

# Current Status and Ethical Considerations of Artificial Intelligence in Medical Imaging

Xinxin Pan<sup>1,a</sup>, Qinshao Wei<sup>2\*</sup>

<sup>1</sup>Clinical Medical College of Tianjin Medical University, Tianjin, 300270, China

<sup>2</sup>Youjiang Medical University for Nationalities, Baise, 533000, Guangxi, China

<sup>a</sup>3488014675@qq.com

\*Corresponding author: 2025161769@qq.com

**Abstract:** Artificial Intelligence (AI) technology is profoundly transforming the diagnostic paradigm of medical imaging. Centered on deep learning and large language models (LLMs), AI technologies have been widely applied in multi-modal imaging tasks, including lesion detection, image segmentation, computer-aided diagnosis, and automated report generation, demonstrating significant value in improving diagnostic efficiency, accuracy, and standardization. However, as these technologies penetrate into core clinical decision-making processes, emerging challenges have become increasingly apparent, including algorithmic “black box” opacity, data privacy breaches, algorithmic bias, ambiguous liability attribution, and the erosion of humanistic care. This paper systematically reviews the technical foundations, clinical practices, and workflow impacts of AI in medical imaging, provides an in-depth analysis of the ethical dilemmas faced by these technologies, and proposes normative pathways and development strategies based on existing research. The aim is to provide reference for the safe, compliant, and sustainable clinical application of medical imaging AI.

**Keywords:** Artificial Intelligence; Medical Imaging; Deep Learning; Large Language Models; Ethical Governance; Clinical Application

## 1. Introduction

As the core visual evidence for clinical diagnosis and treatment, medical imaging accounts for over 90% of total healthcare data<sup>[17]</sup>. Traditional manual analysis has long faced challenges including low diagnostic efficiency, high rates of missed and misdiagnosis, and heavy reliance on operator experience. In primary healthcare institutions, insufficient physician experience further exacerbates the variability and instability of diagnostic quality<sup>[8]</sup>. In recent years, significant breakthroughs in deep learning technologies, particularly convolutional neural networks (CNNs), Transformer architectures, and large language models, have enabled AI to demonstrate performance approaching or even surpassing that of senior physicians in automated lesion recognition and quantitative analysis tasks, becoming a critical force in driving healthcare digital transformation and promote the allocation of high-quality medical resources to lower-level areas<sup>[3][11]</sup>.

As technology penetrates into core clinical decision-making processes, accompanying ethical risks have gradually been exposed. The “black box” nature of deep learning models makes decision-making processes difficult to interpret; tensions exist between medical data sharing and patient privacy protection; algorithmic bias may exacerbate healthcare inequities; and legal frameworks for liability allocation following misdiagnosis remain unclear<sup>[18]</sup>. Seeking balance between technological application and ethical governance has become a core prerequisite for the large-scale clinical deployment of medical imaging AI, holding significant importance for addressing uneven distribution of medical resources and comprehensively improving diagnostic quality<sup>[7][14]</sup>.

## 2. Current Status of AI Applications in Medical Imaging

### 2.1 Core Technical Support Systems

Deep learning serves as the primary technical foundation for current medical imaging AI, with mainstream architectures comprising convolutional neural networks (CNNs) and Transformers<sup>[17]</sup>. CNNs leverage local receptive fields and weight sharing to effectively capture local textural features of

images, demonstrating superior performance in lesion segmentation and small-sample learning scenarios. U-Net has become a classic model in medical image segmentation, achieving a Dice coefficient of 88.7% in CT liver segmentation<sup>[2]</sup>. Transformers, through self-attention mechanisms, model global contextual information and excel in high-resolution pathological whole-slide image classification tasks, effectively capturing long-range dependencies, though their high computational complexity makes them prone to overfitting on small-scale medical datasets<sup>[17]</sup>.

To balance efficiency and performance, CNN-Transformer hybrid architectures have improved segmentation accuracy by integrating the strengths of both approaches<sup>[10]</sup>. The incorporation of Generative Adversarial Networks (GANs) has further expanded technical boundaries; the adversarial training mechanism between generators and discriminators enables cross-modal image synthesis and data augmentation. Wang et al. demonstrated in their research on brain arteriovenous malformation imaging that GAN technology not only helps alleviate data scarcity in medical imaging but also reduces radiation exposure risks associated with CT examinations<sup>[13]</sup>. The rapid development of large language models and multimodal large models has endowed medical imaging AI with enhanced semantic understanding and comprehensive decision-making capabilities. Research by Cheng et al. shows that GPT-4 series models can not only achieve structured generation of imaging reports and multilingual translation but also conduct end-to-end diagnosis by integrating imaging data with clinical textual information, with accuracy in tasks such as tumor TNM staging and imaging report interpretation approaching professional physician levels. Xu et al. further note that multimodal large models, by integrating multi-source information including imaging, text, and time-series data, hold promise for achieving more comprehensive clinical decision support<sup>[15]</sup>.

## **2.2 Multi-modal Imaging Clinical Applications**

At the clinical application level, AI has penetrated multi-modal imaging fields including ultrasound, X-ray, CT, MRI, pathology, and endoscopy<sup>[17]</sup>. In ultrasound imaging, AI helps address challenges of high image noise and strong operator dependence. Yang et al. found in their research on prostate cancer ultrasound diagnosis that AI systems can automatically identify standard planes, measure key parameters, and assist in differentiating benign from malignant nodules<sup>[16]</sup>. In X-ray imaging, AI has improved screening efficiency and accuracy for common diseases including fractures, pneumonia, tuberculosis, and breast cancer; emergency fracture detection sensitivity reaches 92.5% with a 47% reduction in missed diagnosis rates, while chest radiography distinguishes COVID-19 from other pneumonia with an AUC of 0.96<sup>[17]</sup>.

In CT and MRI, AI technology enables high-quality reconstruction of low-dose CT and precise detection of small pulmonary nodules. Research by Li et al. confirmed that AI-assisted assessment of head and neck CT angiography is comparable to senior physicians<sup>[9]</sup>; in MRI, deep learning-accelerated imaging techniques have improved brain tumor segmentation accuracy, with preoperative prediction of glioma genetic mutation status achieving AUCs of 0.88–0.91<sup>[12][17]</sup>. Zhang notes that localized deployment of large models such as DeepSeek in medical institutions provides new technical pathways for privacy protection and real-time application of medical imaging AI<sup>[18]</sup>. In pathology and endoscopic imaging, AI-assisted whole-slide image analysis can predict tumor genetic mutation status, with endoscopic colorectal polyp detection sensitivity reaching 98.6% and missed diagnosis rates reduced by 53%<sup>[4]</sup>. Research by Cheng et al. emphasizes that the value of AI in early screening and treatment of gastrointestinal tumors is particularly prominent, with AUC for early gastric cancer recognition reaching 0.96<sup>[4]</sup>.

## **2.3 Full-process Workflow Reconstruction and Primary Care Empowerment**

AI not only improves efficiency in individual diagnostic processes but also achieves automation and integration of the entire medical imaging workflow from data preprocessing, lesion segmentation, intelligent diagnosis to report generation through workflow reconstruction. This process reconstruction significantly enhances radiology department efficiency, compressing single-case CT pulmonary nodule analysis time from 15 minutes to 4–9 minutes<sup>[2]</sup>. Regional AI quality control platforms have promoted diagnostic standardization across different levels of medical institutions. Li et al. note that regional quality control systems represented by “Yueyi Zhiying” provide real-time intelligent analysis for primary healthcare institutions, shortening physician reading time and improving diagnostic accuracy from 78% to 91%<sup>[8]</sup>. AI-assisted imaging quality control technology achieves intelligent calibration of equipment, real-time quality control, and automatic rejection rate statistics, providing support for the foundational quality of imaging diagnosis<sup>[10]</sup>.

### **3. Ethical Challenges Facing AI in Medical Imaging**

#### ***3.1 Algorithmic “Black Box” and Lack of Explainability***

The “black box” nature of deep learning models represents the primary ethical challenge facing medical imaging AI<sup>[6]</sup>. While visualization techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) can localize lesion regions, they struggle to fully present the internal reasoning logic and decision basis of models, making it difficult for physicians and patients to trace the formation process of diagnostic conclusions<sup>[17]</sup>. This characteristic not only affects clinicians’ trust in AI results but may also lead to misdiagnosis due to blind reliance on AI judgments, contradicting the principles of informed consent and prudence required in medical practice<sup>[19]</sup>. Research by Zhao et al. further emphasizes that in high-risk disease diagnosis scenarios, unexplainable AI results may delay patient treatment and cause adverse outcomes<sup>[19]</sup>. On another level, the “black box” problem reflects the tension between computational rationality and clinical rationality—machines excel at pattern recognition and probabilistic calculation, while clinical decision-making requires causal reasoning and value judgment. Chen et al. note that current regulatory bodies in most countries position AI products as “assistive” in nature, requiring physician confirmation and assumption of legal responsibility; while this institutional design is reasonable at present, it also indirectly confirms the fundamental impact of algorithmic “black boxes” on liability attribution<sup>[1]</sup>.

#### ***3.2 Medical Data Privacy and Security Risks***

Medical imaging data contains substantial sensitive patient information, while AI model training relies on massive multi-center data. Zhang notes that risks of data leakage exist throughout the entire lifecycle of collection, annotation, storage, and sharing<sup>[18]</sup>. Tensions exist between cross-institutional data sharing and patient privacy protection; some medical institutions, in pursuit of model performance, engage in non-standardized data collection and incomplete de-identification. Insufficient data standardization and data heterogeneity caused by equipment differences result in opaque data sources for some model training, potentially infringing on patients’ information rights. From a philosophical perspective, Dai et al. reflect that the transformation of data from “traces” to “evidence” involves processing in specific contexts rather than simple representation of objective facts<sup>[5]</sup>. While localized deployment can reduce data leakage risks to some extent, Zhang emphasizes that without unified data governance systems, it remains difficult to achieve balance between data security and effective utilization<sup>[18]</sup>. From an ethical perspective, the process of patient authorization for data use should not become merely formalistic but should constitute a substantive manifestation of patient autonomy.

#### ***3.3 Algorithmic Bias and Impairment of Healthcare Equity***

The root of algorithmic bias lies in distribution bias in training data. If datasets are overly concentrated on specific populations, equipment, regions, or common diseases, model diagnostic performance will significantly decline when facing rare diseases, ethnic minorities, images from outdated equipment in primary institutions, or elderly patient populations. Diagnostic models for rare diseases, due to insufficient training data, are prone to missed or misdiagnosis; primary healthcare institutions using older equipment models produce image quality differing from high-end equipment, potentially causing model diagnostic inaccuracies. Insufficient model generalization capability may output erroneous judgments with “high confidence” when facing atypical cases or rare complications, potentially exacerbating urban-rural and regional healthcare disparities<sup>[7][8]</sup>. The root of algorithmic bias lies in unequal data power—groups underrepresented in training data are also likely to have their health needs neglected by technical systems.

#### ***3.4 Ambiguous Liability Attribution and Legal Ownership***

The allocation of medical liability arising from medical imaging AI represents a challenge facing current legal systems. Zhao et al. note that when AI systems cause patient harm through misdiagnosis or missed diagnosis, liability boundaries among developers, medical institutions, and clinical physicians remain unclear<sup>[19]</sup>. Complex liability disputes may arise from physicians directly adopting AI results without review, inherent algorithmic defects in AI systems, or non-standard equipment deployment by medical institutions; existing legal provisions have not clarified respective liability proportions and attribution principles. Zhang’s research emphasizes that AI lacks legal subject status and cannot independently bear legal liability; ultimate liability often falls on medical institutions and

physicians, and this accountability mechanism may affect the promotion and adoption of AI technology<sup>[18]</sup>. From the perspective of regulatory innovation, Chen et al. propose that the “value chain” methodology advocated by the World Health Organization could be borrowed to clarify legal relationships and liability attribution according to value chain assessment principles covering design, development, provision, and deployment<sup>[1]</sup>.

### ***3.5 Loss of Humanistic Care and Alienation of Physician-Patient Relationships***

Excessive AI intervention has weakened emotional connections between physicians and patients to some extent. AI systems focus on lesion recognition and data analysis, often neglecting individual patient differences, psychological states, and emotional needs, shifting medical practice from “patient-centered” to “data-centered”<sup>[18]</sup>. If physicians reduce communication with patients due to reliance on AI results, patients’ understanding and trust in the diagnostic process may be affected. From a philosophical perspective, Dai et al. reflect that clinicians’ subjective factors—intuition, experience, and empathy—are not “flaws” to be replaced by algorithms but represent the essence of medicine as a practical science oriented toward “patients”<sup>[5]</sup>. Zhang et al. emphasize that constructing a new “physician-medical AI-patient” relationship requires positioning AI as an assistive tool rather than decision-making subject, preserving emotional connections between physicians and patients<sup>[18]</sup>.

## **4. Ethical Governance and Development Strategies for Medical Imaging AI**

### ***4.1 Strengthening Technical Explainability to Break the “Black Box” Dilemma***

Addressing the lack of algorithmic explainability requires intensified technical research and development efforts to advance explainable AI (XAI) technologies. Duan et al. systematically reviewed progress from traditional concept-based methods to recent advances combining large language models, providing a comprehensive reference framework for technical R&D<sup>[6]</sup>. Visualization methods should be promoted, requiring AI models to simultaneously output diagnostic results while annotating key information including diagnostic basis, lesion location, and confidence levels to help clinicians understand decision logic. Yu et al. recommend that in high-risk disease diagnosis scenarios, the independent use of “black box” models should be restricted, insisting on physicians’ final review authority and constructing an “AI-assisted + manual review” diagnostic model<sup>[17]</sup>. Simultaneously, lightweight and traceable AI models can be developed to achieve full-process recording and tracing of model decision-making.

### ***4.2 Improving Data Governance Systems to Ensure Privacy and Security***

Constructing full-lifecycle medical imaging data governance systems is fundamental to ensuring data security. Zhang recommends establishing unified medical imaging data standards and sharing specifications, clarifying technical requirements and ethical guidelines for data collection, annotation, storage, and sharing<sup>[18]</sup>. Strict adherence to anonymization and de-identification principles is required, employing technologies such as federated learning and blockchain to achieve “data availability without visibility”, completing model training and optimization without leaking raw data. Zhao et al. emphasize improving informed consent processes for data collection and use, safeguarding patients’ rights to know, choose, and withdraw regarding data use<sup>[19]</sup>. Additionally, data security supervision should be strengthened, establishing full-lifecycle traceability mechanisms for data and clarifying data management responsible parties<sup>[18][19]</sup>.

### ***4.3 Eliminating Algorithmic Bias to Safeguard Healthcare Equity***

Optimizing algorithmic fairness requires approaches from both data and model perspectives. At the data level, constructing multi-center, diversified, and balanced training datasets covering different ages, genders, ethnicities, and regions, and incorporating special data such as rare diseases and atypical cases, can improve dataset representativeness; conducting cross-institutional external validation ensures model performance stability under different data distributions. At the model level, establishing algorithmic fairness assessment mechanisms, regularly detecting diagnostic differences across different population groups, and optimizing and iterating models with significant deviations. AI technology should be promoted toward primary and remote areas through technical training and equipment upgrades, enhancing AI application capabilities in primary healthcare institutions and narrowing the

healthcare digital divide<sup>[8]</sup>. Zhang believes that unified procurement of AI technical services by health authorities, unified medical data organization, and unified algorithm model training represents an effective pathway for standardizing medical AI applications and ensuring data security<sup>[18]</sup>.

#### ***4.4 Improving Legal and Ethical Regulation to Clarify Liability Attribution***

Accelerating specialized legislation for medical artificial intelligence is necessary to clarify liability allocation systems for medical imaging AI-related harm, defining liability boundaries among developers, medical institutions, and clinical physicians. Zhao et al. recommend that developers be responsible for model quality, safety, and compliance; medical institutions bear responsibility for AI system deployment, management, and supervision; and clinical physicians retain final clinical decision-making authority and responsibility for diagnostic behaviors adopting AI results<sup>[19]</sup>. Sound approval and post-market surveillance systems for medical imaging AI medical devices should be established, subjecting high-risk AI products to strict approval according to Class III medical device standards and conducting continuous post-market monitoring and evaluation. Chen et al. note that the U.S. FDA's "Software Pre-Certification" model could be borrowed to explore innovative pathways from product certification to institutional pre-certification, promoting "institutional regulation" for high-risk products not yet included in regulatory frameworks<sup>[1]</sup>. Simultaneously, interdisciplinary ethics committees covering multiple fields can be established to conduct ethical review of AI technology applications, ensuring technological development conforms to ethical norms<sup>[1]</sup>.

#### ***4.5 Adhering to Human-Machine Collaboration to Return to the Humanistic Essence of Medicine***

The assistive positioning of AI should be clarified to construct a "human-machine collaborative" diagnostic model. AI should assume data-intensive, standardized tasks, enabling clinicians to focus more on complex case diagnosis, clinical decision-making, humanistic communication, and ethical judgment. Strengthened training for clinicians in AI application should be provided to enhance their ability to interpret, discriminate, and critically evaluate AI results, cultivating professional competence in "using AI wisely without relying on AI"<sup>[18]</sup>. Throughout the diagnostic and treatment process, physician-patient communication should be strengthened, preserving face-to-face interaction between physicians and patients, attending to patients' psychological needs and emotional experiences, and integrating humanistic care into the diagnostic process. From a philosophical perspective, Dai et al. note that clinicians' subjective factors—intuition, experience, and empathy—represent precisely the essence of medicine that AI cannot replace<sup>[5]</sup>. Li et al. emphasize that patients should be guaranteed the freedom from AI "scrutiny" and provided with the option of traditional medical services, representing an important manifestation of safeguarding patient autonomy<sup>[8]</sup>.

## **5. Conclusion**

Artificial intelligence in medical imaging has achieved leapfrog development from technical research to clinical deployment. Research by Yu et al. demonstrates that AI shows application value in multi-modal imaging intelligent diagnosis, workflow optimization, and primary healthcare empowerment<sup>[17]</sup>. However, ethical issues including algorithmic bias, data privacy, informed consent dilemmas, ambiguous liability attribution, and loss of humanistic care have become critical bottlenecks constraining the healthy and sustainable development of this technology<sup>[1][5][8][18][19]</sup>. From a philosophy of technology perspective, these issues are not "side effects" of technological application but "interface tensions" inevitably generated during the process of technology embedding in society—that is, when powerful computational capabilities encounter complex medical practice, the logic of technology and the logic of medicine inevitably collide and adjust. Future development of medical imaging AI must adhere to the principles of technology for good and ethics-first, with strengthening technical explainability, improving data governance systems, eliminating algorithmic bias, perfecting legal regulation, and adhering to human-machine collaboration as primary pathways<sup>[18]</sup>. These pathways require seeking dynamic balance between technological innovation and ethical governance rather than simple zero-sum trade-offs. Through deep integration of medical-engineering interdisciplinary collaboration and multi-stakeholder collaborative governance, promoting safe, compliant, and humanistic applications of artificial intelligence in medical imaging will ultimately provide solid technical support and ethical safeguards for the implementation of the "Healthy China" strategy.

**References**

- [1] Chen B, Wang Y. *Innovation in AI Medical Supervision and Governance: Reflections on Breaking Through Current Regulatory Frameworks [J]*. *Chinese Medical Ethics*, 2024, 37(9): 1030-1036.
- [2] Chen C, Chen J, Xia L. *Artificial Intelligence Promotes Clinical Application and Research of Medical Imaging [J]*. *Radiologic Practice*, 2024, 39(1): 12-16.
- [3] Chen M. *Intelligent Imaging, Integrated Innovation: Development and Prospects of Medical Imaging Discipline in the AI Era [J]*. *Chinese Journal of Medical Imaging*, 2026, 34(1): 1-2.
- [4] Cheng H, Yan H, Yuan Z, et al. *Research Progress on Applications of Large Language Models in Textual Reports and Medical Imaging for Tumor Diagnosis [J]*. *Journal of Southern Medical University*, 2026, 46(1): 231-238.
- [5] Dai J, Jin Y, Li Y, et al. *Practice and Philosophical Reflections on Digital Transformation of Imaging Diagnosis Technology [J]*. *Medicine and Philosophy*, 2025, 46(23): 7-12.
- [6] Duan G, Feng Y, Feng X, et al. *Review of Concept-based Explainability Methods in Medical Image Classification [J/OL]*. *Computer Engineering and Applications*, 2026. (Online First)
- [7] Guo X, Li W, Xia J, et al. *Expert Consensus on Statistical Design Essentials for Clinical Trials of Medical Imaging AI Medical Devices [J]*. *Chinese Journal of Health Statistics*, 2024, 41(3): 471-477.
- [8] Li X, Liu J. *Artificial Intelligence Empowering Primary Healthcare Services: Progress and Challenges [J]*. *Chinese General Practice*, 2026, 29(4): 436-443.
- [9] Li W, Zhang W, Yang J. *Application Status and Trend Analysis of AI Technology in Medical Computer-Aided Diagnosis [J]*. *Chinese Journal of Medical Physics*, 2025, 42(12): 1668-1674.
- [10] Li Z. *Intelligent Control of Medical Imaging Quality: From Subjective Experience to Intelligent Precision [J]*. *Chinese Journal of Medical Imaging Technology*, 2025, 41(8): 1443-1446.
- [11] Tang L, Zhong S. *Artificial Intelligence and Its Applications in the Medical Field [J]*. *Chinese Journal of Practical Surgery*, 2024, 44(9): 961-965.
- [12] Wang H, Zhao P, Lü H, et al. *Medical Imaging in the Digital Intelligence Era Promoting High-Quality Development of General Surgery [J]*. *Chinese Journal of Surgery*, 2026, 64(1): 7-10.
- [13] Wang X, Yang L, Chen G. *Current Status and Prospects of Artificial Intelligence in Brain Arteriovenous Malformation Imaging [J]*. *Chinese Journal of Cerebrovascular Diseases*, 2025, 22(12): 1-7.
- [14] Xu Z, Zheng M. *Era Innovation and Challenges of Imaging AI in the Medical Field [J]*. *Journal of Shandong University (Health Sciences)*, 2023, 61(12): 7-12.
- [15] Xu Y, Zhao X. *Application Progress of Large Language Models in Medical Imaging [J]*. *Radiologic Practice*, 2025, 40(9): 1082-1088.
- [16] Yang Q, Huang J, Liu X. *Research Progress on Artificial Intelligence Ultrasound Image Diagnosis of Prostate Cancer [J]*. *Chinese Journal of Ultrasound in Medicine*, 2026, 42(2): 241-244.
- [17] Yu X, Yuan Y, Li S. *Artificial Intelligence for Medical Imaging: Progress and Challenges [J]*. *Chinese Journal of Medical Imaging Technology*, 2025, 41(12): 1939-1943.
- [18] Zhang R. *Risks and Regulatory Research on Medical Institutions' Application of AI Technology in the Context of DeepSeek Localized Deployment [J]*. *Health Development and Policy Research*, 2025, 28(5): 493-500.
- [19] Zhao M, Sun Y, Yue Y. *Realistic Dilemmas and Countermeasures of Informed Consent in Medical Generative Artificial Intelligence [J/OL]*. *Chinese Medical Ethics*, 2025.