

Mamba-Enhanced 3D Medical Image Reconstruction for Low-Dose Computed Tomography

Rajesh Sorensen

Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.
rajeshsorensen1445@buffalo.edu

Abstract

Low-dose computed tomography has become a central strategy for reducing patient radiation exposure, but it introduces substantial reconstruction challenges including amplified noise, structured artifacts, and degraded tissue contrast. This paper examines the systems-level integration of Mamba-based state space models into three-dimensional medical image reconstruction pipelines for low-dose computed tomography. Rather than treating reconstruction as an isolated algorithmic task, the discussion positions Mamba-enhanced reconstruction within a broader socio-technical infrastructure that includes data governance, computational deployment, regulatory oversight, clinical accountability, and environmental sustainability. The analysis explores how selective state space mechanisms can support efficient long-range volumetric modeling while addressing the architectural trade-offs associated with convolutional neural networks and transformer-based models. The paper further considers the operational requirements of hospital-scale deployment, including interoperability with picture archiving and communication systems, model drift monitoring, failover planning, and postmarket surveillance. Emphasis is placed on fairness across patient populations, robustness under scanner heterogeneity, and the need for transparent validation frameworks that align with evolving regulatory expectations. The discussion argues that the value of Mamba-enhanced reconstruction depends not only on image quality metrics but also on the broader governance structures that ensure safe, equitable, and resilient clinical use. Future directions include federated learning, modular regulatory certification, and energy-aware model design. The paper contributes a cross-disciplinary perspective on how next-generation sequence modeling can be responsibly translated into low-dose computed tomography practice.

Keywords

low-dose computed tomography, Mamba, state space models, medical image reconstruction, regulatory governance, health systems, fairness, sustainability.

1. Introduction

Computed tomography has transformed diagnostic medicine by providing high-resolution volumetric representations of human anatomy, yet the associated radiation burden remains an enduring public health concern. The evolution of reconstruction methods has moved from filtered back projection toward iterative and model-based strategies, each offering distinct trade-offs between image quality, computational expense, and robustness under low photon counts [1]. Population-level evidence has drawn attention to the growing contribution of computed tomography to cumulative medical radiation exposure, making dose reduction an urgent systems priority rather than a purely technical preference [2]. Low-dose protocols reduce radiation but introduce noise statistics that challenge conventional reconstruction

assumptions. Consequently, the field has shifted toward data-driven methods that can learn complex mappings from degraded measurements to higher-quality image estimates.

Deep learning has offered substantial improvements in low-dose computed tomography reconstruction by learning hierarchical representations that suppress noise while preserving diagnostically relevant edges and textures. Residual encoder-decoder architectures have demonstrated that convolutional neural networks can recover structural detail from low-dose acquisitions, but their ability to capture long-range dependencies across entire three-dimensional volumes remains constrained [3]. State space models, and in particular the Mamba architecture, have recently emerged as an alternative sequence modeling paradigm that combines selective state transitions with linear-time computation [4]. Theoretical work has further connected structured state space models with transformer mechanisms, clarifying how efficient global context propagation can be achieved without quadratic attention costs [5]. These developments hold considerable promise for volumetric medical imaging, where dependencies span many axial slices and local context alone may be insufficient for coherent reconstruction.

The application of Mamba to visual data has progressed through bidirectional state space formulations that adapt selective scanning to image patches and feature maps [6]. In medical imaging, self-configuring segmentation frameworks have shown the importance of automated architectural adaptation to heterogeneous datasets, while the U-Net family has provided a robust baseline for biomedical image analysis [7, 8]. These prior systems underscore that performance in clinical translation depends on more than raw model capacity; it depends on how well architectures align with data characteristics, compute budgets, and validation workflows. Deep learning in general has reshaped medical image interpretation, but its integration into healthcare systems introduces ethical responsibilities that extend beyond algorithm design [9, 10, 11]. This paper analyzes Mamba-enhanced three-dimensional reconstruction for low-dose computed tomography from a systems perspective, integrating technical architecture with governance, deployment, fairness, and sustainability concerns.

2. Low-Dose Computed Tomography Reconstruction as a Systems Challenge

Low-dose computed tomography reconstruction is best understood as a systems challenge because decisions made at the scanner acquisition stage propagate through reconstruction, postprocessing, radiologist interpretation, and clinical decision-making. Reducing tube current or exposure time lowers patient dose but increases quantum noise and may introduce electronic noise artifacts that degrade the reliability of attenuation measurements. In helical and multislice acquisitions, the interplay between focal spot dynamics, detector response, and patient motion creates a coupled reconstruction problem that cannot be fully separated into independent two-dimensional slices. Statistical iterative reconstruction methods improved upon filtered back projection by modeling the three-dimensional nature of the acquisition process, but they impose significant computational burdens and require careful regularization [12]. These methods remain important reference points for understanding the trade-offs that any new reconstruction architecture must address.

From an operational perspective, a reconstruction system must produce images rapidly enough to fit within clinical workflows while maintaining consistent quality across scanners, protocols, and patient body habitus. The computational constraints of model-based iterative reconstruction led many sites to accept longer reconstruction times or to restrict use to specific clinical indications. Deep learning approaches promise faster inference, but they introduce new dependencies on training data, hardware acceleration, and model versioning. A

system-level analysis therefore must consider not only peak signal-to-noise ratio or structural similarity but also failure modes under out-of-distribution anatomy, rare pathological presentations, and protocol drift. The need for interpretable quality assurance becomes particularly acute when reconstruction is performed by a learned model whose internal representations are not directly visible to the clinical physicist or radiologist.

The volumetric character of modern computed tomography further complicates the reconstruction problem. A three-dimensional model must preserve continuity across slices, avoid introducing artificial boundaries, and respect anatomical variation along the craniocaudal axis. Processing an entire volume as a sequence of axial slices can be efficient but sacrifices coherence, while fully three-dimensional convolutional processing can exceed memory limits and reduce throughput. The design of Mamba-enhanced reconstruction systems must therefore negotiate a middle path in which long-range context is captured efficiently without incurring the quadratic costs associated with dense attention or the memory costs associated with large three-dimensional kernels. This structural trade-off is central to the architectural discussion that follows, but its implications reach into deployment, sustainability, and clinical validation.

3. State Space Models and Mamba in Medical Imaging

State space models describe how hidden states evolve over time in response to inputs, and when specialized to discrete sequences they can propagate information across long contexts with recurrent efficiency. The Mamba architecture introduced a selective mechanism that allows state transition parameters to depend on the input, thereby maintaining linear-time sequence processing while enabling context-sensitive filtering [4]. This property is particularly relevant to medical image reconstruction, where relevant contextual information may be localized in one region but irrelevant or misleading in another. The connection between structured state space duality and transformer-style sequence mixing has clarified that selective state space models can serve as efficient alternatives to attention for many tasks [5]. In volumetric computed tomography, such efficiency may permit whole-volume processing without downsampling to a degree that destroys small lesions or subtle textural differences.

Vision Mamba adapted selective scanning to image data by processing flattened patches bidirectionally, enabling spatial context accumulation across large feature maps [6]. Medical imaging applications inherit this property but face additional challenges arising from anisotropic voxel sizes, varying slice thicknesses, and the need to preserve anatomical consistency. Self-configuring frameworks such as nnU-Net have demonstrated that biomedical image analysis can benefit from automated configuration rather than manual architecture engineering [7]. Similarly, U-Net has remained a durable backbone because of its ability to combine multi-scale encoder features with decoder upsampling [8]. Integrating Mamba into such frameworks offers a way to replace or supplement bottleneck attention mechanisms with state space layers that capture global context without exceeding memory budgets.

The medical imaging community has begun to explore cross-fusion formulations of Mamba for tasks such as multi-view classification, where different imaging perspectives must be integrated into a coherent representation [13]. These developments suggest that Mamba-based modules can be embedded within larger hybrid architectures that retain convolutional texture encoding while adding selective sequence reasoning over spatially distant regions [14]. For low-dose computed tomography reconstruction, the implication is not that Mamba replaces all

components but that it can serve as a global coordination mechanism. Careful architectural integration is required to ensure that the selective state space pathway does not dominate the local convolutional pathway in ways that suppress fine anatomical detail or introduce hallucinated structures.

4. Architectural Integration of Mamba-Enhanced Volumetric Reconstruction

A Mamba-enhanced reconstruction architecture for low-dose computed tomography can be conceptualized as a hybrid system in which convolutional encoders extract local features from sinogram or image domain representations and state space layers model dependencies across slices and feature channels. The local pathway preserves the strong inductive biases of convolutional networks, which are valuable for detecting edges, textures, and small structures. The global pathway enables the network to reason about anatomical continuity across the entire volume, which is important when noise levels vary from slice to slice or when a lesion spans multiple slices and would be difficult to identify from local context alone. The architectural trade-off involves deciding where to place selective state space layers, how many scanning directions to use, and how to aggregate states across the axial, coronal, and sagittal dimensions.

One critical consideration is memory efficiency. Three-dimensional medical volumes can contain hundreds of slices, and processing all slices simultaneously with dense attention is often infeasible on hospital-grade graphics processing units. Mamba-based sequence modeling offers a potential solution because its complexity scales linearly with sequence length, allowing the model to consume an entire slab or volume without excessive memory growth. However, the selection mechanism must be configured carefully to avoid numerical instability and to ensure that the state transformation remains well conditioned across different scanner calibrations. Architectural decisions therefore include state dimensionality, expansion factors, normalization strategies, and the use of residual connections around the state space blocks. These choices influence not only reconstruction accuracy but also the ease with which the model can be compressed, updated, and deployed across heterogeneous hardware.

The training data used to optimize such models must reflect broad patient populations, anatomical variation, and acquisition protocols. Systematic reviews of algorithmic bias have shown that representation failures in training data can produce disparate performance across demographic groups, leading to downstream harm even when aggregate metrics appear acceptable [13]. In multi-view volumetric learning, cross-fusion approaches based on Mamba have been proposed to integrate complementary views, illustrating how state space models can support medical image classification and feature alignment [14]. For reconstruction, similar fusion principles can be adapted to combine sinogram-domain information with image-domain priors, thereby exploiting the complementary strengths of raw detector data and reconstructed images. This systems-level fusion is important because low-dose artifacts may be more interpretable in one domain than in another, and robust reconstruction requires the model to reconcile conflicting evidence across representations.

Another architectural concern is interpretability. Reconstruction models are often treated as black boxes, but clinical users and regulatory reviewers increasingly expect explanations for automated image processing decisions. Although state space models are not inherently interpretable, their sequential state updates can be instrumented to expose which spatial regions and feature channels influence the reconstructed output. Such instrumentation supports post hoc explanation methods and can help identify failure modes before deployment.

Without this transparency, a Mamba-enhanced system may be difficult to validate in a regulated environment, regardless of its quantitative performance. Thus, architectural integration must be accompanied by validation interfaces that allow medical physicists and radiologists to inspect reconstruction behavior under controlled perturbations.

5. Data Governance and Regulatory Infrastructure

Medical imaging data possess a particularly sensitive combination of personal health information and rich phenotypic detail. The governance of datasets used to train Mamba-enhanced reconstruction models must address consent, de-identification, data provenance, and the potential for leakage across institutions. Federated approaches can reduce the need to centralize raw images, but they introduce coordination costs and require robust aggregation protocols. Regulatory bodies have begun to articulate expectations for artificial intelligence and machine learning software as a medical device, emphasizing lifecycle management and the need for planned updates rather than static approval [15]. A reconstruction model that changes over time, whether through fine-tuning or continuous learning, must be monitored for drift and evaluated against predefined safety thresholds. This regulatory perspective aligns with the broader systems argument that algorithmic performance cannot be separated from the institutional processes that govern data access and model maintenance.

Bias in training data can have direct clinical consequences in low-dose computed tomography because attenuation characteristics vary with patient body habitus, age, sex, and the presence of implants or pathology. Evidence from population health algorithms has shown that unequal access to care can become encoded in learned models, and that auditing for fairness requires more than aggregate calibration checks [16]. In reconstruction, a model that performs well on average but fails for patients with larger body habitus could exacerbate existing disparities in diagnostic accuracy. Data governance policies must therefore require stratified evaluation across clinically meaningful subgroups and mandate documentation of dataset composition. These policies should also define how institutions can contribute data to multi-site training collaborations without violating privacy regulations or creating unmanaged secondary use risks.

The regulatory infrastructure for reconstruction software is evolving alongside technical capabilities. In the United States, the Food and Drug Administration has proposed a framework for adaptive artificial intelligence systems that includes predicate device comparison, change control protocols, and real-world performance monitoring [15]. Similar requirements are emerging in other jurisdictions, and alignment across regulatory regimes is necessary for manufacturers that deploy globally. A Mamba-enhanced reconstruction system may qualify as a software as a medical device component within a larger imaging chain, which means that its outputs must be scrutinized in combination with scanner hardware, dose modulation, and display systems. The regulatory submission should therefore describe not only the architecture but also the operating envelope, including supported scanners, protocol ranges, and patient populations. Such documentation helps ensure that the technology remains safe when used outside the conditions explored during development.

6. Deployment, Scalability, and Operational Resilience

Deployment of Mamba-enhanced reconstruction in hospital settings requires integration with picture archiving and communication systems, radiology information systems, and downstream clinical decision support applications. The reconstruction engine must receive standardized DICOM inputs, produce DICOM-compliant outputs, and preserve metadata that

are essential for dose reporting and longitudinal follow-up. Compatibility with existing workflows is often more important to clinical adoption than marginal improvements in image quality metrics. A system that introduces manual steps or incompatible formats will face resistance even if its reconstructions are technically superior. Therefore, deployment planning should treat interoperability as a first-class design objective rather than an afterthought.

Scalability is shaped by the computational profile of Mamba-enhanced architectures. The linear-time sequence processing of state space models can reduce peak memory consumption relative to transformer-based alternatives, enabling larger volumes to be processed on existing hardware. However, server-side deployment still requires careful batching, mixed-precision inference, and failover mechanisms. Hospitals may run reconstruction on centralized workstations, edge devices adjacent to scanners, or cloud resources with strict data residency requirements. Each topology introduces different latency, reliability, and security trade-offs. A hybrid deployment model may be appropriate, in which local inference handles routine cases while computationally intensive or uncertain cases are routed to centralized resources with additional capacity. Such routing decisions should be informed by automated confidence estimation and by operational metrics that track queue lengths, hardware utilization, and patient scheduling impact.

Operational resilience also requires attention to model updates and version control. A reconstruction model is not a static artifact but a component of a living system that may undergo retraining, calibration, or adaptation to new scanner models. Each update can alter the appearance of reconstructed images in subtle ways that affect radiologist interpretation. Change management must therefore include side-by-side visual comparisons, objective image quality assessments, and clinician feedback mechanisms before new versions are rolled out. Fallback procedures should allow the system to revert to a previous validated version if anomalies are detected. These practices mirror the broader principles of continuous integration and continuous deployment in software engineering, but they must be tempered by the regulatory and clinical consequences of automated medical image processing.

7. Fairness, Robustness, and Clinical Accountability

Fairness in low-dose computed tomography reconstruction encompasses not only demographic fairness but also fairness across scanner vendors, acquisition protocols, and clinical use cases. A model trained primarily on data from a narrow set of institutions may learn scanner-specific noise textures that do not generalize to settings with different detector technologies or reconstruction kernels. This can create a form of distributional unfairness in which patients at some facilities receive lower-quality reconstructions than patients at better-represented facilities. Bias audits should therefore evaluate performance across site-level covariates in addition to patient-level attributes. The computational efficiency of Mamba may facilitate such audits because it allows multiple inference runs with perturbed inputs to be completed within practical time constraints.

Robustness under adversarial or naturally occurring perturbations is another critical concern. Medical images contain structured noise, metal artifacts, contrast timing variability, and motion, all of which can challenge a learned reconstruction model. Although explanation methods such as gradient-weighted class activation mapping have been developed to identify influential regions in classification networks, their extension to reconstruction is less straightforward because the output is a full image rather than a discrete label [17]. Nonetheless, visualization of state space attention maps can provide a complementary view of how the Mamba module attends to different axial positions and feature channels. Such

visualizations may help radiologists understand whether a reconstruction artifact arises from an input anomaly or from a failure of the global context propagation.

There are also deeper limitations of deep learning that should temper expectations for autonomous reconstruction systems. Critical analyses have pointed out that deep networks can learn spurious correlations and may not encode robust causal structures [18]. In the reconstruction setting, a model might learn to suppress noise by smoothing texture in ways that preserve reference image similarity but obscure the very lesions that clinicians need to detect. Clinical accountability therefore requires prospective evaluation with task-specific endpoints, including radiologist sensitivity and specificity, rather than retrospective image quality metrics alone. A system that improves structural similarity but reduces diagnostic accuracy is not acceptable, and accountability mechanisms must make this distinction explicit.

8. Sustainability and Lifecycle Considerations

The environmental footprint of large-scale medical imaging models is an emerging concern. Training and inference both consume electricity, and hospital systems increasingly face pressure to reduce carbon emissions while maintaining diagnostic quality. The efficiency of Mamba-based state space models is relevant here because selective state space processing can reduce the computational burden of whole-volume reconstruction relative to dense attention or fully three-dimensional convolutional alternatives. However, efficiency gains at inference time may be offset by the energy consumed during extensive hyperparameter search and repeated validation runs. Sustainable development practices should therefore include energy-aware model selection, model pruning, distillation, and the reuse of pretrained components across institutions.

Lifecycle sustainability also involves hardware refresh cycles and the avoidable obsolescence of computational infrastructure. A reconstruction system that requires specialized hardware may create dependencies that shorten the useful life of existing imaging equipment or force premature capital expenditure. Mamba-enhanced models should be designed to degrade gracefully across hardware tiers so that lower-resource settings can still benefit from improved reconstruction without requiring state-of-the-art accelerators. This tiering is especially important for global health equity, because many regions cannot rapidly replace scanners or servers. A modular architecture that can disable certain state space layers or reduce state dimensionality at inference time may offer a practical compromise between quality and resource use.

The governance of sustainability should be integrated into the clinical validation process. Institutional review boards and hospital technology assessment committees increasingly consider not only clinical efficacy but also environmental and organizational impacts when approving new software. Manufacturers can support this by publishing standardized benchmarks that report energy consumption per reconstruction, peak memory use, and expected hardware lifetime under typical workloads. Such transparency enables purchasing decisions that account for long-term operational costs rather than only upfront model accuracy. It also aligns with broader movements in artificial intelligence research that call for reporting efficiency alongside performance [19].

9. Policy Implications and Future Directions

Policy for Mamba-enhanced low-dose computed tomography reconstruction must balance innovation with precaution. The history of medical imaging shows that algorithmic advances can be adopted rapidly when they offer operational benefits, but regulatory and

reimbursement structures often lag behind. High-performance artificial intelligence in medicine has been celebrated for its potential to augment clinicians, yet translation remains uneven because health systems differ in governance maturity, data infrastructure, and staff capacity [20]. A coordination mechanism is needed to align manufacturers, regulators, professional societies, and patient advocacy groups around shared evaluation standards. Such standards should define minimum reporting requirements for low-dose reconstruction models, including dataset composition, subgroup performance, failure modes, and postmarket surveillance plans.

Federated learning and privacy-preserving analytics will likely play a growing role in maintaining reconstruction models across institutions. Instead of pooling raw images, participating sites can share model updates or aggregate statistics under protocols that limit disclosure risk. However, federated systems introduce new vulnerabilities related to data poisoning, contribution imbalance, and the difficulty of ensuring that locally trained components remain compatible. Policy should encourage the development of federated governance agreements that specify data use limitations, audit rights, and dispute resolution mechanisms. These agreements are social infrastructure as much as technical artifacts, and their absence can undermine even the most sophisticated reconstruction architecture.

The future trajectory of Mamba-enhanced reconstruction will also be shaped by the convergence of state space models with other sequence modeling paradigms, including hybrid attention mechanisms and structured recurrence [21]. This convergence may produce architectures that adapt their computation to the complexity of the input, allocating more state capacity to ambiguous regions while conserving resources in straightforward cases. From a systems perspective, such adaptivity is valuable because it allows the same software to serve routine screening and complex emergency cases without requiring separate model versions. However, adaptive computation also complicates verification because the model's behavior may vary across patients in ways that are difficult to specify in advance. Regulatory science must therefore evolve toward methods for certifying adaptive systems, including runtime monitors, confidence calibration, and conditional approval pathways.

10. Conclusion

Mamba-enhanced three-dimensional medical image reconstruction represents a promising integration of efficient sequence modeling into the complex socio-technical environment of low-dose computed tomography. Selective state space mechanisms can address the long-range volumetric dependencies that challenge convolutional and attention-based architectures, while linear-time processing offers a pathway toward practical deployment on hospital infrastructure. However, the clinical value of this technology cannot be established by architectural innovation alone. It depends on the governance structures that ensure representative data, stratified validation, regulatory alignment, operational resilience, and transparent accountability. A Mamba-enhanced reconstruction system must be designed as part of a broader imaging pipeline, with attention to interoperability, hardware variability, energy consumption, and the needs of radiologists who interpret its outputs. The field would benefit from stronger collaboration among algorithm developers, medical physicists, radiologists, bioethicists, and regulators to define what constitutes safe and equitable reconstruction performance. As state space models continue to evolve, the central challenge will be to maintain a systems perspective that treats image quality, fairness, sustainability, and clinical effectiveness as interconnected objectives rather than isolated benchmarks. This

interdisciplinary orientation will determine whether next-generation low-dose reconstruction technologies fulfill their promise in real-world health care settings.

References

1. Willemink, M. J., & Noël, P. B. (2019). The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence. *European Radiology*, 29(5), 2185–2195. <https://doi.org/10.1007/s00330-018-5810-7>
2. Brenner, D. J., & Hall, E. J. (2007). Computed tomography—an increasing source of radiation exposure. *New England Journal of Medicine*, 357(22), 2277–2284. <https://doi.org/10.1056/NEJMr072149>
3. Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J., & Wang, G. (2017). Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Transactions on Medical Imaging*, 36(12), 2524–2535. <https://doi.org/10.1109/TMI.2017.2715284>
4. Gu, A., & Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*.
5. Dao, T., & Gu, A. (2024). Transformers are SSMS: Generalized models and efficient algorithms through structured state space duality. *arXiv preprint arXiv:2405.21060*.
6. Zhu, L., Liao, B., Zhang, Q., Wang, X., Liu, W., & Wang, X. (2024). Vision Mamba: Efficient visual representation learning with bidirectional state space model. *arXiv preprint arXiv:2401.09417*.
7. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2), 203–211. <https://doi.org/10.1038/s41592-020-01008-z>
8. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234–241). Springer.
9. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
10. Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. *New England Journal of Medicine*, 378(11), 981–983. <https://doi.org/10.1056/NEJMp1714229>
11. Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28(1), 31–38. <https://doi.org/10.1038/s41591-021-01614-0>
12. Thibault, J.-B., Sauer, K. D., Bouman, C. A., & Hsieh, J. (2007). A three-dimensional statistical approach to improved image quality for multislice helical CT. *Medical Physics*, 34(11), 4526–4544. <https://doi.org/10.1118/1.2789499>
13. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35. <https://doi.org/10.1145/3457607>
14. Zheng, X., Chen, X., Gong, S., Griffin, X., & Slabaugh, G. (2025, September). Xfmamba: Cross-fusion mamba for multi-view medical image classification. In *International*

Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 672-682). Cham: Springer Nature Switzerland.

15. U.S. Food and Drug Administration. (2021). Artificial intelligence/machine learning software as a medical device: Action plan. U.S. Department of Health and Human Services. <https://www.fda.gov/media/145022/download>
16. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
17. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 618–626). IEEE.
18. Marcus, G. (2018). Deep learning: A critical appraisal. *arXiv preprint arXiv:1801.00631*.
19. Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63. <https://doi.org/10.1145/3381831>
20. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
21. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM. <https://doi.org/10.1145/3442188.3445922>