

Applications and Prospects of General Large Language Models in Medical Imaging

Jinzhe Li¹, Haoxiang Guo², Linkai Song³, Wenbo Gong², Chun Yang^{4*}

¹School of Stomatology, Ningxia Medical University, Yinchuan, Ningxia, Hui Autonomous Region, 750004, China

²The Third School of Clinical Medicine, Ningxia Medical University, Yinchuan, Ningxia Hui Autonomous Region, 750004, China

³The First School of Clinical Medicine, Ningxia Medical University, Yinchuan, Ningxia Hui Autonomous Region, 750004, China

⁴The Second School of Clinical Medicine, Ningxia Medical University, Yinchuan, Ningxia Hui Autonomous Region, 750004, China

ABSTRACT

Medical imaging plays a central role in modern medicine, underpinning disease diagnosis, staging, therapeutic evaluation, and longitudinal follow-up. However, conventional image interpretation remains limited by low efficiency, variable consistency, and heavy reliance on clinical expertise. Recent advances in artificial intelligence (AI)—particularly deep learning, radiomics, and general large language models (LLMs)—have opened new avenues for intelligent image analysis and clinical decision-making. Leveraging multimodal data integration and natural language generation, LLMs enable comprehensive interpretation that unifies imaging findings with clinical records, laboratory tests, and molecular profiles, thereby improving the efficiency, consistency, and interpretability of diagnostic and prognostic workflows^[1]. Here, we provide a systematic overview of the evolution and applications of AI in medical imaging, examine key challenges in data quality, algorithmic development, and clinical translation, and highlight future directions including multimodal integration, human–AI collaboration, and privacy-preserving frameworks. Collectively, emerging evidence indicates that LLMs and allied AI technologies are reshaping medical imaging toward more intelligent, individualized, and full-cycle clinical support, offering critical momentum for precision medicine.

KEYWORDS

Medical imaging; artificial intelligence; Deep learning; Large language model (LLM); Multimodal integration; Clinical decision-making

1 Introduction

Medical imaging is a cornerstone of modern medicine, serving as a critical tool for disease diagnosis, staging, therapeutic evaluation, and long-term follow-up. With continued advances in X-ray, CT, MRI, and PET technologies, clinicians are now able to observe structural and functional alterations in the human body with greater precision and comprehensiveness, driving medicine toward increasing precision and individualization. Beyond its role in early disease detection, imaging also provides indispensable support for treatment planning and therapeutic monitoring.

Despite its importance, conventional approaches to image interpretation are hampered by inefficiency, inconsistency, and heavy dependence on physician expertise. The vast volume and complexity of imaging data often exceed the limits of human review within constrained clinical time, raising the risk of missed or inaccurate diagnoses. In recent years, artificial intelligence (AI)—particularly deep learning and radiomics—has provided effective strategies for automated analysis and intelligent interpretation of medical images. More recently, the rapid emergence of general large language models (LLMs) has further expanded the scope of imaging AI. Beyond processing visual data, LLMs exhibit powerful capabilities in natural language generation and multimodal integration, enabling the fusion of imaging results with clinical records, laboratory findings, and molecular information^[1-2]. These capabilities allow LLMs to deliver integrated decision support, thereby advancing the intelligence and efficiency of healthcare services.

In this article, we systematically examine the current applications and future directions of LLMs in medical imaging. We first analyze the clinical value of medical imaging and the key challenges it faces, then review the technological evolution of AI in this domain. We place particular emphasis on the applications of LLMs in disease staging, treatment evaluation, and follow-up management, before addressing data, technical, and translational barriers^[3]. Finally, we outline prospects for multimodal integration and human–AI collaboration as drivers of the next phase of development.

2 The value and challenges of medical imaging

Medical imaging plays an irreplaceable role in disease management, with its value most evident in three key domains: early diagnosis, accurate staging, and therapeutic evaluation. Imaging technologies can directly depict the morphological

* **Corresponding Author:** ChunYang, 18869731392@163.com

characteristics of lesions and, through functional and molecular imaging, reveal underlying physiological and metabolic processes, thereby providing clinicians with a more comprehensive diagnostic foundation. In early diagnosis, imaging enables the detection of abnormalities even when clinical symptoms are atypical or specific biomarkers are absent. For example, low-dose CT has markedly improved the detection rate of early lung lesions in cancer screening ^[4-5], while MRI demonstrates unique advantages in diagnosing central nervous system disorders ^[6]. For disease staging, imaging results can delineate lesion extent, depth of invasion, and metastatic status, thereby offering crucial guidance for individualized treatment planning. In therapeutic evaluation and long-term follow-up, imaging facilitates objective assessment of treatment response through longitudinal comparisons, allowing clinicians to adjust strategies in a timely manner, ultimately improving therapeutic outcomes and patient survival.

Despite these strengths, the clinical application of medical imaging faces considerable challenges. The rapid growth in imaging data has significantly increased the interpretive burden on physicians, making thorough review of all images impractical and raising the risk of missed or inaccurate diagnoses. Moreover, image interpretation is inherently subjective: variations in physician experience and interpretive standards often lead to inconsistent diagnostic conclusions, undermining treatment standardization. With the widespread adoption of multimodal imaging, data complexity has further intensified, spanning anatomical, functional, and molecular dimensions. Efficient and accurate integration of these heterogeneous sources remains a formidable challenge in image analysis. In addition, current evaluation systems fall short in capturing the complexity of treatment responses, and the full potential of imaging for prognostication and long-term follow-up has yet to be realized. Collectively, these issues underscore the urgent need for more intelligent and automated methods in medical image analysis.

3 The evolution of artificial intelligence in medical imaging

The application of artificial intelligence (AI) in medical imaging has undergone a remarkable evolution, progressing from traditional image-processing algorithms to deep learning – based methods. Early approaches largely relied on manually engineered feature extraction techniques, such as edge detection, texture analysis, and morphological operations. While these methods modestly improved the quantification of imaging data, they were limited in representational power and generalizability. The advent of deep learning—particularly convolutional neural networks (CNNs)—has revolutionized medical image analysis, greatly accelerating its automation and intelligence. Through end-to-end training, CNNs can autonomously learn high-level features from raw images, thereby overcoming the constraints of hand-crafted feature selection. In tasks such as lesion detection, organ segmentation, and image classification, deep learning models now perform at levels comparable to, or even surpassing, those of experienced radiologists.

Radiomics has emerged as an important branch of AI in medical imaging. By extracting a large number of quantitative features from images and linking them to disease phenotypes, gene expression, and clinical outcomes, radiomics has advanced the field of precision medicine ^[7]. It enables the discovery of subtle imaging patterns beyond human visual perception and, when integrated with clinical, pathological, and molecular data, facilitates the development of more comprehensive predictive models. In oncology, radiomics has been widely applied to predict molecular subtypes, assess therapeutic response, and guide personalized treatment strategies. The integration of multimodal information further enhances model expressiveness: combining imaging with pathology, genomics, and clinical records improves both interpretability and predictive power, providing clinicians with more reliable decision support.

In recent years, the emergence of general large language models (LLMs) has injected fresh momentum into the field of medical imaging AI. Unlike traditional deep learning models, LLMs possess powerful cross-modal understanding and natural language generation capabilities, enabling them to establish semantic connections between imaging data, textual reports, and structured information ^[8]. These strengths make LLMs particularly valuable in report generation, decision support, and multimodal prediction. For instance, in radiology reporting, LLMs can automatically translate image analyses into clinically standardized narrative descriptions, thereby reducing physician workload and improving report consistency and completeness^[3,9]. In clinical decision support, LLMs can integrate imaging features with patient histories to generate comprehensive diagnostic suggestions, assisting clinicians in formulating individualized treatment plans.

4 Applications of artificial intelligence in medical imaging

Disease diagnosis represents the most mature and clinically impactful application of AI in medical imaging. Imaging data are inherently information-rich, providing not only anatomical structures of lesions but also details regarding morphology, boundaries, density, vascular perfusion, and functional metabolism. These multi-layered features serve as intuitive and quantitative bases for lesion detection, classification, and characterization. Traditional diagnostic workflows rely heavily on physician expertise, with performance influenced by reading volume, subjective judgment, and fatigue—

factors that increase the risk of missed or incorrect diagnoses. By shifting diagnostic interpretation from experience-driven visual inspection to data-driven intelligent analysis, AI significantly improves diagnostic efficiency and consistency, particularly in high-throughput imaging environments.

Deep learning models, especially convolutional neural networks (CNNs), have been widely deployed in tasks such as the automated detection and classification of pulmonary nodules, breast lesions, and brain abnormalities^[10-12]. Through end-to-end training, CNNs can extract higher-order image features, enabling recognition of subtle patterns and microscopic lesions beyond human perception, thereby reducing false negatives and misdiagnoses. Radiomics further enriches diagnostic capacity by extracting high-dimensional, quantitative features from images and linking them with clinical, pathological, and molecular data. This integrative approach bridges visible imaging phenotypes with “invisible” molecular characteristics, providing robust support for early detection and precise disease classification.

The emergence of general large language models (LLMs) has introduced new perspectives into imaging-based diagnosis. Beyond handling visual data, LLMs can integrate imaging findings with patient histories, laboratory results, pathology reports, and both structured and unstructured clinical records. This multimodal capability enables automated lesion detection, diagnostic inference, and report generation^[13]. By extending analysis beyond single-image interpretation to a holistic view of patient health, LLMs enhance both interpretability and clinical utility. Their ability to leverage knowledge transfer also allows strong performance even in low-resource settings or rare disease contexts where annotated data are scarce. Nevertheless, challenges remain, including limited generalizability across centers, devices, and populations, incomplete multimodal integration, and the need for rigorous validation of LLM safety and stability in specialized medical contexts^[12].

Disease staging is another indispensable function of medical imaging, informing treatment strategies, prognosis assessment, and personalized management. Imaging can directly depict lesion location, morphological changes, functional states, and metabolic activity, offering quantifiable and visual evidence for staging. Traditional staging is often subjective and inconsistent across physicians. AI systems, through deep learning and radiomics, can automatically extract high-dimensional features, quantify tumor size, infiltration depth, and metastatic spread, and integrate morphologic and functional data to substantially improve staging accuracy and reliability.

LLMs offer a higher level of intelligence for staging tasks. By simultaneously processing imaging data, electronic health records, laboratory results, and other structured information, they align cross-modal inputs and generate clinically coherent staging conclusions in structured, interpretable formats. This reduces subjective bias and provides unified standards for multicenter studies and treatment planning, thereby enhancing the consistency and scientific rigor of clinical management.

Therapeutic response assessment and prognosis prediction aim to quantify imaging changes to evaluate treatment efficacy and forecast disease trajectories. Traditional methods—such as tumor diameter or volumetric measurements—capture only limited aspects of disease biology and fail to reflect microenvironmental changes or functional heterogeneity. Functional imaging (e.g., PET, fMRI) and molecular imaging capture metabolic and physiological shifts with higher sensitivity, enriching therapeutic evaluation. AI models, leveraging deep learning and radiomics, extract high-throughput features to construct predictive models of treatment response, and temporal sequence analysis enables earlier and more precise detection of subtle changes during therapy^[14]. LLMs further personalize these evaluations by integrating multimodal data to generate individualized therapeutic assessments and prognostic reports. Their natural language explanations clarify the rationale behind predictions, thereby enhancing physician trust and informing rational therapeutic adjustments.

Follow-up and long-term management rely on dynamic imaging surveillance to continuously evaluate disease evolution, enabling early detection of recurrence or metastasis and timely therapeutic adjustments. AI facilitates automated image registration, change detection, and anomaly recognition across time points, increasing sensitivity for disease progression and improving follow-up efficiency. Large-scale data analytics can further identify latent imaging patterns and risk signatures to support individualized management. LLMs, by synthesizing serial imaging examinations with clinical history, treatment records, and laboratory indicators, generate dynamic follow-up reports that summarize disease trajectories and risk stratification^[13,15]. This deep integration of imaging with multidimensional clinical data provides actionable, interpretable tools for long-term management, advancing imaging from one-time diagnosis toward continuous, fine-grained patient care.

5 Challenges and future perspectives

The advancement of AI in medical imaging fundamentally depends on the quality and scale of data, yet multiple challenges remain—including scarcity of annotations, inconsistent quality, lack of standardization, and privacy concerns. Annotation of medical images requires highly trained specialists, making the process costly and time-consuming, while inter-observer variability reduces labeling consistency and impairs model training. Significant heterogeneity across

institutions—stemming from differences in equipment, acquisition protocols, and storage formats—further undermines generalizability. Privacy and data security pose equally critical barriers: imaging contains sensitive patient information, and strict regulations on data sharing across institutions constrain large-scale studies^[16]. Potential solutions include weakly supervised and self-supervised learning to reduce reliance on labels, and domain adaptation or adversarial training to improve cross-center robustness.

On the technical front, challenges center on model generalizability, interpretability, robustness, and computational efficiency. While AI models often achieve high performance in internal testing, their accuracy tends to drop across devices, populations, and centers, limiting clinical deployment. The lack of interpretability reduces physician trust, as “black-box” decisions rarely provide clinically meaningful explanations^[17]. Robustness is equally crucial: image noise, artifacts, or rare pathologies can trigger misclassifications, and adversarial attacks present potential security risks^[18]. Moreover, training and inference for large-scale models demand substantial computational resources, restricting adoption in resource-limited environments. Addressing these issues requires advances in domain adaptation, interpretable AI, robust learning, and efficiency-oriented approaches such as model compression, knowledge distillation, quantization, and edge computing^[19-20]. For LLMs in particular, multimodal alignment and knowledge augmentation represent essential strategies to strengthen medical domain expertise.

Clinical translation remains a decisive step for realizing the value of AI in medical imaging, yet it is hindered by limited physician acceptance, insufficient multicenter evidence, integration complexity, and regulatory gaps. Physicians’ trust depends on model accuracy, interpretability, and seamless workflow integration; opaque models are difficult to embed into diagnostic practice. Most published studies rely on single-center, retrospective data, with conclusions that require prospective, multicenter validation before being endorsed in clinical guidelines. Integrating AI systems into existing hospital information infrastructures also poses practical barriers, requiring smooth data exchange and user-friendly interfaces. From a regulatory perspective, approval standards, liability attribution, and ethical oversight for AI-based medical products remain underdeveloped, necessitating robust frameworks to ensure safety and compliance.

Looking ahead, medical imaging AI is expected to advance toward multimodal integration, privacy preservation, human–AI collaboration, and full-cycle clinical support^[21]. Multimodal foundation models that combine imaging, text, genomics, and clinical data within a unified semantic space will enable more comprehensive reasoning and decision-making. Privacy-preserving techniques such as federated learning may facilitate cross-institutional collaboration, breaking down data silos and supporting large-scale training and validation. Rather than replacing physicians, AI will become an integral partner in collaborative diagnosis and care, enhancing efficiency and consistency. Its applications will extend across screening, staging, treatment planning, therapeutic monitoring, and follow-up, enabling end-to-end intelligent medical support. Establishing robust standards and norms—including technical evaluation, clinical validation, regulatory frameworks, and ethical guidelines—will be essential to ensure safe, reliable, and clinically meaningful deployment of AI in medical imaging.

6 Conclusion

Medical imaging is an indispensable component of modern medicine, with established value in disease diagnosis, staging, therapeutic evaluation, and longitudinal follow-up. The rapid advances in artificial intelligence—particularly deep learning, radiomics, and general large language models (LLMs)—are driving a paradigm shift in imaging, moving from traditional qualitative assessment toward data-driven, intelligent, and individualized analysis. AI markedly enhances efficiency and consistency in image interpretation, while LLMs, through multimodal integration and natural language generation, offer novel solutions for complex clinical tasks.

Nevertheless, the widespread adoption of AI in medical imaging continues to face significant challenges, including limited annotated data, insufficient model generalizability, low physician acceptance, and incomplete regulatory frameworks. Addressing these barriers will require the development of multimodal foundation models, the application of privacy-preserving computational techniques, optimization of human–AI collaboration, and the establishment of robust evaluation and regulatory systems. Together, these efforts will facilitate the transition of medical imaging AI from an auxiliary tool to a knowledge-driven partner in clinical care.

In sum, LLMs and allied AI technologies are reshaping the landscape of medical imaging. Beyond improving accuracy and efficiency, they are infusing precision medicine and smart healthcare systems with new momentum. As the technologies mature and clinical adoption deepens, AI is poised to enable patient-centered, full-cycle management and to serve as a transformative force in the future of medicine.

References

[1] D’Antonoli TA, et al. Large language models in radiology: fundamentals, applications, ethical considerations, risks and future directions. *Diagn*

- Interv Radiol. 2024;30(1). PMID:37789676.
- [2] Tanno R, et al. Collaboration between clinicians and vision – language models for radiology report generation (Flamingo-CXR). *Nat Med.* 2025;31.
- [3] Kim S, et al. Large Language Models: A Guide for Radiologists. *Korean J Radiol.* 2024;25(2):126–133.
- [4] Krist AH, et al. Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA.* 2021;325(10):962–970.
- [5] Fukunaga MI, et al. The 2021 USPSTF lung cancer screening recommendation—context and evidence. *JAMA Oncology.* 2021;7(5):684–686.
- [6] Rocca MA, et al. Current and future role of MRI in the diagnosis and monitoring of multiple sclerosis. *Lancet Reg Health Eur.* 2024;39:100835.
- [7] Li H, et al. Radiomics in precision medicine for colorectal cancer: a decade review (2013–2023). *Front Oncol.* 2024;14.
- [8] Li C, et al. LLaVA-Med: Training a large language-and-vision assistant for biomedicine in one day. *arXiv.* 2023;2306.00890.
- [9] Huang J, et al. Efficiency and quality of generative-AI – assisted radiology reporting in clinical practice: a cohort study. *JAMA Netw Open.* 2025;8.
- [10] Marinovich ML, et al. Artificial intelligence for breast cancer screening: accuracy compared with radiologists in real-world programs. *EBioMedicine.* 2023;94:104686.
- [11] Abeeih EA, et al. Screening mammography and artificial intelligence: a systematic review. *Radiology: Artificial Intelligence.* 2025;7(2):e230144.
- [12] Serapio A, et al. An open-source fine-tuned large language model for radiology impressions across modalities and hospitals. *BMC Med Imaging.* 2024;24.
- [13] Hong EK, et al. Diagnostic accuracy and clinical value of a domain-specific multimodal report-generating model from a single chest radiograph. *Radiology.* 2025.
- [14] Nakajima EC, et al. Tumor size is not everything: advancing radiomics as a biomarker for therapy response. *JCO Precis Oncol.* 2024;8.
- [15] Guan H, et al. Federated learning for medical image analysis: a survey. *Int J Med Inform.* 2024.
- [16] Rehman MH ur, et al. Federated learning for medical imaging and radiology: privacy, heterogeneity and aggregation issues. *Diagnostics.* 2023;13.
- [17] Borys K, et al. Explainable artificial intelligence in medical imaging: a scoping review for clinical practice. *Eur J Radiol.* 2023;162:110786.
- [18] Sorin V, et al. Adversarial attacks in radiology: a systematic review. *Eur J Radiol.* 2023;167:111085.
- [19] Muhammad D, et al. Explainable AI in medical image analysis: a systematic review. *Comput Methods Programs Biomed Update.* 2024;4:100226.
- [20] Dantas PV, et al. A comprehensive review of model compression techniques in machine learning for edge and embedded deployment. *Applied Intelligence.* 2024;54:11804–11844.
- [21] Pati S, et al. Privacy preservation for federated learning in health care: guidelines. *Patterns.* 2024;5(7).