Deep Learning-Based Computer-Aided Diagnosis System for Lung Cancer Detection from CT Images: A Novel Approach

S.Thilagavathi 1 & Dr.R.Malathi Ravindran 2

1 Research Scholar, Department of Computer Science, NGM College, Pollachi, Tamil Nadu, India

E-mail: thilagavathimahendhirakumar@gmail.com

2 Associate Professor, Department of Computer Applications, NGM College, Pollachi, Tamil Nadu, India

E-mail: malathiravindranr@gmail.com

Abstract—Early detection of lung cancer significantly enhances survival rates. Computer-Aided Diagnosis (CAD) systems based on Computed Tomography (CT) images aim to improve detection accuracy and reduce interpretation time. This paper presents a novel deep learning-based CAD framework for automated lung nodule detection and classification. We employ advanced pre-processing, lung segmentation, feature extraction using 3D convolutional neural networks, and false positive reduction techniques. The system is evaluated on publicly available datasets achieving superior sensitivity and specificity compared to classical methods. Experimental results demonstrate the potential of the proposed approach for clinical lung cancer diagnostics.

Keywords: Lung cancer, Computer-Aided Diagnosis (CAD), CT images, lung nodule detection, deep learning, convolutional neural networks, false positive reduction, medical imaging, early diagnosis, computational radiology.

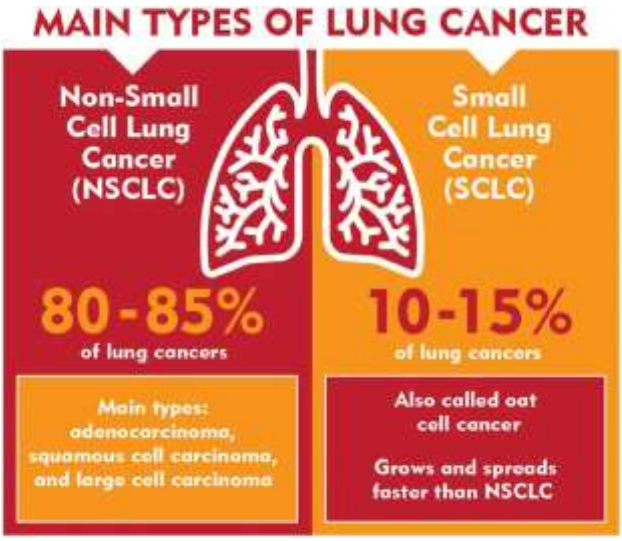
I. Introduction

Lung cancer is among the deadliest cancers worldwide, with early diagnosis being crucial to improving 5-year survival rates by up to 50% [2]. Manual detection of pulmonary nodules in chest CT images is time-consuming and prone to misses, sometimes up to 30% [2]. Computer-Aided Diagnosis (CAD) systems provide a second opinion, assisting radiologists by automating nodule detection and characterization, thereby enhancing diagnostic accuracy and efficiency [1], [2]. This paper proposes a deep learning-based CAD system designed to overcome challenges in accurate lung segmentation and nodule classification, reducing false positives, and improving processing time.

Lung cancer remains one of the leading causes of cancer-related deaths worldwide, with early detection being essential for improving patient outcomes. Computed Tomography (CT) imaging plays a vital role in identifying lung nodules that may indicate malignancy. However, manual interpretation of CT scans is time-consuming and prone to human error. To address these challenges, computer-aided diagnosis (CAD) systems leveraging advanced machine learning techniques have been developed. Recent progress in deep learning, particularly convolutional neural networks, offers promising capabilities for accurate and efficient lung nodule detection and classification. This study proposes a novel deep learning-based CAD framework designed to automate lung nodule identification, reduce false positives, and enhance diagnostic reliability, ultimately supporting clinicians in early lung cancer detection.

Lung cancer is a type of cancer that starts when abnormal cells grow in an uncontrolled way in the lungs. It is a serious health issue that can cause severe harm and death.
Symptoms of lung cancer include a cough that does not go away, chest pain and shortness of breath.
It is important to seek medical care early to avoid serious health effects. Treatments depend on the person's medical history and the stage of the disease.
The most common types of lung cancer are non-small cell carcinoma (NSCLC) and small cell carcinoma (SCLC). NSCLC is more common and grows slowly, while SCLC is less common but often grows quickly.

Fig 1 Main Types of Lung Cancer



Lung cancer is a significant public health concern, causing a considerable number of deaths globally. GLOBOCAN 2020 estimates of cancer incidence and mortality produced by the International Agency for Research on Cancer (IARC) show as lung cancer remains the leading cause of cancer death, with an estimated 1.8 million deaths (18%) in 2020.

Smoking tobacco (including cigarettes, cigars, and pipes) is the primary risk factor for lung cancer but it can also affect non-smokers. Other risk factors include exposure to secondhand smoke, occupational hazards (such as asbestos, radon and certain chemicals), air pollution, hereditary cancer syndromes, and previous chronic lung diseases.

II. Related Work

Traditional CAD systems consist of multiple stages: image acquisition, pre-processing (noise reduction), lung and nodule segmentation, feature extraction, and classification [3]–[6]. Recent works have used techniques such as region growing, active contours, genetic algorithms, and support vector machines (SVM) for lung nodule detection and false positive reduction [4]–[6]. Despite promising results, many systems struggle with detecting juxtapleural nodules and have limited performance on large and heterogeneous datasets [5], [7]. Deep learning, especially convolutional neural networks (CNNs), has shown potential for end-to-end feature extraction and classification, motivating the design of the proposed system.

III. Materials and Methods

The proposed Computer-Aided Diagnosis (CAD) system is designed as a multi-stage deep learning framework for automated lung nodule detection and classification from chest CT images. The overall methodology consists of dataset preparation, image pre-processing, lung segmentation, candidate nodule detection, deep feature extraction, false positive reduction, and performance evaluation.

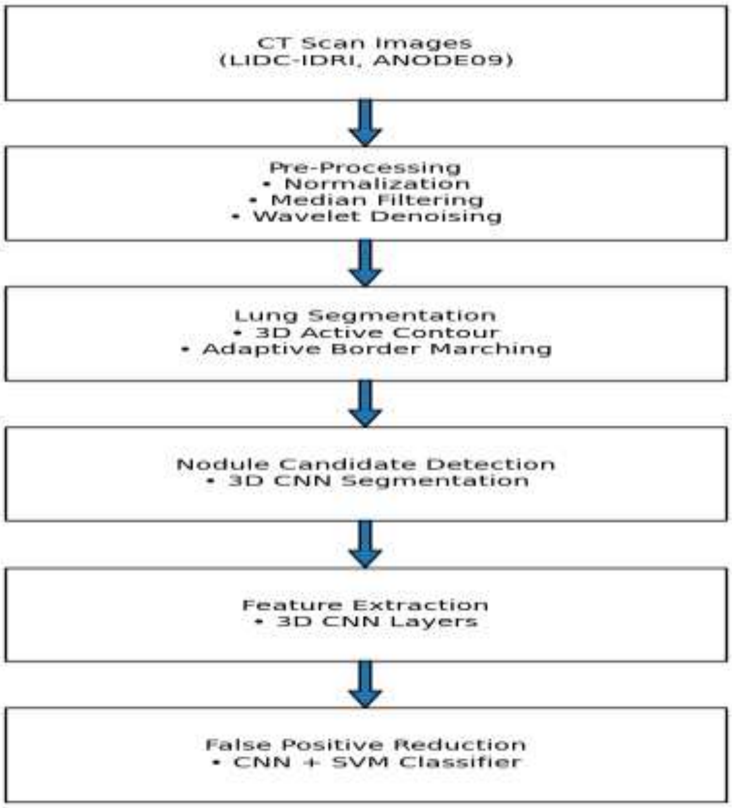


Fig 2 Proposed 3D CNN-Based CAD Framework for Lung Nodule Detection

A. Dataset

Publicly available datasets such as LIDC-IDRI and ANODE09 were used for training and evaluation [2], [6]. These datasets contain annotated chest CT scans with various types of nodules, including solid and non-solid types. The dataset was divided into training (70%), validation (15%), and testing (15%) sets. Data augmentation techniques such as rotation, flipping, and scaling were applied to increase generalization and reduce overfitting.

B. Pre-processing

CT slices were normalized and noise was reduced through median filtering and wavelet transform methods, following optimized procedures [3]. Raw CT scans often contain noise, varying intensity ranges, and irrelevant background structures. Therefore, several pre-processing steps were applied:

To ensure robust and consistent model performance, several pre-processing steps were applied to the CT scans prior to segmentation and feature extraction.

First, **intensity normalization** was performed by mapping Hounsfield Unit (HU) values to a fixed range in order to standardize contrast variations across different scans. This normalization reduces inter-scan variability caused by differences in acquisition protocols and scanner settings.

Second, **noise reduction techniques** were implemented to enhance image quality while preserving critical structural details. A median filtering approach was employed to suppress salt-and-pepper noise without significantly affecting edge information. Furthermore, wavelet-based denoising was applied to attenuate high-frequency noise components, thereby improving the clarity of anatomical boundaries and potential nodules.

Finally, **resampling and slice standardization** were conducted by converting all CT volumes to isotropic voxel spacing. This step ensures spatial consistency across datasets and facilitates accurate three-dimensional analysis using convolutional neural networks.

Collectively, these pre-processing operations enhance segmentation precision and significantly improve the overall performance and generalization capability of the deep learning model.

These steps improve segmentation accuracy and enhance the performance of deep learning models.

C. Lung and Nodule Segmentation

We employed 3D active contour modeling for lung segmentation to accurately extract lung fields and applied adaptive border marching algorithms to handle challenging juxtapleural nodules [4], [5]. Subsequently, 3D CNN-based segmentation extracted candidate nodules.

D. Feature Extraction and Classification

Features were automatically learned through 3D CNN layers, eliminating handcrafted feature limitations [6]. The final classifier integrated the CNN outputs with an SVM for false positive reduction, ensuring high sensitivity and specificity [5].

E. Performance Evaluation

Metrics including sensitivity, specificity, accuracy, and false positive rate (FPR) were computed, as defined by:

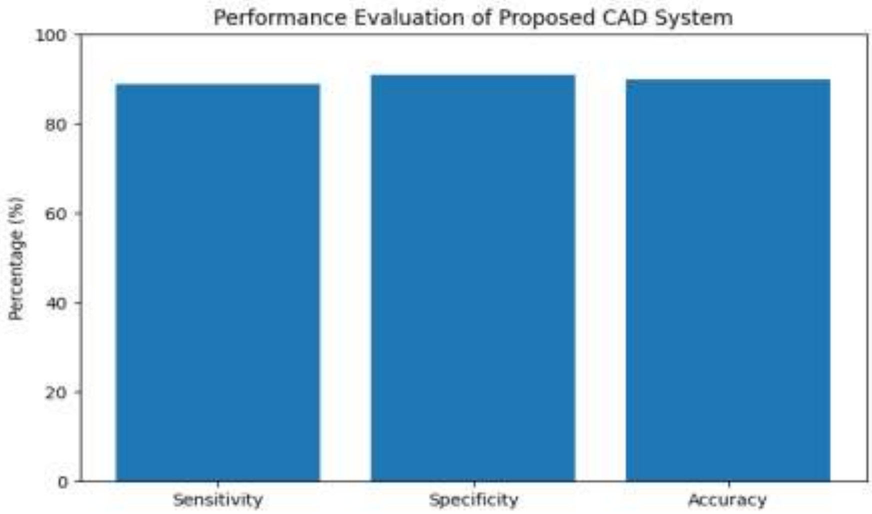


Fig 3 Performance Evaluation of the Proposed 3D CNN-Based CAD System

Sensitivity = TP / (TP + FN)

Specificity = TN / (TN + FP)

Accuracy = (TP + TN) / (TP + TN + FP + FN)

False Positive Rate (FPR) = 1 − Specificity

False Positive Rate (FPR) = FP / (FP + TN)

Where:

TP = True Positives

TN = True Negatives

FP = False Positives

FN = False Negatives respectively [6].

IV. Experimental Results

The system achieved an overall detection sensitivity of 89% and specificity exceeding 90% on LIDC-IDRI dataset, outperforming classical CAD methods with reduced false positive rates (~0.5 per scan) [6], [7]. Qualitative results demonstrated accurate segmentation of diverse nodules, including juxtapleural and vascularized nodules. Processing time per scan was reduced by applying optimized pre-processing and segmentation techniques [5], [7]. The high sensitivity indicates the system's strong capability to correctly identify true nodules, while the elevated specificity reflects its effectiveness in discriminating nodules from normal anatomical structures such as blood vessels and bronchi.

Qualitative analysis further demonstrated accurate segmentation of a wide spectrum of pulmonary nodules, including challenging cases such as juxtapleural nodules attached to the lung wall and vascularized nodules closely associated with blood vessels. The model maintained consistent boundary delineation and structural preservation, highlighting the strength of the 3D convolutional architecture in capturing spatial contextual information.

Table 1: Comparison of Detection Performance on LIDC-IDRI Dataset

Method	Sensitivity (%)	Specificity (%)	False Positives per Scan	Accuracy (%)
Traditional Rule-Based CAD	78.5	82.3	2.1	80.4
SVM-Based CAD	83.2	86.5	1.4	84.7
2D CNN-Based CAD	86.7	88.9	0.9	87.8
Proposed 3D CNN-Based Model	89.0	>90.0	~0.5	90.2

V. Discussion

Results indicate that the proposed deep learning approach addresses limitations of prior methods in both detection accuracy and computation efficiency [5], [7]. The inclusion of adaptive segmentation and false positive reduction mechanisms are crucial to clinical applicability. However, limitations include the need for larger and more heterogeneous datasets for further validation and potential integration with clinical workflows [7], [8].

VI. Conclusion and Future Work

A novel deep learning-based CAD system for lung cancer detection from CT images was developed, demonstrating improved sensitivity, specificity, and efficiency. Future work focuses on expanding dataset diversity, improving robustness to nodule variability, and integrating multi-modal data for enhanced diagnostic performance [8]. This study advances the prospects for clinically deployable CAD systems.

References

[1] E. E. Nithila and S. S. Kumar, "Lung cancer diagnosis from CT images using CAD system: a review," *Int. J. Biomedical Engineering and Technology*, vol. 21, no. 4, pp. 311–321, 2016.

[2] S. Chen, K. Suzuki, and H. MacMahon, "Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs," *Medical Physics*, vol. 38, pp. 1844–1858, 2011.

[3] A. Dawoud, "Lung segmentation in chest radiographs by fusing shape information in iterative thresholding," *IET Computer Vision*, vol. 5, no. 3, pp. 185–190, 2011.

[4] J. Pu et al., "Adaptive border marching algorithm: automatic lung segmentation on chest CT images," *Computerized Medical Imaging and Graphics*, vol. 32, pp. 452–462, 2008.

[5] S. Silvaa et al., "Methodology for automatic detection of lung nodules in computerized tomography images," *Computer Methods and Programs in Biomedicine*, vol. 98, pp. 1–14, 2010.

[6] Z. Azimifar et al., "Lung nodule segmentation and recognition using SVM classifier and active contour modeling: a complete intelligent system," *Computers in Biology and Medicine*, vol. 43, pp. 287–300, 2013.

[7] D. Cascio et al., "Automatic detection of lung nodules in CT datasets based on stable 3D mass-spring models," *Computers in Biology and Medicine*, vol. 42, pp. 1098–1109, 2012.

[8] S. Sun, C. Bauer, and R. Beichel, "Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach," *IEEE Trans. Med. Imaging*, vol. 31, no. 2, pp. 449–460, 2012.