IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155123

AI-Assisted Computer-Aided Diagnosis for Early Detection of Lung Cancer

Ghanashyam Pawar[1], Pranav Shinde[2], Bharati Mahale[3]

K.R.T. Arts,B.H. Commerce and A.M Science (KTHM) College affiliated to Savitribai Phule Pune University, Maharashtra[1-3]

Abstract: Lung cancer is one of the major causes of death across the world, and early detection plays an important role in improving patient survival and treatment success. Computer-aided diagnosis (CAD) systems using CT scan images help doctors identify and classify lung nodules more effectively, supporting early-stage lung cancer diagnosis. Earlier machine learning methods such as Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) were widely used, but they faced difficulties in analyzing large and complex medical image data.

With the growth of deep learning, medical image analysis has improved significantly. Advanced techniques such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Generative Adversarial Networks (GAN) provide better accuracy in detecting, segmenting, and classifying pulmonary nodules. This review highlights recent developments in deep learning approaches for lung cancer diagnosis and explains how these methods perform better than traditional machine learning techniques.

The study also discusses the use of ensemble models and other modern approaches that increase the reliability and efficiency of pulmonary nodule analysis. Overall, deep learning has shown great potential in improving the accuracy of lung cancer detection and diagnosis. Despite some existing challenges, ongoing advancements in artificial intelligence are expected to further enhance early diagnosis and medical decision-making in the future.

Keywords: Lung cancer, Artificial Intelligence, Computer -aided diagnosis(CAD), Pumlonary nodule segmentation and classification, Deep Learning

1.INTRODUCTION

Based on the reviewed research papers and existing AI-based lung cancer detection systems, significant progress has been achieved in pulmonary nodule detection, segmentation, and classification using deep learning techniques such as CNNs, RNNs, GANs, YOLO, U-Net, and hybrid neural networks. However, several important research gaps still exist.

Most existing studies mainly focus on improving classification accuracy using CT scan datasets such as LIDC-IDRI and LUNA16. Current research largely concentrates on binary classification (benign vs malignant) and pulmonary nodule detection. There is limited work on explainable AI, multimodal healthcare integration, real-time deployment, personalized treatment prediction, and lightweight clinical AI systems.

This document presents new research directions, innovative topic ideas, proposed system architectures, and future opportunities that can be explored as novel contributions in AI-based lung cancer diagnosis.

Recent studies have reviewed the use of Artificial Intelligence (AI) in the early detection of lung cancer. The work of Kaulgud et al. mainly focuses on CT image detection techniques and compares different methods used for pulmonary nodule detection. Similarly, Kalkeseetharaman et al. discussed studies based on both X-ray and CT images and emphasized the benefits of deep learning techniques in CT image analysis for pulmonary nodule detection.

However, these review studies do not clearly explain the differences between traditional machine learning methods and deep learning approaches, especially regarding data sources and pulmonary nodule segmentation techniques. In addition, many studies provide only limited details about the algorithms and model structures used, resulting in the absence of a common and well-organized review framework.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155123

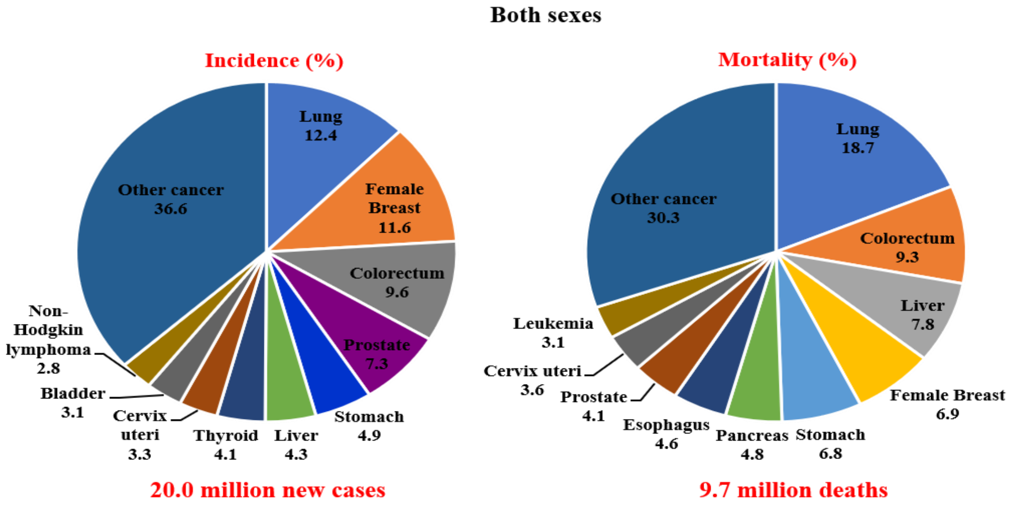

Fig. 1: Pie chart showing the distribution of cases and deaths for the top five cancers for both sexes in 2022. The area of the pie chart reflects the proportion of the total number of cases or deaths

To overcome these limitations, this review focuses on the role of AI technologies in the early detection and diagnosis of lung cancer. It highlights how AI-based approaches can support the fast detection, accurate segmentation, and effective classification of pulmonary nodules, ultimately helping improve diagnostic accuracy and patient care.

2. MEDICAL AI FOR LUNG CANCER DETECTION

Artificial Intelligence (AI) has greatly improved lung cancer detection and diagnosis in recent years. Traditional pulmonary nodule detection mainly depends on radiologists and manual CT image analysis, which can be time-consuming and affected by human error.

AI technology, especially deep learning, helps in automatically detecting and analyzing pulmonary nodules from medical images with better speed and accuracy. AI systems can perform tasks such as image recognition and decision-making, supporting doctors in early lung cancer diagnosis.

These advanced techniques improve diagnostic accuracy, reduce workload, and help physicians make faster and more effective treatment decisions, ultimately increasing patient survival rates.

3. CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional Neural Networks (CNNs) were first introduced in the 1980s through the development of LeNet for handwritten digit recognition. With improvements in computing power and large datasets, CNNs became highly popular after the success of AlexNet in the 2012 ImageNet competition. Later models such as ResNet and DenseNet further improved image classification performance.

Today, CNNs are widely used in medical image analysis, especially for the detection, segmentation, and classification of pulmonary nodules in CT scans. CNNs automatically learn important image features through multiple convolution layers, making them highly effective in identifying small abnormalities in lung images.

Several CNN architectures, including LeNet, AlexNet, VGG, ResNet, U-Net, and DenseNet, have contributed significantly to improving the accuracy and efficiency of lung cancer detection systems.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155123

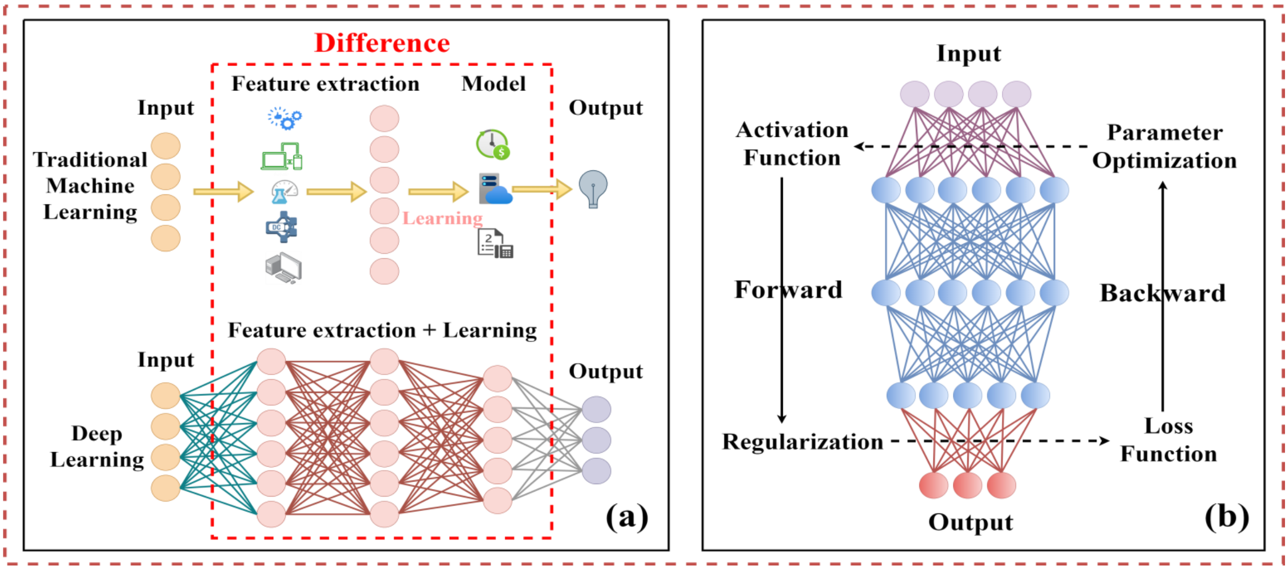

Fig. 2: Basic structure of a Convolutional Neural Network (CNN).

4. RECURRENT NEURAL NETWORK (RNN)

Recurrent Neural Networks (RNNs) are deep learning models designed to process sequential data by remembering information from previous inputs. Because of this capability, RNNs are useful in medical imaging tasks involving continuous CT scan data for pulmonary nodule detection.

Traditional RNNs face problems such as vanishing gradients, which affect long-term learning. To overcome these issues, advanced models like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) were introduced. These models improve the analysis of sequential CT images and help increase detection accuracy.

In lung cancer diagnosis, RNNs are often combined with Convolutional Neural Networks (CNNs) to analyze both spatial and temporal features of pulmonary nodules. This combination improves classification accuracy, sensitivity, and reduces false positive results. Recent studies have shown that RNN-based models provide strong performance in pulmonary nodule detection and classification.

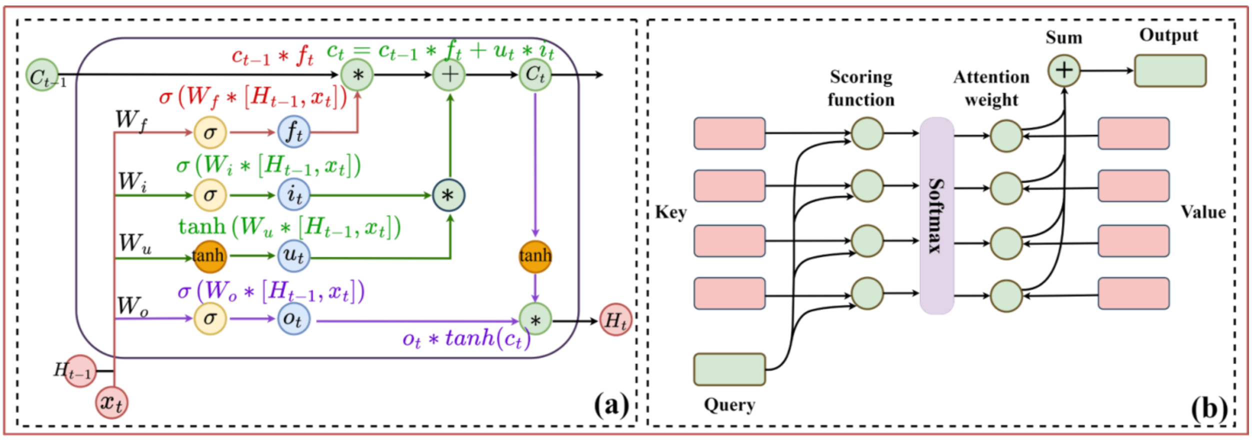

Fig 3:Structure of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) model.

Fig 3: Various deep learning models and structures: (a) Long Short-Term Memory (LSTM) structure is vital for sequence data analysis in medical applications, utilizing forget, input, and output gates to manage information flow, thereby maintaining essential temporal features critical for accurate diagnostic predictions. This mechanism proves

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155123

especially effective in analyzing diagnostic time-series data, capturing crucial changes over time. (b) The Attention Mechanism enhances neural network focus on significant features within diagnostic images, facilitating precise detection and categorization of anomalies such as tumors and lesions by adjusting the model's focus to the most informative regions of the image data.

3. LITERATURE SEARCH AND SCREENING

This review collected research articles from databases including IEEE Xplore, Wiley Online Library, Google Scholar, Web of Science, ACM Digital Library, ScienceDirect, SpringerLink, and Engineering Village, covering studies published from January 2020 to May 2024. The publication distribution is shown in Fig. 4a.

A total of 406 articles related to pulmonary nodule analysis were initially identified, including studies on detection, segmentation, and classification. After applying screening criteria, 131 articles were finally selected for detailed review, as presented in Fig. 4b. These studies mainly focused on AI-based computer-aided diagnosis (CAD) systems using CT images for pulmonary nodule analysis.

This review summarizes the major machine learning and deep learning techniques used in lung cancer detection and discusses their advantages, limitations, and future research directions.

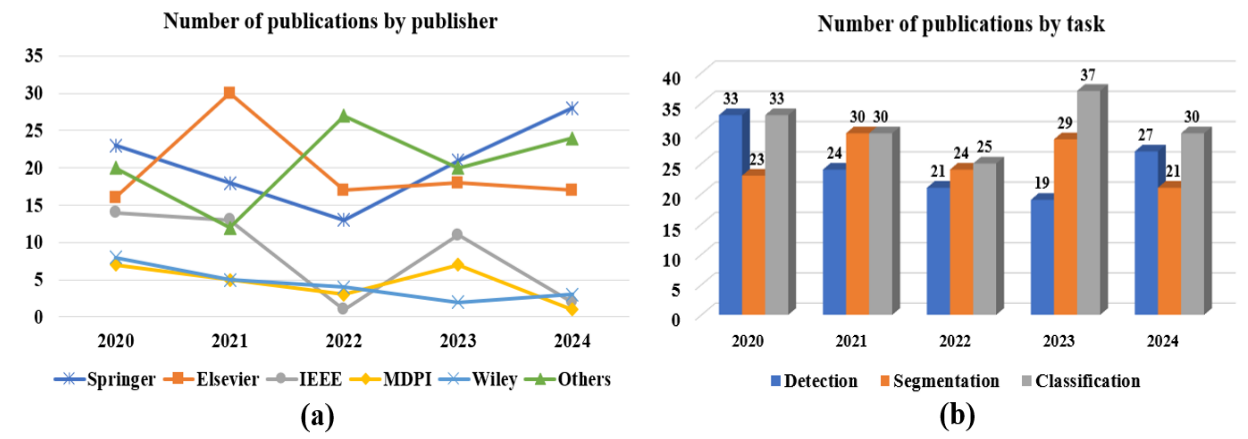

Fig. 4:(a) Number of publications in each database/platform from 2020 to 2024. (b) Number of publications on different pulmonary nodule treatment tasks from 2020 to 2024.

The motivation behind this survey is the critical need for **early and accurate detection** of lung cancer, one of the deadliest cancers globally. Manual interpretation of medical images like CT scans and X-rays often suffers from human error, bias, and high time consumption. AI, especially DL, offers a powerful solution to overcome these limitations.

Unlike previous reviews, which were often general or lacked architecture-level details, this paper specifically focuses on the **design architecture**, **classification models**, and **evolution of CAD systems** used in lung cancer detection.

4.BREAKTHROUGHS AND KEY TECHNIQUES IN LUNG CANCER DETECTION, SEGMENTATION, AND CLASSIFICATION

Recent advancements in Artificial Intelligence (AI) and deep learning have significantly improved lung cancer detection, segmentation, and classification. Modern computer-aided diagnosis (CAD) systems use advanced deep learning models and medical imaging datasets to increase diagnostic accuracy and support early lung cancer diagnosis.

For pulmonary nodule detection, techniques such as Convolutional Neural Networks (CNNs), YOLO, and multi-task learning models have improved the sensitivity and accuracy of identifying nodules in CT scan images. In segmentation, models like UNet, 3D CNNs, and Generative Adversarial Networks (GANs) provide precise identification of nodule boundaries, helping in the analysis of nodule size and growth.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155123

In classification, AI models combined with texture and feature-based analysis help differentiate between benign and malignant nodules with improved reliability. Public datasets such as LIDC-IDRI and LUNA16 have played an important role in training and evaluating these advanced models.

These breakthroughs and key techniques have enhanced the performance of CAD systems, reduced diagnostic errors, and improved clinical decision-making for early lung cancer treatment and patient care.

4.1 Discussion on AI-Driven Advances in Pulmonary Nodule Detection:

Pulmonary nodule detection is an important step in the early diagnosis and treatment of lung cancer. The process mainly includes identifying possible nodules in CT scan images and classifying their malignancy risk for clinical analysis.

Recent AI and deep learning techniques, including multi-task CNNs, YOLOv4, UNet, and GAN-based models, have significantly improved detection accuracy and sensitivity. High-quality datasets such as LIDC-IDRI and LUNA16 have further supported the training and evaluation of these models, improving the effectiveness of CAD systems in pulmonary nodule detection.

Table 1: Pulmonary Nodule Detection Method and Their Performance in CT Images

Authors	Year	2D/3D	Technique / Network Type	Dataset / Split	Main Method	Nodule Types	Performance
Veronica et al.	2020	2D & 3D	ANN, Fuzzy C-Means	LIDC-IDRI, ELCAP	FCM + ANN with OALO	Benign, malignant	Sensitivity: 96.10%
Jain et al.	2023	3D	Two-stage CNN, U-Net	5023 Train/Validation	CNN + U-Net	Nodules ≥ 3 mm	Sensitivity: 83.55%
Gugulothu et al.	2023	2D	CSDR-J-WHGAN	80% Train, 20% Test	SDMMT + WHGAN	Nodule & non-nodule	Sensitivity: 96.98%
Rashid et al.	2024	2D & 3D	HOSNV, LBP, XGBoost	LIDC-IDRI	XGBoost	Vessel, isolated, juxtapleural	Accuracy: 98.49%
Agnes et al.	2024	2D	Wavelet U-Net++	1018 scans	U-Net++	Small irregular nodules	DSC: 93.70%, IoU: 87.80%
Gautam et al.	2024	2D	ResNet-152, EfficientNet-B7	LIDC-IDRI	Ensemble Learning	Benign & malignant	Sensitivity: 98.60%
Gugulothu et al.	2024	2D	Hybrid DL (HDE-NN)	LIDC-IDRI	Differential Evolution NN	Benign & malignant	Sensitivity: 95.25%

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155123

Table 2: Methods and Their Performance in Pulmonary Nodule Classification on CT Images

Authors	Year	Dataset	Classification Type	Technique / Network Type	Main Method	Performance
Savitha et al.	2020	LIDC-IDRI	Nodule / Non-nodule	DCNN + CRF	Feature extraction and classification	ACC: 89.48%
Wang et al.	2021	LIDC-IDRI	Nodule / Non-nodule	CS-LBP, ORT-EOH, H-SVMs	3D feature extraction and classification	Accuracy up to 96.95%
Naveen et al.	2023	LIDC-IDRI	Nodule / Non-nodule	DS, RF, BPNN	Pulmonary nodule classification	Solid: 98.00%, Part-solid: 93.68%, Non-solid: 97.20%
Gugulothu et al	2023	LIDC-IDRI	Nodule / Non-nodule	SDMMT, U-Net, LTrP	CSDR-J-WHGAN classification	97.11% Accuracy
Sengodan et al	2023	LIDC-IDRI	Benign / Malignant	RCNN, Ensemble SVM	Ensemble-based classification	ACC: 98.53%, SEN: 99.30%, AUC: 0.98

5.GAP IN CURRENT RESEARCH

Despite major progress in AI-based lung cancer detection, several research gaps still exist. Most studies mainly focus on pulmonary nodule classification using public CT scan datasets such as LIDC-IDRI and LUNA16, which have limited diversity and may not fully represent real clinical data.

Current research mainly performs binary classification (benign or malignant) and gives less attention to advanced tasks such as tumor growth prediction, treatment response, and recurrence analysis. Most CAD systems also depend only on CT images and rarely combine clinical or genomic data.

Another important limitation is the lack of explainable AI (XAI). Many deep learning models work as "black box" systems, making clinical interpretation difficult. In addition, high computational cost and limited real-world validation reduce the practical use of these models in healthcare environments.

6.CHALLENGES AND FUTURE PROSPECTS

Artificial Intelligence (AI) has greatly improved lung cancer detection, segmentation, and classification. However, several challenges still affect the effective use of AI systems in clinical practice.

6.1 Data Challenges:
Datasets such as LIDC-IDRI and LUNA16 have supported the development of AI models, but they are limited in diversity and standardization. Variations in CT image quality, noise, and annotation inconsistencies can reduce model accuracy and generalization in real-world applications.

6.2 Algorithm Challenges:
Deep learning models such as CNNs and U-Net provide high accuracy, but they often require high computational power and have limited interpretability. Many AI systems work as "black boxes," making it difficult for clinicians to

DOI: 10.17148/IJARCCE.2026.155123

understand the prediction process. Developing accurate and computationally efficient models remains an important challenge.

6.3 Task Complexity:
Most current CAD systems mainly focus on pulmonary nodule detection and classification. Future research should also address advanced tasks such as tumor growth prediction, treatment response analysis, and the integration of 3D imaging, PET-CT, and follow-up scan data for more comprehensive lung cancer analysis.

6.4 Clinical tool and Usability:
There is a need for user-friendly AI platforms that can be easily used by radiologists and clinicians without advanced programming knowledge. Simple and accessible tools for model training, evaluation, and interpretation can help increase AI adoption in healthcare.

6.5 Generalizability and validation:
AI models trained on one dataset may not perform equally well across different hospitals or patient populations due to variations in imaging protocols and equipment. Therefore, large-scale external validation and cross-platform compatibility are necessary to improve the reliability and practical implementation of AI-driven CAD systems.

7.LIMITATIONS

• Lack of Practical Implementation: The current research is primarily theoretical. Practical implementation and testing of the model are yet to be carried out.

• Limited Dataset Availability: Access to high-quality, labeled datasets that combine both symptoms and imaging data is limited, which may impact model training and validation.

• Computational Complexity: Deep learning models like CNN require high computational resources and longer training times, which can be a constraint in real-time clinical settings.

• Complex Data Integration: Combining heterogeneous data types, such as images and clinical symptoms, requires complex preprocessing and alignment techniques.

• Generalizability Issues: The model's effectiveness may vary across different populations, imaging devices, and clinical environments. Extensive clinical validation is necessary before deployment.

8.CONCLUSION

Deep learning has significantly improved medical image analysis, especially in the early detection of lung cancer. Advanced techniques such as CNNs, RNNs, and GANs have shown better performance in pulmonary nodule detection, segmentation, and classification compared to traditional machine learning methods. These approaches help increase diagnostic accuracy and support early treatment decisions.

Despite these advancements, several challenges remain, including limited availability of high-quality annotated datasets and the lack of interpretability in deep learning models. The "black box" nature of many AI systems can reduce trust and transparency in clinical decision-making.

Future research should focus on explainable AI (XAI), data augmentation techniques, and the integration of multimodal data such as clinical and genomic information. Hybrid approaches combining deep learning and traditional machine learning may further improve diagnostic performance and clinical usability.

Overall, deep learning has strong potential to enhance early lung cancer diagnosis and improve patient outcomes, making it an important area for future research and healthcare development.

9.DATA AVAILABILITY

No data was used for the research described in the article.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155123

REFERENCES

[1]. F. Bray, M. Laversanne, H. Sung, J. Ferlay, R. L. Siegel, I. Soerjomataram, and A. Jemal, “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” CA: A Cancer Journal for Clinicians, vol. 74, no. 3, pp. 229–263, 2024.

[2]. H. Jin, C. Yu, Z. Gong, R. Zheng, Y. Zhao, and Q. Fu, “Machine learning techniques for pulmonary nodule computer-aided diagnosis using CT images: A systematic review,” Biomedical Signal Processing and Control, vol. 79, p. 104104, 2023.

[3]. R. V. Kaulgud and A. Patil, “Analysis based on machine and deep learning techniques for the accurate detection of lung nodules from CT images,” Biomedical Signal Processing and Control, vol. 85, p. 105055, 2023.

[4]. P. Kalkeseetharaman and S. T. George, “A bird’s eye view approach on the usage of deep learning methods in lung cancer detection and future directions using X-ray and CT images,” Archives of Computational Methods in Engineering, pp. 1–21, 2024.

[5]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Advances in Neural Information Processing Systems, vol. 25, 2012.

[6]. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[7]. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[8]. W. Zhu, C. Liu, W. Fan, and X. Xie, “DeepLung: Deep 3D dual path nets for automated pulmonary nodule detection and classification,” in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2018, pp. 673–681.

[9]. C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, pp. 273–297, 1995.

[10]. T. Cover and P. Hart, “Nearest neighbor pattern classification,” IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967.

[11]. G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” Neurocomputing, vol. 70, no. 1–3, pp. 489–501, 2006.

[12]. R. E. Bawack, S. Fosso Wamba, and K. D. A. Carillo, “A framework for understanding artificial intelligence research: Insights from practice,” Journal of Enterprise Information Management, vol. 34, no. 2, pp. 645–678, 2021.

[13]. Z. Zhang, K. A. Ahmed, M. R. Hasan, T. Gedeon, and M. Z. Hossain, “A deep learning approach to diabetes diagnosis,” in Asian Conference on Intelligent Information and Database Systems, Springer, 2024, pp. 87–99.

[14]. A. D. Hiwase, C. D. Ovenden, L. M. Kaukas, M. Finnis, Z. Zhang, S. O’Connor, N. Foo, B. Reddi, A. J. Wells, and D. Y. Ellis, “Can rotational thromboelastometry rapidly identify theragnostic targets in isolated traumatic brain injury?” Emergency Medicine Australasia, 2024.

[15]. Y. Zhao, Z. Liao, Y. Liu, K. O. Nijhuis, B. Barvelink, J. Prijs, J. Colaris, M. Wijffels, M. Reijman, Z. Zhang, et al., “A landmark-based approach for instability prediction in distal radius fractures,” in 2024 IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2024, pp. 1–5.

[16]. J. Ge, Z. Zhang, M. H. Phan, B. Zhang, A. Liu, and Y. Zhao, “ESA: Annotation-efficient active learning for semantic segmentation,” arXiv preprint arXiv:2408.13491, 2024.

[17]. Z. Zhang, N. Yi, S. Tan, Y. Cai, Y. Yang, L. Xu, Q. Li, Z. Yi, D. Ergu, and Y. Zhao, “MedDet: Generative adversarial distillation for efficient cervical disc herniation detection,” arXiv preprint arXiv:2409.00204, 2024.

[18]. B. Wu, Y. Xie, Z. Zhang, J. Ge, K. Yaxley, S. Bahadir, Q. Wu, Y. Liu, and M.-S. To, “BHSD: A 3D multi-class brain hemorrhage segmentation dataset,” in International Workshop on Machine Learning in Medical Imaging, Springer, 2023, pp. 147–156.

[19]. Z. Zhang, X. Qi, B. Zhang, B. Wu, H. Le, B. Jeong, Z. Liao, Y. Liu, J. Verjans, and M.-S. To, “SegReg: Segmenting OARs by registering MR images and CT annotations,” in 2024 IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2024, pp. 1–5.

[20]. Z. Zhang, B. Zhang, A. Hiwase, C. Barras, F. Chen, B. Wu, A. J. Wells, D. Y. Ellis, B. Reddi, A. W. Burgan, M.-S. To, I. Reid, and R. Hartley, “Thin-Thick Adapter: Segmenting thin scans using thick annotations,” OpenReview, 2023.