ISSN 2616-5775 Vol. 9, Issue 5: 77-82, DOI: 10.25236/AJCIS.2026.090509

# Artificial Intelligence in Pulmonary Nodule Assessment: From Radiomics to Multimodal Integration

**Yueying Zhou[1,a,*], Jianming Dong[2,b]**

[1]*College of Cyberspace Security, Xizang Minzu University, Xianyang, 712082, China*
[2]*College of Cyberspace Security, Xizang Minzu University, Xianyang, 712082, China*
[a]*1446179182@qq.com*, [b]*jmdong@xzmu.edu.cn*
** Corresponding author*

***Abstract:*** *Accurate characterization of pulmonary nodules is critical for early lung cancer diagnosis and treatment planning. Recent advances in artificial intelligence (AI) have demonstrated substantial potential in automating nodule detection, segmentation, malignancy classification, and invasiveness prediction. This review summarizes key developments across radiomics, deep learning, pathomics, vision-language models, and liquid biopsy, highlighting the transition from single-modality analysis to multimodal integration. Representative studies are discussed to illustrate the performance gains and remaining challenges in clinical translation, including generalizability, interpretability, and implementation feasibility.*

***Keywords:*** *Artificial intelligence; Pulmonary nodules; Radiomics; Deep learning; Multimodal integration; Vision-language models; Liquid biopsy; Pathomics; Malignancy classification; Computed tomography*

## 1. Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide, responsible for approximately 1.8 million deaths annually[1]. Despite therapeutic advances, the overall 5-year survival rate remains around 19%, underscoring the importance of early detection. Detecting malignant pulmonary nodules at a localized stage dramatically improves outcomes, yet accurately differentiating benign from malignant nodules and assessing the invasiveness of subsolid lesions—including ground-glass and part-solid nodules—remains a formidable challenge. Overlapping morphological features on CT contribute to substantial inter-observer variability, high false-positive rates, and unnecessary invasive procedures.

Artificial intelligence (AI), particularly radiomics and deep learning, has emerged as a powerful complement to conventional radiological evaluation. Radiomics extracts high-throughput quantitative features from medical images, while deep learning automatically learns hierarchical feature representations from raw imaging data. More recently, the integration of complementary modalities—including histopathology, liquid biopsy, and vision-language models—has further expanded AI-assisted diagnostic capabilities.

Figure 1 outlines the evolutionary trajectory of AI-driven pulmonary nodule assessment in three phases: (1) an early stage driven by radiomics and conventional machine learning, (2) a current stage propelled by deep learning architectures such as convolutional neural networks, vision transformers, and multi-task learning, and (3) an emerging phase characterized by multimodal integration, combining vision-language models, circulating tumor DNA, and stacked ensemble strategies to enhance diagnostic accuracy, generalizability, and interpretability. This review synthesizes recent progress across these methodological approaches and critically discusses remaining challenges that will shape clinical translation.

ISSN 2616-5775 Vol. 9, Issue 5: 77-82, DOI: 10.25236/AJCIS.2026.090509

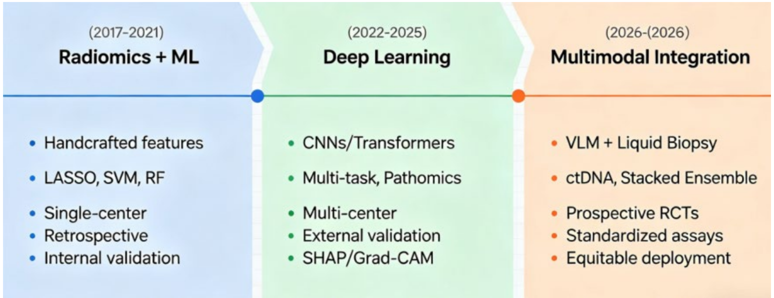


*Figure 1. Evolutionary roadmap of AI-driven pulmonary nodule assessment: from single-modality radiomics to multimodal integration.*

*DL, deep learning; VLM, vision-language model; ctDNA, circulating tumor DNA.*

## 2. Radiomics and Conventional Machine Learning

Radiomics extracts high-dimensional quantitative features from medical images to characterize tissue heterogeneity. Liu et al. developed a combined model integrating radiomic and CT semantic features to differentiate malignant from benign subcentimeter solid pulmonary nodules (SSPNs), achieving an area under the receiver operating characteristic curve (AUC) of 0.930 in the testing cohort [1]. Digumarthy et al. applied radiomics to subsolid nodules and found that temporal changes in features strongly favored malignancy, while baseline radiomic features had limited discriminative power [2]. Zhu et al. constructed a radiomic nomogram incorporating intra- and peri-nodular features for distinguishing invasive from preinvasive ground-glass opacity nodules, attaining an AUC of 0.835 [3]. A review by Kim emphasized that despite promising AUCs, clinical utility and prospective validation of radiomics-based computer-aided diagnosis tools remain insufficient[4]. Selvam et al. further demonstrated that a radiomics-based support vector machine performed comparably to expert radiologists in classifying lung nodule malignancy, with no statistically significant difference in accuracy[5]. These studies underscore the potential of radiomics while highlighting the need for standardized validation.

## 3. Deep Learning Approaches

Deep learning, especially convolutional neural networks (CNNs) and vision transformers, has enabled automated feature learning from raw images. Xie et al. proposed an anthropomorphic diagnosis system trained with weak annotations that achieved an AUC of 0.912 for nodule differential diagnosis on the LIDC-IDRI dataset, reducing the reliance on labor-intensive manual segmentation[6]. Du et al. combined deep learning-based tumor region identification with pathomics to predict invasiveness of lung adenocarcinoma manifesting as ground-glass nodules, yielding an accuracy of 0.814 on the test set and demonstrating that pathomics can effectively assist junior pathologists[7]. He et al. designed a pre- and post-fusion strategy integrating clinical, radiomic, and deep learning features for predicting infiltration degree of ground-glass nodules, achieving an AUC of 0.865[8]. Rao et al. introduced LungDxFormer, a hybrid CNN-transformer with dynamic spatial attention, achieving 97.35% overall accuracy for three-class malignancy classification on LIDC-IDRI[9]. In a multi-task framework, Liu et al.'s LNMSNet jointly performed nodule segmentation and malignancy classification using multi-scale convolutions, outperforming several state-of-the-art models on external validation cohorts[10]. Li et al. employed an explainable Tab-Transformer-ResMLP hybrid architecture for differential diagnosis of pulmonary space-occupying lesions and pathological subtyping, and interpreted feature contributions using SHAP analysis[11]. These models illustrate the trend toward end-to-end learning and explainability.

## 4. Histopathology-Based Deep Learning and Pathomics

Histopathology slides capture morphological features of tumor tissue at microscopic resolution, offering complementary prognostic information that may not be fully captured by radiological imaging alone. Zhang et al. developed a pathomics signature (PathoSig) from hematoxylin and eosin (H&E)-stained images of small cell lung cancer (SCLC) by applying unsupervised deep learning with contrastive clustering to tile-level histomorphological phenotypes[12]. Among 50 identified phenotype clusters, HPC19 and HPC39 showed independent prognostic value and were integrated into a Cox regression-based PathoSig, which significantly stratified patients into high-, intermediate-, and low-risk groups for overall survival and disease-free survival across multicenter validation cohorts. Moreover, PathoSig was able to identify patients likely to benefit from postoperative or preoperative chemoradiotherapy and provided prognostic stratification beyond conventional TNM staging and molecular subtyping. In the context of lung adenocarcinoma, Du et al. proposed a pipeline combining deep learning-based tumor region identification with pathomics; a ResNet18 model first localized tumor areas on whole slide images, and a random forest classifier built on deep learning-extracted features subsequently assessed tumor invasiveness, achieving an accuracy of 0.814 and an AUC of 0.807 on the test set[7]. Notably, when pathomics output was provided as a diagnostic aid, the average accuracy of both junior and intermediate pathologists in distinguishing invasive from non-invasive adenocarcinoma increased substantially (e.g., from 0.566 to 0.797 for junior pathologists), highlighting the potential of pathomics to enhance human diagnostic performance. Together, these studies demonstrate that histopathology-based deep learning and pathomics can serve as valuable complements to imaging-based models, enriching prognostic evaluation and supporting more precise clinical decision-making.

## 5. Vision-Language Models and Multimodal Fusion

A notable trend is the fusion of multiple data sources to improve robustness.

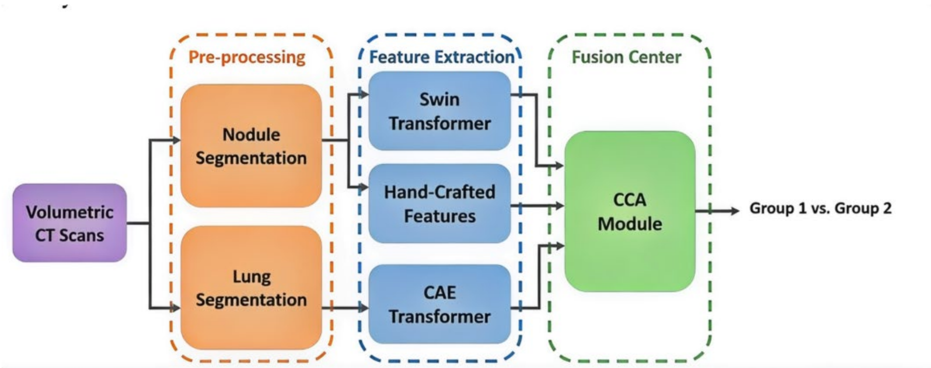


*Figure 2. I-VISTA architecture*

Zong et al. constructed a stacked ensemble model that combined clinical variables, radiomic features, deep learning visual classifications, and serum autoantibodies for solitary pulmonary nodule diagnosis, achieving an AUC of 0.823 on an independent test set[13].A representative example is the I-VISTA framework proposed by Khademi et al., which integrates three complementary feature sets through a criss-cross attention fusion module. As illustrated in Figure 2, the framework employs a Swin Transformer path to capture intra-slice spatial features, a CAE-Transformer path to model inter-slice temporal variations, and a 3D radiomics path that extracts handcrafted texture features at multiple spatial scales. These three sets of features are projected into a shared space and fused via criss-cross attention, enabling the model to dynamically weigh heterogeneous information for the classification of subsolid nodule invasiveness. On an in-house dataset of 114 biopsy-proven adenocarcinomas, the fused I-VISTA model achieved an AUC of 0.93, outperforming both the standalone deep learning models and the radiomics-only model[14]. Zhuang et al. developed a vision-language model based on contrastive language-image pretraining (CLIP) to align CT imaging features with radiologist-annotated semantic descriptions, achieving an AUC of 0.901 for one-year lung cancer prediction while offering zero-shot inference of semantic features[15]. Xu et al. demonstrated that GPT-4o, a large vision-language model, could perform ternary classification of pure ground-glass nodule invasiveness using few-shot learning, improving radiologists' diagnostic accuracy when used as an assistant [16]. Khovanova et al. evaluated several multimodal large language models for detecting pulmonary nodules on chest radiographs and found that the MedRAX agent-based framework and BiomedCLIP achieved the highest accuracy, though still insufficient for clinical practice [17]. These studies highlight the promise of vision-language models in providing interpretable, feature-level reasoning.

ISSN 2616-5775 Vol. 9, Issue 5: 77-82, DOI: 10.25236/AJCIS.2026.090509

## 6. Liquid Biopsy and Multi-Omics

Liquid biopsy has emerged as a promising, minimally invasive strategy for obtaining molecular biomarkers that may aid in the non-invasive differentiation of malignant from benign pulmonary nodules. Caballero-Vazquez et al. conducted a comprehensive meta-analysis of 16 eligible studies evaluating the diagnostic capacity of various circulating biomarkers [18]. Among the analyzed biomarkers, circulating free DNA (cfDNA)—particularly methylation profiling and mutation detection—demonstrated the highest diagnostic accuracy, yielding a pooled sensitivity of 0.81, specificity of 0.88, and AUC of 0.86. In a hypothetical clinical model based on the Brock risk calculator, a positive cfDNA result raised the post-test probability of malignancy from 18% to 70.9%, while a negative result reduced it to 2.5%, underscoring the potential of cfDNA to meaningfully refine risk stratification. The analysis also revealed that small non-coding RNAs (miRNAs and protein functional effector RNAs) achieved even higher pooled specificity of 0.94 and sensitivity of 0.82, with a pooled AUC of 0.90, although the limited number of available studies warrants cautious interpretation. Circulating tumor cells (CTCs) showed more modest performance, with a pooled AUC of 0.73 and sensitivity of 0.65, while glycan signatures on extracellular vesicles exhibited high specificity but lacked sufficient data for robust meta-analytic synthesis. Notably, the integration of multi-omics data with machine learning classifiers—including XGBoost and backpropagation neural networks—was shown to further enhance diagnostic accuracy, highlighting the value of combining multiple circulating biomarker classes and computational algorithms. Although liquid biopsy biomarkers have not yet been systematically integrated with imaging-based AI models in clinical workflows, these findings suggest that combined ctDNA, microRNA, and protein-based panels hold considerable promise as complementary modalities that could augment imaging-derived risk estimates in future multimodal diagnostic systems. Realizing this potential will require standardization of pre-analytical protocols, prospective validation across diverse populations, and the development of interpretable frameworks for combining molecular and imaging features within unified clinical decision support tools.

## 7. Challenges and Future Directions

Despite the impressive diagnostic performance reported across numerous studies, several persistent barriers continue to impede the clinical translation of AI-based tools for pulmonary nodule assessment. The 2025 lung cancer landscape review by Fu et al. emphasized that rigorous prospective validation, proper model calibration, and seamless integration into existing clinical workflows are essential prerequisites before AI tools can be widely adopted [19]. A fundamental limitation of the current evidence base is the predominance of retrospective, single-institution study designs with limited external validation across diverse patient populations, scanner vendors, and CT acquisition protocols. For instance, while Zong et al. evaluated their stacked ensemble model across three CT vendors, they acknowledged the absence of geographically independent external validation [13], and Liu et al. similarly noted that the lack of external cohorts constrained the generalizability assessment of their multi-task LNMSNet framework[10]. Such gaps raise legitimate concerns about performance deterioration when models encounter real-world data distributions that differ from those seen during training.

The inherent opacity of deep learning models, often characterized as the “black-box” problem, further undermines clinical trust and adoption. Although post-hoc interpretability techniques such as SHAP (SHapley Additive exPlanations) and Grad-CAM attention maps have been increasingly employed to reveal which image regions or features drive model predictions—as demonstrated by Li et al. in their explainable Tab-Transformer-ResMLP framework [11] and by Zhuang et al. through zero-shot semantic inference in vision-language models [15]—these methods may incompletely capture the complex reasoning underlying model decisions, particularly in ambiguous or diagnostically uncertain cases. Additional obstacles include domain shift between training and deployment environments, pronounced class imbalance in clinical datasets, and the substantial human expertise and annotation costs required to construct large, high-quality, pathologically confirmed datasets. The challenges inherent in weak-annotation-based nodule detection systems and multi-center data collection efforts underscore these practical constraints[6].

As noted in the focused review by Kim, the overwhelming majority of studies on radiomics-based CAD tools have concentrated on analytical and clinical validation, yet robust clinical utility studies—ideally prospective randomized controlled trials evaluating patient-centered outcomes such as diagnostic yield, time to diagnosis, and rates of unnecessary invasive procedures—remain

ISSN 2616-5775 Vol. 9, Issue 5: 77-82, DOI: 10.25236/AJCIS.2026.090509

conspicuously absent [4]. Consequently, whether AI-augmented workflows can genuinely improve patient outcomes or merely add computational overhead has not been definitively established. For liquid biopsy, the meta-analysis by Caballero-Vazquez et al. revealed significant heterogeneity in biomarker selection and detection methodologies, highlighting the need for standardization of pre-analytical protocols before molecular biomarkers such as cfDNA, microRNAs, and circulating tumor cells can be meaningfully integrated into multimodal diagnostic systems[18]. Future work should prioritize prospective implementation trials conducted across varied healthcare settings, the development of transparent, interpretable, and resource-appropriate AI frameworks that maintain robust performance across diverse populations and imaging protocols, and the establishment of clear regulatory guidelines for the validation, monitoring, and post-market surveillance of AI-based medical devices.

## 8. Conclusion

In summary, artificial intelligence has substantially advanced the assessment of pulmonary nodules, driving a progressive evolution from handcrafted radiomic features and conventional machine learning toward end-to-end deep learning architectures and, most recently, multimodal integration frameworks that combine imaging, histopathology, molecular biomarkers, and even semantic text descriptions. Radiomics-based models have demonstrated the value of quantitatively capturing tumor heterogeneity and perinodular microenvironment characteristics, while deep learning approaches—including convolutional neural networks, vision transformers, and hybrid architectures—have enabled automated feature learning directly from raw images, achieving performance levels that approach or match those of expert radiologists in specific tasks. The emergence of vision-language models and liquid biopsy technologies represents a particularly promising frontier, offering complementary diagnostic and prognostic information that may enhance both accuracy and interpretability: for instance, CLIP-based models can provide zero-shot semantic predictions of clinically relevant nodule features, and circulating biomarkers such as cfDNA methylation profiles have shown high specificity for malignancy risk stratification. However, the translation of these research advances into routine clinical practice remains contingent upon several critical prerequisites. Rigorous prospective validation in diverse, multicenter cohorts is essential to establish generalizability across different populations, scanner vendors, and acquisition protocols. The development and adoption of standardized evaluation frameworks—including robust calibration metrics, uncertainty quantification, and transparent reporting of failure modes—are necessary to build clinician trust and enable safe deployment. Importantly, clinical utility must be demonstrated through studies that assess patient-centered outcomes, such as reductions in unnecessary invasive procedures, improvements in diagnostic timeliness, and cost-effectiveness, rather than through retrospective performance metrics alone. Achieving these goals will require sustained, multidisciplinary collaboration among clinicians, data scientists, regulatory agencies, and healthcare systems to ensure that AI-driven innovations are not only technically sound but also equitable, interpretable, and aligned with the practical realities of diverse clinical environments.

## Acknowledgment

This work was supported by the Scientific Research Innovation and Practice Project of Xizang Minzu University (Grant No. Y2026136).

## References

*[1] Liu J, Qi L, Wang Y, et al. Development of a combined radiomics and CT feature-based model for differentiating malignant from benign subcentimeter solid pulmonary nodules. Eur Radiol Exp. 2024;8:8.*
*[2] Digumarthy SR, Padole AM, Rastogi S, et al. Predicting malignant potential of subsolid nodules: can radiomics preempt longitudinal follow up CT? Cancer Imaging. 2019;19:36.*
*[3] Zhu M, Yang Z, Wang M, et al. A computerized tomography-based radiomic model for assessing the invasiveness of lung adenocarcinoma manifesting as ground-glass opacity nodules. Respir Res. 2022;23:96.*
*[4] Kim RY. Radiomics-based computer-aided diagnosis for pulmonary nodule risk stratification: a narrative review. Cancer Biomark. 2025. doi:10.3233/CBM-230360.*

ISSN 2616-5775 Vol. 9, Issue 5: 77-82, DOI: 10.25236/AJCIS.2026.090509

*[5] Selvam M, Ramesh S, Sadanandan A, et al. Radiomics-derived classifier performance evaluation in lung nodule characterization compared with expert radiologists. Sci Rep. 2026. doi:10.1038/s41598-026-47656-7.*

*[6] Xie L, Xu Y, Zheng M, et al. An Anthropomorphic Diagnosis System of Pulmonary Nodules using Weak Annotation-Based Deep Learning. Comput Med Imaging Graph. 2024;117:102438.*

*[7] Du H, Wang X, Wang K, et al. Identifying invasiveness to aid lung adenocarcinoma diagnosis using deep learning and pathomics. Sci Rep. 2025;15:4913.*

*[8] He L, Li Z, Duan Y, et al. Prediction of infiltration degree of ground-glass nodules using a fusion of CT radiomics and deep learning. Sci Rep. 2026. doi:10.1038/s41598-026-50328-1.*

*[9] Rao KS, Kavitha H, Juluru TK, et al. LungDxFormer: a transformer-CNN hybrid model with dynamic spatial attention for accurate lung cancer detection and classification. Sci Rep. 2026. doi:10.1038/s41598-026-50449-7.*

*[10] Liu Y, Tang Z, Tang Z, et al. LNMSNet: a multi-task deep learning network for pulmonary nodules segmentation and malignancy classification. Front Med. 2026;13:1773338.*

*[11] Li H, Wang Y, Wang Y, et al. Application of explainable deep learning in differentiating benign from malignant pulmonary space-occupying lesions and classifying pathological subtypes of lung cancer. Chin J Lung Cancer. 2025;28(10):738-750.*

*[12] Zhang Y, Yang Z, Chen R, et al. Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. npj Digit Med. 2024;7:15.*

*[13] Zong J, Jiang B, Li H, et al. Stacked CT radiomics, deep learning and clinical feature models for differentiating benign and malignant solitary pulmonary nodules. Sci Rep. 2026. doi:10.1038/s41598-026-49720-8.*

*[14] Khademi S, Heidarian S, Afshar P, et al. I-VISTA: Integrating Visual, Spatial, and Temporal features with Attention for classifying invasiveness of lung adenocarcinoma subsolid nodules on CT. J Imaging. 2025;11:360.*

*[15] Zhuang L, Tabatabaei SMH, Salehi-Rad R, et al. Vision-language model-based semantic-guided imaging biomarker for lung nodule malignancy prediction. J Biomed Inform. 2025. doi:10.1016/j.jbi.2025.104947.*

*[16] Xu N, He Q, Wang L, et al. Predicting Invasiveness of Lung Adenocarcinoma from Chest CT with Few-shot Vision-Language Ternary Classification Model. npj Digit Med. 2026;9:60.*

*[17] Khovanova D, Vasilev Y, Vladymyrskyy A, et al. A comparative accuracy study of multimodal LLMs, VLM and agent-based framework for pulmonary nodule detection on chest radiographs. Front Digit Health. 2026;7:1674835.*

*[18] Caballero-Vazquez A, Garcia-Cerezo M, Fernandez-Castro LR, et al. Liquid biopsy biomarkers for accurate detection of malignant pulmonary nodules: a meta-analytic approach. Discov Oncol. 2026;17:178.*

*[19] Fu W, Zou X, He T, et al. The 2025 lung cancer landscape: advances in screening, molecular taxonomy and therapeutic strategy: a narrative review. Transl Lung Cancer Res. 2026;15(3):62.*