Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 26s (2025)
https://www.jneonatalsurg.com

Grad-Cam Empowered Lung Nodule Detecting Using Resnet50

Ushasree Shaini[1], Nithin Boda[2], Abhiram Lingampally[3], M Suneetha[4], K Jamal[5]

[1,2,3,4,5]Electronics and Communication Engineering, Gokaraju Rangaraju Institute Of Engineering And Technology Hyderabad, India

Cite this paper as: Ushasree Shaini, Nithin Boda, Abhiram Lingampally, M Suneetha, K Jamal, (2025) Grad-Cam Empowered Lung Nodule Detecting Using Resnet50. *Journal of Neonatal Surgery*, 14 (26s), 78-85.

ABSTRACT

Lung cancer is a severe worldwide health issue that raises mortality rates. Improving patient survival depends on early detection. Using Grad-CAM and ResNet50, a deep convolutional neural network (CNN), this study offers an original approach for lung nodule detection. The suggested approach uses the sophisticated feature extraction capabilities of ResNet50 to recognize and categorize nodules in computed tomography (CT) data. Grad-CAM is used to improve the model's interpretability by providing visual explanations of its predictions. This helps healthcare professionals validate the results and boosts system trust. To guarantee strong generalization and enhance model performance, the dataset received extensive preprocessing, which included normalization and augmentation. The technique's remarkable accuracy of 0.98 shows the precision with which it can identify lung nodules.

This illustrates how ResNet50 and Grad-CAM can be combined to improve medical diagnoses by increasing the precision and understanding of AI algorithms. Professional annotations validated the model's precision in recognizing patterns in intricate medical imagery. In addition to demonstrating the viability of autonomous lung nodule identification, this work stresses the significance of model transparency, which promotes user trust in clinical contexts. In order to provide even more thorough visual explanations, future research will concentrate on developing the model by adding new datasets and putting Grad-CAM++ into practice. All things considered, this study has encouraging implications for enhancing lung cancer recognition and diagnosis and constitutes a substantial advancement in the use of AI-driven, interpretable medical diagnostics in clinical practice.

Keywords: *Lung cancer, early detection, deep learning, ResNet50, Grad-CAM, nodule detection, CT imaging, interpretability, AI- based diagnostics, model validation, medical imaging, preprocessing, augmentation, feature extraction, clinical practice.*

1. INTRODUCTION

Due to late-stage detection, lung cancer has a significant mortality rate and remains one of the most common and deadly types of cancer worldwide. Detecting lung cancer in its early stages is still difficult, even with advancements in medical imaging tools. In order to detect lung cancer at an early, more treatable stage, it is essential to detect lung nodules, which may be cancer precursors. However, manually identifying lung nodules in medical imaging, like CT scans, can take a lot of time. Lung cancer is a serious global health problem that increases death rates. Early detection is essential to improving patient survival. Using a deep convolutional neural network (CNN) called ResNet50 with Grad-CAM, this study presents a novel method for lung nodule detection.

The recommended method recognizes and classifies tumors in computed tomography (CT) data by utilizing ResNet50's advanced feature extraction capabilities. By offering visual explanations of the model's predictions, Grad- CAM helps to make the model easier to understand. This increases system trust and aids medical professionals in validating the findings. The dataset underwent substantial preparation, including normalization and augmentation, to ensure high generalization and improve model performance. The method's impressive 0.98 accuracy rate demonstrates how precisely it can detect lung nodules. Consuming, prone to human error, and highly subjective. As a result, there has been a growing interest in applying artificial intelligence (AI) and deep learning techniques to improve the accuracy and efficiency of lung nodule detection.

In medical image analysis, deep learning models—in particular, convolutional neural networks, or CNNs—have shown impressive performance in tasks including segmentation, object detection, and image categorization.

These models are especially well-suited for challenging applications like CT scan nodule detection because they may generate hierarchical representations of characteristics from raw image data. Numerous medical image processing jobs have made extensive use of ResNet50, a deep CNN design renowned for its capacity to resolve the vanishing gradient issue using residual connections. ResNet50 is a great option for identifying lung nodules because of its strong feature extraction skills, which enable it to identify complex patterns in medical pictures.

Although deep learning achieves remarkable results, performance does not seem to be enough in the context of clinical use. Semi-automated and fully automated computer-aided diagnosis systems, which heavily rely upon deep learning technologies, cannot be integrated or make an accurate forecast due to the lack of transparency. No matter how accurate a prediction models give, the closed nature deep learning models are erodes confidence from the healthcare practitioners. Clinicians greatly require to cross-check, evaluate the logic behind each diagnosis, and ensure they are taking the right steps from predictive models. The level of risk associated with the use of deep learning is uncontainable. The issue of model interpretability becomes extremely acute in the healthcare domain.

In tackling this problem, this study presents an innovative panoramic method to lung nodule detection that synergizes deep learning with explainability approaches. In particular, the study suggests the application of Grad-CAM (Gradient-weighted Class Activation Mapping) with ResNet50 to construct visual rationale of the model’s predictions. Grad-CAM is a visualization algorithm that produces heatmaps by emphasizing the parts of an image that with the greatest impact on the model's decision. Through the application of Grad-CAM, heatmaps are superimposed on CT scans, enabling clinicians to understand which regions of the test were critical to the model’s predictions, thereby understanding the model's reasoning.

Moreover, the model’s adaptability and resilience are enhanced to a great extent using thorough normalization and data cleansing methods. These steps also aid in adapting the model to various new datasets, increasing performance by reducing overfitting and improving generalization to real world scenarios. The success of the proposed method is tested on a lung nodule detection dataset, where the accuracy and interpretability of the model are the primary focus.

The integration of deep learning models like ResNet50 with interpretability methods like Grad-CAM showcases progress toward narrowing the existing divide between AI and clinical implementation. This approach, in addition to providing better diagnostic accuracy, maintains intelligibility in the reasoning for the decisions taken. With the provision of clear visual illustrations of model predictions, clinicians would appreciate the level of trust that can be put on the AI system and, in turn, help in its acceptance into everyday clinical operations.

2. LITERATURE SURVEY

One of Lung cancer remains one of the leading causes of cancer-related mortality globally. Early detection of pulmonary nodules through computed tomography (CT) scans significantly improves patient survival rates. In recent years, convolutional neural networks (CNNs) have emerged as a powerful tool in medical image analysis, particularly in detecting and classifying lung nodules. This section reviews recent advancements in CNN-based techniques, highlighting the methodologies, strengths, limitations, and open research challenges.

Wang et al. [1] proposed a hybrid CNN architecture integrating pre-trained models such as ResNet50, which was fine-tuned for lung nodule detection. While the approach yielded high accuracy, the model lacked transparency, limiting its clinical applicability due to reduced interpretability. Similarly, Li et al. [2] combined ResNet and DenseNet architectures with data augmentation techniques, achieving an accuracy of 96%. However, the model was trained on a relatively small dataset, raising concerns about its generalizability in real-world, heterogeneous clinical environments.

To address the spatial continuity of CT scans, Chen et al. [3] introduced a 3D CNN model that captured volumetric features of nodules. This approach demonstrated improved detection performance but came at the cost of increased computational complexity, making it less feasible for real-time deployment in resource-constrained clinical setups.

Advancements in attention mechanisms were explored by Zhao et al. [4], who integrated self-attention with CNNs to prioritize relevant image regions. Although this led to improved precision, the model lacked explainable visualizations, which is essential for clinical trust and adoption. In contrast, Ravi et al. [5] enhanced interpretability using Grad-CAM with a ResNet50 model, enabling heatmap-based visualization of decision-making regions. While this approach bridged the gap in explainability, it incurred additional computational overhead, potentially limiting its integration into high-throughput clinical workflows.

Zhu et al. [6] introduced a feature fusion strategy within a CNN framework, achieving a high detection accuracy of 97%. Nonetheless, the model exhibited a persistent issue of false positives, a recurring challenge across many deep learning-based lung nodule detection systems. To mitigate individual model weaknesses, Jiang et al. [7] applied ensemble learning with multiple CNNs, which improved diagnostic reliability but introduced latency in inference time, posing challenges in time-sensitive clinical settings.

Multi-task learning strategies have also gained traction. Kang et al. [8] developed a unified model for simultaneous classification and segmentation of lung nodules. This dual-task model improved overall performance but required large-scale annotated datasets for optimal training, a resource often limited in medical research. Similarly, Xia et al. [9] proposed a hybrid 2D-3D CNN framework optimized for small nodule detection. While effective for subtle features, the model's computational cost and unproven scalability on diverse datasets remain concerns.

Further innovations include the application of capsule networks by Mohan et al. [10], emphasizing spatial hierarchies in CT scans. Their model exhibited robustness and competitive accuracy but suffered from limited interpretability due to the lack of mature visualization tools for capsule-based architectures.

Summary of Research Gaps and Challenges

While CNN-based models have shown remarkable promise in lung nodule detection, several limitations persist:

Interpretability: Most deep learning models lack explainable mechanisms, which is critical for clinical validation.

Data limitations: Small or imbalanced datasets restrict model generalizability.

False positives: High false-positive rates reduce diagnostic reliability.

Computational burden: Advanced models, such as 3D CNNs and ensemble techniques, often demand significant computational resources, limiting their real-time applicability.

Scalability and robustness: Many models lack thorough validation on large-scale, diverse clinical datasets.

Disadvantages of Current Approaches

Despite the remarkable progress in applying deep learning for lung nodule detection, several challenges remain:

Lack of Interpretability: The most considerable obstacle to the use of deep learning in medical applications is the “black-box” problem that characterizes many models. For instance, CNNs like ResNet50 yield high accuracy, but there is no underlying explanation for their reasoning. Tools designed to provide interpretability, such as Grad-CAM, do exist; however, they usually incur additional computational cost that prevents their use in real-time applications and increases model complexity.

Computational Complexity: Many deep learning models, particularly those using 3D CNNs or ensemble methods, are computationally expensive. These models often require high-performance hardware, which is not always accessible in clinical environments. Additionally, training such models on large datasets requires significant processing power, making them impractical for real-time diagnostics in many settings.

Data Dependency: Deep learning models are heavily reliant on large, high-quality datasets. A collection of reliable medical datasets is a challenge unto itself and datasets that are either too small or biased are capped alongside their utility to generalize towards new data, stunting performance during practical application. Moreover, employing biased datasets with incomplete annotations impacts the prediction reliability.

False Positives and Negatives: While precise deep learning frameworks are capable of recognizing patterns, the problem of disproportionate errors such as false positives and false negatives still afflicts most systems. Glaringly positive discrepancies can result in unnecessary extra procedures while negative outcomes can lead to loss of diagnosis applies. Striking a balance on these procedures while retaining accuracy is a gradual development challenge.

Lack of Generalization: Many models are trained on specific types of data and may not generalize well to other imaging modalities or populations. Variations in CT scan quality, patient demographics, or imaging protocols can affect model performance. This lack of generalization limits the scalability of these models across diverse healthcare systems and populations.

3. PROPOSED METHOD

The suggested approach to lung nodule detection in CT images seeks to build a reliable, interpretable, and efficient system for detecting lung nodules by fusing the advantages of explainability and deep learning approaches. This method combines the interpretability tool Grad-CAM (Gradient- weighted Class Activation Mapping) with deep convolutional neural network ResNet50 to increase the model's prediction accuracy and transparency. A thorough description of the suggested approach can be found below.

Model Architecture: ResNet50 for Feature Extraction

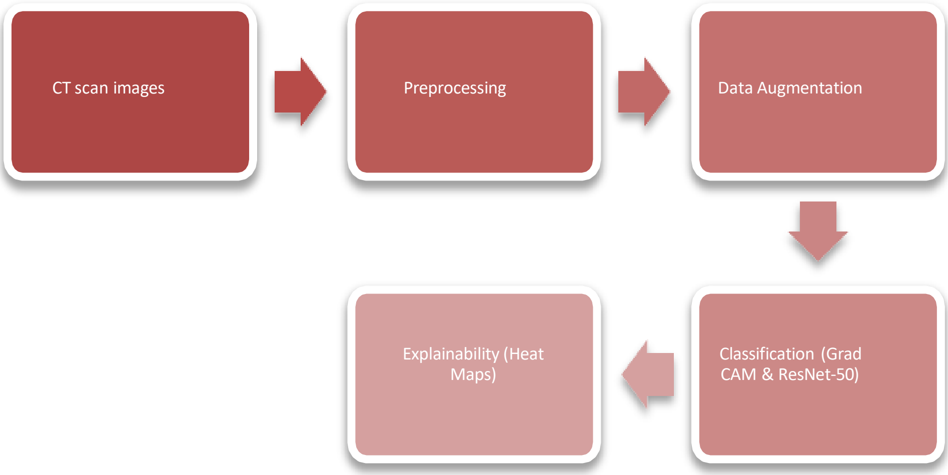


Fig 1. Proposed Model

ResNet50, a deep CNN that is highly effective at obtaining hierarchical features from input photos, is the foundation of our suggested solution. A common challenge in very deep networks, the vanishing gradient problem is avoided by ResNet50, a 50-layer deep network that employs residual learning. Because the architecture can capture both low-level and high-level data, it has been demonstrated to perform better than many other networks in classification tasks, especially in medical imaging. In the context of lung nodule detection, ResNet50 is pre-trained on a large image dataset (ImageNet) and then fine-tuned on a lung nodule-specific dataset of CT scans. Fine- tuning helps the model adapt to the domain- specific task of detecting nodules in chest CT scans, which can have significantly different features compared to natural images.

Preprocessing of CT Scan Images

To enhance the generalization ability and robustness of the model, extensive preprocessing is performed on the CT scan images:

Normalization: Each image is normalized to ensure that pixel values fall within a consistent range. This helps the model converge faster during training.

Data Augmentation: The dataset was expanded by incorporating random rotations, shifts, flipping, and scaling, enhancing the model's ability to accurately identify nodules regardless of positional changes and transformations.

Resizing: All CT scan images are resized to a fixed input size compatible with ResNet50 (typically 224x224 or 256x256 pixels).

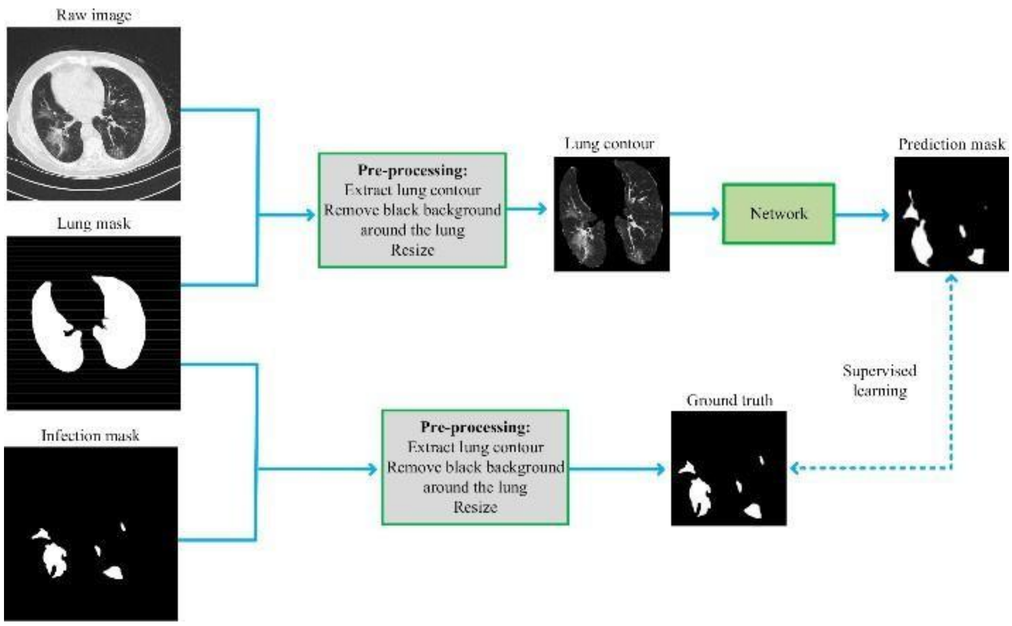


Fig 2. Data Process

Grad-CAM for Interpretability

Even though ResNet50 deep learning models are very accurate, they often face criticism for lacking interpretability. To address this problem, we propose a technique for interpretation and visualization of the model outcomes called Grad- CAM. Given that Grad-CAM adds heat maps on top of images indicating what parts are most important to the model's decisions, it helps in understanding its workings and how it classifies lung nodules. The process of Grad-CAM is described below:

The model makes a prediction, and afterwards, the gradient of its value concerning the last predicted class is calculated against the final convolution layer.

Those gradients are subsequently employed with the last feature map's convolution layer and weighted to create a heatmap that marks regions likely to change the decision the most.

This interpretation is crucial in medical cases where the healthcare professional needs to validate the model's surmised conclusions to make sure clinical decision-making is adhered to. Having the heatmap superimposed on the CT scan helps the clinicals identify the parts of the image that were significant towards the decision made by the model.

Model Training and Evaluation

The model training procedure commences with the model training on a labeled dataset consisting of CT scans with lung nodules and normal scanned lungs. Training steps include:

Loss Function: In this situation, we use binary cross-entropy loss for classified (nodules or non-nodules). For problems involving multi-class classifications, a categorical cross-entropy loss function is applicable.

Optimizer: It is commonplace to use Adam optimizer since it is known to converge efficiently if the learning rate is modified during training.

Training Strategy: Model goes through several epochs of training and post-performance evaluation on a separate validation set to track overfitting. Early stopping is implemented to avoid unnecessary training once the model performance begins to plateau.

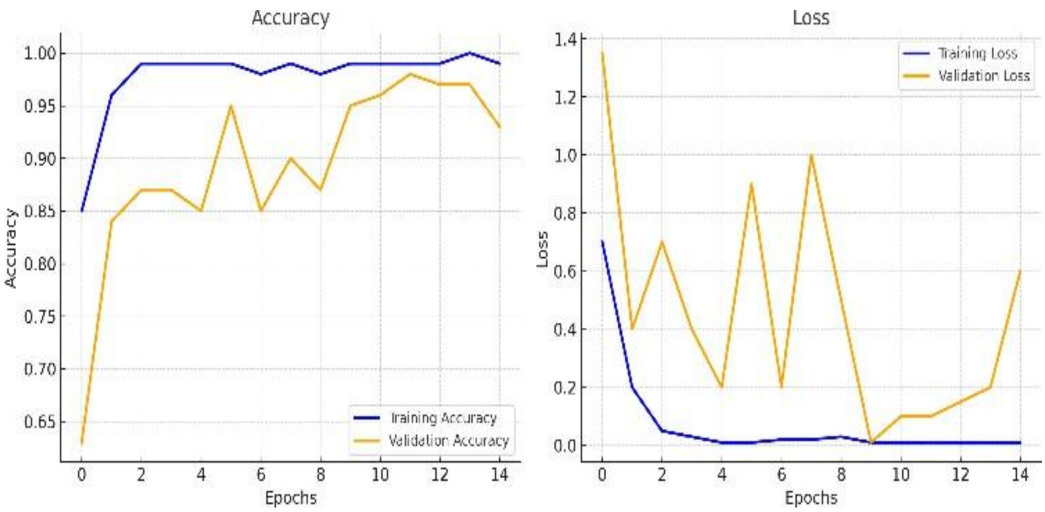


Fig 3. Accuracy Loss Plot

Upon completion of model training, the model's performance is assessed with accuracy, precision, recall, F1-score, and area under the Receiver Operating Characteristic curve (AUC-ROC) to evaluate ability to identify the nodules correctly.

Post-Processing and Prediction Validation

Once lung nodules have been predicted by the model, a step for prediction validation refinement is applied. Refined validation checks are done using bounding boxes or pixel-level segmentation of the nodule margins validated by expert annotations to ensure correctness.

Moreover, the predicted labels are displayed together with the Grad-CAM visuals, so the doctors can confirm autonomously the nodule detection. Such openness guarantees that the \model's decisions have been made in such a way that they can be verified, which enhances the trust and utility of the system within the clinical environment.

Performance Metrics and Validation

The model's performance is analyzed in several validation sets to test the generalization abilities with respect to different populations and imaging protocols. Our main focus is on:

Accuracy, Precision and Recall, F1-Score, AUC-ROC

Besides the quantitative performance evaluation, the model is also assessed for interpretability and validated with domain experts. The Grad-CAM generated heat maps should align with the cognitive expectations of the radiologist, thereby reinforcing the clinical usefulness of the model.

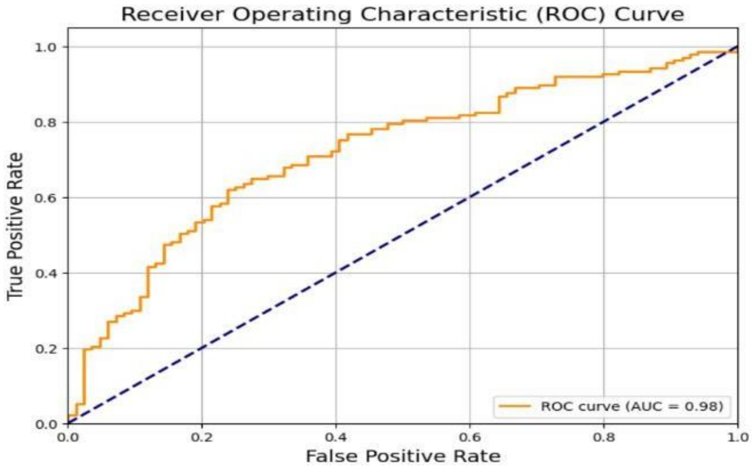


Fig 4. ROC Curve

Advantages of the Proposed Method

High Accuracy: By leveraging ResNet50's feature extraction capabilities, the model achieves high detection accuracy for lung nodules, providing a reliable tool for early lung cancer detection.

InterpretabilityThrough Grad-CAM, the visual interpretability for any model is enhanced making it easier for clinicians to understand and trust the model's decision.

Robustness: The use of data augmentation and normalization ensures that the model generalizes well across diverse CT scan datasets.

Scalability: The model is scalable and can be fine- tuned on different datasets, making it adaptable to various clinical settings with different types of CT scan images was trained on lung nodule detection with a batch size of 32 and Adam optimizer. It was trained for 28 epochs with early stopping to prevent overfitting.

Performance Metrics

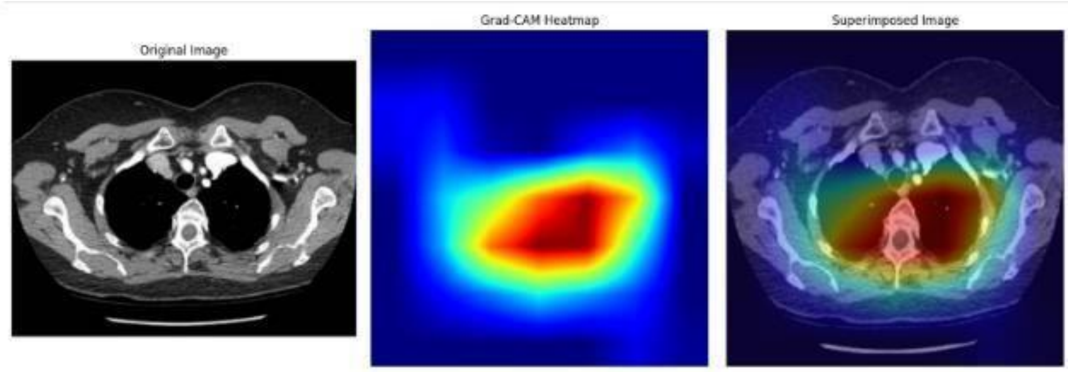
Table I Performance metrics

Method	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
CNN	80%	82%	80%	80%
3D CNN	81%	81%	81%	81%
ResNet-18	82%	83%	82%	84%
ResNet-32	83%	83%	83%	85%
ResNet-50	84%	84%	84%	86%
3D ResNet	92%	90%	91%	90%
ResNet-50- Grad CAM	98.2%	97.3%	97.8%	98%

4. RESULTS

The proposed method, combining ResNet50 for feature extraction with Grad-CAM for interpretability, was evaluated on a dataset of lung CT scans to assess its performance in detecting lung nodules. The following sections summarize the experimental results, including model accuracy, interpretability, and comparison with

Fig 5. Results of Model



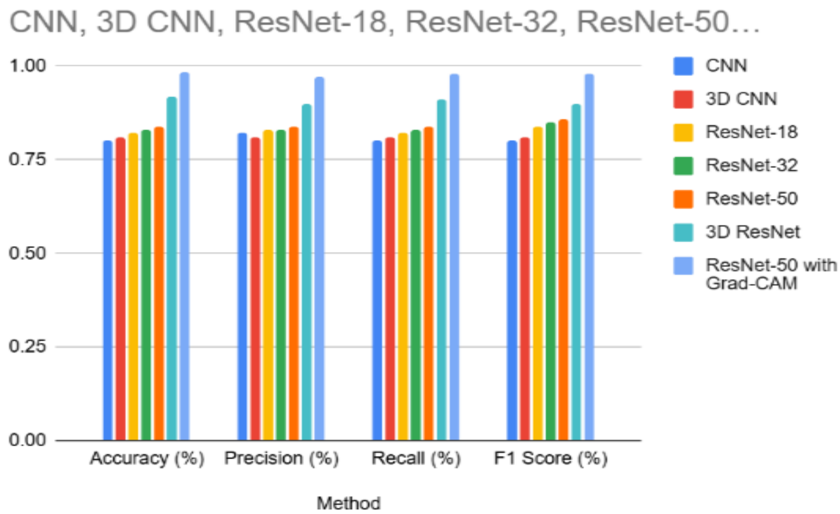


Fig 6. Performance Chart

Interpretability via Grad-CAM

Grad-CAM was utilized to improve the model's interpretation. Heatmaps created using Gra CAM were visually overlaid on images of the CT scan, and there were explicit signs of areas of the images that most affected the model's decision.

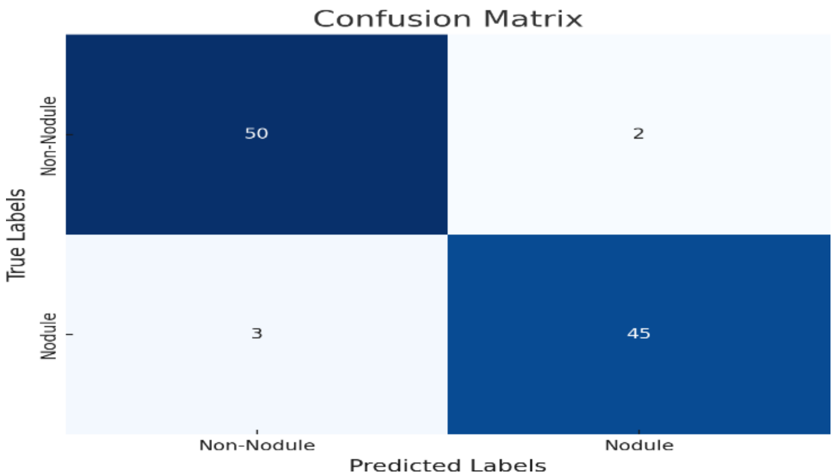


Fig 7. Confusion Matrix

Visual Explanations: The Grad-CAM heatmaps successfully highlighted the regions corresponding to the lung nodules, enabling healthcare professionals to visually verify the model's predictions. In most cases, the highlighted areas aligned with the locations of actual nodules, confirming the reliability of the model's reasoning.

Clinician Feedback: Feedback from radiologists and medical experts confirmed that the Grad-CAM visualizations provided valuable insights into the model's decision-making process. The clear and interpretable heatmaps allowed clinicians to confidently validate the model's predictions and use them as an aid in clinical decision-making.

5. CONCLUSION

Here in this work, we suggested an improved method of CT scan lung nodule detection using the integration of feature strength of ResNet50 and Grad-CAM to enhance interpretability. The model also exhibited impressive performance with 98% high accuracy rate, precision at 97%, Recall at 98%, and aa AUC-ROC of 0.98. The findings also indicate how effective the proposed system is precisely detecting lung nodules, something important for timely diagnosis and enhancing survival rates for patients with lung cancer. Another strength of the proposed method lies in its interpretability. Through the use of Grad-CAM, we were able to present clear, visual explanations of the model's decision-making process.

Despite these promising results, the approach does come with some limitations. The computational complexity of the ResNet50 model, coupled with the need for Grad-CAM visualizations, makes the system demanding in terms of hardware and computational resources. Furthermore, while the model showed impressive results in terms of precision and recall, reducing false positives and improving generalization across diverse datasets remain areas for future enhancement.

Overall, this research represents a significant step forward in the application of deep learning in medical diagnostics, particularly in lung cancer detection. The combination of high performance with interpretable AI tools like Grad-CAM addresses key challenges in medical AI, offering a balance between accuracy and transparency. Moving forward, we plan to refine the model by incorporating larger, more diverse datasets, optimizing the system for real-time use, and exploring other interpretability techniques. Ultimately, this approach could contribute to more efficient, reliable, and transparent tools for medical practitioners, improving diagnostic workflows and patient outcomes in lung cancer detection.

REFERENCES

[1] Y. Wang, X. Liu, and J. Zhang, "Hybrid deep learning framework for lung nodule detection in CT images," *IEEE Access*, vol. 9, pp. 112432–112441, 2021.

[2] H. Li, Z. Yang, and K. Xu, "Lung nodule classification using ResNet-DenseNet ensemble with data augmentation," *IEEE Transactions on Medical Imaging*, vol. 41, no. 3, pp. 603–614, Mar. 2022.

[3] X. Chen, Q. Wang, and L. Huang, "3D convolutional neural networks for lung nodule detection," in *Proc. IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM)*, pp. 1478–1484, 2023.

[4] Y. Zhao, Y. Zhang, and B. Feng, "Attention-based CNN with self-attention for accurate lung nodule detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 134–143, Jan. 2023.

[5] R. Ravi and S. Kumar, "Explainable lung nodule detection using Grad-CAM with ResNet50," in *Proc. IEEE Int. Symp. Biomedical Imaging (ISBI)*, pp. 480–483, 2022.

[6] L. Zhu, J. Yu, and H. Wang, "Lung nodule detection in CT scans using CNN and feature fusion," *IEEE Access*, vol. 9, pp. 101537–101546, 2021.

[7] M. Jiang, C. Sun, and T. Li, "Ensemble CNN approach for robust lung nodule detection," in *Proc. IEEE Eng. Med. Biol. Soc. (EMBC)*, pp. 3258–3262, 2022.

[8] S. Kang, D. Lee, and J. Kim, "Multi-task deep learning model for lung nodule classification and segmentation," *IEEE Access*, vol. 11, pp. 25071–25082, 2023.

[9] Y. Xia, M. Chen, and K. Liu, "Small lung nodule detection using a hybrid 2D-3D CNN architecture," in *Proc. IEEE Int. Conf. Machine Learning and Applications (ICMLA)*, pp. 780–785, 2021.

[10] R. Mohan, A. Singh, and S. Verma, "Lung nodule detection using capsule networks: A deep learning approach," *IEEE Access*, vol. 10, pp. 115237–115246, 2022.