Medicine Plus 2 (2025) 100081

Contents lists available at ScienceDirect

Medicine Plus

journal homepage: www.sciencedirect.com/journal/medicine-plus

Perspective

Computational biomedical imaging: AI innovations and pitfalls

He Sun[a,b,*], Jinzhuo Wang[a,*]

[a] College of Future Technology, Peking University, Beijing 100871, China

[b] National Biomedical Imaging Center, Peking University, Beijing 100871, China

Computational biomedical imaging lies at the intersection of physics, computer science, and biomedicine, aiming to produce visual representations of biological or physiological phenomena that may be otherwise imperceptible to measuring instruments. Over the last few decades, breakthroughs in imaging physics—as evidenced by modalities like magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, optical microscopy, and endoscopy—have profoundly impacted the way clinicians visualize and understand living systems. Yet, the task of producing accurate and meaningful images from raw sensor data remains non-trivial, often necessitating advanced algorithms capable of dealing with noise, incomplete measurements, motion artifacts, or other confounding factors.

In recent years, artificial intelligence (AI) has emerged as a catalyst for revolutionizing this domain. Following two significant trends in AI developments—the shift from discriminative to generative models and the transition from single-modality to multi-modal foundation models—many innovative AI technologies for natural image and language processing have been quickly adapted to biomedical imaging, offering new methodologies for both image reconstruction and image analysis, thus enabling the development of innovative imaging technologies that can visualize previously invisible biomarkers of diseases and support automated medical diagnosis.

In this paper, we examine how emerging AI trends are reshaping computational biomedical imaging, from enhancing image reconstruction under challenging conditions to revolutionizing diagnostic pipelines. At the same time, we also critically assess the pitfalls of directly adopting AI methodologies without adapting them to the unique features of biomedical images. We aim to offer both an optimistic perspective on the potential of these innovations to revolutionize healthcare and a cautionary note on the risks of their misapplication, emphasizing the need to balance these opportunities and risks to fully unlock AI's potential in this field.

---

* Corresponding authors.
E-mail addresses: hesun@pku.edu.cn (H. Sun), wangjinzhuo@pku.edu.cn (J. Wang).

https://doi.org/10.1016/j.medp.2025.100081
Received 31 December 2024; Received in revised form 21 February 2025; Accepted 17 March 2025

AI innovations in medical image reconstruction. Biomedical images often require reconstruction from indirect measurement data, such as sinograms in CT, k-space data in MRI, or acoustic wavefields in ultrasound. Unlike photography, where images directly represent captured light, biomedical imaging involves decoding hidden visual signals through advanced algorithms. Mathematically, this process can be formulated as a Bayesian inverse problem, commonly known as low-level computer vision tasks:

$$x \sim p(x|y) \propto p(y|x)p(x), \tag{1}$$

where y represents the measured data, x is the latent image to be reconstructed, $p(y \mid x)$ models the physical process of image formation, and $p(x)$ encodes prior knowledge about the image. Traditional reconstruction methods often make strong assumptions or simplifications about imaging physics $p(y|x)$ and image priors $p(x)$. For example, they use linearized models and handcrafted priors such as sparsity or total variation, which can result in inaccuracies and limit their applicability for complex scenarios. Recent AI advancements have revolutionized this process by introducing tools that better model both the physics of imaging and the complexities of real-world anatomical structures.

Deep generative models are powerful tools in approximating the image prior $p(x)$, which is critical for reconstructing high-quality images. Generative models such as variational autoencoders (VAEs), generative adversarial networks (GANs), and diffusion models can learn high-dimensional distributions of medical images from large datasets. These models act as learned priors, guiding the reconstruction process to compensate for incomplete or noisy measurements. In low-dose and sparse-view CT,[1] these models reconstruct images with reduced artifacts, enhancing diagnostic accuracy while reducing radiation exposure without compromising image quality. Similarly, in compressed sensing MRI,[2] generative models recover fine anatomical details from undersampled k-space data, reducing scan times, improving patient comfort, and increasing clinical throughput. Beyond CT and MRI, these approaches have been extended to other modalities, including ultrasound, optical coherence tomography (OCT), and PET-CT (Supplementary Text 1 online), demonstrating their ability to push the limits of reconstruction under many constrained measurement conditions. These advances mark a transformative shift in medical imaging, leveraging data-driven priors to achieve reconstructions that surpass traditional methods.

Another key advancement in modeling image priors is the transition from traditional discrete pixel or voxel representations to implicit neural representations, also known as neural fields. Unlike conventional methods that store images as grids of intensity values and rely on explicit regularization techniques, neural fields utilize continuous neural networks to parameterize the entire image domain. This approach inherently enforces smoothness in reconstructions, producing more coherent results without the need for additional constraints. Furthermore, neural fields offer compact storage for high-dimensional data, such as 3D or 4D volumes, making them efficient for both reconstruction and memory management. Their ability to render images at arbitrary resolutions or interpolate temporal data enables flexible applications, such as time-lapse imaging. In biomedical imaging, neural fields have been successfully applied to tasks such as reconstructing 3D refractive indices in quantitative phase microscopy, where they provide unprecedented detail of biological tissues.[3] Similarly, they are used in fluorescent microscopy and dynamic CT to achieve long-term, high-resolution time-lapse imaging, capturing dynamic physiological changes with exceptional accuracy.[4,5] By combining implicit regularization, memory efficiency, and flexible rendering capabilities, neural fields have expanded the scope of biomedical imaging applications, enhancing both accuracy and versatility.

In addition to complete modeling of image priors, accurate modeling of imaging physics is also essential for decoding signals from complex measurement data. Biomedical imaging modalities like ultrasound and optical imaging typically rely on partial differential equations (PDEs), such as acoustic equations for wave propagation or Maxwell's equations for electromagnetic scattering, to describe the underlying physics. However, traditional PDE solvers are computationally expensive, often requiring simplifying assumptions to ensure practical runtimes. These approximations can result in the loss of important details in reconstructed images. AI-driven neural surrogate models provide an innovative alternative by approximating PDE-based physics with deep neural networks. Trained on simulated or real measurement data, these models bypass the iterative, computationally intensive nature of classical solvers while preserving fine details often missed by linearized physical models. For instance, in ultrasound, these models enable faster and more precise wavefield reconstructions, resolving complex structures like bone-tissue interfaces.[6] In optical imaging, they enhance resolution and fidelity in biological samples while delivering over 20 times faster reconstruction.[7] By integrating physics-based modeling with AI-driven learning, neural PDE models represent a significant step forward in real-time imaging systems, paving the way for faster and more precise medical imaging solutions.

AI innovations in medical image analysis. Medical image analysis involves interpreting visual data to identify patterns, detect anomalies, and make diagnostic predictions, often corresponding to high-level computer vision tasks such as classification, segmentation, detection, and style transfer. Unlike traditional visual tasks, medical image analysis must account for variability in imaging protocols, data quality, and population diversity, necessitating highly specialized algorithms. Recent advancements in AI, spanning vision foundation models, multimodal learning, and synthetic data generation, are reshaping the field. By addressing challenges such as data variability, integration, and scarcity, these innovations have the potential to significantly enhance diagnostics, treatment planning, and scientific discovery, paving the way for more robust and comprehensive approaches to medical image analysis.

Vision foundation models (FMs) have introduced a paradigm shift in medical image analysis. These large-scale AI models are trained on extensive and diverse datasets using unsupervised, weakly supervised, or self-supervised learning, enabling them to acquire general-purpose representations. Once trained, these models can be fine-tuned for a variety of downstream tasks, from image segmentation and anomaly detection to diagnosis and prognosis. For example, CHIEF,[8] pretrained on 44 TB of pathology images, demonstrated exceptional diagnostic and prognostic accuracy across 19 cancer types, even when tested on datasets with domain shifts caused by diverse population sources and slide preparation techniques. Similarly, RETFound,[9] a retinal FM trained on 1.6 million unlabeled retinal images, has achieved success in diagnosing sight-threatening eye diseases while predicting systemic disorders like heart failure and myocardial infarction. More ambitious FMs, such as BME-x,[10] a magnetic resonance FM, and BiomedParse,[11] a model spanning nine imaging modalities, go beyond image-level interpretations to enable detailed intra-image analysis, offering robust capabilities for tasks like tissue segmentation and image registration. By leveraging architectures and training methodologies adapted from general computer vision, vision FMs have proven remarkably effective in addressing the unique challenges of medical image analysis.

Beyond vision-only innovations, multimodal AI approaches have further advanced medical image analysis by integrating information from multiple data modalities, such as clinical records, omics data, and biomedical images (Supplementary Text 2 online). These data inherently involve complementary information to biomedical images, so combining these can provide deeper insights into both disease mechanisms and patient care. For instance, BiomedGPT,[12] the first open-source biomedical multimodal FM, was pretrained on 2D/3D biomedical images and texts, achieving state-of-the-art results in tasks such as visual question answering and report generation. Similarly, PathChat,[13] a vision-language FM fine-tuned on over 4 million visual-language instructions, outperformed GPT-4 in pathology-specific diagnostic tasks and generated accurate, clinically relevant responses preferred by pathologists. By synthesizing information across diverse modalities, multimodal AI offers a more comprehensive perspective on complex medical problems, advancing both scientific discovery and healthcare delivery.

Despite these successes, medical image analysis faces a major limitation: the scarcity of large-scale datasets compared to fields like general computer vision or natural language processing. While models like Google Gemini are trained on billions of image-text pairs, medical datasets are significantly smaller, constraining the scale of training and limiting AI performance. This challenge is particularly acute in multimodal learning, where missing data modalities are common in biomedical datasets. Synthetic data offers a promising solution to address these limitations.[14] Generative models, such as GANs and diffusion models, can create realistic synthetic images to augment datasets, impute missing modalities, and enhance model performance across tasks. (Supplementary Text 3 online) However, the use of synthetic data still raises some concerns. Its quality must be rigorously validated, as low-quality synthetic data can hinder, rather than improve, AI models in biomedicine. Ethical and regulatory safeguards are essential to ensure responsible use. Moreover, excessive reliance on synthetic data could diminish efforts to collect high-quality real-world data, potentially compromising the reliability and applicability of AI models in real-world scenarios.

AI pitfalls in biomedical imaging. The rapid progress in computational biomedical imaging largely owes to the progress in deep learning-based natural image/language processing. However, directly borrowing ideas from these fields also causes problems, as the characteristics of natural images and biomedical images can be fundamentally different. As a result, many aspects such as evaluation metrics, algorithmic development, and practical deployments should be carefully tailored to adapt the biomedical applications. However, these aspects have not attracted enough attention from the community, leading to somewhat reporting biases that may bring over-optimism in machine learning for computational biomedical imaging. These issues often relate to unrealistic performance claims, misuse of advanced AI terminology, and inadequate safety considerations. Below, we outline the three most significant examples of the pitfalls we have observed.

(1) Inadequate evaluation: average performance vs. rare events

A pervasive concern is the reliance on inadequate evaluation metrics. Currently, most evaluation metrics in biomedical imaging are directly borrowed from computer vision research, however, these typical metrics (e.g., SSIM: structural similarity index measure, PSNR: peak signal-to-noise ratio, LPIPS: learned perceptual image patch similarity) mainly focus on evaluating a statistical average performance. They compress reconstruction quality in large amounts of test data into a single number. Such metrics mask the ability—or inability—of models to handle rare but clinically significant events. For instance, if a small fraction of the population presents with atypical anatomical variations or rare pathologies, a model can score highly overall yet fail these minority cases. This scenario is especially grave in oncology or neurology, where early detection of a small tumor or lesion can dramatically alter patient outcomes. Generative models, when employed as priors in medical imaging, are especially susceptible to this issue, known as mode collapse—a phenomenon where the model only learns the most common patterns in the training data while ignoring less frequent ones. This shortfall can lead to reconstructions or analyses that hallucinate missing details or segment biological tissues in statistically plausible ways, inadvertently erasing or neglecting unique anomalies—precisely the features clinicians and researchers need to observe. A systematic evaluation regime must, therefore, measure performance not just in typical cases but also in “long-tail” or outlier scenarios, and incorporate robust uncertainty quantification to flag areas of low model confidence.

(2) Overhyping the “foundation model” concept

While FMs are undeniably powerful, their potential hinges on access to vast datasets and substantial computational resources. True FMs, like GPT-4 and contrastive language–image pre-training (CLIP), are trained on billions of diverse data points, enabling broad generalization across tasks. However, the term “foundation model” is often misapplied to architectures trained on relatively small, domain-specific datasets. While these models may perform well within their specific contexts, labeling them as FMs creates confusion and raises unrealistic expectations, implying they can generalize across diverse tasks without significant retraining or adaptation—an assumption that rarely holds. Models trained on limited data typically lack the breadth and depth required for true generalization, restricting their applicability beyond their original scope. This issue is particularly pronounced in biomedical imaging, where data scarcity is a persistent challenge. Unlike general-domain FMs trained on massive datasets, biomedical models often lack the diversity needed for wide-ranging applicability, leading to limitations in flexibility and scalability. Mislabeling such models as FMs risks undermining trust and setting unrealistic expectations among researchers, practitioners, and end users.

To address this, the research community must establish clearer criteria for defining FMs, emphasizing the scale of training data, computational resources, and generalization capabilities required to earn this designation. Researchers should also critically assess whether an FM approach is suitable for their specific use case. In many biomedical imaging scenarios, domain-specific models tailored to specific tasks may be more practical and effective.[15] For example, architectures designed explicitly for medical imaging or microscopy, integrating domain knowledge such as anatomical structures or cellular patterns, often deliver superior results. Hybrid approaches that leverage techniques like attention mechanisms, rule-based systems, or physics-informed constraints, further enhance performance by embedding domain expertise. By promoting a deeper understanding of FMs and their limitations, the biomedical imaging field can better align expectations with reality, ensuring that model development and deployment remain both responsible and impactful.[16]

(3) Lack of post-deployment monitoring

AI models, once trained and deployed, are vulnerable to performance degradation over time, particularly in dynamic, real-world environments.[17] In biomedical imaging, changes in data distribution—such as shifts in disease prevalence, advancements in imaging technologies, or evolving clinical practices—can significantly impact model accuracy. Similarly, in biological imaging, variations in experimental protocols, staining techniques, or the introduction of new imaging modalities can alter data characteristics. Without timely updates, these shifts can result in diminished model performance, unreliable analyses, and potential misuse in clinical and research settings.

Robust post-deployment monitoring is critical for ensuring model reliability and continued relevance. Continuous monitoring enables early detection of dataset shifts and allows for proactive measures to mitigate their effects. Regularly updating training data and applying continual learning techniques to retrain models ensures they remain responsive to emerging trends and maintain accuracy over time (Supplementary Text 4 online). In regulated environments like healthcare, updating deployed models can be challenging due to strict approval processes and institutional policies. To address this, clear protocols for model evaluation and retraining should be established during the design phase, ensuring compliance while safeguarding performance. The involvement of domain experts—such as clinicians, radiologists, or biologists—is equally critical. Their insights can identify edge cases, uncover areas where models struggle, and provide targeted feedback for improvements, particularly in complex biomedical imaging tasks. This collaborative approach not only enhances model performance but also strengthens real-world applicability.

In summary, this paper delves into the transformative impact of AI on computational biomedical imaging, emphasizing groundbreaking innovations in image reconstruction, analysis, and integration across modalities. These advancements have opened new frontiers in visualizing complex biological and physiological phenomena, enabling unprecedented diagnostic capabilities and facilitating the automation of critical tasks. While we remain optimistic about AI's potential to revolutionize biomedical research and healthcare delivery, we also highlight the risks associated with its misapplication. Striking a careful balance between these opportunities and risks is essential to fully unlock AI's potential while ensuring its reliability, safety, and fairness in real-world biomedical applications.

CRediT authorship contribution statement

He Sun: Writing – review & editing, Writing – original draft, Conceptualization. **Jinzhuo Wang:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2022YFC3401100, 2024YFC3406402, and 2024YFF0507400), the National Natural Science Foundation of China (62371007 and 6220071694), and the Beijing Natural Science Foundation (Z240010).

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.medp.2025.100081.

References

1. Xia W, Cong W, Wang G. Patch-based denoising diffusion probabilistic model for sparse-view CT reconstruction. arXiv:221110388. 2022.
2. Jalal A, Arvinte M, Daras G, Price E, Dimakis AG, Tamir J. Robust compressed sensing MRI with deep generative priors. *Adv Neural Inf Process Syst.* 2021;34:14938–14954.
3. Liu R, Sun Y, Zhu J, Tian L, Kamilov US. Recovery of continuous 3D refractive index maps from discrete intensity-only measurements using neural fields. *Nat Mach Intell.* 2022;4(9):781–791.
4. Cao R, Divekar NS, Nuñez JK, Upadhyayula S, Waller L. Neural space–time model for dynamic multi-shot imaging. *Nat Methods.* 2024;21(12):2336–2341.
5. Reed AW, Kim H, Anirudh R, et al. Dynamic CT reconstruction from limited views with implicit neural representations and parametric motion fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. pp. 2258–2268.
6. Ren J, Li J, Liu C, Chen S, Liang L, Liu Y. Deep learning with physics-embedded neural network for full waveform ultrasonic brain imaging. *IEEE Trans Med Imaging.* 2024;43(6):2332–2346.
7. Dove C, Boondicharern J, Waller L. Physics-guided training of neural electromagnetic wave simulators with time-reversal consistency. NeurIPS 2023 Workshop on Deep Learning and Inverse Problems. 2023.
8. Wang X, Zhao J, Marostica E, et al. A pathology foundation model for cancer diagnosis and prognosis prediction. *Nature.* 2024;634(8035):970–978.
9. Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images. *Nature.* 2023;622(7981):156–163.
10. Sun Y, Wang L, Li G, Lin W, Wang L. A foundation model for enhancing magnetic resonance images and downstream segmentation, registration and diagnostic tasks. *Nat Biomed Eng.* 2024.
11. Zhao T, Gu Y, Yang J, et al. A foundation model for joint segmentation, detection and recognition of biomedical objects across nine modalities. *Nat Methods.* 2025;22(1):166–176.
12. Zhang K, Zhou R, Adhikarla E, et al. A generalist vision-language foundation model for diverse biomedical tasks. *Nat Med.* 2024;30(11):3129–3141.
13. Lu MY, Chen B, Williamson DF, et al. A multimodal generative AI copilot for human pathology. *Nature.* 2024;634(8033):466–473.
14. Wang J, Wang K, Yu Y, et al. Self-improving generative foundation model for synthetic medical image generation and clinical applications. *Nat Med.* 2025;31(2):609–617.
15. Wainberg M, Merico D, Delong A, Frey BJ. Deep learning in biomedicine. *Nat Biotechnol.* 2018;36(9):829–838.
16. Wang H, Fu T, Du Y, et al. Scientific discovery in the age of artificial intelligence. *Nature.* 2023;620(7972):47–60.
17. Ackerman S, Raz O, Zalmanovici M, Zlotnick A. Automatically detecting data drift in machine learning classifiers. arXiv:211105672. 2021.