OPEN ACCESS

Artificial Intelligence EDITORIAL

Bridging physics and intelligence: The evolutionary landscape of PET image reconstruction

Heng Jiang,[1,2] Wanwan Guo,[1,2,3] Qiyang Zhang,[1,2] Yuxi Jin,[1,2] Haizhou Liu,[1,2] Zixiang Chen,[1,2] Hairong Zheng,[1,2] Dong Liang,[1,2] and Zhanli Hu[1,2,*]

[1]Research Center for Medical AI, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[2]Key Laboratory of Biomedical Imaging Science and System, Chinese Academy of Sciences, State Key Laboratory of Biomedical Imaging Science and System, Shenzhen 518055, China

[3]Lab of Molecular Imaging and Medical Intelligence, Department of Radiology, Longgang Central Hospital of Shenzhen, Shenzhen 518116, China

*Correspondence: zl.hu@siat.ac.cn

Received: January 23, 2026; Accepted: June 2, 2026; Published Online: June 4, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100052

Citation: Jiang H., Guo W., Zhang Q., et al. (2026). Bridging physics and intelligence: The evolutionary landscape of PET image reconstruction. *The Innovation Informatics* **2**:100052.

Positron emission tomography (PET) is a cornerstone of molecular imaging, critical for oncology and neurology. Because PET scanners record raw coincidence events rather than direct images, the chosen reconstruction algorithm fundamentally determines final image quality and diagnostic accuracy. The field has successfully evolved from analytical to iterative methods and deep learning is currently driving the next major transition. While data-driven approaches offer significant potential, they also present challenges regarding reliability, generalization across scanner platforms, and adherence to physical imaging principles. Consequently, the field faces a fundamental challenge: bridging the analytical rigor of physics-based modeling with the versatility of artificial intelligence. This editorial examines the evolving landscape of PET image reconstruction, weighing the inherent trade-offs between established physics-driven algorithms and emerging deep learning techniques. To chart a course toward reliable clinical intelligence, we discuss key future directions. Specifically, we highlight the integration of physical constraints into generative models, the potential of scalable foundation models to harmonize multi-center data, and the critical need to transition toward task-oriented evaluation frameworks. By synthesizing these perspectives, we aim to ensure that increasing algorithmic complexity genuinely translates into robust, real-world clinical utility (Figure 1).

CURRENT METHODS & CHALLENGES: THE TUG-OF-WAR

Bottlenecks of traditional iterative methods and physics-based corrections

Currently, statistical iterative reconstruction remains the dominant paradigm in clinical PET, with ordered subsets expectation maximization (OSEM) serving as the quintessential method. By explicitly modeling the statistical properties of radioactive decay (specifically the Poisson process), the OSEM algorithm ensures robust stability and reliability. Nevertheless, this methodology has long been constrained by the fundamental trade-off between spatial resolution and image noise. While fine structural details can be recovered through an increased number of iterations, noise is concomitantly amplified. Although post-reconstruction smoothing filters are frequently employed for noise suppression, high-frequency information is inevitably sacrificed. This results in image blurring and consequently impairs the detectability of small lesions. To mitigate this inherent conflict, physics-based corrections, most notably time-of-flight (TOF) and point spread function (PSF) modeling, have been introduced. The TOF exploits the arrival-time difference of annihilation photons to constrain coincidence events to a restricted segment along the line of response (LOR), thereby yielding a substantial improvement in the signal-to-noise ratio (SNR). Complementarily, PSF modeling enhances spatial resolution by compensating for detector scattering and system blurring effects. Despite the significant enhancements in image quality afforded by these corrections, the reconstruction process remains fundamentally rooted in conventional mathematical frameworks. Consequently, under challenging conditions such as ultra-low-dose acquisition, reconstruction performance remains markedly compromised.

Deep learning intervention and multi-modal fusion

In recent years, PET image reconstruction has witnessed rapid advancements driven by deep learning, with architectures like convolutional neural networks (CNNs), Transformers, and Mamba demonstrating significant superiority in noise reduction and the recovery of fine details.[1] Concurrently, multi-modal fusion reconstruction (e.g., PET/MR, PET/CT) utilizes high-resolution anatomical priors to successfully preserve boundary sharpness. However, pure data-driven approaches face critical bottlenecks regarding physical consistency and clinical safety. To elucidate this issue, we categorize existing deep learning interventions, with a detailed comparison presented in Table S1.

Importantly, the risk of structural hallucinations, such as the fabrication of non-existent lesions or spurious textures, is not uniform. Instead, it must be carefully stratified by the specific deep learning pathway employed.

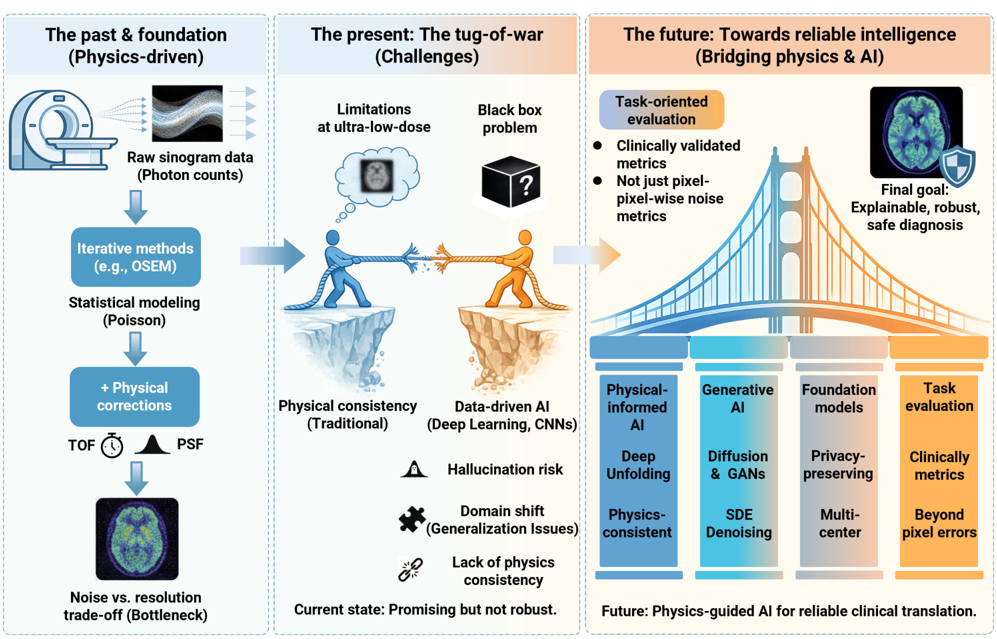

Figure 1. Bridging physics and intelligence: The trajectory and future blueprint of PET image reconstruction This figure illustrates the evolution of PET reconstruction. It contrasts the physical foundations of traditional methods (left) with the generalization bottlenecks of pure deep learning (middle), ultimately highlighting physics-informed AI as the future paradigm for robust clinical imaging (right).

Post-processing networks operate entirely within the image domain and carry a high inherent clinical risk. By disregarding raw measurement data, these models are prone to over-smoothing subtle lesions or generating false structures driven by training biases. Direct mapping approaches bypass imaging physics by translating sinograms directly into images, posing a moderate-to-high risk. Their "black-box" nature neglects fundamental Poisson noise mechanisms, making them highly susceptible to cross-scanner domain shifts. Deep unfolding architectures, however, offer a more rigorous and controlled risk profile. By explicitly incorporating forward and backward physical projection models at each unrolled iteration, they mathematically bound the solution space. This physical integration ensures strict data fidelity and effectively suppresses unconstrained hallucinations. Without this physical anchoring, hallucination risks escalate significantly during extreme photon starvation. This vulnerability is particularly evident in ultra-low-dose scans utilizing less than 5% of standard clinical dose or during ultra-fast dynamic imaging. In these severely degraded scenarios, the extreme sparsity of reliable projection data forces unconstrained networks to infer critical structures based solely on learned priors, ultimately introducing severe diagnostic risks into clinical applications.

FUTURE DIRECTIONS: TOWARDS RELIABLE INTELLIGENCE

Physics-driven deep learning

To overcome the inherent black-box limitations of pure deep learning, recent research has increasingly focused on integrating physical models with neural networks. This model-based paradigm bridges the gap between the interpretability of traditional imaging physics and the expressive power of data-driven representations. Among these techniques, deep unfolding has emerged as a highly effective strategy.[2] By explicitly mapping each iterative update of a conventional optimization algorithm to a distinct network layer, deep unfolding creates an architecture that directly reflects the physical reconstruction workflow, ensuring that essential constraints are preserved. A fundamental component of this approach is the incorporation of the exact system matrix, which comprehensively models physical factors such as TOF and PSF to mitigate reconstruction artifacts. Furthermore, strictly enforcing data consistency via the Poisson likelihood function provides a critical safeguard against structural hallucinations. Within this unfolded framework, parameters such as regularization coefficients are learned entirely from data, eliminating the need for empirical manual tuning. Consequently, this methodology successfully synergizes the analytical rigor of physical modeling with the flexible adaptability of advanced neural networks.

Exploration of generative models and SDEs

Recent advancements in generative adversarial networks (GANs) and diffusion models have shown significant potential for detailed medical image synthesis. Emerging frameworks like optimal transport flow matching and score-based models are increasingly being adapted for PET imaging. When applied to low-dose reconstruction, these architectures can approximate standard-dose image quality while reducing patient radiation exposure. Although they excel at texture completion, models driven purely by data introduce substantial risks to structural authenticity. Under severe Poisson noise conditions, relying solely on learned priors often leads to anatomical fabrications. This limitation necessitates the mathematical integration of physical constraints directly into the generative framework.[3] Moving beyond conventional synthesis tasks, modern physics-informed networks explicitly embed the Poisson log-likelihood gradient within the reverse stochastic differential equation (SDE) or Langevin dynamics. This physical anchoring ties the reverse sampling steps strictly to the measured projection data. Consequently, the resulting high-frequency details are driven by authentic underlying physical signals rather than unconstrained algorithmic hallucinations.

Federated learning and PET foundation models

For prospective clinical deployment, scaling PET reconstruction algorithms for multicenter clinical deployment requires overcoming privacy and data heterogeneity challenges. Federated learning (FL)ensures privacy by aggregating model updates instead of raw images,[4] thereby enabling secure distributed networks guided by physical principles. While this framework drives the development of PET foundation models (FMs),[5] overcoming profound data variations across diverse scanners, imaging protocols, and patient demographics requires robust harmonization strategies. To build truly robust anatomical representations, foundational architectures must incorporate data harmonization strategies to prevent bias from site-specific confounders. Furthermore, processing massive volumetric spatial tensors necessitates scalable designs, such as sparse architectures or mixture of experts, to resolve memory bottlenecks and ensure computational efficiency.

To embed physical priors into FMs, we advocate for a decoupled, two-stage approach. Initially, pre-training should incorporate universal physical principles (e.g., Poisson statistics) to build broadly physics-compliant representations. Subsequently, downstream fine-tuning must enforce scanner-specific constraints, like precise system matrices. This strategy balances the broad generalizability of FMs with the strict data fidelity required for clinical deployment.

Task-oriented evaluation frameworks

As these advanced algorithms evolve, traditional evaluation metrics like PSNR, which primarily quantify global pixel-wise differences, are insufficient for capturing clinical utility. An adaptive and rigorous evaluation framework is therefore essential, calling for advanced methodologies that extend beyond conventional image quality assessment. For example, the task-based channelized hotelling observers (CHO) can model human visual perception for tumor detection tasks. Furthermore, region-of-interest (ROI)-specific quantitative metrics, such as the contrast recovery coefficient (CRC) and background variance, help ensure diagnostic reliability in clinical settings. Shifting towards these task-oriented metrics ensures that algorithmic improvements genuinely translate into enhanced diagnostic accuracy.

CONCLUSION

PET image reconstruction is transitioning toward physically informed artificial intelligence. While deep learning enhances image quality, clinical deployment strictly demands explicit physical constraints to ensure interpretability and diagnostic safety. Furthermore, escalating model complexity requires hardware specific optimization to maintain standard workflows. Although advanced processors excel in training, field programmable gate arrays (FPGAs) offer superior low latency and power efficiency for edge inference directly within scanner gantries. Ultimately, translating these innovations into routine practice necessitates robust interdisciplinary collaboration among physicists, computer scientists, and clinicians to deliver mathematically precise, computationally efficient, and reliable technologies for patient care.

REFERENCES

1. Bousse A., Kandarpa V.S.S., Shi K., et al. (2024). A review on low-dose emission tomography post-reconstruction denoising with neural network approaches. *IEEE Trans. Radiat. Plasma Med. Sci.* 8:333–347. DOI:10.1109/TRPMS.2023.3349194
2. Jiang H., Zhang Q., Hu Y., et al. (2025). Memory-enhanced and multi-domain learning-based deep unrolling network for medical image reconstruction. *Phys. Med. Biol.* 70: 175008. DOI: 10.1088/1361-6560/adf939
3. Hein D., Bozorgpour A., Merhof D., et al. (2025). Physics-inspired generative models in medical imaging. *Annu. Rev. Biomed. Eng.* 27:499–525. DOI:10.1146/annurev-bioeng-102723-013922
4. Sun M., Yang Z., Huang Y., et al. (2025). Federated learning for large models in medical imaging: A comprehensive review. *arXiv* preprint. DOI: arXiv:2508.20414
5. Huang B., Chen K., Li B., et al. (2025). ALL-PET: A low-resource and low-shot PET foundation model in projection domain. *arXiv* preprint. DOI: arXiv: 2509.09130

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2024YFE0202400), the National Natural Science Foundation of China (82372038), the China Postdoctoral Science Foundation (2025M782965), the Natural Science Foundation of Guangdong Province in China (2023B1515120007, 2024B1515040018, 2026A1515011709), the Shenzhen Science and Technology Program of China (KJZD20240903101307010). Shenzhen-Hong Kong-Macao Science and Technology Plan Project (Category C Project) under Shenzhen Municipal Science and Technology Innovation Commission (SGDX20230821092359002). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

Hairong Zheng is the Editor-in-Chief of The Innovation Informatics and was blinded from reviewing or making final decisions on the manuscript. Peer review was handled independently of this member and their research group. The other authors declare no conflicts of interest.

SUPPLEMENTAL INFORMATION

It can be found online at https://doi.org/10.59717/j.xinn-inform.2026.100052