

The Reform and Application of AI-Driven Multimodal Data Fusion in Undergraduate Medical Imaging Education

Rui Yang*, Qingwei Chen, Ruibo Zhang, Xiaoming Yu, Yuan Yuan

Second Affiliated Hospital of Baotou Medical College, Baotou, Inner Mongolia Autonomous Region, China

*Correspondence Author

Abstract: *Medical imaging serves as a critical pillar of clinical diagnosis. Traditional teaching methods predominantly rely on single-modality images, inadequately preparing students for comprehensive decision-making in complex clinical scenarios. In recent years, the rapid advancement of artificial intelligence (AI), particularly multimodal data fusion technologies, has introduced new paradigms for reforming medical imaging education. This paper systematically reviews cutting-edge developments in AI-driven multimodal fusion for disease prediction, precision diagnosis, and clinical management while exploring feasible pathways for integrating these innovations into undergraduate teaching. By incorporating cross-scale data integration, interpretability analysis, and deep learning architectures, we propose a teaching reform framework centered on “clinical problem-driven learning, multimodal data immersion, and AI-assisted reasoning” to enhance medical students’ abilities in synthesizing complex imaging data and advancing clinical thinking.*

Keywords: Artificial Intelligence, Multimodal Data Fusion, Medical Imaging Education, Deep Learning, Teaching Reform.

1. Introduction

Medical imaging, serving as the critical bridge between foundational theories and clinical practice, has long confined its pedagogical approach to single-modality image interpretation (e.g., CT, MRI, or ultrasound) supplemented by textual descriptions. However, real-world clinical decision-making is an intricate process requiring the integration of patient histories, laboratory biochemical markers, genomic features, and multisource imaging data. The current undergraduate education system exhibits significant structural deficiencies in cultivating students’ ability to synthesize such heterogeneous multisource information.

Recent advancements in artificial intelligence, particularly deep learning and multimodal data fusion algorithms, have demonstrated cross-scale and cross-modality analytical advantages across numerous medical subspecialties. Empirical studies have revealed that in diabetic complication prediction tasks, multimodal models integrating clinical phenotypes, molecular omics, and medical imaging significantly outperform single-modality architectures [1]. Similarly, in holistic liver disease management, deep learning-based imaging-clinical data fusion strategies have successfully achieved precise diagnosis and optimized treatment planning [2]. These breakthroughs not only establish a robust technological foundation for pedagogical reform in medical imaging but also provide new conceptual guidance: educational paradigms should transition comprehensively from static image recognition training to dynamic, multidimensional, and clinically task-oriented integrated learning models.

2. Theoretical Foundation and Technological Evolution of Multimodal Data Fusion

The core objective of multimodal data fusion lies in aggregating complementary information from heterogeneous

data sources such as imaging, text, genomic, and physiological signals, thereby significantly enhancing the models’ discriminative capabilities. Depending on the stage of fusion implementation, this technical approach can be categorized into early fusion, intermediate fusion, late fusion, and attention mechanism-based fusion strategies [6]. Within deep learning frameworks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), transformer architectures, and their hybrid variants (e.g., CNN-LSTM) have been extensively deployed for cardiovascular disease risk assessment and medical image analysis, with certain models achieving classification accuracy surpassing 99% [6].

It is noteworthy that despite the strong momentum of deep learning, traditional machine learning methods still retain irreplaceable unique value in model interpretability, few-shot learning scenarios, and multimodal fusion mechanisms [4]. Particularly in the pedagogical context of medical imaging analysis, traditional algorithms prove highly suitable for rapid prototyping and intuitive demonstration of core principles due to their high transparency and computational efficiency [4]. This insight reveals that when constructing educational models, we should adopt a progressive teaching strategy: “consolidating theoretical foundations with traditional machine learning while enhancing prediction accuracy through deep learning.”

3. Current Status Review and Reform Imperatives in Undergraduate Medical Imaging Education

Currently, undergraduate-level imaging education faces three major bottlenecks: First, teaching case libraries overly rely on single-modality typical cases, severely lacking practical training in multimodal joint analysis; second, students struggle to establish deep logical connections between imaging features and clinical prognosis/molecular mechanisms; and finally, existing teaching tools predominantly involve static presentations, lacking

interactive and intelligent support systems. To address these pain points, implementing AI-driven multimodal data integration concepts has become imperative.

Existing research confirms that by deeply exploring the intrinsic relationships among medical imaging, electronic health records, and genomic data, deep learning can not only effectively predict high-risk complications but also optimize diagnostic pathways and curb healthcare costs [5]. This logic similarly applies to the restructuring of educational scenarios, where virtual case-based teaching compels students to simultaneously process multidimensional information such as imaging materials, laboratory test indicators, and genetic backgrounds, thereby simulating real clinical decision-making chains to authentically cultivate clinical reasoning and comprehensive decision-making skills.

4. Design Concept of the Teaching Reform Framework

Based on the aforementioned analysis, this study constructed a “three-stage” teaching reform framework:

The first phase focuses on solidifying the data foundation. It aims to establish a multimodal teaching case library encompassing CT/MRI/ultrasound imaging, structured clinical records, laboratory test indicators, and genomic biomarkers (e.g., SNP or proteomic features). Following the data integration paradigm for diabetes complication prediction outlined in reference [1], a dedicated “virtual patient” database for teaching purposes will be created.

The second phase focuses on the integration of model tools. Develop or introduce an explainable AI-assisted teaching system: On the one hand, utilize traditional machine learning algorithms such as support vector machines (SVMs) and random forests to guide students in mastering disease characteristics and model logic [4]; on the other hand, deploy pretrained deep learning models such as CNN-LSTM or attention fusion networks to demonstrate the auxiliary diagnostic process for complex cases [6]. Simultaneously, explainability techniques such as gradient-weighted class activation mapping (Grad-CAM) and SHAP can be integrated to elucidate the rationale behind AI decisions [6].

Phase Three: Implementation of Teaching Practices and Evaluation. The “Case-Driven + Flipped Classroom” model is implemented: Students work in groups to process multimodal case data, utilize AI tools for preliminary analysis, and then engage in discussions by comparing results with clinical gold standards. Evaluation metrics include diagnostic accuracy, completeness of reasoning logic, and critical assessment capabilities of AI outputs. This model draws on the “from pixels to patients” deep learning evolution concept from the literature [3], emphasizing a shift from single-image recognition to patient-centered comprehensive decision-making capabilities.

5. Applications and Challenges

The AI teaching model driven by multimodal data integration not only enhances medical students’ higher-order thinking but also lays a solid foundation for future clinical practice. For

example, in liver disease imaging instruction, integrating imaging, liver function indicators, and pathological data can visually demonstrate disease progression and treatment responses [2]; in cardiovascular training, combining wearable device data with imaging helps students master dynamic risk assessment [5]. However, challenges remain: 1) High-quality multimodal teaching datasets are difficult to obtain, requiring multicenter collaboration and data standardization; 2) While AI tools are powerful, their potential “black box” risks demand an emphasis on critical usage in teaching [3]; 3) High costs for faculty training and hardware infrastructure depend on policy and funding support.

6. Conclusion

AI-driven multimodal data fusion has brought revolutionary opportunities to undergraduate medical imaging education. By integrating multisource information such as clinical data, molecular omics, and medical imaging, combined with explainable deep learning and traditional machine learning methods, a teaching system highly aligned with real clinical environments can be constructed. Future research should further explore the educational applications of large language models and AI agents, promote interdisciplinary collaboration, and achieve integrated reform in “teaching, learning, and assessment” [13]. Although current studies primarily focus on clinical prediction, their methodologies and core concepts can be fully extended to medical education, cultivating future physicians’ core competencies in this data-intensive era.

Fund Project

This paper is supported by the fund Exploring the Application Value of AI Multimodal Data Fusion in Medical Imaging Education, Project Number: 2025BYJG-LC-05.

References

- [1] Xie X, Wu C, Huang Z, et al. Artificial intelligence-powered prediction of diabetic complications: from clinical data to molecular omics. *Brief Bioinform.* 2026. PMID: 41744227. 1
- [2] Zhang P, Gao C, Huang Y, et al. Artificial intelligence in liver imaging: methods and applications. *Hepatol Int.* 2024. PMID: 38376649. 2
- [3] Yang Y, Shen H, Chen K, et al. From pixels to patients: the evolution and future of deep learning in cancer diagnostics. *Trends Mol Med.* 2025. PMID: 39665958. 3
- [4] Chechekhina E, Voloshin N, Solopov M, et al. Traditional machine learning in biomedical image analysis: before you go too deep. *Front Artif Intell.* 2026. PMID: 41696042. 4
- [5] Chen H, Zeng Y, Cai D. Deep learning for cardiovascular management: optimizing pathways and cost control under diagnosis-related group models. *Front Artif Intell.* 2025. PMID: 40959774. 5
- [6] Ganeshan N, Magesh G. Deep learning for cardiovascular disease: a comprehensive review of detection and risk forecasting. *Front Artif Intell.* 2026. PMID: 42434360. 6