Graph-Mamba Networks for Integrating Anatomical Relationships in Medical Image Analysis

Cody Greene
Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA.
cody1975@missouri.edu

Brendan Makinen
School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
brendanmail@uc.edu

Sagar Tandon
Department of Computer Science, University of Central Florida, Orlando, FL, USA.
tandonsagar@ucf.edu

Abstract

Medical image analysis increasingly requires computational systems that can reason over spatially distributed anatomical structures rather than isolated pixels or patches. Graph-Mambanetworks integrate the relational expressiveness of graph neural networks with the linear-time sequence modeling capabilities of state-space models. This paper presents a system-level examination of Graph-Mamba architectures for medical image analysis, with particular attention to how anatomical relationships are encoded, propagated, and constrained across heterogeneous imaging modalities. Rather than focusing narrowly on benchmark performance, the discussion interrogates structural trade-offs between local anatomical fidelity and global contextual reasoning, the implications of selective state-space propagation for long-range dependencies, and the infrastructural demands of deploying graph-structured state-space systems in clinical environments. The paper further explores governance mechanisms for safety assurance, robustness against distributional shift, fairness across patient populations, and sustainability concerns arising from model scale and data movement. A cross-domain comparative lens is adopted to contrast Graph-Mamba designs with convolutional, transformer, and attention-based graph systems, highlighting where hybrid relational-state-space models offer distinct advantages for organ segmentation, disease staging, and longitudinal anatomical modeling. The analysis identifies key architectural decision points including graph construction granularity, bidirectional state propagation, memory gating, and regularization of anatomical priors. It argues that the value of Graph-Mamba in medical imaging resides not only in representational capacity but also in its potential to support more interpretable, resource-aware, and governable learning systems when coupled with explicit anatomical constraints. Throughout, emphasis is placed on the socio-technical embedding of such models in clinical workflows, diagnostic accountability, and the broader regulatory landscape. The paper concludes with forward-looking perspectives on federated learning, uncertainty quantification, and policy-aligned evaluation frameworks for next-generation medical image analysis infrastructures.

Keywords

Graph-Mamba networks, medical image analysis, anatomical graphs, state-space models, clinical deployment, robustness, fairness, governance.

1. Introduction

Medical image analysis has undergone a fundamental transformation through deep learning, moving from handcrafted feature extraction to end-to-end representation learning. Convolutional architectures such as U-Net established strong baselines for semantic segmentation by combining local feature extraction with symmetric skip connections [1]. Subsequent systems like nnU-Net demonstrated that automated configuration pipelines could adapt these convolutional backbones to diverse biomedical imaging tasks without task-specific manual tuning [2]. A broad survey of deep learning in medical image analysis further confirmed that models trained on large imaging datasets could achieve expert-level performance in selected diagnostic tasks, while also revealing persistent challenges in generalization and interpretability [3]. These foundational developments established the technical substrate upon which more structurally expressive models have been built. However, convolutional systems remain limited by their local receptive fields, which constrain their ability to reason explicitly about long-range anatomical dependencies that are often critical in radiological and pathological interpretation.

The introduction of self-attention mechanisms represented a significant shift toward global context modeling. The transformer architecture redefined sequence modeling by replacing recurrence with parallelizable attention over all positions [4]. Vision transformers subsequently adapted this principle to image classification by treating image patches as tokens, demonstrating that pure attention could rival convolutional networks when sufficient data were available [5]. In medical imaging, hybrid architectures such as UNETR integrated transformer encoders with convolutional decoders to capture long-range spatial correlations in volumetric data [6]. Extensions like Swin UNETR leveraged hierarchical shifted-window attention to improve computational efficiency while preserving multi-scale representational capacity for three-dimensional segmentation [7]. Despite these advances, attention mechanisms scale quadratically with sequence length, creating substantial memory and latency bottlenecks for high-resolution volumetric images or densely sampled anatomical graphs. This limitation has motivated renewed interest in alternative sequence modeling paradigms that preserve global context without incurring prohibitive computational costs.

State-space models have emerged as a compelling alternative to attention. The Mamba architecture introduced selective state-space sequence modeling with linear-time complexity and strong empirical performance on long sequences [8]. Unlike conventional recurrent networks, selective state-space models dynamically adapt their transition parameters based on the input, enabling context-dependent memory gating while retaining efficient parallel scanning. This property is particularly relevant for medical image analysis, where anatomical structures exhibit highly variable spatial extent and relational organization. When combined with graph representations, state-space propagation can traverse anatomical connectivity patterns in a structured manner, potentially capturing both local organ morphology and global physiological relationships. The Graph-Mamba paradigm arises precisely at this intersection, seeking to integrate explicit relational structure with efficient long-range sequence modeling for clinically meaningful image interpretation.

Simultaneously, graph neural networks have matured as tools for learning on irregular relational domains. Graph convolutional networks extended convolution to graph-structured data by aggregating information from local neighborhoods [9]. Spectral graph convolutions

provided a frequency-domain perspective on graph signal processing, enabling learned filters over graph Laplacians [10]. Graph attention networks further introduced self-attention over graph neighbors, allowing the model to weigh relational importance adaptively [11]. A comprehensive review of graph neural network methods underscored their capacity to model dependencies in social networks, molecular graphs, and knowledge graphs, while noting unresolved challenges in scalability, over-smoothing, and dynamic graph learning [12]. In medical image analysis, graphs can naturally represent anatomical regions, organs, lesions, or tissue boundaries, with edges encoding spatial adjacency, vascular connectivity, or functional interactions. A recent cross-fusion Mamba architecture for multi-view medical image classification highlights the growing interest in combining state-space efficiency with structurally informed fusion across views [13]. The Graph-Mamba approach extends this logic by making the graph itself the substrate over which selective state-space updates propagate.

This paper examines Graph-Mamba networks from a systems perspective, emphasizing architectural trade-offs, deployment feasibility, robustness, fairness, and governance rather than incremental benchmark improvements. The central argument is that Graph-Mamba should be understood not as a single model but as a family of computational designs in which the choice of graph topology, state transition mechanism, and propagation direction jointly determine both technical performance and socio-technical properties. By foregrounding infrastructure and policy, the paper aims to connect algorithmic innovation with the institutional realities of clinical translation.

2. Background and Related Work

The evolution of medical image analysis systems has been shaped by recurring tensions between local detail preservation and global context aggregation. Early convolutional networks excelled at extracting hierarchical features but struggled to model distant structures without deep architectures or large receptive fields. nnU-Net addressed part of this challenge through automated configuration, yet its underlying convolutional inductive bias remained local [2]. Transformer-based systems such as UNETR and Swin UNETR expanded context modeling through self-attention, but their quadratic complexity posed practical obstacles for three-dimensional imaging at native resolution [6, 7]. Hybrid approaches often resorted to patch-based processing, which reintroduced boundary artifacts and required careful overlap strategies. These trade-offs are not merely technical; they influence whether a model can be deployed in real-time interventional settings, integrated into retrospective research pipelines, or trained within constrained computational budgets in public hospitals.

Graph-based methods offer a structurally different solution by decoupling the representation of entities from the representation of their relations. In medical imaging, anatomical structures can be segmented first and then connected into a graph, or graph nodes can be learned implicitly from image features and updated through message passing. Graph convolutional networks provide a flexible mechanism for aggregating information along edges, but repeated layers can lead to over-smoothing, where node representations become indistinguishable [9]. Graph attention networks mitigate this by learning edge-specific weights, yet attention over large graphs can become computationally expensive [11]. Reviews of graph neural network applications have emphasized the importance of graph construction as a modeling decision, because the choice of edges determines which relational hypotheses can be represented at all [12]. This observation is particularly acute in medicine, where anatomical relationships are

both spatial and functional, and may vary across patients due to surgical history, congenital anomalies, or disease progression.

State-space models entered the deep learning mainstream through advances in efficient sequence modeling. The Mamba architecture introduced a selective mechanism that allows the continuous-time state matrix to depend on the input, thereby enabling the model to ignore irrelevant tokens and focus on salient context [8]. Unlike standard recurrent networks, Mamba uses a hardware-aware parallel scan that achieves linear-time training while maintaining recurrent-style inference. This efficiency profile has attracted interest from the medical imaging community, especially for high-dimensional and long-sequence applications such as whole-slide pathology, four-dimensional cardiac flow, and longitudinal volumetric studies. Vision Mamba variants adapted the selective state-space mechanism to two-dimensional images by flattening patches into sequences, demonstrating competitive accuracy with reduced memory consumption compared to vision transformers [14]. SegMamba extended this design to three-dimensional medical segmentation by modeling long-range spatial dependencies across volumetric slices [15]. These systems confirm that state-space models can serve as efficient backbones for medical image analysis, but they do not explicitly incorporate relational structure.

The Graph-Mamba paradigm combines the two threads by embedding selective state-space updates within graph-structured propagation. Instead of flattening all image patches into a single sequence, the model maintains node-specific state variables and updates them along graph edges. This allows the system to respect anatomical connectivity while retaining the linear-time computational benefits of selective state-space models. In principle, such a design can integrate multi-organ adjacency, vascular networks, or functional connectivity into the learning process without requiring full pair-wise attention. The architectural challenge lies in determining how many propagation steps are necessary, whether the graph should be static or learned, and how to initialize states from image features in a manner that preserves spatial resolution. These choices have direct implications for model capacity, interpretability, and robustness.

3. Graph-Mamba Architectural Principles

A Graph-Mamba architecture can be decomposed into three interacting subsystems: the graph construction layer, the selective state propagation layer, and the readout mechanism. The graph construction layer maps image regions or anatomical entities into node embeddings and defines an adjacency structure that may be generated from spatial proximity, organ segmentation, or learned attention similarities. This layer functions as an explicit inductive prior, constraining subsequent computation along anatomically meaningful edges. When graph edges are derived from prior anatomical knowledge, the system acquires a form of structural regularization that purely local or purely global methods lack. However, rigid prior graphs may fail to capture patient-specific variations, motivating hybrid designs in which a base anatomical graph is augmented by learned dynamic edges. The balance between fixed and learned connectivity is a central governance decision because it determines how much anatomical knowledge is hard-coded into the model and how much remains open to data-driven revision.

The selective state propagation layer applies a state-space update at each node based on the node's current representation and the aggregated messages from neighboring nodes. A key architectural choice is whether the state transition parameters are shared globally, conditioned on node features, or conditioned on edge features. Global sharing reduces parameter count

and improves sample efficiency but may be insufficient for heterogeneous anatomical regions with different dynamic properties. Node-specific or edge-specific conditioning allows the model to adapt its memory horizon to local context, mimicking the selective attention mechanism of Mamba [8]. However, this flexibility increases the risk of overfitting and complicates interpretability, because the effective receptive field of a node becomes a learned function of both graph topology and input-dependent gating. Bidirectional propagation along graph edges can capture symmetric relationships such as anatomical adjacency, while directed propagation may be more appropriate for physiological flows such as blood circulation or cerebrospinal fluid movement.

The readout mechanism aggregates node-level representations into task-specific predictions, such as pixel-level segmentation maps, organ-level classification labels, or graph-level disease scores. In segmentation tasks, node features must be mapped back to the original image grid, often through convolutional decoders or cross-attention layers. This inverse mapping introduces additional design trade-offs, particularly in boundary regions where the graph representation may lose fine spatial detail. Graph-Mamba systems that operate purely at the graph level may excel at semantic reasoning but underperform on voxel-level accuracy, whereas systems that retain high-resolution convolutional pathways may achieve better perceptual fidelity at the cost of greater memory usage. A senior systems perspective therefore emphasizes that Graph-Mamba should be evaluated not only by Dice scores or classification accuracy but also by the alignment between its internal relational structure and the clinical topology of the anatomy under study.

The computational advantages of selective state-space propagation are most pronounced when the graph is large and irregular. Unlike graph attention, which requires pair-wise compatibility functions, selective state-space updates can be computed with linear complexity in the number of edges. This makes Graph-Mamba particularly attractive for whole-body imaging, where the number of anatomical nodes may range from dozens to thousands and the number of edges may scale accordingly. Nevertheless, the parallel scan implementation of state-space models assumes a sequential order over tokens. In graph settings, the absence of a natural sequence order requires either a learned traversal strategy or a message-passing reformulation that preserves the linear-time property. Several approaches use bidirectional graph scanning by sorting nodes according to spatial coordinates or spectral embeddings, but this risks introducing order-dependent artifacts. The resolution of this issue remains an active research area with significant implications for deployment consistency and reproducibility.

4. Integrating Anatomical Relationships

Anatomical relationships in medical imaging are fundamentally relational, hierarchical, and context-dependent. Organs are not independent entities; they are connected through fascial planes, vascular supply, lymphatic drainage, and neural innervation. Pathological processes often respect or disrupt these relationships in characteristic ways. For example, the spread of abdominal malignancies frequently follows peritoneal reflections and vascular routes, while neurodegenerative diseases propagate along anatomical networks in the brain. Graph-Mamba networks provide a natural computational framework for representing such dependencies because the graph topology itself can encode anatomical prior knowledge. By initializing edges according to canonical anatomical adjacency, the model is encouraged to reason about disease processes in ways that align with clinical understanding. This is not merely a performance issue; it is an issue of epistemic alignment between machine representations and medical knowledge systems.

A concrete illustration can be found in multi-organ segmentation of abdominal CT. A naive patch-based convolutional model may segment the liver, spleen, and kidneys independently, but it may miss the consistent anatomical relationships that constrain their relative positions. A graph-based model can connect organ nodes according to adjacency and propagate state updates that enforce relational consistency. If the graph is constructed after an initial coarse segmentation, the model can refine boundary predictions by aggregating contextual information from neighboring organs. In this setting, Graph-Mamba can operate as a post-processing refinement stage or as an integrated end-to-end model with differentiable graph construction. The latter is more powerful but also more difficult to train, because the graph topology itself becomes a latent variable that must be optimized jointly with the state transition parameters.

Longitudinal imaging presents another compelling use case. In chronic disease monitoring, images are acquired at multiple time points, and the change in anatomical structure over time is often more informative than any single scan. Graph-Mamba can represent time as an additional dimension of sequential propagation, allowing state variables to carry forward information about prior anatomy. Because selective state-space models are designed for long sequences, they are well suited to handling irregularly sampled longitudinal data without requiring the fixed-window assumptions of recurrent neural networks or the quadratic cost of temporal attention. This capability may enable more nuanced trajectory modeling for conditions such as multiple sclerosis, cancer surveillance, or cardiac remodeling after myocardial infarction. However, the clinical deployment of such longitudinal models requires careful handling of missing visits, scanner variation, and registration errors, all of which can propagate through the graph and distort downstream state updates.

Cross-view integration is another domain where relational state-space models show promise. Multi-view medical imaging, such as paired mammography views, orthogonal chest radiographs, or multiparametric MRI, provides complementary information that cannot be fully exploited by treating each view independently. A recent cross-fusion Mamba approach for multi-view classification demonstrates that selective state-space fusion can effectively combine features across views [13]. Graph-Mamba extends this idea by treating each view as a subgraph and connecting corresponding anatomical landmarks across views, enabling relational fusion that respects both intra-view structure and inter-view correspondence. This design can improve diagnostic consistency and reduce the likelihood of contradictory predictions across views. From a governance standpoint, such cross-view relational fusion is valuable because it encourages the model to produce coherent outputs, which is a necessary though not sufficient condition for clinical trust.

5. System Infrastructure and Deployment Considerations

The deployment of Graph-Mamba networks in clinical environments is inseparable from the broader computational and data infrastructure in which they operate. Unlike laboratory benchmarks, clinical systems must handle heterogeneous data sources, variable acquisition protocols, missing metadata, and strict latency requirements. The graph construction step introduces additional preprocessing overhead, because anatomical nodes must be identified and connected before the state-space model can be applied. In some settings, this can be performed by a separate segmentation model, but that introduces a dependency chain in which errors from the upstream segmentation propagate into the graph representation. A more integrated infrastructure may use a shared backbone to produce both segmentations and graph nodes, but this increases model complexity and may require specialized training procedures.

System designers must therefore weigh the benefits of explicit anatomical structure against the fragility of multi-stage pipelines.

Compute resource constraints are particularly acute in public hospitals and low-resource settings. While selective state-space models reduce asymptotic complexity compared to attention, the practical memory footprint of a Graph-Mamba model depends on the number of nodes, the dimensionality of node embeddings, the number of propagation steps, and the precision used during training and inference. Mixed-precision training can reduce memory usage, but it may destabilize the selective state transition if the learned gating parameters exhibit large dynamic ranges. Model compression techniques such as distillation and pruning can be applied, but they must be validated for graph-structured state-space models, where pruning edges may inadvertently remove anatomically critical connections. Infrastructure decisions about whether to deploy on-premise, in the cloud, or through edge devices must account for data privacy, network reliability, and the need for audit logs that document model versions and input characteristics.

Interoperability with existing clinical systems is another critical factor. Picture archiving and communication systems, electronic health records, and radiology information systems have their own data models and workflow constraints. A Graph-Mamba model that requires specialized graph construction software may not integrate easily with standard DICOM pipelines. Containerized deployment using standards-compliant interfaces can mitigate some of these issues, but the semantic gap between graph-based model outputs and the structured reporting formats used by clinicians remains significant. Outputs must be translated into actionable visualizations, structured measurements, or narrative summaries without losing the relational context that the model was designed to capture. This translation layer is often underappreciated in academic research but is essential for real-world adoption.

Monitoring and maintenance present further deployment challenges. Medical imaging models are subject to data drift as acquisition protocols evolve, patient populations change, and new scanners are introduced. Graph-Mamba networks may be particularly sensitive to changes in graph topology that result from segmentation failures, because the state propagation depends on the adjacency structure. Continuous monitoring of node-level and edge-level statistics can detect early signs of distributional shift, but such monitoring requires additional infrastructure for logging model activations and comparing them across deployment sites. Governance frameworks should require not only aggregate performance metrics but also disaggregated monitoring by anatomical region, patient subgroup, and imaging modality. This level of observability is necessary to ensure that the model remains safe and effective over time.

6. Robustness, Fairness, and Governance

Robustness in medical image analysis encompasses several distinct notions: robustness to image noise, robustness to scanner variation, robustness to adversarial perturbation, and robustness to missing or corrupted graph structures. Graph-Mamba networks introduce a new failure mode because the graph itself can be corrupted by segmentation errors or incomplete anatomical priors. If an edge is missing, state information cannot propagate between two anatomically connected nodes, potentially leading to contextually incoherent predictions. Conversely, spurious edges may create shortcuts that propagate noise across unrelated regions. Adversarial attacks could target the graph construction layer by perturbing image features in ways that change the learned adjacency, even if the final segmentation appears superficially correct. Defenses may include graph structure regularization, edge dropout, and uncertainty-

aware aggregation, but these mechanisms must be validated across diverse patient populations and imaging conditions.

Fairness considerations are equally important. Medical imaging models have been shown to exhibit performance disparities across demographic groups, often reflecting imbalances in training data [16, 17]. Graph-Mamba is not immune to these disparities, and the use of anatomical graphs may introduce new sources of bias if the prior graph is derived from a population with limited anatomical variation. For example, a standard atlas of abdominal anatomy may not accurately represent patients with situs inversus, organ transplantation, or congenital malformations. If the model penalizes deviations from the prior graph, it may systematically underperform for rare anatomical variants. Fairness-aware training strategies and subgroup-level evaluation are therefore essential, not as optional add-ons but as integral components of the model development lifecycle. The governance of medical AI systems must demand evidence of equitable performance before deployment [18].

Explainability and interpretability are closely tied to governance. Graph-Mamba networks offer a potential advantage because their relational structure can be visualized as a network of anatomical nodes with propagated state information. Unlike black-box attention maps, the graph topology provides an explicit representation of which regions are considered related by the model. However, this potential is realized only if the graph construction layer is transparent and the propagation dynamics can be summarized in clinically meaningful terms. Techniques such as gradient-based saliency and class activation mapping can be extended to graph-structured state-space models, but they require careful adaptation [19]. Regulators and clinical users should be able to interrogate why a prediction was made, which edges contributed most strongly, and whether the model's relational reasoning aligns with known anatomy. This level of transparency supports accountability and may facilitate regulatory approval under evolving frameworks for artificial intelligence in medicine.

Bias in labels and annotations also interacts with graph structure. If training labels are generated by automated systems or by clinicians with variable expertise, errors may be spatially correlated and propagate through the graph. For example, a systematic under-segmentation of a particular organ in the training data may lead to missing graph nodes for that organ, reducing the model's ability to reason about it. Governance mechanisms should include annotation quality audits, inter-rater reliability assessments, and explicit documentation of the anatomical coverage and limitations of training datasets. Federated learning offers one path toward more diverse and representative training data without centralizing patient information, but it introduces its own governance challenges related to data heterogeneity, model aggregation, and security [20]. Graph-Mamba models designed for federated settings must be robust to non-identically distributed local graphs.

7. Sustainability and Policy Implications

Sustainability in medical AI encompasses both environmental and institutional dimensions. The environmental cost of training large models has drawn increasing scrutiny, and medical imaging models are no exception. Graph-Mamba networks may reduce energy consumption relative to transformer-based models because of their linear-time complexity, but this benefit can be offset by the added cost of graph construction and multi-stage preprocessing [8, 14]. Lifecycle assessments should account for the full pipeline, including data storage, preprocessing, training, inference, and periodic retraining. In resource-constrained healthcare systems, the carbon footprint of model development may be less visible than the financial cost, but both must be weighed against clinical benefit. Efficient architectures should be favored

not merely for their computational elegance but because they enable broader access to advanced diagnostic support.

Policy implications are profound. Regulatory bodies in multiple jurisdictions are developing frameworks for software as a medical device, with particular attention to artificial intelligence and machine learning. Graph-Mamba models that incorporate anatomical priors may be viewed as lower risk if their relational constraints reduce the space of implausible outputs, but this does not eliminate the need for rigorous clinical validation. Model updates, including retraining on new data or modification of the graph topology, may trigger regulatory review depending on the significance of the change. A governance framework that treats the graph structure as a controlled component, with versioning and documentation requirements, could facilitate safer iterative improvement. Transparent reporting of model architecture, training data composition, and evaluation protocols is consistent with emerging best practices for machine learning in health.

Cross-domain comparisons offer valuable lessons. In autonomous driving, graph-based scene understanding systems have been subject to extensive safety analysis because relational reasoning about traffic participants is essential for avoiding collisions. In drug discovery, graph neural networks are used to model molecular interactions, and their failure modes have led to the development of uncertainty-aware screening protocols. Medical imaging can adapt these lessons by treating anatomical relationship errors as safety-relevant failures even when aggregate segmentation metrics remain high. Policy frameworks should require scenario-based testing that probes the model's behavior on rare anatomical variants, unusual disease presentations, and corrupted inputs. Such testing goes beyond average performance and aligns with the safety culture of high-reliability industries.

The institutional sustainability of Graph-Mamba deployments depends on workforce readiness and interdisciplinary collaboration. Radiologists, technologists, data scientists, and IT administrators must be able to understand and maintain the system. This requires documentation that is accessible to non-specialists, training programs that bridge clinical and computational domains, and clear lines of accountability for model outputs. Without such institutional support, even technically sophisticated models may fail to achieve sustained clinical use. Policy should therefore incentivize not only model development but also the creation of shared governance infrastructure, open evaluation benchmarks, and data-sharing agreements that respect patient privacy. The long-term value of Graph-Mamba networks will be determined as much by these socio-technical factors as by the underlying algorithmic innovations.

8. Conclusion

Graph-Mamba networks represent a promising architectural direction for integrating anatomical relationships into medical image analysis. By combining the relational expressiveness of graph neural networks with the linear-time efficiency of selective state-space models, they offer a structured approach to modeling the complex dependencies that characterize human anatomy and disease. This paper has argued that the significance of Graph-Mamba extends beyond benchmark accuracy to encompass structural trade-offs in graph construction, propagation dynamics, and readout design. The choice of anatomical prior, the directionality of state updates, and the granularity of graph nodes all have direct consequences for model behavior, interpretability, and robustness.

From a systems perspective, Graph-Mamba should be evaluated as a socio-technical infrastructure rather than an isolated algorithm. Deployment requires careful attention to computational constraints, data pipelines, interoperability with clinical systems, and long-term maintenance. Robustness and fairness cannot be treated as post hoc considerations; they must be designed into the graph structure and the training procedure. Governance mechanisms, including monitoring, documentation, and subgroup-level evaluation, are essential to ensure that these models are safe and equitable across diverse patient populations. Sustainability concerns further motivate the development of efficient architectures that can be deployed in resource-limited settings without prohibitive environmental costs.

The forward-looking agenda for Graph-Mamba research includes federated learning across institutions, uncertainty quantification tailored to graph propagation, and tighter integration with structured radiology reporting. Policy frameworks should evolve in parallel to address the unique risks of graph-based relational reasoning, particularly the possibility of anatomically implausible outputs arising from corrupted graph topologies. By treating anatomical relationships as first-class citizens in both model architecture and governance, the field can move toward medical image analysis systems that are not only accurate but also clinically aligned, transparent, and institutionally sustainable.

References

1. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention (pp. 234–241). Springer.

2. 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.

3. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88.

4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.

5. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations.

6. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H. R., & Xu, D. (2022). UNETR: Transformers for 3D medical image segmentation. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 574–584.

7. Tang, Y., Yang, D., Li, W., Roth, H. R., Landman, B., Xu, D., Nath, V., & Hatamizadeh, A. (2022). Self-supervised pre-training of Swin transformers for 3D medical image analysis. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 20730–20740.

8. Gu, A., & Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752.

9. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. International Conference on Learning Representations.

10. Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. Advances in Neural Information Processing Systems, 29, 3844–3852.

11. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph attention networks. International Conference on Learning Representations.

12. Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. AI Open, 1, 57–81.

13. 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.

14. Zhu, L., Liao, B., Zhang, Q., Wang, X., Liu, W., & Wang, X. (2024). Vision Mamba: Efficient visual representation learning with bidirectional state space model. International Conference on Machine Learning.

15. Xing, Z., Ye, T., Yang, Y., Liu, G., & Zhu, L. (2024). SegMamba: Long-range sequential modeling Mamba for 3D medical image segmentation. Medical Image Computing and Computer-Assisted Intervention, 578–588.

16. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225.

17. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.

18. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707.

19. 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. Proceedings of the IEEE International Conference on Computer Vision, 618–626.

20. Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M., & Cardoso, M. J. (2020). The future of digital health with federated learning. npj Digital Medicine, 3, 119.

21. 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.