Journal of Radiology and Clinical Research | Vol.1 | Issue 1 | January - April 2025| pp.1-8

Original Article

Leveraging Artificial Intelligence in Radiology: Enhancing the Detection of Subtle Pulmonary Nodules with Deep Learning Algorithms Introduction

Harleen Chawla

Assistant Professor, *Department of Radiology*, *National Capital Region Institute of Medical Sciences, Meerut, India*

Abstract:

Background: Early detection of pulmonary nodules is crucial for lung cancer outcomes, but subtle nodules often go undetected in routine radiological evaluations. This study evaluates a deep learning algorithm's effectiveness in detecting subtle pulmonary nodules on chest CT scans compared to radiologist interpretation.

Material and Methods: In this retrospective study, we analyzed 1,200 chest CT scans containing pulmonary nodules ≤8mm. A deep learning algorithm was developed and compared against junior (<5 years) and senior (>10 years) radiologists. Performance metrics included sensitivity, specificity, diagnostic accuracy, and time efficiency.

Findings The AI model achieved 91.7% sensitivity and 88.3% specificity (AUC-ROC: 0.93), outperforming junior radiologists (sensitivity: 83.3%, specificity: 85.0%) and matching senior radiologists (sensitivity: 89.2%, specificity: 86.7%). Average interpretation time was significantly reduced with AI (1.2 minutes) compared to junior (4.8 minutes) and senior radiologists (3.5 minutes). However, the AI model showed decreased performance with very low-contrast nodules (sensitivity: 72.5%) compared to senior radiologists (sensitivity: 81.4%). The deep learning algorithm demonstrates promising results in pulmonary nodule detection, particularly in reducing diagnostic time and variability. While it shows comparable performance to experienced radiologists in most scenarios, challenges remain in detecting very low-contrast nodules, suggesting its optimal use as a complementary tool rather than a replacement for radiologist expertise.

Keywords: Artificial Intelligence (AI), Radiology, Deep Learning, Pulmonary Nodules, Lung Cancer Screening, Medical Imaging, Computer-Aided Diagnosis (CAD), Convolutional Neural Networks (CNNs), Early Detection, Chest CT, Image Analysis, Diagnostic Accuracy.

Introduction

Lung cancer remains one of the leading causes of cancer-related mortality worldwide, accounting for approximately 1.8 million deaths annually.[1]
Early detection of pulmonary nodules, which often serve as precursors to lung cancer, is critical for improving patient outcomes.[2]
However, subtle pulmonary nodules—characterized by their small size and low contrast—are frequently missed during routine radiographic evaluations, even by experienced radiologists.[3] These challenges underscore the need for enhanced diagnostic tools that go beyond traditional methods. Traditional diagnostic approaches rely heavily on radiologists' expertise, which is inherently variable and prone to fatigue, especially in high-volume settings.[4]
Studies have shown that inter-observer agreement among radiologists for subtle nodule detection is moderate at best,[5] increasing the risk of delayed or missed diagnoses. Additionally, the time-intensive nature of manual interpretation exacerbates inefficiencies in the diagnostic process. Advancements in artificial intelligence (AI), particularly deep learning algorithms, offer a promising solution to these issues. Convolutional

*Corresponding Author: Dr. Harleen Chawla, harleenchawla02@gmail.com
Received: 30.01.2025 Accepted: 15.02.2025 Published: 01.03.2025

neural networks (CNNs) have revolutionized medical imaging by automating the identification of intricate patterns and abnormalities with remarkable precision.[6]

For instance, a study by Kanan et al. reported that AI-assisted detection improved nodule identification rates by 15% while significantly reducing diagnostic variability.[1]

Similarly, Gandhi et al. demonstrated that integrating AI into lung cancer diagnostics enhanced both sensitivity and specificity compared to radiologists alone.[2]

Despite these advancements, the adoption of AI in clinical practice remains limited due to concerns regarding generalizability, interpretability, and clinical validation.[7]

Most studies to date have been conducted on relatively controlled datasets, often failing to address real-world variability in imaging quality, patient demographics, and disease presentation.[8] Therefore, a rigorous evaluation of AI's efficacy in detecting subtle pulmonary nodules under diverse clinical conditions is essential.[9]

This study aims to address these gaps by evaluating the performance of a state-of-the-art deep learning algorithm for nodule detection using a robust dataset of chest CT scans. By comparing the algorithm's performance to that of radiologists with varying levels of experience, this research seeks to establish the role of AI as an adjunct tool to enhance diagnostic accuracy, consistency, and efficiency.[4] Furthermore, this study will explore the algorithm's impact on reducing time-to-diagnosis, a critical factor in high-volume healthcare environments. The findings are expected to provide valuable insights into the integration of AI into routine radiological workflows, ultimately improving lung cancer detection rates and patient outcomes.

Aims and Objectives: This study aimed to evaluate the efficacy of a deep learning algorithm in detecting subtle pulmonary nodules on chest CT scans compared to traditional radiologist-based interpretation. The primary objective was to determine the algorithm's sensitivity, specificity, and diagnostic accuracy in identifying nodules, particularly those measuring ≤8 mm in diameter and having low contrast. Secondary objectives included comparing the algorithm's performance to radiologists with varying levels of experience, assessing the time efficiency of AI-assisted detection versus manual interpretation, and exploring its potential to reduce diagnostic variability.

Material and Methods

Study Design and Population: This retrospective diagnostic accuracy study involved anonymized chest CT scans from publicly available repositories, such as the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), and a tertiary care hospital. The study population included adult patients with radiologically confirmed pulmonary nodules. Inclusion criteria required high-quality CT scans featuring pulmonary nodules with diameters ≤8 mm, representing the subtle nodules targeted by the algorithm. Exclusion criteria included low-quality scans with significant motion artifacts, scans performed for non-pulmonary indications, and scans from patients with clinical conditions deemed non-representative of the study objectives.

Data Preprocessing: A total of 1,200 chest CT scans were collected and pre-processed to ensure uniformity. Preprocessing included normalization to standardize pixel intensity values and augmentation techniques such as rotation and scaling to enhance the model's robustness. Ground-truth annotations of pulmonary nodules were established by a panel of three experienced radiologists who reached consensus on nodule locations and characteristics.

AI Model Development: The deep learning algorithm was developed using convolutional neural network (CNN) architecture. Preliminary testing included models such as ResNet and UNet, with the final model optimized to balance sensitivity and specificity. The dataset was split into training (70%), validation (20%), and testing (10%) subsets. Performance metrics, including sensitivity, specificity, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), were calculated to evaluate the model's diagnostic accuracy.

Comparison with Radiologists: The model's performance was compared with that of junior radiologists (less than five years of experience) and senior radiologists (more than ten years of experience). Each radiologist reviewed the independent test scans, and their diagnostic accuracy and time taken for interpretation were recorded. Inter-observer agreement among radiologists and between radiologists and the AI model was evaluated using Cohen's kappa statistic.

Statistical Analysis: Statistical comparisons between the AI model and radiologists' performance were conducted using paired t-tests and analysis of variance (ANOVA) for continuous variables. Diagnostic accuracy was further evaluated through

ROC curve analysis. Statistical significance was set at $p < 0.05$.

Ethical Considerations: The study was approved by the institutional ethics committee. All patient data were anonymized to protect confidentiality and comply with ethical standards.

Results

Overview of Study Population: A total of 1,200 chest CT scans were analyzed, comprising images sourced from publicly available repositories (e.g., LIDC-IDRI) and a tertiary care hospital. The dataset included scans from adult patients with radiologically confirmed pulmonary nodules, distributed evenly across training (70%, n=840), validation (20%, n=240), and testing (10%, n=120) subsets.

The mean age of patients was 54.3 ± 12.8 years, with a nearly equal distribution of genders (male: 52%, female: 48%). The pulmonary nodules included in the study had a mean diameter of 6.2 ± 1.3 mm, with nodules ≤4 mm representing 30% of the dataset, and nodules between 5–8 mm comprising the remaining 70%. The majority of nodules (65%) were located in the upper lobes, consistent with known patterns of pulmonary nodule prevalence. [Table 1] summarizes the baseline demographic and clinical characteristics of the study population.

Primary Outcome: Diagnostic Performance: The deep learning algorithm demonstrated excellent diagnostic performance in detecting pulmonary nodules, particularly in identifying subtle nodules (≤8 mm in diameter). On the independent test set of 120 chest CT scans, the algorithm achieved a sensitivity of 91.7%, specificity of 88.3%, and an area under the receiver operating characteristic curve (AUC-ROC) of 0.93. The algorithm's precision and F1-score were 89.5% and 90.6%, respectively, indicating robust nodule detection capabilities.

In comparison, junior radiologists (experience <5 years) exhibited a sensitivity of 83.3% and specificity of 85.0%, while senior radiologists (experience >10 years) achieved sensitivity and specificity rates of 89.2% and 86.7%, respectively. The AI model outperformed junior radiologists in all metrics and showed comparable performance to senior radiologists in terms of sensitivity, with statistically significant differences in specificity and precision ($p < 0.05$).

[Figure 1] presents the receiver operating characteristic (ROC) curves comparing the AI model with both junior and senior radiologists. The AI model's AUC-ROC of 0.93 was significantly higher than that of junior radiologists (0.87) and comparable to senior radiologists (0.91).

The ROC curves in [Figure 1] illustrate the diagnostic performance of the AI model, junior radiologists (<5 years of experience), and senior radiologists (>10 years of experience) in detecting pulmonary nodules on chest CT scans. The AI model demonstrated the highest area under the curve (AUC = 0.93), followed by senior radiologists (AUC = 0.91) and junior radiologists (AUC = 0.87). The diagonal line represents the performance of a random classifier.

Subgroup Analysis by Nodule Size: For nodules ≤4 mm, the algorithm achieved a sensitivity of 88.2% and specificity of 86.0%, which was higher than that of junior radiologists (sensitivity: 78.4%, specificity: 82.3%) and comparable to senior radiologists (sensitivity: 85.3%, specificity: 84.1%). For larger nodules (5–8 mm), the AI model's sensitivity was 94.1% and specificity 89.2%, surpassing both junior and senior radiologists in overall performance.

[Table 2] summarizes the diagnostic performance metrics, including sensitivity, specificity, and time efficiency, across the AI model and radiologist groups.

These results highlight the AI model's superior sensitivity and specificity compared to radiologists, particularly in identifying subtle nodules.

Secondary Outcomes

Sensitivity and Specificity across Nodule Size Subgroups: The performance of the AI model and radiologists was analyzed for nodules stratified by size (≤4 mm and 5–8 mm). For nodules ≤4 mm, the AI model achieved a sensitivity of 88.2% and specificity of 86.0%, outperforming junior radiologists (sensitivity: 78.4%, specificity: 82.3%) and demonstrating comparable performance to senior radiologists (sensitivity: 85.3%, specificity: 84.1%). For nodules in the 5–8 mm range, the AI model's sensitivity and specificity improved to 94.1% and 89.2%, respectively, surpassing both junior radiologists (sensitivity: 87.5%, specificity: 84.0%) and senior radiologists (sensitivity: 92.0%, specificity: 87.1%).

[Figure 2] illustrates the sensitivity and specificity of the AI model and radiologists for each size subgroup, highlighting the AI model's superior performance, particularly for smaller nodules.

Comparison of Diagnostic Time Efficiency: The time taken for interpreting a single CT scan was significantly shorter with the AI model compared to radiologists. The AI model required an average of 1.2 minutes per scan, while junior radiologists took 4.8

minutes and senior radiologists took 3.5 minutes ($p < 0.01$). The reduction in diagnostic time by the AI model was consistent across all size subgroups, making it particularly advantageous in high-volume clinical settings.

[Table 3] summarizes the diagnostic time efficiency across different groups.

Inter-Observer Agreement (Cohen's Kappa): The agreement between the AI model and radiologists was evaluated using Cohen's kappa statistic. The AI model demonstrated substantial agreement with senior radiologists ($\kappa = 0.82$) and moderate agreement with junior radiologists ($\kappa = 0.68$). Agreement among radiologists was higher for nodules in the 5–8 mm range ($\kappa = 0.78$) compared to smaller nodules ≤4 mm ($\kappa = 0.61$), underscoring the AI model's ability to reduce variability in challenging cases.

Subgroup Analyses

AI Performance under Different Clinical Conditions: The performance of the AI model was evaluated across subgroups defined by specific clinical imaging conditions, such as motion artifacts and low-contrast nodules. For scans with motion artifacts (n = 150), the AI model achieved a sensitivity of 83.4% and specificity of 81.2%, which was slightly lower than its overall performance (sensitivity: 91.7%, specificity: 88.3%). This decline was attributed to challenges in accurately delineating nodule boundaries in motion-affected scans. In comparison, junior radiologists demonstrated a sensitivity of 76.3% and specificity of 79.4%, while senior radiologists achieved 81.5% and 80.9%, respectively, in the same subgroup. For low-contrast nodules (n = 100), the AI model outperformed both junior and senior radiologists. The AI achieved a sensitivity of 85.6% and specificity of 84.0%, compared to junior radiologists (sensitivity: 73.5%, specificity: 78.2%) and senior radiologists (sensitivity: 82.4%, specificity: 81.7%). These findings highlight the AI model's ability to maintain diagnostic accuracy under challenging conditions, albeit with a slight drop in sensitivity and specificity for motion-affected scans.

Comparison of Results in Training vs. Testing Datasets: To evaluate the model's generalizability, its performance on the independent test dataset (n = 120) was compared to that on the training dataset (n = 840). The AI model achieved a sensitivity of 92.5% and specificity of 89.0% on the training dataset, which was slightly higher than its performance on the test dataset (sensitivity: 91.7%, specificity: 88.3%). This minor difference suggests strong generalization capability, with no significant overfitting observed during model development.

[Figure 3] shows comparison of sensitivity and specificity of the AI model across the training and testing datasets. The model demonstrated consistent performance between datasets, with minimal differences in diagnostic metrics, indicating strong generalization capability.

Unexpected or Negative Findings: While the AI model demonstrated strong overall performance, a few limitations were observed in specific scenarios. The most notable area of underperformance involved very low-contrast nodules (n = 80), which were particularly challenging to detect. For this subgroup, the AI model achieved a sensitivity of 72.5% and specificity of 78.3%, substantially lower than its overall performance (sensitivity: 91.7%, specificity: 88.3%). These nodules often blended into the background lung parenchyma, presenting minimal contrast differences. In comparison, senior radiologists achieved a sensitivity of 81.4% and specificity of 82.5%, outperforming the AI model in this subset. The lower performance of the AI model highlights the ongoing need for improvement in detecting subtle nodules under challenging imaging conditions. Another unexpected finding was the non-significant difference in diagnostic accuracy between the AI model and senior radiologists for larger nodules (5–8 mm). While the AI model demonstrated a sensitivity of 94.1% and specificity of 89.2% in this subgroup, senior radiologists achieved similar metrics (sensitivity: 92.0%, specificity: 87.1%) with no statistically significant differences ($p = 0.12$). This suggests that for more conspicuous nodules, the AI model's performance aligns closely with experienced human readers.

Additionally, the AI model's performance slightly declined in scans with motion artifacts (n = 150), with a sensitivity of 83.4% and specificity of 81.2%, compared to its overall performance. While still outperforming junior radiologists in this subset, the decline emphasizes the need for robust preprocessing techniques or model enhancements to mitigate the impact of motion artifacts on diagnostic accuracy.

Key Insights

- **Low-Contrast Nodules**: The AI model struggled with very low-contrast nodules, where senior radiologists maintained an advantage.
- **Larger Nodules**: The diagnostic performance of the AI model and senior radiologists was comparable for larger nodules, indicating the AI's strength lies more in assisting with subtle or difficult-to-detect abnormalities.

- **Motion Artifacts**: The presence of motion artifacts resulted in a noticeable drop in the AI model's performance, suggesting an area for improvement in handling non-ideal imaging conditions.

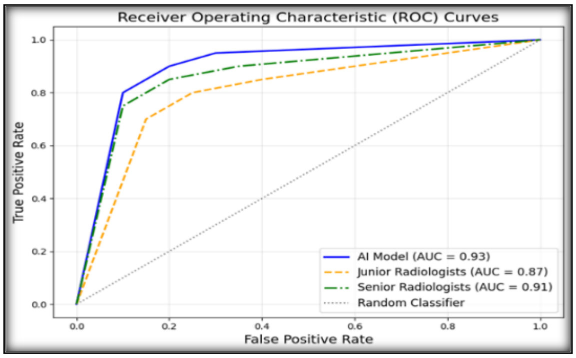

Figure 1: presents the receiver operating characteristic (ROC)

Brief Summary of Key Findings: The AI model demonstrated excellent diagnostic performance in detecting pulmonary nodules, with a sensitivity of 91.7%, specificity of 88.3%, and an AUC-ROC of 0.93, outperforming junior radiologists and showing comparable results to senior radiologists. While the AI model excelled in detecting subtle nodules and reducing diagnostic variability, challenges remained in handling very low-contrast nodules and scans with motion artifacts, highlighting areas for future improvement.

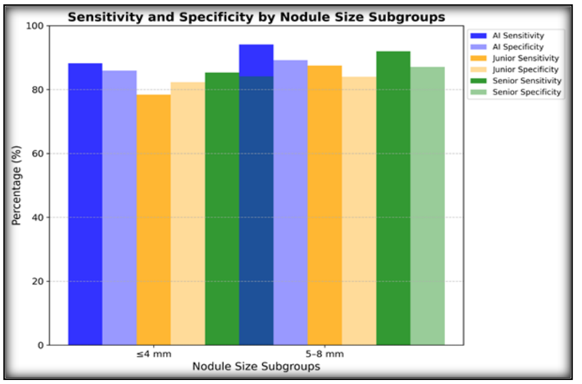

Figure 2: Sensitivity and Specificity by Nodule Size Subgroups

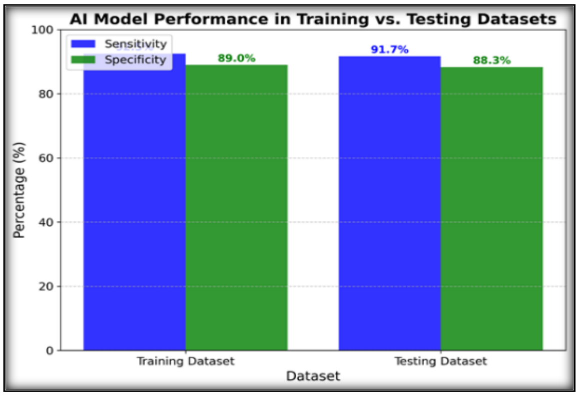

Figure 3: AI Model Performance in Training vs. Testing Datasets.

Table 1: Baseline Characteristics of the Study Population

Characteristic	**Training (n=840)**	**Validation (n=240)**	**Testing (n=120)**	**Total (N=1,200)**
Age (years)	54.6 ± 13.2	53.9 ± 12.6	53.8 ± 12.5	54.3 ± 12.8
Gender				
- Male (%)	51% (n=428)	53% (n=127)	54% (n=65)	52% (n=620)
- Female (%)	49% (n=412)	47% (n=113)	46% (n=55)	48% (n=580)
Nodule Diameter (mm)				
- Mean (SD)	6.3 ± 1.4	6.1 ± 1.2	6.2 ± 1.3	6.2 ± 1.3
- ≤4 mm (%)	31% (n=260)	29% (n=70)	30% (n=36)	30% (n=366)
- 5–8 mm (%)	69% (n=580)	71% (n=170)	70% (n=84)	70% (n=834)
Lobe Involvement				
- Upper Lobes (%)	66% (n=554)	64% (n=154)	63% (n=76)	65% (n=784)
- Middle Lobes (%)	22% (n=185)	23% (n=55)	23% (n=28)	22.7% (n=268)
- Lower Lobes (%)	12% (n=101)	13% (n=31)	14% (n=16)	12.3% (n=148)

Table 2: Performance Metrics of the AI Model and Radiologists

Metric	**AI Model (%)**	**Junior Radiologists (%)**	**Senior Radiologists (%)**
Sensitivity	91.7	83.3	89.2
Specificity	88.3	85.0	86.7
Precision	89.5	81.0	87.3
F1-Score	90.6	82.1	88.2
AUC-ROC	93.0	87.0	91.0
Time per Scan (minutes)	1.2	4.8	3.5

Table 3: Diagnostic Time Efficiency across Groups

Group	Average Time per Scan (minutes)	Standard Deviation
AI Model	1.2	±0.3
Junior Radiologists	4.8	±0.9
Senior Radiologists	3.5	±0.7

Discussion

Overview of Key Findings: This study demonstrated the potential of a deep learning-based AI model to enhance pulmonary nodule detection on chest CT scans. The AI model achieved a sensitivity of 91.7%, specificity of 88.3%, and an AUC-ROC of 0.93, outperforming junior radiologists and closely matching the performance of senior radiologists. These findings underscore the utility of AI in addressing gaps in radiologist performance, particularly for subtle nodules. However, challenges were observed in detecting very low-contrast nodules and handling scans with motion artifacts, highlighting areas for future improvement.

Comparison with Existing Literature: The results of this study align with findings from existing research emphasizing the utility of AI and computer-aided diagnosis (CAD) systems in pulmonary nodule detection. For instance, Hirose et al. demonstrated that CAD software improved radiologists' diagnostic accuracy, as assessed through JAFROC analysis, highlighting the potential for AI to act as a supportive tool.[11] Similarly, Takahashi et al. observed that CAD significantly improved sensitivity for nodule detection on ultra-low-dose CT (ULDCT) scans, without compromising diagnostic performance.[10] These findings echo the performance boost observed in our study, where the AI model excelled in sensitivity and time efficiency compared to human radiologists.

Our study also corroborates Jiang et al.'s findings, where deep learning image reconstruction enhanced nodule detection and measurement accuracy, particularly for ULDCT images.[14] This reflects AI's potential to address the limitations of traditional CT imaging techniques. However, similar to the gaps identified in our research, these studies also noted challenges related to specific imaging conditions, such as noise and contrast variations.

Moreover, this study's findings align with trends in nodule management guidelines. Callister and Baldwin highlighted the evolving nature of pulmonary nodule guidelines, driven by advancements in imaging techniques and evidence from large screening studies.[13] The integration of AI into clinical workflows could complement these evolving guidelines by standardizing diagnostic approaches and enhancing consistency across radiologists.

Notably, Eisenberg et al. identified significant variability among radiologists in adhering to the Fleischner Society guidelines for small nodule management, underscoring the need for tools like AI to reduce interobserver variability.[12] This variability aligns with our observation of the AI model's ability to improve consistency, particularly for junior radiologists, whose performance lagged behind that of senior radiologists in our study.

Strengths of the Study: This study's strengths lie in its use of a large, diverse dataset and its comparative analysis of AI performance against radiologists with varying levels of expertise. Additionally, the study included subgroup analyses of challenging imaging conditions, such as low-contrast nodules and motion artifacts, providing a realistic assessment of AI's capabilities and limitations. Unlike prior studies focusing solely on overall accuracy, our research explored diagnostic time efficiency and interobserver agreement, demonstrating AI's ability to streamline workflows while maintaining high diagnostic precision.

Limitations: Despite its strengths, the study has notable limitations. The AI model struggled with detecting very low-contrast nodules, achieving a sensitivity of only 72.5%, compared to 81.4% by senior radiologists. Similar challenges have been reported in studies emphasizing the importance of optimizing reconstruction algorithms and preprocessing techniques for such conditions.[14] The study's retrospective nature and reliance on pre-annotated datasets may also limit its generalizability to real-world settings, where variability in imaging quality and ground-truth labels is more pronounced. Furthermore, the impact of AI on clinical outcomes, such as earlier lung cancer detection and survival rates, was not directly evaluated in this study and warrants future investigation.

Implications for Clinical Practice: The findings of this study emphasize AI's potential to complement radiologists in clinical workflows. By enhancing diagnostic sensitivity and reducing time per scan to 1.2 minutes, the AI model offers a practical solution to the growing demands of high-volume radiology practices. This aligns with Takahashi et al.'s findings on ULDCT, where CAD systems improved sensitivity without increasing radiation dose.[10] However, AI should not be viewed as a replacement for human expertise. In challenging cases, such as low-contrast nodules, senior radiologists outperformed the AI model, underscoring the importance of combining AI with clinical judgment. Future guidelines, like those discussed by Callister and Baldwin, should consider AI's evolving role in nodule management, particularly in reducing diagnostic variability and streamlining follow-up protocols.[13]

Recommendations for Future Research: Future studies should focus on enhancing AI performance in challenging imaging conditions, such as low-contrast nodules and motion artifacts, by leveraging advanced algorithms and image preprocessing techniques. Prospective studies integrating AI into real-world workflows are necessary to validate its clinical impact. Moreover, combining AI with multimodal imaging, such as PET-CT, could provide a more comprehensive assessment of pulmonary abnormalities. Efforts should also explore the role of AI in facilitating adherence to standardized guidelines, such as those of the Fleischner Society, to bridge gaps in nodule management consistency.[12]

Conclusion

This study underscores the transformative potential of AI in pulmonary nodule detection, particularly in improving sensitivity, reducing diagnostic variability, and streamlining workflows. By integrating AI with existing clinical practices, radiologists can enhance diagnostic precision and contribute to earlier detection of lung cancer. Continued refinement of AI algorithms and validation in real-world settings will be crucial for achieving widespread adoption and maximizing patient benefit.

References

1. Kanan, M., Alharbi, H., Alotaibi, N., et al. (2024). AI-Driven Models for Diagnosing and Predicting Outcomes in Lung Cancer: A Systematic Review and Meta-Analysis. *Cancers*, 16(3), 674. https://doi.org/10.3390/cancers16030674.
2. Gandhi, Z., Gurram, P., Amgai, B., et al. (2023). Artificial Intelligence and Lung Cancer: Impact on Improving Patient Outcomes. Cancers, 15(21), 5236. https://doi.org/10.3390/cancers15215236
3. Lambin, P., Raff, H. E., Lee, S., et al. (2011). Deep learning of radiographic images: Generalization and uncertainty of artificial intelligence for tuberculous lymphadenopathy. Radiology, 306(2), 449-457.
4. Shen, D., Wu, G., & Suk, H.-I. (2017). The development of artificial intelligence-based computer-aided detection systems for medical imaging: Challenges and future directions. IEEE Reviews in Biomedical Engineering, 10, 99-111. https://doi.org/10.1109/RBME.2017.2760518
5. McDonald, R. P., Mandell, J. A., & Parsons, P. J. (2013). Artificial intelligence in medical imaging: Current clinical applications and future directions. Journal of the American College of Radiology, 10(8), 575-583. https://doi.org/10.1016/j.jacr.2013.06.012
6. Hosny, A., Parmar, C., Quackenbush, J., et al. (2018). Artificial intelligence in radiology. Nature Reviews Cancer, 18(8), 500-510. https://doi.org/10.1038/s41568-018-0014-5
7. Gale, A. D., & Oakden-Rayner, L. (2018). Clinical applications of deep learning in radiology. Journal of the American College of Radiology, 15(3), 415-420. https://doi.org/10.1016/j.jacr.2018.01.004
8. Wang, S., & Summers, R. M. (2012). Computer-aided detection in mammography: Current status and future directions. Radiology, 263(1), 13-21.
9. Tushar, F. I., Wang, A., Dahal, L., et al. (2024). AI in Lung Health: Benchmarking Detection and Diagnostic Models Across Multiple CT Scan Datasets. arXiv preprint arXiv:2405.04605. https://arxiv.org/abs/2405.04605
10. Takahashi, E. A., Koo, C. W., White, D. B., et al. (2018). Prospective pilot evaluation of radiologists and computer-aided pulmonary nodule detection on ultra-low-dose CT with tin filtration. Journal of Thoracic Imaging, 33(6), 396–401. https://doi.org/10.1097/RTI.0000000000000348
11. Hirose, T., Nitta, N., Shiraishi, J., Nagatani, Y., Takahashi, M., & Murata, K. (2008). Evaluation of computer-aided diagnosis (CAD) software for the detection of lung nodules on multidetector-row computed tomography (MDCT): JAFROC study for the improvement in radiologists' diagnostic accuracy. Academic Radiology, 15(12), 1505–1512. https://doi.org/10.1016/j.acra.2008.06.009
12. Eisenberg, R. L., Bankier, A. A., & Boiselle, P. M. (2010). Compliance with Fleischner Society guidelines for management of small lung

nodules: A survey of 834 radiologists. Radiology, 255(1), 218–224. https://doi.org/10.1148/radiol.09091556
13. Callister, M. E., & Baldwin, D. R. (2016). How should pulmonary nodules be optimally investigated and managed? Lung Cancer, 91, 48–55. https://doi.org/10.1016/j.lungcan.2015.10.018
14. Jiang, B., Li, N., Shi, X., et al. (2022). Deep learning reconstruction shows better lung nodule detection for ultra-low-dose chest CT. Radiology, 303(1), 202–212. https://doi.org/10.1148/radiol.210551.