Accurate Detection of Pulmonary Nodule and False Positive Reduction with Faster R-CNN and ResNet Models

Taghreed Alzahrani, Suhare Solaiman

Department of Computer Sciences, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia

Summary

The annual incidence of lung cancer has significantly increased. The detection of pulmonary nodule is critical for the diagnosis of lung cancer. In recent years, deep learning models have demonstrated encouraging results, surpassing traditional machine learning methods across different fields Researchers have used several deep learning methods to enhnace the effectiveness of computer-aided detection (CAD) systems that use computed tomography (CT) images for lung cancer diagnosis. Despite the capability of deep-learning models to detect pulmonary nodules, they still encounter significant challenges in terms of a high rate of false positive. In this study, we propose a model for pulmonary nodule detection and false positive reduction using CT images. This model combines a faster region-based convolutional neural network (Faster R-CNN) for accurate detection and a residual network (ResNet50) to reduce the false positive rate. Our model was evaluated using the public LUng Nodule Analysis (LUNA16) dataset, and the experimental results revealed that our model outperformed existing methods in the literature regarding competition performance metric (CPM) at 95.1%. An additional evaluation, the10-Fold Cross-Validation approach, was used to evaluate the proposed model, which demonstrated the reliability and capability of the generalization of our model, with an average accuracy of 91.1%.

Keywords:
Pulmonary nodule detection; false positive reduction; Faster R-CNN model; ResNet model; CT images; CAD

1. Introduction

Lung cancer is a major global health issue affecting a substantial number of individuals around the world [1]. Based on a recent report by the world health organization (WHO), lung cancer continues to be the leading cause of cancer-associated deaths, accounting for roughly 18.4% of all such fatalities in 2020 [2]. The low survival rate can be credited to the fact that the clinical manifestations of lung cancer usually appear in the advanced stages. Therefore, early detection is crucial, as it can raise the five-year survival rate to over 90% and increase the likelihood of treatment.

Early-stage lung cancer typically appears as lung nodules. Pulmonary nodules are spherical or irregular growths in the lungs with a maximum diameter of under 3 cm that can display benign or malignant features, which can be assessed using different medical imaging modalities.

Radiologists face a continuous diagnostic challenge owing to the complex morphologies and multiple structures of certain nodules, which often adhere to the trachea, blood vessels, and other organs [3]. The morphological features of the candidate nodules are similar to those found in the chest area, such as lymph nodes, airways, and blood vessels [4]. Figure 1, shows examples of pulmonary nodules and non-nodules with complex and similar morphological features. This makes reducing false positive in candidate nodules significantly more challenging.
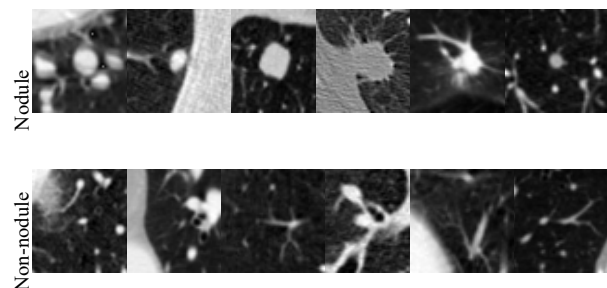

Fig. 1 Illustrations of pulmonary nodules and non-nodules.

Computed tomography (CT) is a traditional diagnostic imaging method that facilitates the observation and analysis of lung nodules within CT images by radiologists. It is frequently used for identifying pulmonary nodules because of its faster imaging speed, reduced cost, and better density resolution in comparison to positron emission tomography (PET) and magnetic resonance imaging (MRI) [5]. In clinical practice, radiologists are tasked with reviewing numerous CT scans for a single patient. The use of manual evaluation alone can be difficult for radiologists. This could lead to disastrous results, such as false positive (identifying normal tissue as nodules) or false negatives (not finding any nodules at all) in pulmonary nodule diagnosis [6]. As a result, the development of computer-aided detection (CAD) systems has emerged as an effective solution for the automatic identification of pulmonary nodules in CT images. Common CAD systems for lung nodule diagnosis consist of two main stages: (1) candidate nodule detection and (2) false positive reduction (FPR). In the first stage, the CAD system scans the patient's CT images to identify as many

potential nodule candidates as possible, highlighting their locations. While this stage typically exhibits high sensitivity, it also leads to a significant number of false positive detections. To mitigate this issue, the second stage involves inputting the potential nodule candidates into a binary classifier to reduce the false positive rate [6].

Typically, traditional CAD systems rely on manual generation of nodule characteristics including morphological and intensity thresholding [7], to generate initial nodule candidates. Subsequently, they employed traditional classifiers, such as the k-nearest neighbor (KNN) [8] or support vector machine (SVM) [9], to reduce false positive, which involves a combined analysis of the location, size, shape, density, texture, and gradient features. However, these approaches frequently yield unsatisfactory results, mainly due to the significant variability in the appearance of pulmonary nodules. This variability makes it challenging to differentiate nodules from other lung structures, such as blood vessels or lung tissue.

The rapid growth of artificial intelligence (AI), availability of extensively annotated datasets, and graphic processing unit (GPU) computing efficiency have played crucial roles in the advancement of deep learning algorithms, especially deep convolutional neural networks (CNNs), which excel in computer vision and medical image processing. Several deep learning methods have been employed in medical image processing field, CNN [10] and transfer learning (TL), which are used for disease detection [11], classification [12], and segmentation [13]. The ability of a CNN to automatically extract features of pulmonary nodules stems from its capacity to learn and recognize image characteristics from training data compared to traditional machine learning (ML) approaches, CNNs outperform automated lung nodule diagnosis. Consequently, an increasing number of studies demonstrating the efficacy of CNN architectures have made them the primary method for diagnosing pulmonary nodules [14].

Deep-learning-based object detection and classification methods are now widely used in CAD systems to detect and classify candidate nodules. Typically, nodule detection involves two phases, namely, regional proposal generation and false-positive reduction. The faster region-based convolutional neural network (Faster R-CNN) [15] is a two-stage algorithm that identifies numerous regions of interest (ROIs) and subsequently classifies and refines the bounding box for each region. As a consequence, the Faster R-CNN outperformed existing approaches such as SPPnet, R-CNN, and Fast R-CNN to detect pulmonary nodule [16]. For false-positive reduction, recent studies have utilized different approaches in the literature based on transfer learning, such as the Residual Network (ResNet) model [17], which exhibits exceptional results in computer vision and medical imaging applications and can distinguish complex patterns of pulmonary nodules [18].

Researchers [12][19] developed a CAD system based on a deep residual network to classify different lung nodules instead of using a complex method that relies on traditional CT image processing to manually extract the features of nodules. The results demonstrated the effectiveness of this approach to classify different lung nodules. Although DL models can detect pulmonary nodules, they still have a significant issue in terms of high false positive rates [20].

In this study, to solve the significant challenges related to accurate nodule detection and the high rate of false positive in deep learning models, we proposed a model to detect pulmonary nodules on CT images and minimize the false positive rate. This approach combines the strengths of a Faster R-CNN for accurate pulmonary nodule detection and a ResNet50 model for false positive reduction by classifying candidate nodules into nodules or non-nodules. The proposed model is assessed using the publicly available lung nodule analysis (LUNA16) dataset [21]. The key contributions of our study are outlined as follows:

- Comprehensive details are provided for the data preparation process, highlighting the necessary procedures to ensure that high-quality data are prepared for the training model.

- A framework is proposed for accurate nodule detection using the Faster R-CNN. For this task, we adapt the Faster R-CNN architecture by including the ResNet network as the backbone for feature extraction. We designed five anchors of different sizes in the RPN, each with three aspect ratios, to accommodate different nodule sizes.

- A residual network based on the ResNet50 architecture is proposed to reduce false positive rates while maintaining a high sensitivity.

- Our model achieved excellent results compared with state-of-the-art methods in reducing false positive, with a high competition performance metric (CPM) score of 95.1% on the LUNA16 dataset.

The rest of this paper is organized as follows. Section 2 presents an overview of the relevant literature and previous research on pulmonary nodule diagnoses. Section 3 provides a detailed explanation of the materials and methods of the proposed model. Section 4 outlines the experimental design and findings of this study, while Section 5 discusses the conclusions, limitations, and potential avenues for future research.

2. Related Work

In the last decade, there have been notable advancements in identifying pulmonary nodules, reducing false positive results, and classifying them using CT

images. These advancements have primarily concerned on integrating machine learning and deep learning models into computer-aided detection (CAD) systems. In this literature review, we discuss notable studies and advancements that have shaped the role of both traditional ML and deep-learning models in the development of CAD systems.

2.1 Traditional ML in CAD Systems

Traditional ML methods were the cornerstone of pulmonary nodule CAD systems before the ascent of deep learning as a dominant force in the field. These ML techniques are part of artificial intelligence, which develops algorithms and statistical models to enhance the abilities of computer systems by learning from data without explicit programming [22]. In the following section, we present CAD systems developed based on traditional classifiers.

Subsolid pulmonary nodules are less common than solid lung nodules but have a much greater incidence of malignancy. To address this issue, [23] presented a system based on CAD for subsolid nodules in CT images with a focus on lung cancer screening trials. The system uses 128 characteristics such as intensity, shape, texture, and context, to enhance the classification performance. The system obtained a sensitivity of 80% with a low false positive rate and could identify subsolid nodules that were not found in the screening database, thereby demonstrating its potential for accurate detection. Current CAD systems are prone to potentially missing true nodules and generating excessive false positive. To address the issue of identifying actual lung nodules, [24] presented a cascaded SVM classifier for lung nodule detection that sequentially performs two classification tasks to select the most likely candidates from a large pool of potential candidates through thresholding, morphological analysis, and feature selection, which aims to reduce the chance of falsely rejecting true nodules.

2.2 Learning-Based CAD Systems

Deep learning is a branch of machine learning and artificial intelligence that employs artificial neural networks to address and model complex problems. It utilizes multiple layers to automatically learn and hierarchically represent data. Existing literature demonstrates that deep learning has produced significant achievements in medical imaging, particularly in the detection and classification of lung nodules [25][26]. In this section, we present a common type of deep learning known as convolutional neural networks (CNNs), which serve various functions, including computer vision tasks. Convolutional layers are automatically employed to extract spatial patterns in images, which enables them especially effective for tasks such as image classification, object detection, and image generation. Several studies

have proven that CNNs are crucial for attaining cutting-edge performance in diverse computer vision applications, particularly in medical imaging.

A novel pulmonary nodules CAD system employed multi-view convolutional networks (ConvNets) to learn recognizable features from the training data. The system extracted 2-D patches from different planes of each candidate by combining the outputs using a fusion method [27]. In addition, [20] proposed an innovative automated framework for spotting pulmonary nodules that utilizes a 2D CNN for CT scans. The framework adjusts the Faster R-CNN structure, trains models for different slices, and uses a boosting architecture to reduce false positive rate. Significant experiments on LUNA16 show a sensitivity of 86.42% for nodule candidate spotting.

To tackle the difficulty of candidate nodule detection due to significant variations in nodule morphology and the potential for confusing them with nearby organs, introduced a new multi-scale gradient integration CNN (MGI-CNN) [4]. The network employs multi-scale inputs that provide diverse contextual information, extracts abstract features from various input scales, and integrates multi-stream features in an end-to-end manner. Moreover, a study in [28] was introduced to address the significant variation in the morphological characteristics of nodules. This study presented an innovative multi-scale heterogeneous 3D CNN (MSH-CNN) specifically designed for CT scanned images. The framework incorporates multi-scale 3D nodule blocks that capture diverse situational characters, extracts feature representations from two distinct branches of the 3D CNN and integrates these features in the second step using back-propagation weights. Additionally, [29] proposed an anchor-free approach for reliable pulmonary nodule spotting, which features a novel representation of spotted nodules based on their 3D center locations. This approach employs a two-stream network that minimizes false positive candidates by integrating information from the streams of different images and the motion-history. In addition, [18] proposed a new CAD system to detect lung nodules based on the Faster R-CNN model and utilized an adaptive anchor box.

In general, CAD systems can generate more false positive results because of unbalanced datasets used for testing and training. Based on this, the study [30] indicates that a filtering step was implemented to eliminate irrelevant images out of the dataset, demonstrating the effectiveness of this method in detection. The method involves screening pulmonary nodules from comprehensive lung CT images and employing Faster R-CNN to accurately detect the location of the nodules. This approach may assist in the early diagnosis of lung cancer and demonstrates a reduction in false positive. To overcome the issue of network overfitting, [31] presented a 3D automated detection method for pulmonary nodules

that utilizes multi-scale attention networks. This approach uses the multi-scale properties of the nodules. [32] proposed an approach to spot lung nodules in three-dimensional CT scanned images. This approach is based on parallel pooling and dense blocks, and he study comprises two components, namely candidate nodule extraction and false positive reduction.

Table 1 summarizes related works on the development of a CAD system for pulmonary nodule diagnosis using traditional ML and DL techniques, most of these studies consist of two stages: (1) for pulmonary nodule detection and (2) for false s reduction and use a CPM as primary metrics to evaluate their work.

Table 1: The summary of the related works.

Ref.	Approach	Dataset	No. of Cases	Sensitivity at false positive value	CPM score
[23]	Detection pipeline, Multiple ML classifier	NELSON trail	7557	Sen: 80% at 1 FP/scan Sen: 88% at 4 FP/scan	-
[24]	Cascaded SVM	Private Dataset	3278	Sen: 85.9% at 2.5 FP/volume	-
[27]	ConvNets	LIDC-IDRI	888	Sen: 85.4% at 1 FP/scan Sen: 90.1% at 4 FP/scan	0.824
[20]	Faster R-CNN with 2RPN	LUNA16	888	Sen: 73.4 % at 0.125 FP/scan Sen: 74.4 % at 0.25 FP/scan	0.790
[4]	MGI-CNN	LUNA16	888	Sen: 91.2% at 1 FP/scan	0.908
[28]	MSH-CNN	LUNA16	888	Sen: 91.7% at 2 FP/scan	0.874
[29]	3D-CNN	LUNA16	888	Sen: 96.1% at 8 FP/scan	0.906
[18]	Faster R-CNN with adaptive anchors	LUNA16	888	Sen: 93% at 1 FP/scan	0.882
[31]	3D deep CNN	LUNA16	888	Sen: 94.1% at 1 FP/scan	0.927
[32]	PPD-UNet + DBHA-NET	LUNA16	888	Sen: 9.21% at 1 FP/scan	0.91

3. Proposed Work

The proposed model composes of two stages for accurate nodule detection: (1) pulmonary nodule detection by a Faster R-CNN model and (2) false positive reduction by utilizing the residual convolutional network (ResNet50) to distinguish true nodules from non-nodules. Figure 2, presents the general framework of the proposed model that utilizes faster R-CNN and ResNet models for accurate detection of pulmonary nodule and false positive reduction.
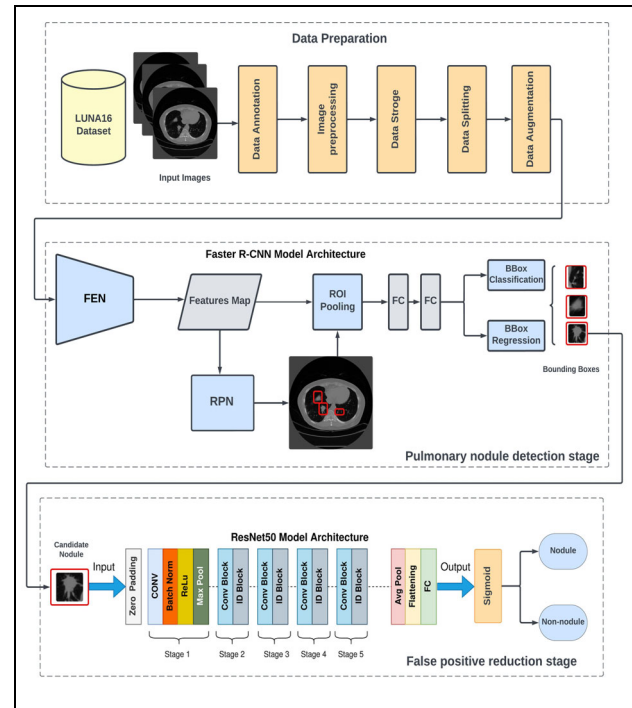

Fig. 2 The general framework of the proposed model.

3.1   Pulmonary Nodule Detection

During the detection stage, the Faster R-CNN model is trained for detecting pulmonary nodules using CT images. Faster R-CNN [15] is a target detection technique that identifies regions of interest (ROI) from the provided images. In contrast to R-CNN, SPP-NET, and Fast R-CNN, these methods utilize selective search algorithms to retrieve the target regions [30]. Faster R-CNN directly calculates the candidate boxes using the RPN, which greatly enhances target detection speed compared to earlier methods. The network architecture of a Faster R-CNN comprises three components:

3.1.1   Feature Extractor Network (FEN):

This section provides the basis for image feature extraction, including convolutional, ReLU, and pooling layers. The feature maps produced by the convolution layers are shared with both RPN and ROI Pooling. In this study, we utilized the ResNet module as the backbone for the feature extractor.

3.1.2 Region Proposal Network (RPN):

RPN accepts images with different sizes as input and produces a collection of rectangular object proposals, each linked to an objectness score. To create region proposals, a 3 × 3 sliding window traverses the feature map produced by the final shared convolutional layer. The parameterization for each region proposal relies on a reference box called an anchor box, which is centered within the sliding window and defined by two parameters: scale and aspect ratio. A feature vector is extracted for each region proposal during processing. This vector is then inputted into a pair of fully connected sibling layers: a classification (Cls) layer and a regression (Reg) layer, as shown in Figure 3. The Reg layer has $4k$ coordinates for the candidate nodule. The Cls layer has $2k$ scores to determine whether the proposal is a candidate nodule or background [15].
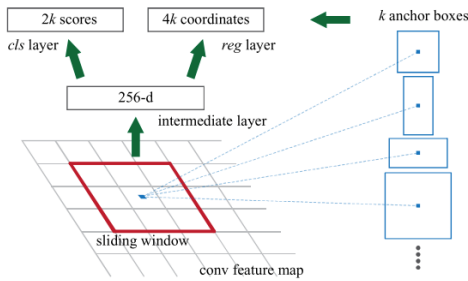

Fig. 3 The Region Proposal Network (RPN).

3.1.3 The ROI pooling layer

The ROI pooling layer maps region proposals from the RPN output, which vary in size, into feature maps of uniform dimensions. These feature maps are then transformed into vectors using a fully connected layer. The resulting vector is forwarded to a regression layer that predicts the coordinates of the bounding boxes, as well as a classification layer that distinguishes candidate nodules from background tissues (non-nodule). The loss function is expressed as:

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}\ (p_i,p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i,t_i^*)$$

(1)

The formula for the anchor index in a mini-batch involves i representing the anchor index, p_i representing the positive softmax probability, and, p_i^* representing the ground-truth prediction probability. Anchors that have the highest intersection over union (IoU) with a ground-truth box (an anchor with an IOU > 0.5) are considered positive, while those with an IOU <0.5 were considered negative. The predicted bounding box is denoted as t_i. The ground-truth box is denoted as t_i^*, respectively. The overall loss is split into two components,

with the regression loss activated only for the positive anchors. The two components are normalized using N_{cls} and N_{reg}, and weighted by a balancing parameter λ.

3.2 False Positive Reduction

After identifying the candidate nodules within the bounding boxes from the previous stage, it is necessary to classify them as either nodules or npon-nodules to minimize the chances of false positive. At this stage, an established CNN architecture called ResNet, specifically ResNet50, was utilized in the proposed model. He et al. [17] developed the deep residual network ResNet in 2015 to address the challenges of deep learning networks, such as gradient vanishing or exploding as the network depth increases. Traditional solutions involve data initialization and regularization; however, these approaches may degrade performance. Residual networks, on the other hand, seek to enhance network performance while addressing the vanishing gradient problem.

Deep residual network design involves skipping connections or shortcuts, thereby allowing a deeper network to alleviate the vanishing gradient problem [12]. Skip connections on two or three layers are required to implement the model along with ReLU and batch normalization across architectures, as illustrated in Figure 4. This method uses shortcut connections to implement identity maps in a network and reduces computational complexity.
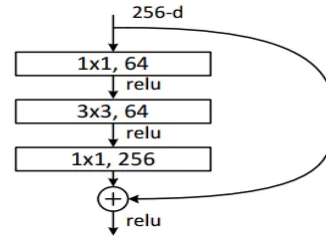

Fig. 4 Residual block in ResNet50 Model.

The network processes a candidate nodule as input and first applies a 7 × 7 convolutional layers with 64 channels, then applies a max-pooling operation with a size of 3 × 3 and a stride of two. Next, it employs multiple residual blocks—both identity and convolutional—to obtain top-tier features from the image. Each block is composed of three convolutional layers and a direct link with varying kernel sizes. Lastly, the network executes an average pooling layer, employes a pair of fully connected layers, and concludes with a sigmoid activation function.

4. EXPERIMENTAL AND RESULTS

4.1 Dataset and Preprocessing

4.1.1 Data Acquisition:

The LUNA16[1] dataset was employed to train, evaluate, and assess our proposed model. LUNA16 is a notable international dataset for pulmonary nodules detection and provides a publicly available dataset of lung nodules. LUNA16 is a subset of the dataset lung image database consortium image collection (LIDC-IDRI) [21], which contains heterogeneous scans filtered according to several standards. It is best to select a thin slice because pulmonary nodules might be very small. Consequently, images with slice thickness ≥3 mm were excluded from the examination. In addition, scans displaying unreliable slice spacings or lost slices were removed. Consequently, a final set of 888 scans were obtained. In these scans, radiologists made 36,378 annotations, classifying nodules that were ≥ 3 mm as relevant. Conversely, annotations for nodules < 3 mm and non-nodules were not considered significant within the lung cancer screening protocols. Images in the LUNA16 dataset were formatted as files with extensions (.mhd) and (.raw), each having a size of 512×512 pixels. Images were obtained using the LUNA-16 website. Table 2 provides essential details regarding the LUNA16 dataset, including the number of CT scans, patients, nodules, and annotations.

Table 2: Characteristics of the LUNA16 Dataset used in Lung Cancer Diagnosis Studies.

Dataset	CT scans	Patients	Nodules	Annotations
LUNA16	888	888	1186	36,378

4.1.2 Data Annotation:

To train an object detection model, two steps are required to create the bounding box coordinates. First, images and masks were extracted by transforming the raw files into simpler formats, such as NumPy arrays. An open-source script 2 is used to process the images and generate the associated mask positions. This script automates the process of producing NPY files for each image and calculating the mask using the x-, y-, and z-coordinates and diameter values of the nodules that are offered in the annotation file. Second, labels (ground truth) were prepared in the required format by utilizing OpenCV's built-in contour method, which was used to identify the boundaries of the nodules or other significant structures within the scan.

This method uses a mask as an input and returns the corresponding coordinates, including x1, y1, x2, and y2. These coordinates are then used to draw a rectangular bounding box around the lung nodule in the image, which helps locate the nodule position [33].

4.1.3 Image Preprocessing:

Image preprocessing is essential for achieving the specific criteria for the input data. These include techniques that boost the accuracy of the model, such as image resizing and data normalization. In this study, the original size of the image was 512×512 pixels, and it was resized to 640×640 pixels, as required by the detection model. The pixel and voxel of medical imags were rescaled into values ranging between [0–1] or [-1–1], which has become an essential step in data normalization prior to training a deep learning network. In the training phase, this process enhances the convergence speed of gradient descent and guarantees that the model is fed with uniformly scaled input data. If the input pixel values are excessively large, they could impede optimal performance Therefore, the pixel values of the grayscale images were normalized to values ranging between [0 – 1].

4.1.4 Data Storge:

To train object detection models, researchers must set up data storage using either Pascal-VOC or Microsoft-COCO, which are two commonly used formats [33]. In our study, to train the Faster R-CNN module, label files (. xml) were created for each image in Pascal-VOC format that contain the bounding boxes for the nodule part of the image and other related information like class name, image size, etc. The provided formats simplify the process and provide a standardized method for organizing and storing the images and annotations. We saved the CT images of lung nodules in the PNG format to preserve full image information, leading to a more accurate preservation of the original pixel values compared to the JPEG/JPG format.

4.1.5 Data Split:

In this study, we implemented a fixed ratio split of 70:10:20. This means that 70% of the data was allocated for model training, 10% was used to validate the optimization of model parameters, and the remaining 20% was reserved for testing and evaluating the model's performance on unseen data.

4.1.6 Data Augmentation:

The LUNA16 dataset contains 1186 nodules, and 551,065 candidates that are labeled with a class of 0 for non-nodules and 1 for nodules. However, there is a significant imbalance between non-nodule and true nodules. To solve this problem, data augmentation approaches, which involve image translation, rescaling, and horizontal reflection, are utilized to produce new

---

[1] https://luna16.grand-challenge.org/Download/

images from the original nodules. Additionally, to maintain more balanced data for the candidate nodules, the number of non-nodule instances is limited to five for each seriesUID in the candidate file.
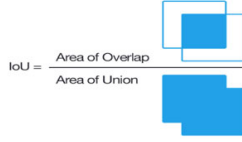

Fig. 5 Intersection over union (IOU).

We utilized the curve of free-response receiver operating characteristic (FROC) and the score of competition performance metric (CPM) to provide a more comprehensive assessment of the model's performance in minimizing false s rates. The curve of FROC illustrates sensitivity in relation to the average number of false positive cases per CT scanned image (FP/scan). Additionally, the CPM assessment was calculated by averaging the sensitivity over seven predefined false positive ratios: 1/8, 1/4, 1/2, 1, 2, 4, and 8 FP/scan [35]. To evaluate the performance of the false positive reduction stage, we employed a 10-fold cross-validation approach, using accuracy as the primary metric across 10 iterations for both the training and testing processes.

Table 3: Key performance metrics used to evaluate proposed model.

Metric	Definition	Note
mAP	$mAP = \frac{1}{N} \sum_{l=1}^{N} AP_l$	Mean average precision
Sensitivity	Sen = TP / (TP + FN)	True positive rate TPR, (Recall)
Accuracy	Acc = (TP + TN) / (TP + TN + FP + FN)	Total true results.

TP=true positive; TN=true negative; FP=false positive; FN=false negative; APi = AP of class i; N=no. of classes.

5. Results and Discussion

5.1 Pulmonary Nodule Detection Result:

Table 4 presents the effectiveness of the nodule detection model (Faster R-CNN) over 100 epochs in terms of mAP values at different IoU thresholds, mAP@0.5:0.95, and mAP@0.5. The mAP@0.5:0.95 is approximately 0.70, which indicates that the model achieves good accuracy, and the ability to locate nodules accurately is acceptable. Meanwhile, the mAP over a threshold of IoU 0.5 (mAP@0.5) reach to excellent value of 0.92, which means that the model is very accurate in nodules localization. The model can obtain its predicted bounding box to overlap by 50% with the ground truth, which is easier than accurate localization in the range of IoU thresholds up to 0.95. Therefore, mAP@0.5 was always higher than mAP@0.5:0.95. Figure 6, shows that the two curves increase in mAP as the epochs numbers increase, showing that the detection model has been learned, and as training progresses, the model enhances its ability to detect

pulmonary nodules. Additionally, Table 4 shows the overall loss during training in terms of training loss with a value of 0.0223, which is relatively low. The train cls loss was too low with a value of 0.0053, which means that the model is very accurate in classifying whether a region of interest has a nodule. In addition, the value of the train box reg loss was 0.00849, which was also very low, indicating that the model accurately predicted the exact location and size of the nodules. Finally, the train obj loss and train rpn loss were extremely low, with values of 0.00073 and 0.00097, respectively. These values indicate that the RPN performs well in proposing the candidate nodule bounding box, and the abjectness score correctly distinguishes between the nodule and background. Figure 7, shows the different loss plots after training the detection model based on the LUNA16 dataset.

Table 4: Performance metrics of the Faster R-CNN model for pulmonary nodule detection.

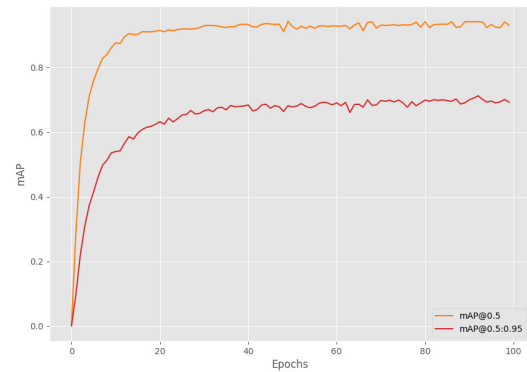| mAP@0.5:0.95	mAP 05@0.5	train loss	train cls loss	train box reg loss	train obj loss	train rpn loss
0.696	0.921	0.0223	0.0053	0.00849	0.00073	0.00097

Fig. 6 mAP plot after training the Faster R-CNN model.

Figure 8 shows that the proposed detection model can correctly identify pulmonary nodules, particularly when there is a high level of intersection between the ground-truth and prediction boxes. However, a prediction with lower confidence indicates that the predicted boxes do not align exactly with the ground-truth nodule but still provide high confidence (0.69 and 0.99) for true positive results. Furthermore, there is a case, particularly in the bottom-right scan, where the model detected two distinct areas as nodules with different levels of confidence (0.99 and 1.0), despite the existence of only one actual nodule. These findings suggest that the model was generally accurate and exhibited a high level of sensitivity. However, this also implies that the model may produce false positive by detecting nodules that may not exist.

a) train loss   b) train loss iter

c) train rpn loss   d) train obj loss

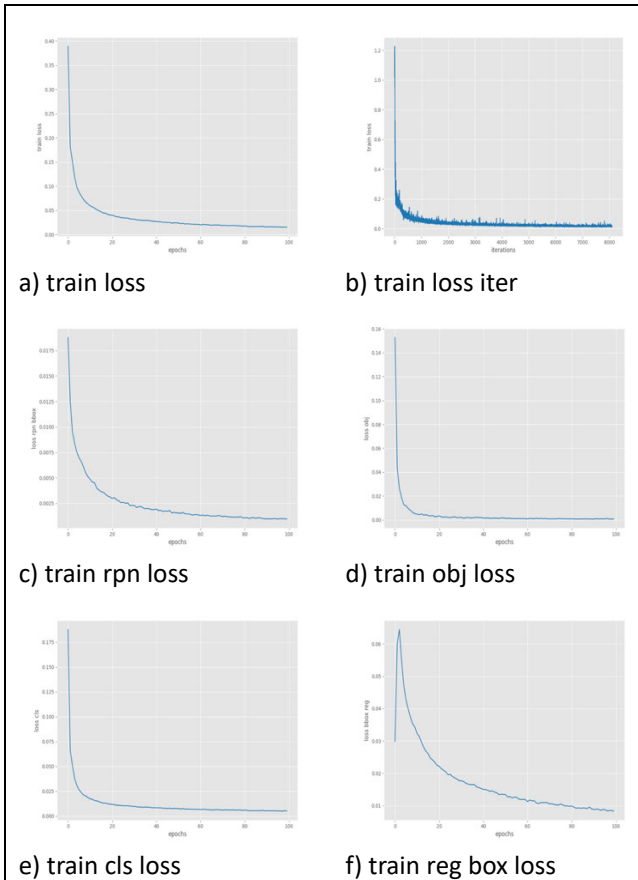e) train cls loss   f) train reg box loss

Fig. 7 Loss plots after training the detection model.

5.2 False Positive Reduction Result

The proposed residual model ResNet50 shows the best sensitivity results at various false positive rates per scan, as well as CPM score as results in Table 5, which shows that the proposed model has a strong ability to differentiate between nodules and non-nodules.

5.3 Comparison Between Different Methods

Table 5 presents the numerical findings of the proposed model and other methods for the LUNA16 dataset. The best performances of the different methods for every false positive rate are emphasized in underline font. A perfect model scores 1 and the lowest possible score is 0 [36]. The table indicates that our model exhibits lower sensitivity compared to [31] at false positive rates of 0.125 and 0.25, respectively. However, based on false positive rates of 0.5 to 8, our model achieved the highest sensitivity performance compared to other methods, reaching an impressive 99.9% sensitivity at eight false positive per scan. According to the table, it is clear that our proposed model surpassed the other four methods with a CPM score of 95.1%, emphasizing its superiority, which is crucial for achieving the clinical requirements of CAD systems, particularly in terms of a higher sensitivity rate. Figure 9,

illustrates the FROC curves of the proposed model and four other methods during the false positive reduction stage.

Table 5: Effectiveness comparison between different methods on LUNA16 dataset.

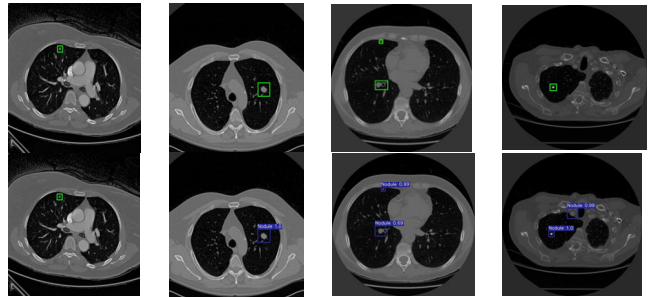| Methods	Sensitivity (%) at False Positive Rates							CPM (%)
	0.125	0.25	0.5	1	2	4	8	
Faster R-CNN with 2RPN[20]	0.734	0.744	0.763	0.796	0.824	0.832	0.834	79.0
3D-CNN [29]	0.784	0.847	0.906	0.938	0.950	0.955	0.961	90.6
3D deep CNN [31]	0.836	0.898	0.930	0.945	0.953	0.962	0.962	92.7
PPD-Unet + DBHA-NET[32]	0.833	0.87	0.898	0.921	0.938	0.951	0.96	91.0
Proposed Model (Our)	0.835	0.897	0.959	0.981	0.988	0.995	0.999	95.1

Fig. 8 Samples of CT images with a ground truth box and Faster R-CNN model predictions. Green boxes represent the verified locations of nodules, while blue boxes show the model's nodule predictions with their respective.
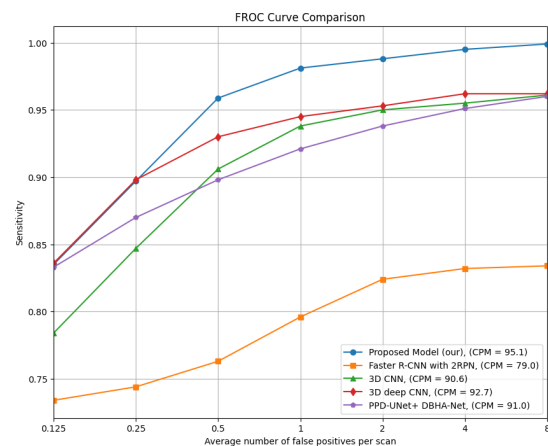

Fig. 9 FROC curves for the proposed model and other methods.

5.3.1 K-Fold Cross Validation Results:

Table 6 presents a detailed overview of the performance of the training Faster R-CNN model based on 10-fold cross-validation with the best result for fold #6, with a high mAP of 0.92 at an IoU threshold of 0.5. This

shows the strong capability of the model to correctly localize nodules. However, the mAP across an IoU thresholds ranging between 0.5 to 0.95 is lower at 0.708, indicating more precise localization of the expected bounding boxes with respect to the ground truth. The model had a low classification loss of 0.0049 and a low bounding box regression loss of 0.0083, which means that it learned well at distinguishing the difference between nodules and other lung structures (lung tissues) that are considered as background, and correctly predicting the bounding boxes around the nodules. The low loss values demonstrated the effective performance of our model in accurately identifying nodules and reducing errors in both classification and localization.

Table 6: Performance metrics for the detection model with results of best-fold.

mAP@0.5:0.95	mAP_05@0.5	train loss	train cls loss	train box reg loss	train obj loss	train rpn loss
0.7087	0.9224	0.0174	0.00495	0.00833	0.00069	0.00094

The accuracy results of the 10-fold cross-validation process is illustrated in Figure 10. The results were obtained by evaluating the performance of the ResNet50 model during the stage of false positive reduction. The model consistently achieved high training accuracy, averaging 98.2%. The seventh-fold model demonstrated perfect accuracy. However, the test accuracy, which is a more reliable indicator of the performance of the model on unseen or new data, averaged 91.1%. The test accuracy remained consistent across folds, ranging from 89.1% to 93.2%, indicating the strong ability of the model to generalize. The variability in test accuracy can be attributed to the variations in the data distribution for each fold.
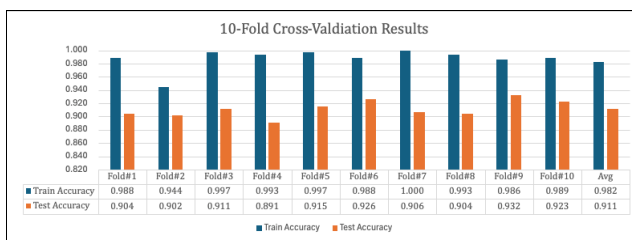

	Fold#1	Fold#2	Fold#3	Fold#4	Fold#5	Fold#6	Fold#7	Fold#8	Fold#9	Fold#10	Avg
Train Accuracy	0.988	0.944	0.997	0.993	0.997	0.988	1.000	0.993	0.986	0.989	0.982
Test Accuracy	0.904	0.902	0.911	0.891	0.915	0.926	0.906	0.904	0.932	0.923	0.911

Fig. 10 Performance of the proposed model in terms of accuracy in 10-Ford cross-validation.

6. Conclusion

There has been a significant annual increase in lung cancer incidence. Preemptive detection and immediate treatment may significantly enhance patient survival. Since pulmonary nodules are often the initial indicators of lung cancer, preemptive testing for the disease is recommended. This study proposed a model to detect pulmonary nodules and minimizing false positive rates by utilizing two well-known pre-trained CNN models: Faster R-CNN to detect pulmonary nodules and ResNet50 to reduce false positive.

We applied a fine-tuning process with the necessary adjustments. The experimental numerical findings demonstrated that the proposed framework, which incorporates a Faster R-CNN network followed by a residual convolutional network on the LUNA16 dataset, outperformed other models in the literature regarding sensitivity across various predefined false positive thresholds, achieving a CPM of 95%. Moreover, testing our model with 10-fold cross-validation confirmed its robustness, reliability, and strong conclusion performance. The test accuracy of the proposed model, averaging over 91%, emphasizes its efficacy in minimizing the rate of false positive.

7. Limitations and Future Works

Although this study yielded promising results, there are some limitations worth noting. First, we trained and tested our model using 1,186 nodules from the LUNA16 dataset, even when applying data augmentation techniques. The dataset was still limited in size. Second, our proposed model is based on 2D CNN models, and it is important to note the potential for further exploration of 3D CNN networks to improve the results by feeding models with more spatial information. However, this could lead to a more complex network architecture as a trade-off to improve the overall model performance. In addition, the overall diagnostic process can be improved by integrating automated diagnosis systems with patients' electronic health records (EHRs) to provide a precise diagnosis that assists radiologists in their decision-making and provides recommendations for an appropriate treatment plan.

8. References

[1] N. W. Schluger and R. Koppaka, "Lung disease in a global context: A call for public health action," Annals of the American Thoracic Society, vol. 11, no. 3. American Thoracic Society, pp. 407–416, 2014. doi: 10.1513/AnnalsATS.201312-420PS.

[2] "Cancer." Accessed: Nov. 07, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer

[3] A. R. Larici et al., "Lung nodules: Size still matters," European Respiratory Review, vol. 26, no. 146, Dec. 2017, doi: 10.1183/16000617.0025-2017.

[4] B. C. Kim, J. S. Yoon, J. S. Choi, and H. Il Suk, "Multi-scale gradual integration CNN for false positive reduction in pulmonary nodule detection," Neural Networks, vol. 115, pp. 1–10, Jul. 2019, doi: 10.1016/j.neunet.2019.03.003.

[5] J. Yanase and E. Triantaphyllou, "A systematic survey of computer-aided diagnosis in medicine: Past and present developments," Expert Systems with Applications, vol. 138. Elsevier Ltd, Dec. 30, 2019. doi: 10.1016/j.eswa.2019.112821.

[6] Y. Su, D. Li, and X. Chen, "Lung Nodule Detection based on Faster R-CNN Framework," Comput Methods Programs Biomed, vol. 200, p. 105866, Mar. 2021, doi: 10.1016/J.CMPB.2020.105866.

[7] T. Manikandan and N. Bharathi, "Lung Cancer Detection

Using Fuzzy Auto-Seed Cluster Means Morphological Segmentation and SVM Classifier," J Med Syst, vol. 40, no. 7, Jul. 2016, doi: 10.1007/s10916-016-0539-9.

[8]   A. Bhattacharjee, R. Murugan, and T. Goel, "A hybrid approach for lung cancer diagnosis using optimized random forest classification and K-means visualization algorithm," Health Technol (Berl), vol. 12, no. 4, pp. 787–800, 2022.

[9]   N. Guo, R.-F. Yen, G. El Fakhri, and Q. Li, "SVM based lung cancer diagnosis using multiple image features in PET/CT," in 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2015, pp. 1–4.

[10] S. Garud and S. Dhage, "Lung cancer detection using CT images and CNN algorithm," in 2021 International Conference on Advances in Computing, Communication, and Control (ICAC3), 2021, pp. 1–6.

[11] M. S. AL-Huseiny and A. S. Sajit, "Transfer learning with GoogLeNet for detection of lung cancer," Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 2, pp. 1078–1086, Apr. 2021, doi: 10.11591/ijeecs.v22.i2.pp1078-1086.

[12] P. Wu, X. Sun, Z. Zhao, H. Wang, S. Pan, and B. Schuller, "Classification of Lung Nodules Based on Deep Residual Networks and Migration Learning," Comput Intell Neurosci, vol. 2020, 2020, doi: 10.1155/2020/8975078.

[13] X. Liu, L. Song, S. Liu, and Y. Zhang, "A review of deep-learning-based medical image segmentation methods," Sustainability (Switzerland), vol. 13, no. 3, pp. 1–29, Feb. 2021, doi: 10.3390/su13031224.

[14] X. Rafael Palou, M. Angel González Ballester, G. Piella Fenoy, and V. Ribas Ripoll, "Detection, quantification, malignancy prediction and growth forecasting of pulmonary nodules using deep learning in follow-up CT scans."

[15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." [Online]. Available: https://github.com/

[16] X. Huang, W. Sun, T.-L. Tseng, C. Li, and W. Qian, "Fast and Fully-Automated Detection and Segmentation of Pulmonary Nodules in Thoracic CT Scans Using Deep Convolutional Neural Networks," 2019.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." [Online]. Available: http://image-net.org/challenges/LSVRC/2015/

[18] C. C. Nguyen, G. S. Tran, V. T. Nguyen, J. C. Burie, and T. P. Nghiem, "Pulmonary Nodule Detection Based on Faster R-CNN with Adaptive Anchor Box," IEEE Access, vol. 9, pp. 154740–154751, 2021, doi: 10.1109/ACCESS.2021.3128942.

[19] J. Lyu, X. Bi, and S. H. Ling, "Multi-level cross residual network for lung nodule classification," Sensors (Switzerland), vol. 20, no. 10, May 2020, doi: 10.3390/s20102837.

[20] 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, doi: 10.1016/j.patcog.2018.07.031.

[21] "Home - Grand Challenge." Accessed: Nov. 12, 2023. [Online]. Available: https://luna16.grand-challenge.org/

[22] B. Mahesh, "Machine Learning Algorithms-A Review," International Journal of Science and Research, 2018, doi: 10.21275/ART20203995.

[23] C. Jacobs et al., "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images," Med Image Anal, vol. 18, no. 2, pp. 374–384, Feb. 2014, doi: 10.1016/j.media.2013.12.001.

[24] M. Bergtholdt, R. Wiemker, and T. Klinder, "Pulmonary Nodule Detection Using a Cascaded SVM Classifier." [Online]. Available: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI

[25] R. Manickavasagam, S. Selvan, and M. Selvan, "CAD system for lung nodule detection using deep learning with CNN," Med Biol Eng Comput, vol. 60, no. 1, pp. 221–228, Jan. 2022, doi: 10.1007/s11517-021-02462-3.

[26] D. Riquelme and M. A. Akhloufi, "Deep Learning for Lung Cancer Nodules Detection and Classification in CT Scans," AI (Switzerland), vol. 1, no. 1. Multidisciplinary Digital Publishing Institute (MDPI), pp. 28–67, Dec. 01, 2020. doi: 10.3390/ai1010003.

[27] A. A. A. Setio et al., "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks," IEEE Trans Med Imaging, vol. 35, no. 5, pp. 1160–1169, May 2016, doi: 10.1109/TMI.2016.2536809.

[28] Z. Xiao, N. Du, L. Geng, F. Zhang, J. Wu, and Y. Liu, "Multi-scale heterogeneous 3D CNN for false-positive reduction in pulmonary nodule detection, based on chest CT images," Applied Sciences (Switzerland), vol. 9, no. 16, Aug. 2019, doi: 10.3390/app9163261.

[29] Z. Gong, D. Li, J. Lin, Y. Zhang, and K. M. Lam, "Towards accurate pulmonary nodule detection by representing nodules as points with high-resolution network," IEEE Access, vol. 8, pp. 157391–157402, 2020, doi: 10.1109/ACCESS.2020.3019104.

[30] J. Liang, G. Ye, J. Guo, Q. Huang, and S. Zhang, "Reducing False-Positives in Lung Nodules Detection Using Balanced Datasets," Front Public Health, vol. 9, May 2021, doi: 10.3389/fpubh.2021.671070.

[31] H. Zhang, Y. Peng, and Y. Guo, "Pulmonary nodules detection based on multi-scale attention networks," Sci Rep, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-05372-y.

[32] M. Liao, Z. Chi, H. Wu, S. Di, Y. Hu, and Y. Li, "Pulmonary Nodule Detection from 3D CT Image with a Two-Stage Network," International Journal of Intelligent Systems, vol. 2023, 2023, doi: 10.1155/2023/3028869.

[33] J. Wang et al., "Preparing CT imaging datasets for deep learning in lung nodule analysis: Insights from four well-known datasets," Heliyon, vol. 9, no. 6, Jun. 2023, doi: 10.1016/j.heliyon.2023.e17104.

[34] R. Padilla, S. L. Netto, and E. A. B. Da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms."

[35] A. A. A. Setio et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," Med Image Anal, vol. 42, pp. 1–13, Dec. 2017, doi: 10.1016/j.media.2017.06.015.

[36] "Evaluation - Grand Challenge." Accessed: Apr. 28, 2024. [Online]. Available: https://luna16.grand-challenge.org/Evaluation/