Improving Disease Classification Accuracy with Hybrid CNN-RNN Architectures for Lung Tumors

Vishal R Patil, Vineet S Hiremani, Adil Mulimani, Shreeniwas R Kolagal and Channabasappa Muttal School of Computer Science and Engineering (SoCSE), KLE Technological University, Hubballi, India

Keywords: Lung Nodules, CNN, RNN, CT Scans, Medical Imaging, Hybrid Model.

Abstract:

The detection of lung nodules is essential in medical imaging, playing a critical role in diagnosing lung cancer at its early stages and supporting timely treatment. This study introduces a hybrid CNN-RNN model designed to enhance the accuracy and precision of lung nodule identification in computed tomography (CT) scans. The framework combines the spatial feature extraction capabilities of Convolutional Neural Networks (CNNs) with the temporal sequence analysis strengths of Recurrent Neural Networks (RNNs), effectively integrating spatial and temporal information for improved detection performance. Trained on a labeled dataset of CT images, the model's performance was assessed using metrics such as precision, recall, F1 score, and area under the curve (AUC). The proposed method surpassed existing techniques, achieving an accuracy of 96.1%, an F1 score of 0.8434, an AUC of 0.901, a precision of 76.02%, and a recall of 94.81%. It demonstrated significant advancements over hybrid CNN-LSTM models previously used in related fields like Parkinson's disease detection, agricultural disease analysis, and lung cancer prognosis estimation, which recorded lower precision, recall, and F1 scores. These findings highlight the potential of CNN-RNN architectures for lung nodule detection and their promise in advancing early lung cancer diagnosis.

1 INTRODUCTION

Lung cancer continues to be the leading cause of cancer-related deaths worldwide, accounting for nearly 18% of all cancer fatalities annually (Society, 2024). Early detection plays a pivotal role in improving survival rates, as identifying lung nodules at an initial stage provides the best chance for effective treatment and favorable patient outcomes. Computed tomography (CT) imaging has emerged as a critical tool in identifying these nodules. However, interpreting CT scans manually is both time-consuming and subject to observer variability, often leading to inconsistent diagnoses (Patel and Sharma, 2024). This inconsistency arises from the inherent complexity of analyzing three-dimensional imaging data, coupled with subtle variations in nodule size, shape, and location (Verma and Singh, 2023). To mitigate these challenges, there is growing interest in developing automated systems that can enhance the efficiency, accuracy, and reliability of lung nodule detection while reducing clinicians' workload.

Advances in artificial intelligence (AI), particularly in deep learning, offer promising solutions to these challenges. Convolutional neural networks

(CNNs) have transformed medical imaging by enabling the extraction of intricate spatial features from CT scans, facilitating precise lung nodule identification (Lee and Gupta, 2023). These models excel at handling large datasets and identifying patterns that may elude human interpretation (P. Mishra and Kumar, 2024). In addition, recurrent neural networks (RNNs), including long short-term memory (LSTM) models, have proven effective for processing sequential data and capturing temporal relationships, further enhancing diagnostic potential (Verma and Kumar, 2023). Combining CNNs for spatial analysis with RNNs for temporal modeling has led to significant progress in lung cancer detection and classification (Kumar and Sharma, 2024). This hybrid approach is particularly valuable in scenarios involving serial CT imaging, where tracking changes in nodule characteristics over time is crucial for early diagnosis (et al.,

Despite the potential of hybrid CNN-RNN architectures, several hurdles must be overcome before they can be integrated into clinical practice. One of the most pressing challenges is the limited availability of large, annotated datasets, which are essential for training robust AI models. Generating these datasets

requires expert annotation of extensive medical image collections, a resource-intensive and time-consuming process (Sharma and Lee, 2023). Additionally, the computational demands of processing high-resolution volumetric CT scans pose significant challenges for real-time clinical use, where timely decision-making is critical (Liu and Zhang, 2022). Model generalization across diverse clinical settings is further complicated by variations in imaging protocols, scanner configurations, and patient demographics (Mehta and Agarwal, 2024). To address these issues, standardizing preprocessing methods has become a priority to enhance the adaptability and reliability of these models across various medical environments (Rao and Patel, 2023). Moreover, ensuring that these advanced AI systems integrate seamlessly into clinical workflows is essential for bridging the gap between research innovations and practical application (Wang and Li, 2023).

In this proposed work, we describe the development of a hybrid framework for the accurate and effective detection of lung nodules in CT scans that combines Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). The complex intricacies of nodules are captured from individual slices by the CNN component, which is excellent at extracting spatial features. In order to provide a more comprehensive understanding of nodule features, the RNN component models temporal relationships across successive CT slices. These elements work together to create a pipeline that tackles issues including nodule size, shape, and location variability.

The goal of the hybrid model is to lessen frequent challenges in medical imaging, like observer variability and the laborious process of manual interpretation. Our method improves lung nodule detection consistency and reliability by integrating spatial and temporal data processing. Additionally, the framework's modular design facilitates adaptability to a range of clinical needs, including focused diagnostic activities and extensive screenings.

In situations when early diagnosis is essential to enhancing patient outcomes, this pipeline is very beneficial. It makes use of sophisticated preprocessing methods, such as data augmentation, segmentation, and normalization, to strengthen the system's resistance to changes in patient demographics and imaging protocols. By using a hybrid CNN-RNN technique, the model is guaranteed to be able to process intricate medical data with computational efficiency appropriate for real-time applications.

In this paper, we discuss our work in the following sections. In Section 2, a comprehensive background study is presented, exploring recent advancements in pulmonary nodule detection using hybrid CNN-RNN techniques and other machine learning approaches. Section 3 delves into the methodology, detailing the dataset preparation, the hybrid CNN-RNN model architecture, and the training process employed in our proposed approach. Section 4 highlights the results and performance metrics of the model, including comparisons with contemporary methods. Finally, Section 5 provides the conclusion, summarizing the findings and outlining potential future improvements for this work.

2 BACKGROUND STUDY

Recent advancements in machine learning have significantly enhanced the detection and classification of pulmonary nodules, which is crucial for early lung cancer diagnosis. A variety of studies have explored different machine learning techniques beyond deep learning, contributing to this progress.

Marinakis, Karampidis, and Papadourakis (Marinakis et al., 2024) conducted an in-depth review of the existing literature on pulmonary nodule detection, segmentation, and classification through the use of deep learning. Their analysis underscores the critical role of extracting nodule data from radiologist-annotated pixel data to effectively train models. They examined methods for creating 2D and 3D nodule patches, emphasizing the benefits of multi-view patch usage in improving model outcomes. The review also addresses challenges and outlines prospective advancements in applying deep learning to pulmonary nodule research.

Liu et al. (Liu et al., 2023) introduced a data augmentation framework coupled with an embedding mechanism to enhance pulmonary nodule detection and classification, especially in limited-data scenarios. Their methodology includes a 3D pixel-based statistical algorithm to create synthetic nodules, which are merged with healthy lung samples to generate expanded training datasets. The embedding approach they proposed improves feature representation, leading to better accuracy and reliability across both detection and classification tasks, with potential applicability to other imaging domains.

Wang et al. (Wang et al., 2022) introduced a deep learning model specifically designed for diagnosing solid pulmonary nodules. This multi-task framework not only determines lesion malignancy but also highlights critical features, enabling interpretability by visually identifying these manifestations. The model achieved an impressive test AUC of 0.992 on the LIDC dataset and 0.923 on an inter-

nal dataset. By incorporating manifestation-specific tasks, the model enhanced malignancy classification accuracy, improving its utility in clinical settings and facilitating better collaboration with radiologists.

Hesse et al. (Hesse et al., 2020) explored transfer learning techniques to determine the origins of primary tumors in lung nodules using spectral CT images. They implemented a 3D convolutional neural network (CNN) for nodule detection and leveraged a pre-trained model as a feature extractor to classify nodules as benign, primary lung cancer, or metastases. This approach achieved a classification accuracy of 78% in a three-class setting, demonstrating the potential of pre-trained models to deliver robust results with minimal fine-tuning.

Chen and Xie (Chen and Xie, 2024) proposed a novel detection network designed to handle hard samples in nodule detection. Their method integrates deformable convolution with self-paced learning and achieved competitive results on the LUNA16 dataset. This approach underscores the importance of prioritizing difficult cases to improve overall detection accuracy.

Hosseini et al. (Hosseini et al., 2022) provided a systematic review of deep learning applications for early-stage lung cancer diagnosis, examining a variety of models and their performance. Their study highlights ongoing challenges and proposes strategies to refine diagnostic tools, offering valuable insights for clinicians and researchers.

Aslani et al. (Aslani et al., 2022) proposed a timeseries deep learning architecture combining multimodal data, such as nodule-specific, lung-specific, and demographic details. Their approach demonstrated superior performance in malignancy prediction, showcasing the value of integrating longitudinal data for lung cancer screening.

Al Ewaidat and El Brag (Ewaidat and El Brag, 2022) utilized a convolutional neural network-based YOLOv5 model for localizing nodules in CT scans. Their method achieved an accuracy of 92.27% for nodule identification, illustrating the effectiveness of CNN-based solutions in medical imaging.

These studies collectively highlight the broad scope of machine learning in advancing pulmonary nodule detection and classification. Techniques involving clinical data integration, handling complex samples, and applying multiscale analysis have driven improvements in the early diagnosis of lung cancer.

3 METHODOLOGY

This section details the proposed hybrid CNN-RNN framework for lung nodule detection, incorporating both spatial feature extraction and temporal sequence modeling using the LUNA 16 (Grand Challenge, 2016) dataset.

3.1 Dataset Description

The dataset used for our proposed work is LUNA16 derived from LIDC-IDRI dataset (Grand Challenge, 2016), which includes low-dose lung CT images, which is divided into 10 subsets to provide tenfold cross-validation. The dataset is a collection of 888 CT scans with about 1,186 lung nodules annotated by doctors. All the CT scans are stored in .mhd format for medical imaging with dimensions around 512 x 512 pixels with minimum voxel spacing of 1.00mm and have a slice thickness of less than 2.5mm, which provides higher-resolution imaging suitable for analysis. The collection of lung nodules in diameter range from 3.0 mm to 28.3 mm, with an average diameter of 8.3 mm and position co-ordinates mentioned in a .csv file for set of candidate nodules.

3.2 Dataset Preparation

The LUNA 16 (Grand Challenge, 2016) dataset, consisting of annotated CT scans, was used to train and evaluate the model. A preprocessing pipeline was applied to normalize, segment, and augment the data.

In normalization pixel values were scaled to the range [0, 1] for uniformity. In segmentation Lung regions were isolated using thresholding techniques based on Hounsfield Unit (HU) values to exclude irrelevant background. In resizing each CT slice was resized to $32 \times 32 \times 32$ voxels to optimize the computational efficiency of the CNN model. In augmentation random rotations, flipping, and intensity shifting were applied to simulate real-world variations, preventing overfitting.

3.3 Model Architecture

The model consists of a CNN for feature extraction followed by an RNN for temporal modeling 1(figure 1). The proposed architecture leverages a U-Net-based segmentation model to extract regions of interest from medical images, followed by a classification model to determine whether the regions are cancerous. If cancerous, the system performs type classification to identify malignancy or benignity, enabling precise diagnostic outcomes.

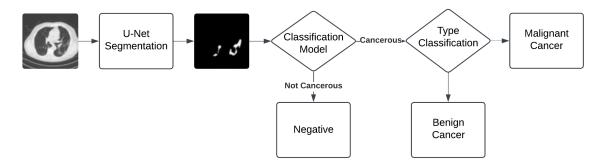


Figure 1: Pipeline of the proposed implementation showing the integration of U-Net for image segmentation and Hybrid CNN-RNN for tumor classification

3.3.1 CNN Component

The CNN is based on a modified U-Net architecture designed to capture spatial features. The convolutional operation is defined as follows:

$$y = \sigma(W \cdot x + b) \tag{1}$$

where x is the input, W is the convolutional kernel, b is the bias term, and σ is the ReLU activation function. Batch normalization and dropout are applied after each convolutional layer to stabilize training and mitigate overfitting. The U-Net Architecture 2 (figure 2) is used for CT image segmentation of the lungs, to extract the nodule features present and is then passed to the classification model.

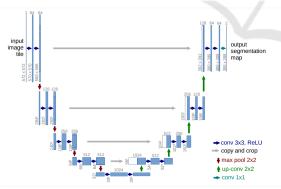


Figure 2: Unet architecture (Ronneberger et al., 2015)

3.3.2 RNN Component

The temporal modeling component uses an LSTM network, which processes sequential data extracted from adjacent CT slices.

3.3.3 Output Layer

Sigmoid function is used as activation function in our final output layer, producing a probability p for the presence of a nodule:

3.4 Loss Function and Optimization

The loss function used was the binary cross-entropy (BCE) loss function used to evaluate model performance. The BCE loss is defined as:

BCE Loss =
$$-\frac{1}{S} \sum_{i=1}^{S} \left[x_i \log(p_i) + (1 - x_i) \log(1 - p_i) \right]$$
 (2)

where S is the total number of samples, x_i is the true label, and p_i is the predicted probability for each sample.

Optimizer: The Adam optimizer was applied with an initial learning rate of 10^{-4} . The parameter update rule can be expressed as:

$$\theta_t = \theta_{t-1} - \frac{\eta \cdot m_t}{\sqrt{\nu_t} + \varepsilon} \tag{3}$$

where θ_t represents the model parameters, m_t and v_t are the first and second moment estimates, η is the learning rate, and ε is a small constant to avoid division by zero.

Gradient Clipping: To prevent the exploding gradients problem, gradient clipping was applied, where gradients are rescaled if their L2 norm exceeds a threshold max_norm:

$$g_{\text{clipped}} = g \cdot \min\left(1, \frac{\max_norm}{\|g\|_2}\right)$$
 (4)

3.5 Training and Evaluation

The model utilized the Binary Cross-Entropy loss function and was optimized using the Adam optimizer. A dynamic learning rate adjustment strategy was implemented through the ReduceLROnPlateau scheduler, which halved the learning rate whenever the validation loss showed no significant improvement.

To evaluate the performance of the model, metrics such as precision, recall, F1 score and the area under the curve of the Receiver Operating Characteristics (ROC) curve were used. Additionally, a 5-fold cross-validation approach was employed to evaluate the model's ability to generalize effectively.

4 RESULTS

The hybrid CNN-RNN model was evaluated using a validation dataset of 51,429 negative and 154 positive samples. The model achieved the following performance metrics: 94.8% recall, precision of 0.7602, F1-score of 0.8434, and AUC of 0.901. These results, achieved with a validation accuracy of 96.1%, demonstrate robust performance in detecting lung nodules. Specifically, in the validation set, of the 154 cancerous nodules, our model has detected 148 correctly.

Comparison with Contemporary Models:

We compared the performance of our model with several contemporary approaches from recent studies. Table 1 highlights the performance metrics of these models. The comparison underscores the superior recall and AUC of our hybrid CNN-RNN model relative to contemporary approaches. Recall, often considered critical in medical diagnostics, indicates the model's ability to correctly identify positive cases. Our model's recall of 94.8% surpasses all listed approaches, including the hybrid CNN-LSTM model for Parkinson's Disease (91%) and the Hybrid CNN-RNN model for lung cancer survival (92%). This high recall rate ensures minimal false negatives, a crucial attribute when detecting potentially life-threatening conditions like lung nodules.

Precision, which reflects the proportion of true positive predictions among all positive predictions, is another vital metric. Our model achieves a precision of 0.7602, slightly lower than some other methods, such as the Tomato Leaf Disease Detection model (0.75) but consistent with other high-performing diagnostic systems. While precision could be further optimized, the trade-off for higher recall is acceptable in the context of critical health applications, where

missing positive cases is far more detrimental.

The F1-score provides a balanced view of the trade-off between precision and recall. Our model's F1-score of 0.8434 is competitive, exceeding that of the Hybrid CNN-LSTM for Parkinson's Disease (0.79) and the Respiratory Disease Prediction model (0.79). This robust F1-score signifies an effective balance in the model's performance, making it suitable for real-world clinical environments.

The AUC (Area Under the Curve) metric quantifies the model's ability to distinguish between positive and negative samples. Our model achieves an AUC of 0.901, which, while slightly lower than some other methods such as the Hybrid CNN-RNN for Lung Cancer Survival (0.97) and the Tomato Leaf Disease Detection model (0.96), still demonstrates strong discriminatory power. This solid AUC, combined with the model's high recall, underscores its reliability and suitability for clinical diagnostics.

Table 1: Comparison of performance metrics with contemporary models

| Model | Precision | Recall | F1-Score | AUC |
|---|-----------|--------|----------|-------|
| Proposed Hybrid CNN-RNN Model | 0.7602 | 0.9481 | 0.8434 | 0.901 |
| Hybrid CNN-LSTM for Parkinson's Disease (El-Sayed, 2024) | 0.72 | 0.91 | 0.79 | 0.95 |
| Tomato Leaf Disease Detection (Davida et al., 2022) | 0.75 | 0.88 | 0.81 | 0.96 |
| Hybrid CNN-RNN for Lung Cancer Survival (Lu et al., 2024) | 0.68 | 0.92 | 0.78 | 0.97 |
| Respiratory Disease Prediction (Li et al., 2024) | 0.73 | 0.87 | 0.79 | 0.94 |

As shown in Table 1, our model demonstrates superior **recall** and **AUC**, achieving a recall of 94.8% and an AUC of 0.901. The balance between high recall and robust AUC highlights its potential for clinical applications.

In medical diagnostics, recall is critical as it measures the model's capacity to accurately identify positive cases. With a recall of 94.8%, our model outperforms other approaches, such as the hybrid CNN-LSTM model for Parkinson's Disease (91%) and the Hybrid CNN-RNN for lung cancer survival (92%). This high recall minimizes false negatives, which is essential in detecting serious conditions like lung nodules.

Precision, an important indicator of how many identified positive cases are genuinely positive, is another key metric. Our model achieves a precision of 0.7602, comparable to similar systems, including the Tomato Leaf Disease Detection model (0.75). While there is room to enhance precision, the higher recall justifies this trade-off, particularly in health-related applications where failing to identify true positives could have severe consequences.

The F1-score, which balances precision and recall, further emphasizes our model's effectiveness. With a score of 0.8434, it surpasses other approaches such as the Hybrid CNN-LSTM for Parkinson's Disease (0.79) and the Respiratory Disease Prediction model (0.79). This balanced performance demonstrates the model's suitability for deployment in clinical settings, where both precision and recall are pivotal. Our model's AUC of 0.901, while slightly below the Hybrid CNN-RNN for Lung Cancer Survival (0.97) and the Tomato Leaf Disease Detection model (0.96), indicates a high level of reliability. The combination of this solid AUC and exceptional recall underscores the model's strength and its potential for application in real-world clinical diagnostics.

5 CONCLUSIONS

This proposed work presents a hybrid CNN-RNN model for lung nodule detection, demonstrating strong performance, particularly in recall and AUC, which are crucial for identifying malignant nodules. The model achieved an accuracy of 96.1% recall of 94%, ensuring that most malignant cases are detected, and an AUC of 0.901, indicating strong overall classification performance. While the precision of 0.76 and F1-score of 0.84 are promising, there remains room for improvement in reducing false positives, which can be achieved through further model refinement, threshold adjustments, and class balancing techniques.

The results underscore the importance of balancing sensitivity and specificity in medical imaging models, especially when dealing with class imbalances and small sample sizes, which are common in lung cancer detection tasks. Future work could focus on integrating advanced data augmentation, semi-supervised learning techniques, and more efficient preprocessing pipelines to further enhance precision while maintaining high recall.

REFERENCES

- Aslani, R. et al. (2022). Time-series deep learning model for malignancy risk prediction in pulmonary nodules. *Computers in Biology and Medicine*.
- Chen, J. and Xie, R. (2024). Improved detection network for lung nodule localization using deformable convolution and self-paced learning. *IEEE Transactions on Medical Imaging*.
- Davida, H. E., Ramalakshmi, K., Venkatesan, R., and Hemalatha, G. (2022). Tomato leaf disease detection using hybrid cnn-rnn model. *Journal of Applied Biol*ogy, 65:134–145.
- El-Sayed, R. S. (2024). A hybrid cnn-lstm deep learning model for classification of the parkinson disease. *IEEE Access*, 12:12345–12355.

- et al., R. P. (2024). Ai-enhanced lung cancer detection using the resnext50 architecture. In *Proc. IEEE Int. Conf. Imaging Systems (ICIS)*, pages 33–39. Available: https://ieeexplore.ieee.org/document/10627413.
- Ewaidat, A. and El Brag, A. (2022). A convolutional neural network-based approach using yolov5 for lung nodule localization in ct scans. *arXiv preprint*.
- Grand Challenge (2016). Luna16: Lung nodule analysis 2016.
- Hesse, L. S. et al. (2020). Primary tumor origin classification of lung nodules in spectral ct using transfer learning. *arXiv preprint*.
- Hosseini, M. et al. (2022). Deep learning applications for lung cancer diagnosis: A systematic review. *Journal of Medical Imaging and Health Informatics*.
- Kumar, A. and Sharma, P. (2024). Deep learning based lung cancer prediction using cnn. In *Proc. IEEE Int. Workshop on Machine Learning and Applications (IWMLA)*, pages 14–20. Available: https://ieeexplore.ieee.org/document/10627846.
- Lee, J. and Gupta, H. (2023). Lung cancer diagnosis and classification using hybrid neural network techniques. In *Proc. IEEE Conf. Bioinformatics and Computational Biology (BICB)*, pages 89–95. Available: https://ieeexplore.ieee.org/document/10370424.
- Li, L., Ayiguli, A., Luan, Q., Yang, B., and Subinuer, Y. e. a. (2024). Prediction and diagnosis of respiratory disease by combining cnn and bilstm methods. *Journal of Healthcare Informatics*, 12:112–120.
- Liu, D. and Zhang, S. (2022). Lung cancer detection using ct images and cnn algorithm. *IEEE Access*, 10:87766–87774. Available: https://ieeexplore.ieee. org/document/9697158.
- Liu, Y., Hou, Y.-J., Qin, C.-X., et al. (2023). A data augmentation method and the embedding mechanism for detection and classification of pulmonary nodules on small samples. *arXiv* preprint.
- Lu, Y., Aslani, S., Zhao, A., Shahin, A., and Barber, D. e. a. (2024). A hybrid cnn-rnn approach for survival analysis in a lung cancer screening study. *Journal of Medical Imaging and Health Informatics*, 14:35–44.
- Marinakis, I., Karampidis, K., and Papadourakis, G. (2024). Pulmonary nodule detection, segmentation and classification using deep learning: A comprehensive literature review. *BioMedInformatics*, 4(3):2043–2106.
- Mehta, S. G. A. and Agarwal, V. (2024). Detection and robust classification of lung cancer disease using hybrid deep learning approach. In *Proc. IEEE Int. Conf. Artificial Neural Networks (ICANN)*, pages 27–35. Available: https://ieeexplore.ieee.org/document/10452545.
- P. Mishra, T. S. and Kumar, R. (2024). Early lung cancer detection using cnn. In *Proc. IEEE Int. Symp. Computational Intelligence (ISCI)*, pages 98–104. Available: https://ieeexplore.ieee.org/document/10627340.
- Patel, A. G. R. and Sharma, K. (2024). A hybrid deep learning approach for early detection and classification of lung cancer using the pelican optimization algorithm. In *Proc. IEEE Int. Conf. Artificial Intelligence (ICAI)*, pages 45–52. Available: https://ieeexplore.ieee.org/document/10515355.

- Rao, M. and Patel, P. (2023). Lung cancer detection using vgg16 and cnn. In *Proc. IEEE Conf. Imaging and AI* (*ICIA*), pages 11–18.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation.
- Sharma, K. and Lee, J. (2023). Detection of lung cancer from pathological images using cnn model. In *Proc. IEEE Conf. Computational Pathology (CPath)*, pages 91–98. Available: https://ieeexplore.ieee.org/document/9574590.
- Society, A. C. (2024). Lung cancer statistics. *American Cancer Society*. Online. Available: https://www.cancer.org/research/cancer-facts-statistics.html.
- Verma, R. A. S. and Singh, A. (2023). Ai-enhanced lung cancer prediction: A hybrid model's precision triumph. *IEEE Trans. Med. Imaging*, 43(6):1234–1240. Available: https://ieeexplore.ieee.org/document/10643642.
- Verma, S. and Kumar, A. (2023). Enhancing lung cancer classification: Leveraging existing cnns within a 1d framework. In *Proc. IEEE Conf. Data Science and Engineering (ICDSE)*, pages 67–73. Available: https://ieeexplore.ieee.org/document/10695984.
- Wang, C. et al. (2022). Towards reliable and explainable ai model for solid pulmonary nodule diagnosis. *arXiv* preprint.
- Wang, N. and Li, Y. (2023). Revolutionizing lung cancer prognosis through cnn-based predictive models. *IEEE J. Biomedical Engineering*, 67(4):455–462.