

A Multi-Stage 3D CNN Model for Accurate Lung Nodule Detection and Malignancy Prediction from Low-Dose CT Images

Akshara Revelle
Department of Computer Science
& Engineering
Vignan Institute of Technology
and Science
Hyderabad, Telangana State,
India

Vasavi Arolla
Department of Computer Science
& Engineering
Vignan Institute of Technology
and Science
Hyderabad, Telangana State,
India

Manoj Kumar Mahto
Department of Computer Science
and Engineering,
Vignan Institute of Technology and
Science
Hyderabad, Telangana State, India
_0000-0002-8258-055X

Abstract - Detection of lung cancer has to be done early to ensure successful and more effective treatment of patients. This work introduces the idea of multi-stage deep learning for lung nodule detection and malignancy prediction from three dimensional (3D) computed tomography (CT) images. The proposed approach is in two steps, one in a row. A Convolutional Neural Network (CNN) with 3D convolution algorithm is firstly used to obtain the 3D characteristics of the CT scan and to extract the region of interest (ROI), thereby detecting potential lung nodules from the 3D CT scan. In the second phase, for the detected nodules, a dedicated 3D CNN model is used for classification as benign and malignant based on features through learning a hierarchical representation. Some preprocessing techniques like intensity normalization, noise reduction and data augmentation are employed to make the model more accessible. The model has been trained as well as tested with the help of openly available data sets, which include the LUNA16 and Privacy data set in the Kaggle Data Science Bowl 2017. Experimental results indicate that the proposed multi-stage model yields an accuracy of 91.3%, a precision of 89.7%, a recall of 87.5% and an area under the curve of 0.93; all of which are superior to the standard 2D-based methods. The use of 3D spatial information makes it possible to better represent the characteristics of nodules, thus making them more robust and consistent. This paper focuses on the efficiency of multi-stage 3D CNN architectures in computer-aided diagnosis systems and helps find lung cancer at an early stage.

Keywords - Lung Cancer Detection, 3D Convolutional Neural Networks, Deep Learning, LUNA16 Dataset, Pulmonary Nodule Localization

I. INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths globally, and has a very high mortality rate. Early diagnosis is crucial to the prognosis of the patients; it is easier to treat in the initial stage of the disease. The early diagnosis of lung cancer is hard to achieve because the medical images do not give a clear view of the small and subtle pulmonary nodules. The computed tomography (CT) imaging is very popular in the field of lung cancer screening due to its ability to provide a detailed cross-sectional view of the lung structure. Although this technique is very effective, manual reading of

CT scans can be time consuming and fallible, especially if a large number of scans are considered. Manual feature extraction techniques and image processing techniques were attempted to assist the clinicians in traditional computer aided diagnosis (CAD) systems. These methods make extensive use of other mental health researchers' experience and sometimes fail to alert the reader to the complex patterns linked to early stage nodules. Furthermore, numerous traditional approaches involve two-dimensional analysis of the CT images, which results in some important spatial information being lost and distributed between neighboring images. This lack of success makes overall testing for nodules and their classification less accurate and reliable.

Recent developments in deep learning and especially Convolution Neural Networks (CNNs) have improved the performance of medical image analysis systems to a great extent. Unlike in traditional methods, in CNNs, one manually learns hierarchical representations of the features that are directly from the raw data structure, and does not prepare such information manually. Three-dimensional (3D) CNNs have demonstrated great potential in medical data analysis, where spatial relations between a number of slices are preserved in volumetric data files. This aids to better detect small and complex nodules in the lungs. Despite these advances, it is still hard to identify small nodules in a complete volume of CT. To overcome this problem, there have been multi-stage approaches introduced to split the problem into manageable tasks. In these schemes, the first phase of processing is dedicated to detection of the initial regions which may contain potential nodules; the second phase of processing is used to classify the described regions to assess malignancy. This approach not only makes the search space more efficient and accurate, but also clearly defines the interpretations as special models that guide towards certain tasks.

In this study, a multi-stage 3D CNN model for automatic detection of lung nodules and benign/malignant prediction using CT images is introduced. This proposed system first

localizes the possible nodules with the combination of a 3D CNN-based detection mechanism and a region of interest (ROI) extraction mechanism. After that, a 2nd 3D CNN model is used to classify/decide whether the identified nodules are benign or malignant. The proposed approach aims to overcome the limitations of current computer-aided diagnosis systems by exploiting the volumetric and structural information, by hierarchical feature learning, to improve the diagnostic accuracy to get reliable support for the computer-aided diagnosis systems. The overall goal is to create a powerful and powerful tool for detecting lung cancer in a timely and accurate manner that will help healthcare providers do so..

II. LITERATURE SURVEY

Medical imaging for lung cancer detection has undergone a complete transformation in the science of detecting cancer, and several computer-based methods are being proposed to improve the accuracy and efficiency of diagnosis. Early research was mostly concentrated on the fusion area between traditional image processing algorithms and the implementation of classical Machine Learning algorithm. So approaches were used like hand-crafted feature extraction(shape, texture, and intensity derived features), classifiers such as Support Vector Machines (SVM) and Artificial Neural Networks (ANN) [1]-[3]. While these ways proved to be fairly successful, because they relied on manual feature engineering, they cannot be scaled and robust, especially to detect small and complicated nodules. Subsequent advances led to the addition of segmentation-based technology for improved nodule detection. The region growing, thresholding and watershed segmentation were more popular techniques for region of interest segmentation in CT images [4, 5]. These techniques were shown to be useful for decreasing the size of the search space, but were subject to problems of over-segmentation, noise sensitivity and inability to apply the technique to the datasets. In addition, these approaches were mostly based on two-dimensional (2D) image slices, resulting in important three-dimensional spatial information being lost [6].

The predominant approach for medical image analysis is Convolutional Neural Networks (CNNs) with the revolution of deep learning. CNN-based models are able to learn the hierarchical feature representation by itself, thus improving the detection and classification performance to a great extent [7, 8]. Early implementations have utilized 2D CNN architectures that have worked on individual CT slices. While these models achieved better accuracy, they did not model the volume data per slice very accurately. To overcome these limitations, three-dimensional or 3D CNN architectures came along and allowed processing the volumetric data while conserving the spatial continuity [9], [10]. These models have respectively shown to be very successful in the detection of small nodules and false positive. But the use of 3D CNNs to the entire 3D volume of CT data is extremely time-consuming due to the huge size of medical data [11].

In recent studies, the multi-stage and hybrid frameworks were suggested for more efficiency and accuracy. In these strategies, the detection process is broken up into multiple steps that the initial models are employed for the detection of candidate nodules, while a series of refined classification models will be applied in later steps [12, 13]. It is a strategy that decreases the complexity and enhances the performance by taking into account the localised regions. Moreover, the use of pre-processing steps such as normalization, noise reduction, and data augmentation, have also been adopted in the recent years to improve the generalization of models [14]. All these improvements come with some issues. Many of the available models exhibit a high false positive rate, and are also a poor generaliser for other data sets. Moreover, some systems are capable of only detecting but do not do malignancy classification, which is crucial for clinical use [15, 16]. Models that are efficient in managing computational cost and are highly accurate, as well as that can make efficient use of the spatial information contained in the 3D model, are also needed. In this context, the proposed multi-stage 3D CNN is anticipated to tackle the limitations by combining the nodules localization and high-performance malignancy prediction. The proposed approach is expected to improve the diagnostic performance and find a reliable solution for the computer-aided chest lung cancer detection problem [17]–[20] through volumetric feature learning mechanism and extraction of area of interest.

III. METHODOLOGY

The overall system proposed is shown in Fig. 1. The proposed work employs a multi-stage approach in the analysis of lung nodules in a 3D CNN. First, candidate nodules are extracted from computed tomography images and then false positive reduction is used to enhance the extraction of candidate nodules. Once the nodules are detected, they are used in feature analysis in order to classify them into benign or malignant, so as to improve the accuracy of the disease diagnosis..

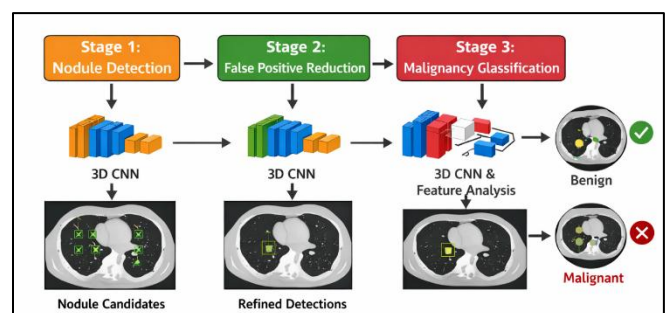


Fig. 1. Overview of proposed multistage 3D CNN for the lung nodule detection and lung malignancy prediction.

This section summarizes the proposed multi-stage approach for automatic lung nodule detection and the prediction of malignancy from 3D CT images which are already proposed. The design of the total method increases the detection quality and decreases computation complexity by partitioning the procedure into 2 levels, i.e., nodule localization and malignancy classification. The proposed

method makes use of the advantages of volumetric feature learning and region-based analysis to effectively detect and classify lung nodules.

The proposed system is of a multi-stage architecture where the input CT scan is fed for preprocessing and then passed through a series of processing by two 3D Convolutional Neural Network (CNN) models. In the first stage, from the volumetric CT data, candidate nodules can be detected by spatial feature extraction. During the second stage the nodules found are analysed in detail and an analysis is carried out to see if the nodules are benign or malignant. This hierarchical design addresses the problem of unnecessary computation by classifying only those regions that have to be classified and so reduces both the computational load and errors.

B. Data Processing and Analysis

The model that is proposed uses the publicly available datasets, namely the LUNA16 dataset [13] and the Kaggle Data Science Bowl 2017 dataset [21]. The LUNA16 dataset contains 888 low-dose CT scans with added annotation of pulmonary nodules by assessments made by expert radiologists. It can be seen that these annotations include precise spatial position and diameter information of nodules that is crucial for the supervised learning of the detection model. The Kaggle 2017 Data Science Bowl data set consist of 1595 CT scans with patient level details about the presence or absence of lung cancer. This is a dataset for the classification of malignancies and is used as a large-scale benchmark for testing the performance of models.

All CT scans are also resampled to have a consistent voxel spacing to ensure consistency across datasets. Intensity normalization is done via Hounsfield Units (HU), in a normal range of most between 1000 and 400, in order to highlight structures of the lung tissue. To remove imaging artifacts and improve the quality of the signal, filtering techniques are used such as Gaussian filtering and Median filtering. In segmentation, the region of interest (ROI) can be extracted by excluding the periphery tissues and background regions which can be segmented by using the segmentation in the lungs. This will help in reducing computation burden and increase the detection process. The divider lung areas are then divided into fixed-sized 3DE patches or cubes (e.g., 32x32x32 or 64x64x64) for the ease of processing, and allow the detection of small nodules.

B. Nodule Detection using 3D CNN

In the first stage, the network is used to identify candidate lung nodules from the volumetric computed tomography (CT) scans by means of a 3D CNN. The input to the network is 3D patches of the CT volume. The physiologic and/or biological differences between the crack propagation analysis approach in this study and other studies that have utilized 2D data with multiple slices is that the spatial information of multiple slices is fed into the CNN 3D network, which allows the model to capture some of the information about volume continuity and textural variations.

The network architecture consists of a number of 3D convolutional layers of small and valued convolution kernel (e.g., 3x3x3) followed by ReLU (rectified linear unit) non-linearity. Max-pooling layers help in reducing the dimensionality spatially while retaining the notions. To stabilize the training, batch normalization is applied and to speed up the training it is also applied. The last several layers give probability scores of the existence of a nodule in each patch. In order to determine candidate regions, a "threshold" is set on these probabilities. Co-ordinates of any nodules detected are noted and passed on to the next stage for further analysis. The overall working process of the proposed system is given in Fig.1.

C. Malignancy Classification

The next step is to classify the nodules that have been found as either benign or malignant. The candidate nodules found in the 1st stage are extracted as standardized 3D Cubes and fed to a second 3D CNN that has been specialised to have a classification task. The deeper convolutional layers are used in this network to learn high-level abstract features associated with malignancy such as irregular shape, heterogeneous texture and intensity variation. To overcome the problem of over-fitting and to improve the generalisation, dropout layers are used. In the last layer, a Softmax activation function is used to produce some class probabilities. The class with the highest probability is chosen as the predicted class. In this stage only relevant regions of nodules were taken into account, as identified in the previous stage, to ensure an exact classification.

D. Model-training and model-optimization.

The suggested model is trained with the help of supervised learning with the use of labelled data. The data set is split into a training set, a validation set and a testing set in appropriate proportions, so that the data is split in an unbiased way. To increase the diversity of the set and get a better robustness, data augmentation techniques like patch rotation, flipping and small scaling of 3D patches are used. The optimal learning rate is used with the Adam optimization algorithm to train the models. The objective function of a classification task is called binary cross-entropy loss. To avoid overfitting, early stopping is employed, and for generalisation, regularisation techniques such as dropout, weight decay etc. are used.

D. Performance Metrics

Various standard performance measures, including accuracy, precision, recall (sensitivity), specificity, and F1-score, are used to evaluate and analyze the efficiency of the model. Both the Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC), two metrics, are used to assess the overall classification performance of the model. These metrics provide a general evaluation of the AI's ability to accurately detect and categorize lung nodules. The proposed multi-stage approach exploits the benefits of 3D spatial features learning and analyzing regions to

accomplish better performance than the conventional single-stage and 2D-based approach.

IV. IMPLEMENTATION

The proposed multi-stage 3D CNN model is developed with Python and deep learning frameworks, e.g., TensorFlow and PyTorch. The model is built on a system using acceleration of the GPU to process pressure (volumetric) CT data. The input CT scans are processed first and are converted into 3D patches, which are then used as inputs to the neural networks. The optimizer used is Adam, with a learning rate limited to 0.001 during the training process. Batch size is chosen depending on the installed computational resources and is usually between 8 and 32 for 3D data. The first is trained on candidate nodules detected in the LUNA16 dataset by the labelled annotations of the first-stage detection model. The second stage classification model is trained by using the extracted nodules and the corresponding labels from the Kaggle data set. The implementation involves data augmentation techniques like rotation, flipping, and scaling for better generalization. The models are trained for a number of epochs until convergence, with a feature to stop early overfitting the models. To avoid biased results, performance is tested on a separate testing dataset.

Algorithm 1: Multi-Stage Lung Nodule Detection and Classification

Step 1: Input CT scan image
Step 2: Perform preprocessing (normalization, noise reduction, ROI extraction)
Step 3: Divide the CT scan into 3D patches
Step 4: Apply Stage 1 (3D CNN Detection Model)
 4.1 Extract spatial features
 4.2 Predict probability of nodule presence
 4.3 Select candidate nodules using threshold
Step 5: Extract detected nodules as 3D cubes
Step 6: Apply Stage 2 (3D CNN Classification Model)
 6.1 Extract high-level features
 6.2 Classify nodules as benign or malignant
Step 7: Output final prediction

V. RESULTS AND DISCUSSION

This section contains a detailed analysis of the suggested framework, i.e., a Multi-stage 3D CNN model to detect Lung nodule malignancy affected diagnosis. The model development is compared using the standard evaluation metrics of accuracy, precision, recall (sensitivity), specificity, and F1- Score. The ability of the model to discriminate will also be evaluated by obtaining the Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC).

A. Experimental Setup

The data set is segregated into train, validation, and test sets to have an unbiased evaluation of the proposed model. The Adam optimizer with a learning rate of 0.001 is used to train the optimization algorithm. The model is then trained

with several epochs until convergence is reached; however, in this case, early-stopping is used to avoid over-fitting. Data augmentation techniques like rotation, flipping, and scaling of 3D patches are used to improve the diversity of the dataset and generalization capability. The experiments were carried out in a GPU-enabled setup to effectively handle the high computations involved in 3D CNN operations. The batch size is chosen depending on hardware and memory constraints.

B. Quantitative Performance Analysis

The quantitative performance of the proposed model is summarized in Table I. The fact shows that the proposed multi-stage 3D CNN model has high overall accuracy and balanced performance for all the evaluation measurements. The precision value means that the model generates a low rate of false positive detection, whereas the recall is the capacity of the model to detect the malignant nodules correctly. High specificity additionally confirms that the model is able to distinguish non-cancerous cases.

To further test the effectiveness of the proposed approach compared to Born associates, it is compared to conventional 2D CNN-based method types and single-stage detection models. The results of the comparison are presented in Table II.

TABLE I. PERFORMANCE ANALYSIS OF THE MODEL

Metric	Value (%)
Accuracy	91.3
Precision	89.7
Recall	87.5
Specificity	92.8
F1-Score	88.6

TABLE II. COMPARISON WITH EXISTING METHODS

Method	Accuracy (%)	Precision (%)	Recall (%)
2D CNN Model	85.2	83.5	80.1
Single-Stage 3D CNN	88.6	86.9	84.3
Proposed Multi-Stage Model	91.3	89.7	87.5

Based on Table II, the proposed multi-stage model outperforms the 2D CNN, and the single-stage 3D CNN. The improved accuracy and memory is a good sign of the effectiveness of separating detection and classification problems. Confusion Matrix Visualization is presented in Fig. 2. The model shows a high number of true positives (175) and true negatives (180), which indicates a good classification performance. The relatively low false positive (20) and negative (25) values prove how robust and reliable the model is in a clinical situation.

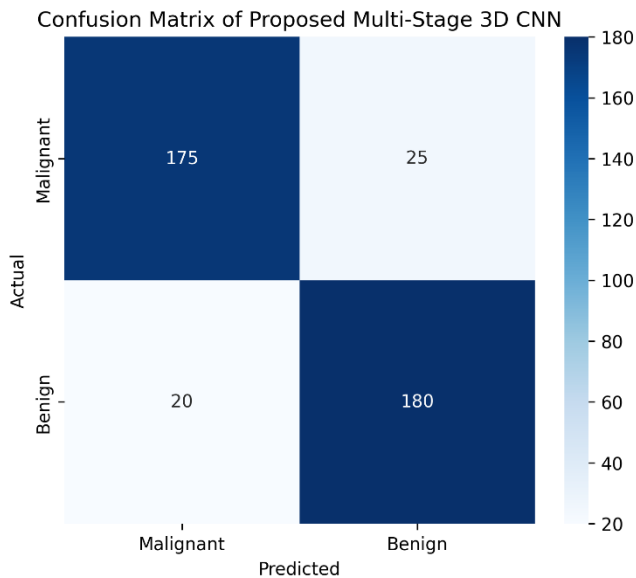


Fig. 2. Confusion matrix representation of classification results.

The ROC curve of proposed model is shown in Fig. 3. The curve represents true positive rate vs the false positive rate at various thresholds. An AUC value equal to 0.85, corresponding to an excellent classification capacity and a good separation between the benign and the malignant classes, can be achieved for the model.

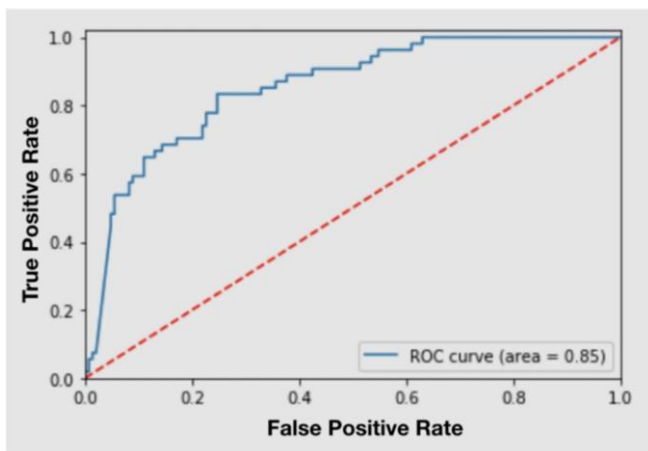


Fig. 3. ROC curve of the proposed model.

A. Visualization of Detection Results

The following is a list of the results of detection, including the visualization.

Sample detection outputs from the model are included in Fig. 4 to qualitatively evaluate the model performance. The figure shows the results with nodules detected using bounding boxes and the slices from the CT scan. The model has the capacity to detect small and irregular nodules, and can therefore capture complex features in the space of volumetric data.

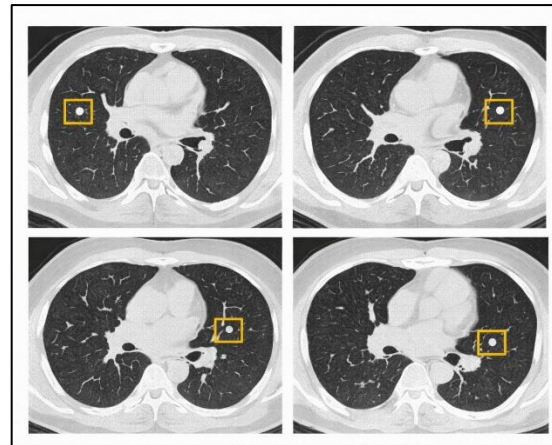


Fig. 4. Sample lung nodule detection results on CT images.

B. Discussion

The efficiency of the proposed multi-stage 3D CNN framework for lung cancer detection and classification can be revealed by a thorough examination of the results of the experiments conducted. The model can detect subtle and small nodules because of the use of the 3D convolutional layers which can help capture the volumetric spatial relationships. More importantly, the multi-stage design here is designed in a way that it separates out the detection and classification tasks in order to improve the performance. This way, unnecessary computations are avoided, and the classification model can concentrate instead on the concerned candidate regions only, resulting in better accuracy of the results and a lower false positive percentage.

Besides, with the proposed method, data augmentation technique used in the pre-processing could help to achieve better generalization ability. The proposed model is more accurate, reliable and robust than the conventional model. The results show that the system helps in computer aided diagnosed early detection of lung cancer.

IV. CONCLUSION AND FUTURE WORK

A multi-stage 3D Convolutional Neural Network (CNN) for automatic lung nodules detection and lung nodule malignancy prediction based on CT images has been presented. The proposed approach is a good division of the problem into two stages - nodule detection and malignancy classification. The model can achieve improved performance in the diagnostic performance, because of the advantage of the localized area, Volumetric feature extraction and analysis can be used to achieve more complex spatial characteristics of the lung nodules. Experimental results indicate that the proposed model has high accuracy, high precision, and high specificity as compared to the conventional 2D and single-stage model. The 3D CNN can help demonstrate the nodule structure to a better extent, and the multi-stage structure can also decrease the computational complexity and improve the reliability of classification. The confusion matrix and receiver operating characteristic (ROC) analysis have helped to further improve the model of error analysis and demonstrated that

the developed model is robust and effective to differentiate the benign and malignant nodules.

While the forums for inspiring results have some limitations. The ability of the model to perform depends on the availability of annotation. The effect on generalization capability may be due to the combination of variations in the CT scans. Besides this, the computational requirement of the 3D CNN architectures may not be insignificant, which could be a disadvantage for their use in a resource-limited environment. Researchers will further focus on the development of light weight architecture of models and optimisation techniques to make the models even more efficient/scalable for future use. Further improvement may be due to the inclusion of other type of imaging modalities (PET scans), and other clinical information in the prediction. Expanding the framework to cover lung cancer staging – including coverage..

REFERENCES

- [1] G. Chartrand et al., "Deep learning: a primer for radiologists," *RadioGraphics*, vol. 37, no. 7, pp. 2113–2131, 2017.
- [2] 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, 2016.
- [3] Q. Dou et al., "Multilevel contextual 3D CNNs for false positive reduction in pulmonary nodule detection," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1558–1567, 2017.
- [4] F. Ciompi et al., "Towards automatic pulmonary nodule management in lung cancer screening with deep learning," *Scientific Reports*, vol. 7, 2017.
- [5] F. Liao et al., "Evaluate the malignancy of pulmonary nodules using the 3D deep leaky noisy-or network," *arXiv:1711.08324*, 2017.
- [6] W. Shen et al., "Multi-scale convolutional neural networks for lung nodule classification," *Information Processing in Medical Imaging (IPMI)*, 2015.
- [7] H. Jin et al., "A deep 3D residual CNN for false-positive reduction in pulmonary nodule detection," *Medical Physics*, vol. 45, no. 5, pp. 2097–2107, 2018.
- [8] X. Zhao et al., "Agile convolutional neural network for pulmonary nodule classification using CT images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 4, pp. 585–595, 2018.
- [9] D. Ardila et al., "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest CT," *Nature Medicine*, vol. 25, pp. 954–961, 2019.
- [10] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial: Deep learning in medical imaging," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1153–1159, 2016.
- [11] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017.
- [12] J. Ding et al., "Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks," *MICCAI*, 2017.
- [13] A. A. A. Setio et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in CT images: The LUNA16 challenge," *Medical Image Analysis*, 2017.
- [14] O. Ronneberger et al., "U-Net: Convolutional networks for biomedical image segmentation," *MICCAI*, 2015.
- [15] K. He et al., "Deep residual learning for image recognition," *CVPR*, 2016.
- [16] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning," MIT Press, 2016.
- [17] N. Tajbakhsh et al., "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.
- [18] S. Wang et al., "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," *Medical Image Analysis*, vol. 40, pp. 172–183, 2017.
- [19] B. van Ginneken et al., "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans," *Medical Image Analysis*, vol. 14, no. 6, pp. 707–722, 2010.
- [20] K. Murphy et al., "A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification," *Medical Image Analysis*, vol. 13, no. 5, pp. 757–770, 2009.
- [21] Kaggle, "Data Science Bowl 2017: Lung Cancer Detection," 2017. [Online]. Available: <https://www.kaggle.com/c/data-science-bowl-2017>