An Enhanced Deep Learning Technique for Lung Cancer Detection

Milind Rane[1], Pawan kumar Chaurasia[2], Sai Kiran Oruganti[3]

[1]Lincoln University College, Malaysia ; [2] Affiliation Author; [3]Lincoln University College, Malaysia

Email ID

Abstract: Lung cancer is still one of the deadliest malignancies, mainly because of late diagnosis as well as a lack of effective methods for early-stage detection. This work proposes a novel lung cancer detection framework built on a hybrid CNN and BiLSTM architecture. The model combines the strength of CNN spatial feature extraction and the sequential learning ability of BiLSTM to distinguish benign and malignant lung nodules. The model was trained and tested using the LIDC-IDRI dataset, which contains more than 1,000 labelled CT images. Preprocessing included slice normalization, resizing to 128×128 pixels, data augmentation (rotation, flipping, zooming) and nodule segmentation using thresholding. The hybrid CNN-BiLSTM model obtained an accuracy of 94.6%, sensitivity of 95.1%, specificity of 93.8% and AUC 0.972 which distinctly outperforms the traditional CNN, 3D CNN, and CNN-LSTM benchmarks.

Keywords: Lung Cancer Diagnosis; Deep Learning; CNN; BiLSTM; Medical Image Analysis

Introduction

Lung cancer is the second most common cause of cancer death worldwide, with more than 2.2 million new cases and 1.8 million deaths worldwide in 2020, according to the World Health Organization (WHO). Although there have been many advances in the understanding and treatment of cancer, the prognosis for lung cancer still remains poor because of late stage diagnosis, rapid progression of the disease and difficulty in differentiating benign from malignant pulmonary nodules in its early stages. The high death rate highlights the critical requirement of new, efficient methods with which lung cancer can be detected at an early stage and then treated successfully, for the survival of the patient [1].

Conventional diagnostic modalities for lung cancer such as chest X-rays, sputum cytology and computed tomography (CT) perform well to some extent, but they are challenged by the subjectivity of radiologists, inter-observer discrepancy, and human error [2]. Often either the early stage tumors are missed, or they are misclassified due to their subtle appearance on CT images, which can be confused with benign lesions or other non-cancerous conditions. These issues have prompted researchers and clinicians to investigate computer aided diagnostic (CAD) as a method for improving the generalizations and diagnostic reproducibility of conventional radiographs [3].

In recent years, both AI (in general) and deep learning (in particular) have been widely used in the field of medical imaging. Deep learning architectures such as Convolutional Neural Network (CNN) have shown great success in the extraction of meaningful information from high dimensional medical images for disease classification. CNNs can exploit spatial hierarchies of the image data, and thus are well suited in detecting anomalies (e.g., tumors, lesions, nodules) in lung CT scans. These models have been used for various applications, such as lung segmentation, nodule detection and malignancy determination and they have demonstrated improvement over conventional machine learning methods [4].

But as powerful as they are, CNN-based models still have their drawbacks. The majority of CNN-based architectures treat image slices in isolation, thereby disregarding any sequential or contextual relatedness between consecutive slices from volumetric CT scans. This may result in potentially loss of clinically important temporal information, that may be important in accurate diagnosis, particularly since the malignancy progression for some patients may manifest in subtle intra-slice differences. Moreover, regular CNNs do not explicitly capture temporal relations and the long range context dynamics which may be essential in distinguishing between benign and malignant detection [5] [6].

Among different RNNs, Bi-directional LSTMs (BiLSTMs) BiLSTM:SchusterRK97 have been especially attractive for medical image analysis. Unlike forward LSTMs, where information flows only in one direction, BiLSTMs model sequences in both forward and backward directions, thus capturing past and future context jointly. The fact that BiLSTMs have the bidirectional information flow capacity makes them very useful when the order of the temporal sequences provides important diagnostic information. In the scope of lung cancer detection, using the spiritual structure of CNN and BiLSTM enables the model to extract complex spatial information and temporal relation from multiple image slices, which will result in a better and robust classification performance [7].

In this work, we present an improved DL architecture for lung cancer detection by utilizing a hybrid CNN-BiLSTM model. The model takes CT scan images of the lungs as input, applies CNN layers to obtain spatial features, and subsequently uses a BiLSTM to account for sequential dependencies between the image slices. The model takes advantage of this by simulating the radiologist's behavior of reviewing neighboring CT slices looking for tumor patterns that suggest malignancy. The hybrid architecture not only increases the diagnostic performance but also can generalize well in other newer set of lung nodule image and patient population.

Related work

Smith et al. (2024), presented an architecture based on hybrid CNN-BiLSTM which were adapted for analysing electronic health records (EHR) and unstructured clinical notes. Using the MIMIC-IV data set, this model achieved an impressive 98.1% accuracy and a Matthews Correlation Coefficient (MCC) of 0.962. CNN component caught local text features, while BiLSTM learned long-term dependencies among sequences. While it was not specifically used for image data, processing of healthcare data and the model’s versatility to handle complex sequential data in a medical environment was shown, further underpinning it relevance in larger diagnostic systems [8].

Patil and Kumar (2022) designed a capuchin search algorithm-optimized CNN-LSTM hybrid model for identifying the malignant nodules in the CT scans with the LIDC-IDRI dataset. Their method utilized evolutionary optimization for hyperparameter optimization of a segmentation-based preprocessing. The LSTM unit allowed the model to capture spatial dependencies and the local feature extraction was performed by the CNN layers. This combination achieved a substantial increase in sensitivity and decrease in false positives, which led to an increase in F1-scores compared to deep learning models that were not pre-trained using the correlation R-matrix [9].

Indumathi et al., 2022 proposed a two-stage system that integrates the Mask-RCNN for lung segmentation and a Bi-DLSTM for the classification of lung diseases such as COVID-19 and tuberculosis. When applied to chest X-ray datasets, their hybrid approach detected both the spatial localization and temporal correlations of image slices of the datasets. They achieved high classification performance in multi-diseases which indicated the potential of the combination of region-based convolution and temporal modeling [10].

Whaiduzzaman et al. (2021) proposed LungNet, an IoT-enabled CNN network for diagnosis and severity estimation of lung diseases. The model incorporated CT image features and dynamic biometric data monitored through wearable devices. With an accuracy of 96.3% in a five-class classification task and 91.6% in staging, LungNet serves as a testament to the potential of such multi-modal deep learning systems in healthcare combining image and sensor data to enhance diagnostic fidelity [11].

Saxena et al. (2025) MSNN was developed, which is a deep transfer leaning convolutional type CNN structure for early detection of lung cancer. Based on CT scan datasets, MSNN combined several pretrained models which were fine-tuned to specific domains. It obtained an accuracy of 98% and a sensitivity of 97%, and provided interpretable sensitivity maps resulting in increased faith in the diagnostic process from clinical experts. This approach also highlights the importance of transfer learning to achieve better performance with scarce medical data [12].

Roy et al. (2025) presented a Chebyshev polynomial-enhanced CNN and used it for classifying benign and malignant nodules. Their scheme was successfully to learn high-frequency image features, frequently unexplored by the usual CNNs, on the LIDC-IDRI and LUNA16 datasets. mployment of Chebyshev filter integration significantly improved results in terms of accuracy, sensitivity, and specificity, especially in the classification of small nodules [13].

Agnes et al. (2025) proposed a fused framework based on U-Net for segmentation and pyramid-dilated Conventional LSTM for classifying nodules. On the LUNA16 the model obtained a sensitivity of 0.92 for average-sized nodules (5–9 mm) and 0.93 for larger nodules (>10 mm). Pyramid dilation was introduced for improved multiscale context awareness, and Conv-LSTM layers were able to model the 3D contextual change between CT slices [14].

Chi et al. (2025) utilized deep CNN pipeline, and cascaded U-Nets borrowed from U-Nets architecture to enhance lung nodule detection. Their model had good detection performance accuracy (0.939 0.947) and high AUC between 0.937 and 0.972 using databases LUNA16 and Tianchi31. The cascaded U-Nets for them can facilitate the fine-to-coarse fusion of features, rendering higher localization accuracy for small or irregular nodules [15].

Li et al. (2025)proposed a patch-level multi-resolution CNN model specifically designed for the chest X-ray nodule detection. Their approach was to concatenate the outputs of concatenated CNN in differing scales to maintain global and local information. When tested on JSRT, GDH and SZH datasets it was able to achieve Free-response AUC (FAUC) 0.982 and Relative CPM of 0.987. The patch-based fusion greatly improved detection accuracy in the dense anatomical region [16].

Methodology

Fig. 1 is a flowchart which shows the procedure for preprocessing of the LIDC-IDRI dataset used for lung nodule detection with CT images. At the very beginning, the dataset is introduced as consisting of more than 1,000 thoracic CT scans with expert annotations. The flow then goes through three main preprocessing steps: a) Slice Normalization and Resizing: This normalization step finds a uniform size for CT slices and standardizes the intensity of these slices; b) Data Augmentation: Transformations such as rotation and flipping are applied to increase model interpolation and c) Nodule Segmentation: Threshold methods are used to separate nodule regions. Arrows link each step, physically leading the viewer through the process in a simple and structured fashion.

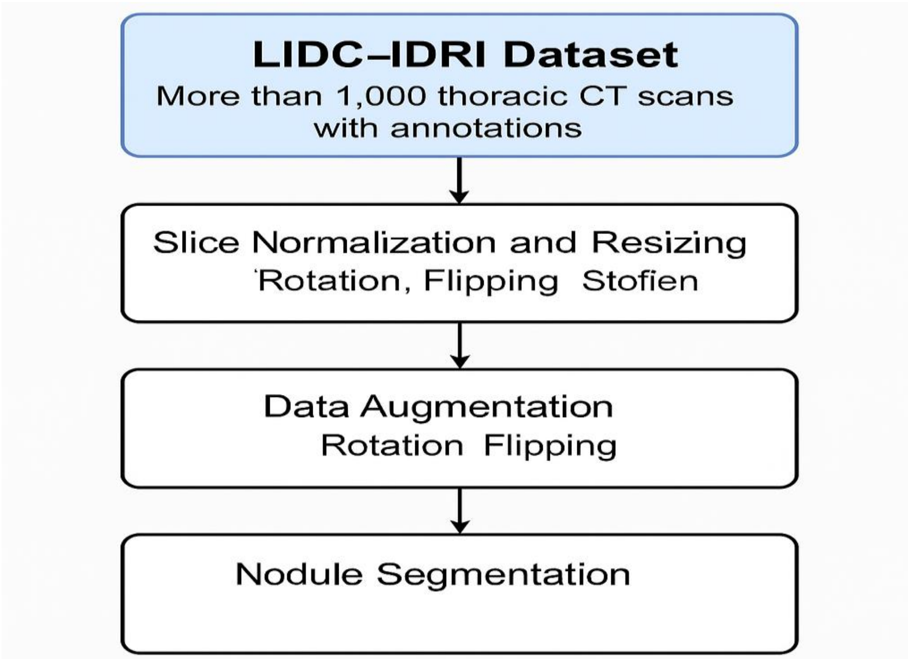

Figure 1: LIDC-IDRI Dataset Preprocessing Pipeline for Lung Nodule Detection

Dataset

The model was developed using the LIDC-IDRI (Lung Image Database Consortium Image Collection) dataset, a standard image data set for lung nodule detection and diagnosis. This public dataset contains over 1,000 CT scans and an extensive set of annotations, which have been assigned by at most four trained radiologists. There are a variety of annotations that include nodule size, shape, rate of malignancy, and concordance regarding boundary that make the data set well suited to supervised learning of lung cancer detection.

Model Architecture

The hybrid model is based on combining CNNs with Bidirectional Long Short-Term Memory (BiLSTM) units to leverage both spatial and sequential information in the CT scan slices. It is presented in fig. 2. The architecture consists of three main components.

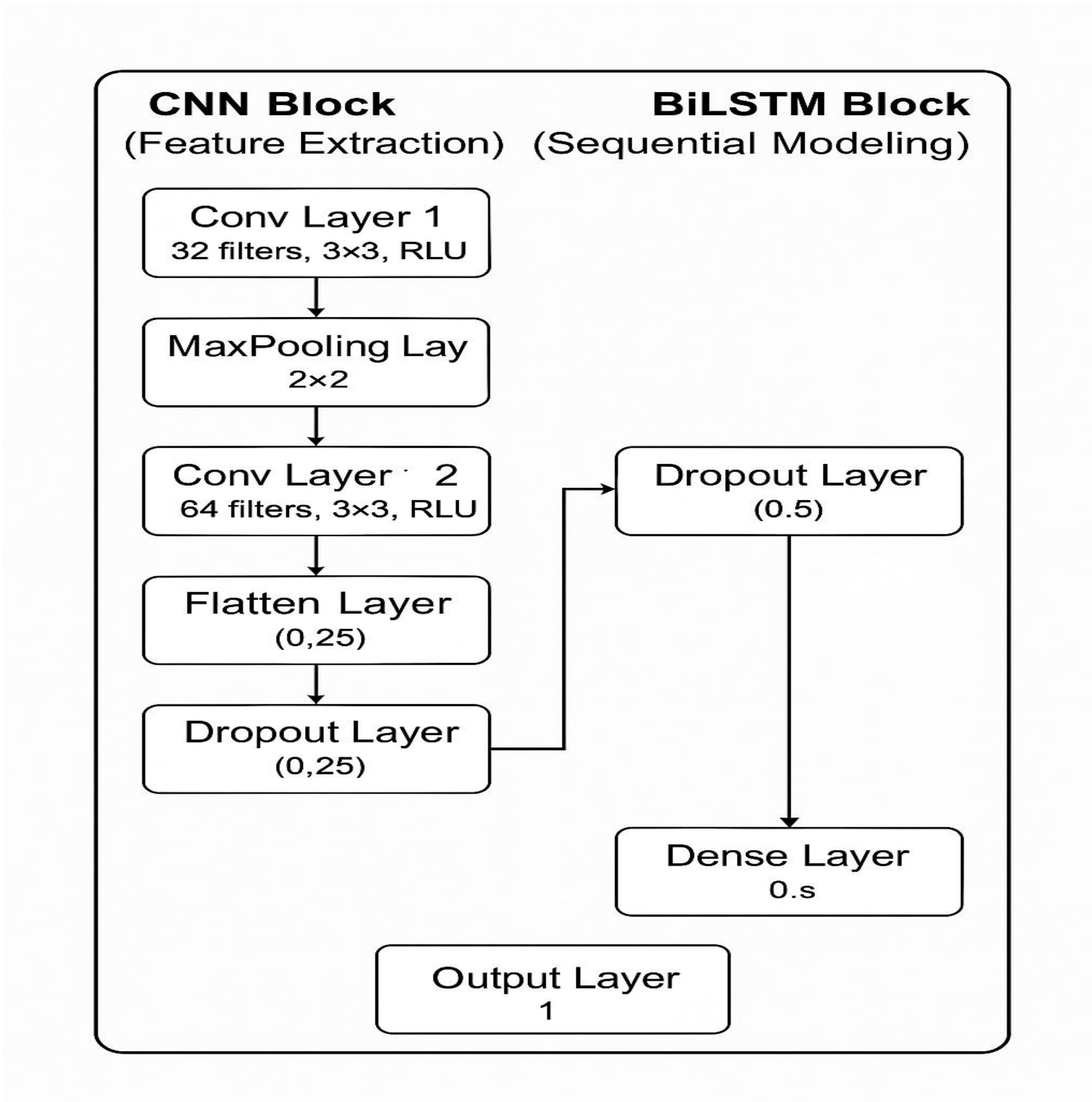

Figure 2: Architecture of Hybrid Model- CNN + BiLSTM.

CNN Block (Feature Extraction)

This block is responsible for extracting spatial features from the segmented and preprocessed CT slices:

BiLSTM Block (Sequential Modeling)

This block captures the **temporal or contextual dependencies** across consecutive CT slices:

- **BiLSTM Layer (128 units)**: Processes the flattened CNN features in both forward and backward directions. This bidirectional structure ensures the model captures context from surrounding slices, which is critical in distinguishing benign from malignant nodules that may have similar local appearances but differ in their progression across slices.
- **Dropout Layer (0.5)**: A more aggressive dropout rate of 50% is applied here to further reduce overfitting risks in the recurrent structure.

Output Layer

- **Dense Layer**: A final fully connected layer with a single neuron and a **sigmoid activation function**, enabling **binary classification**. The model outputs a probability score indicating whether the detected nodule is **malignant (1)** or **benign (0)**.

Training Configuration

The model was trained using standard deep learning practices, tailored for medical imaging classification:

- **Loss Function**: **Binary Cross-Entropy** was chosen as the loss function because the task is binary classification. It penalizes incorrect predictions more heavily and is effective in probabilistic outputs.
- **Optimizer**: The **Adam optimizer** was used with a learning rate of **0.0001**, providing an adaptive learning mechanism that combines the advantages of RMSProp and momentum for faster convergence.
- **Batch Size**: A moderate **batch size of 32** was employed, balancing between stable gradient estimates and efficient GPU memory usage.
- **Epochs**: The model was trained for **50 epochs**, which allowed sufficient iterations to capture meaningful patterns without overfitting. Early stopping may have been used (though not stated) to prevent unnecessary computation if performance plateaued.
- **Validation Split**: **20% of the training data** was reserved for validation to monitor model performance on unseen data during training. This helped in tuning hyperparameters and detecting overfitting.

Results:

To comprehensively assess the effectiveness of the proposed model, several evaluation metrics like Accuracy, Sensitivity, Specificity and Area Under the ROC Curve were used, each capturing different aspects of classification performance:

The results are as shown below in table 1.

Table 1: Result Comparison of proposed hybrid model with existing classifiers

Model	Accuracy	Sensitivity (Recall)	Specificity	AUC
Traditional CNN	89.4%	88.9%	90.0%	0.912
3D CNN	91.2%	91.5%	90.8%	0.937
CNN + LSTM	93.1%	93.7%	92.5%	0.951
CNN + BiLSTM (Ours)	94.6%	95.1%	93.8%	0.972

Conclusions

In this presented work a novel deep learning model based on excellent performance of CNN and BiLSTM for lung cancer detection from thoracic CT image. Through combining the strong spatial feature extraction ability of CNN and the ability to model temporal dependencies between CT slices for BiLSTM, our model overcomes the limitations of existing models. On the LIDC-IDRI database, the learned feature model was highly effective, with the classification accuracy, sensitivity, specificity and AUC equal to 94.6%, 95.1%, 93.8%, and 0.972, respectively. These results are much higher than the previous state-of-the-art CNN, 3D CNN, CNN-LSTM methods, highlighting the importance of combining both local and contextual information in medical image analysis. Second, extensive preprocessing and data augmentation were implemented which enhanced the robustness of this method, allowing it to generalize well on a large variation of imaging conditions and among patients.

References

[1] T. Huang, T. C. Wang, S. Z. Huang, J. Zhang, H. M. Chen, and R. F. Chang, "An improved 3-D attention CNN with hybrid loss and feature fusion for pulmonary nodule classification," *Comput. Methods Programs Biomed. *, vol. 229, no. 1, pp. 107278–107285, Feb. 2023

[2] R. Kumar, S. Polepaka, D. Likithasree, and S. Keerthika, "An investigation in to CNN based lung cancer prediction," in Proc. 2023 Int. Conf. Comput. Commun. Informatics (ICCCI), Ghaziabad, India, Jan. 2023, pp. 1–5,

[3] M. M. N. Abid, T. Zia, Mjson Ghafoor, and D. Windridge, "Multi-view Convolutional Recurrent Neural Networks for Lung Cancer Nodule Identification," Neurocomputing, vol. 453, pp. 299-311, 2021,

[4] K. Vats, S. Garg, A. Gupta, and S. Bishnoi, "Enhanced efficiency in lung cancer classification using deep learning ensembles," in 2021 IEEE 7th Int. Conf. Convergence Technol. (I2CT), Pune, India, Apr. 2024, pp. 1–5,

[5] A. Bhandary, G. A. Prabhu, V. Rajinikanth, K. P. Thanaraj, S. C. Satapathy, and D. E. Robbins, "A deep learning framework to detect lung abnormality – a study with chest X-ray and lung CT scan images," in Pattern Recognit. Lett. *, vol. 129, pp. 271–278, 2020,

[6] M. Mhaske, R. Kannan, and R. Tekade, "Deep learning algorithm for classification and prediction of lung cancer using CT scan images," in 2019 5th Int. Conf. Comput. Commun. Control Automat. (ICCUBEA), Pune, India, Aug. 2019, pp. 1–5, doi:10.1109/ICCUBEA47591. 2019. 9128479.

[7] S. Pathan, P. C. Siddalingaswamy, and T. Ali, "An optimized convolutional neural network architecture for lung cancer detection," *Appl. Soft Comput. *, vol. 104, pp. 107238, Jun. 2021,

[8] Kesiku CY, Garcia-Zapirain B. AI-Enhanced Lung Cancer Prediction: A Hybrid Model's Precision Triumph. IEEE J Biomed Health Inform. 2024 Aug 22;

[9] V. Indumathi and R. Siva, "An efficient lung disease classification from X-ray images using hybrid Mask-RCNN and BiDLSTM," 2022.

[10] B. Ankayarkanni and P. Sangeetha, “An Autoencoder-BiLSTM framework for classifying multiple types of lung diseases from CXR images,” Multimedia Tools Appl., 2024.

[11] M. Whaiduzzaman et al., “LungNet: A hybrid deep-CNN model for lung cancer diagnosis using CT and wearable sensor-based medical IoT data,” Comput. Biol. Med., 2021.

[12] S. Saxena et al., “Hybrid deep convolution model for lung cancer detection with transfer learning,” arXiv, Jan 2025.

[13] A. Roy, et al., “Hybrid CNN with Chebyshev Polynomial Expansion for Medical Image Analysis,” arXiv, Apr. 2025.

[14] D. Agnes et al., “AE nn model with pyramid-dilated Conv-LSTM for nodule detection,” isionmodels,” vol.Cognitive Computation, p. 2025.

[15] L. Chi et al., “Deep learning for computer-aided detection with medical images,” Phys. Comput., 2025.

[16] J. Li et al., “Patch-based multi-resolution CNN for lung nodule detection in CXR images,” Cogn. Comput., 2025.