.1	98.8

We selected RetinaNet for the lung cancer detection as it offers several benefits over other models. Firstly, it achieves superior accuracy on several object detection benchmarks, even when dealing with severe class imbalance. This is due to its design, which enables it to identify objects with high accuracy. Secondly, RetinaNet boasts faster inference speeds compared to traditional two-stage detection models. It uses a single-stage detection process, which means that it requires only one forward pass through the network to produce object detections. Thirdly, RetinaNet is easy to train due to its simple architecture and the use of focal loss, which simplifies the learning process by focusing on hard examples. Fourthly, RetinaNet is highly effective at detecting small objects due to the use of sub networks and usage of the context block that provides multi-scale representations of the input image.

Lastly, RetinaNet has a high recall rate, meaning that it can identify most of the objects in the image, even when they are small or partially occluded. Overall, RetinaNet is a powerful model for object detection that provides high accuracy, speed, and recall, making it a popular choice for a wide range of applications.

I. COMPUTATIONAL COST

In the lung tumor segmentation task, our one-stage RetinaNet detector has proven to outperform previous methods, which included one- and two-stage detectors such as faster R-CNN, R-CNN, and SSD321. Faster R-CNN had an average precision of 36.4% at the first five scale (400-800) pixels, with an inference time of 192ms, while R-CNN and SSD321 had an average precision of 35% and 39%, respectively, with inference times of 95ms and 82ms. By comparison, our proposed methods, which included RetinaNet with ResNet-50 and ResNet-101, achieved average precisions of 25% and 31%, respectively, at the first five scale (400-800) pixels. Additionally, we achieved an inference time of 57ms and 51ms, respectively, as illustrated in Figure 6.
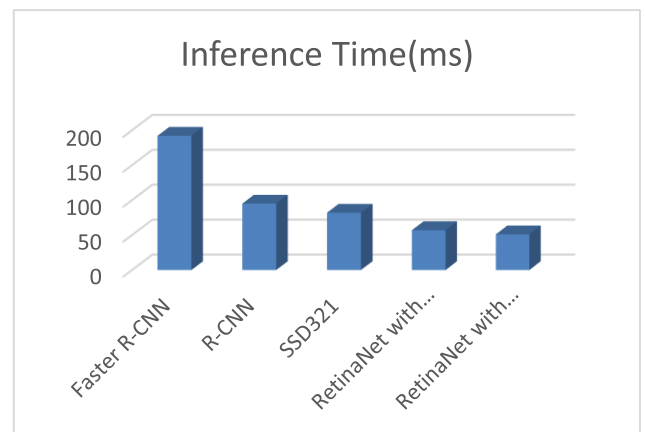
FIGURE 6. Inference time for several models on simba dataset.

V. CONCLUSION

This work suggests a novel and robust lung tumor detection method based on fused RetinaNet. Our proposed model utilizes CT scans for the training and testing of the model. We have introduced a context aggregation module instead of FPN in the original RetinaNet. Our proposed model comprises a ResNet as a backbone network for extracting features from input along with dilated contextual block-based fusion, categorization and regression sub-networks. Due to an improved structure of RetinaNet, it precisely detects tiny lung tumors. More specifically, a fusion block is down-sampled and combined with a context-dilated module at all networks level. This connection enhances the capability of the proposed network to extract valuable features. Moreover, it also improves the localization performance of the proposed network. Our proposed methodology achieved 99.8% accuracy, 99.3% recall, 99.4% precision, 99.5% F1 score,

and 0.989 AUC. We evaluated our proposed method and compared its results with state-of-the-art DL-based methods, which revealed that our technique outperforms existing systems.Thus, we believe that our proposed system can be utilized by medical experts to identify the tumor at early stages.

Our proposed method performed significantly for lung tumor detection, however, it faced some challenges which are required to be addressed in the future. First, the ability to identify tiny tumors is reduced due to the limited resolution of the input images. Second, if the images had a lot of clutter or noise, the system was not able to differentiate the background tissues from tiny tumors.

In the future, our objective is to employ our suggested method for the multi-classification of various cancers, such as skin, bone, etc. Moreover, the training time of the proposed model was considerably high; however, it could be improved using a multi-GPU-based training system. This will also improve the inference time of the proposed system to deploy it clinically. Furthermore, we will add multi-modal information as input to our system such as combining imaging data with other forms of clinical data, (genomic or proteomic data) to improve the accuracy of tumor detection.

DATA AVAILABILITY
The data used for this research is available publically.

REFERENCES

[1] Q. Abbas and M. M. Q. A. Yasin, "Lungs cancer detection using convolutional neural network," *Int. J. Recent Adv. Multidisciplinary Topics*, vol. 3, no. 4, pp. 90–92, 2022.

[2] A. Khan, "Identification of lung cancer using convolutional neural networks based classification," *Turkish J. Comput. Math. Educ. (TURCOMAT)*, vol. 12, no. 10, pp. 192–203, 2021.

[3] C. Thallam, A. Peruboyina, S. S. T. Raju, and N. Sampath, "Early stage lung cancer prediction using various machine learning techniques," in *Proc. 4th Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Nov. 2020, pp. 1285–1292.

[4] Y. Xie, W.-Y. Meng, R.-Z. Li, Y.-W. Wang, X. Qian, C. Chan, Z.-F. Yu, X.-X. Fan, H.-D. Pan, C. Xie, Q.-B. Wu, P.-Y. Yan, L. Liu, Y.-J. Tang, X.-J. Yao, M.-F. Wang, and E. L.-H. Leung, "Early lung cancer diagnostic biomarker discovery by machine learning methods," *Translational Oncol.*, vol. 14, no. 1, Jan. 2021, Art. no. 100907.

[5] X. Liu, K.-W. Li, R. Yang, and L.-S. Geng, "Review of deep learning based automatic segmentation for lung cancer radiotherapy," *Frontiers Oncol.*, vol. 11, p. 2599, Jul. 2021.

[6] R. Mahum, S. U. Rehman, O. D. Okon, A. Alabrah, T. Meraj, and H. T. Rauf, "A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging," *Electronics*, vol. 11, no. 1, p. 26, Dec. 2021.

[7] R. Mahum, "A novel framework for potato leaf disease detection using an efficient deep learning model," *Hum. Ecol. Risk Assessment, An Int. J.*, vol. 29, no. 2, pp. 303–326, 2022.

[8] M. Masood, R. Maham, A. Javed, U. Tariq, M. A. Khan, and S. Kadry, "Brain MRI analysis using deep neural network for medical of internet things applications," *Comput. Electr. Eng.*, vol. 103, Oct. 2022, Art. no. 108386.

[9] M. J. Akhtar, R. Mahum, F. S. Butt, R. Amin, A. M. El-Sherbeeny, S. M. Lee, and S. Shaikh, "A robust framework for object detection in a traffic surveillance system," *Electronics*, vol. 11, no. 21, p. 3425, Oct. 2022.

[10] N. Kalaivani, "Deep learning based lung cancer detection and classification," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 994, no. 1, 2020, Art. no. 012026.

[11] L. Ale, N. Zhang, and L. Li, "Road damage detection using RetinaNet," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 5197–5200.

[12] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[13] J. Chaki and M. Woźniak, "Deep learning for neurodegenerative disorder (2016 to 2022): A systematic review," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104223.

[14] M. Woźniak, M. Wieczorek, and J. Siłka, "BiLSTM deep neural network model for imbalanced medical data of IoT systems," *Future Gener. Comput. Syst.*, vol. 141, pp. 489–499, Apr. 2023.

[15] M. Woźniak, J. Siłka, and M. Wieczorek, "Deep neural network correlation learning mechanism for CT brain tumor detection," *Neural Comput. Appl.*, 2021, doi: 10.1007/s00521-021-05841-x.

[16] R. Mahum and S. Aladhadh, "Skin lesion detection using hand-crafted and DL-based features fusion and LSTM," *Diagnostics*, vol. 12, no. 12, p. 2974, Nov. 2022.

[17] A. C. Borczuk, "Benign tumors and tumorlike conditions of the lung," *Arch. Pathol. Lab. Med.*, vol. 132, no. 7, pp. 1133–1148, Jul. 2008.

[18] Q. Song, L. Zhao, X. Luo, and X. Dou, "Using deep learning for classification of lung nodules on computed tomography images," *J. Healthcare Eng.*, vol. 2017, pp. 1–7, Aug. 2017.

[19] Y.-J. Yu-Jen Chen, K.-L. Hua, C.-H. Hsu, W.-H. Cheng, and S. C. Hidayati, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," *OncoTargets Therapy*, vol. 8, p. 2015, Aug. 2015.

[20] E. Kurniawan, P. Prajitno, and D. S. Soejoko, "Computer-aided detection of mediastinal lymph nodes using simple architectural convolutional neural network," *J. Phys., Conf. Ser.*, vol. 1505, no. 1, Mar. 2020, Art. no. 012018.

[21] P. Rao, N. A. Pereira, and R. Srinivasan, "Convolutional neural networks for lung cancer screening in computed tomography (CT) scans," in *Proc. 2nd Int. Conf. Contemp. Comput. Informat. (IC3I)*, Dec. 2016, pp. 489–493.

[22] F. Ciompi, K. Chung, S. J. van Riel, A. A. A. Setio, P. K. Gerke, C. Jacobs, E. T. Scholten, C. Schaefer-Prokop, M. M. W. Wille, A. Marchianò, U. Pastorino, M. Prokop, and B. van Ginneken, "Towards automatic pulmonary nodule management in lung cancer screening with deep learning," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, Apr. 2017.

[23] Y. Gordienko, "Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer," in *Proc. Int. Conf. Comput. Sci., Eng. Educ. Appl.* Cham, Switzerland: Springer. 2018, pp. 638–647.

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

[25] D. Ardila, A. P. Kiraly, S. Bharadwaj, B. Choi, J. J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D. P. Naidich, and S. Shetty, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Med.*, vol. 25, no. 6, pp. 954–961, Jun. 2019.

[26] S. Chen, Y. Han, J. Lin, X. Zhao, and P. Kong, "Pulmonary nodule detection on chest radiographs using balanced convolutional neural network and classic candidate detection," *Artif. Intell. Med.*, vol. 107, Jul. 2020, Art. no. 101881.

[27] A. Hosny, C. Parmar, T. P. Coroller, P. Grossmann, R. Zeleznik, A. Kumar, J. Bussink, R. J. Gillies, R. H. Mak, and H. J. W. L. Aerts, "Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study," *PLOS Med.*, vol. 15, no. 11, Nov. 2018, Art. no. e1002711.

[28] Y. Xu, "Deep learning predicts lung cancer treatment response from serial medical ImagingLongitudinal deep learning to track treatment response," *Clinical Cancer Res.*, vol. 25, no. 11, pp. 3266–3275, 2019.

[29] W. Zhao, Y. Sun, K. Kuang, J. Yang, G. Li, B. Ni, Y. Jiang, B. Jiang, J. Liu, and M. Li, "ViSTA: A novel network improving lung adenocarcinoma invasiveness prediction from follow-up CT series," *Cancers*, vol. 14, no. 15, p. 3675, Jul. 2022.

[30] K. Kuan, M. Ravaut, G. Manek, H. Chen, J. Lin, B. Nazir, C. Chen, T. C. Howe, Z. Zeng, and V. Chandrasekhar, "Deep learning for lung cancer detection: Tackling the kaggle data science bowl 2017 challenge," 2017, *arXiv:1705.09435*.

[31] W. Ausawalaithong, A. Thirach, S. Marukatat, and T. Wilaiprasitporn, "Automatic lung cancer prediction from chest X-ray images using the deep learning approach," in *Proc. 11th Biomed. Eng. Int. Conf. (BMEiCON)*, Nov. 2018, pp. 1–5.

[32] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, "Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning," *Nature Med.*, vol. 24, no. 10, pp. 1559–1567, Oct. 2018.

[33] S. Xu, J. Guo, G. Zhang, and R. Bie, "Automated detection of multiple lesions on chest X-ray images: Classification using a neural network technique with association-specific contexts," *Appl. Sci.*, vol. 10, no. 5, p. 1742, Mar. 2020.

[34] B. S. A. El-Wahab, M. E. Nasr, S. Khamis, and A. S. Ashour, "BTC-fCNN: Fast convolution neural network for multi-class brain tumor classification," *Health Inf. Sci. Syst.*, vol. 11, no. 1, p. 3, Jan. 2023.

[35] A. Su, F. R. Pp, A. Abraham, and D. Stephen, "Deep learning-based BoVW–CRNN model for lung tumor detection in nano-segmented CT images," *Electronics*, vol. 12, no. 1, p. 14, Dec. 2022.

[36] K. S. Pradhan, P. Chawla, and R. Tiwari, "HRDEL: High ranking deep ensemble learning-based lung cancer diagnosis model," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 118956.

[37] Z. Zhu, X. He, G. Qi, Y. Li, B. Cong, and Y. Liu, "Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI," *Inf. Fusion*, vol. 91, pp. 376–387, Mar. 2023.

[38] I. U. Haq, H. Ali, H. Y. Wang, C. Lei, and H. Ali, "Feature fusion and ensemble learning-based CNN model for mammographic image classification," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 6, pp. 3310–3318, Jun. 2022.

[39] M. Jiang, F. Zhai, and J. Kong, "Sparse attention module for optimizing semantic segmentation performance combined with a multi-task feature extraction network," *Vis. Comput.*, vol. 38, no. 7, pp. 2473–2488, Jul. 2022.

[40] M. Ju, M. Luo, P. Zhang, M. He, and H. Luo, "A simple and efficient network for small target detection," *IEEE Access*, vol. 7, pp. 85771–85781, 2019.

[41] S. Chen, W. Ma, and L. Zhang, "Dual-bottleneck feature pyramid network for multiscale object detection," *J. Electron. Imag.*, vol. 31, no. 1, Jan. 2022, Art. no. 013009.

[42] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[43] S. G. Armato, L. Hadjiiski, G. D. Tourassi, K. Drukker, M. L. Giger, F. Li, G. Redmond, K. Farahani, J. S. Kirby, and L. P. Clarke, "LUNGx challenge for computerized lung nodule classification: Reflections and lessons learned," *J. Med. Imag.*, vol. 2, no. 2, Jun. 2015, Art. no. 020103.

[44] A. P. Reeves and A. Biancardi, "The SIMBA image management and analysis system," Cornell Univ., Ithaca, NY, USA, Tech. Rep., 2012. Accessed: May 30, 2023. [Online]. Available: https://www.via.cornell.edu/visionx/v4/vbasics/simba-api.pdf

[45] H. Xie, D. Yang, N. Sun, Z. Chen, and Y. Zhang, "Automated pulmonary nodule detection in CT images using deep convolutional neural networks," *Pattern Recognit.*, vol. 85, pp. 109–119, Jan. 2019.

[46] C. Zhang, "Toward an expert level of lung cancer detection and classification using a deep convolutional neural network," *Oncologist*, vol. 24, no. 9, pp. 1159–1165, 2019.

[47] E. S. N. Joshua, D. Bhattacharyya, M. Chakkravarthy, and Y.-C. Byun, "3D CNN with visual insights for early detection of lung cancer using gradient-weighted class activation," *J. Healthcare Eng.*, vol. 2021, pp. 1–11, Mar. 2021.

[48] M. Al-Shabi, K. Shak, and M. Tan, "ProCAN: Progressive growing channel attentive non-local network for lung nodule classification," *Pattern Recognit.*, vol. 122, Feb. 2022, Art. no. 108309.

[49] M. M. N. Abid, T. Zia, M. Ghafoor, and D. Windridge, "Multi-view convolutional recurrent neural networks for lung cancer nodule identification," *Neurocomputing*, vol. 453, pp. 299–311, Sep. 2021.

[50] A. Nibali, Z. He, and D. Wollersheim, "Pulmonary nodule classification with deep residual networks," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 10, pp. 1799–1808, Oct. 2017.

RABBIA MAHUM received the B.Sc. degree in computer science from COMSATS University Islamabad, Wah Campus, in 2015, and the M.S. degree in computer science from the University of Engineering and Technology (UET) at Taxila, Taxila, in 2018, where she is currently pursuing the Ph.D. degree. Her research interests include computer vision, deep fake audio detection, and medical imaging.

ABDULMALIK S. AL-SALMAN received the B.S. degree in computer science from King Saud University, Saudi Arabia, in 1988, the master's degree in computer science from Georgia University, Athens, GA, USA, in 1992, and the Ph.D. degree from the Department of Computer and Information Sciences, Oklahoma State University, Stillwater, OK, USA, in 1996. He was the Head of the Department of Computer Science, the Computing College's Vice Dean, and the Scientific Council General Secretary with King Saud University, Riyadh, Saudi Arabia, where he is currently a Full Professor with the Department of Computer Science. He has more than 130 publications. His current research interests include AI applications, natural language processing, assistive technologies, and mobile computing.

• • •