State Space Models for Long-Range Dependency Modeling in Histopathological Image Analysis

Brooks Steele
Department of Computer Science, George Mason University, Fairfax, VA, USA.
b.steele@gmu.edu

Randy Richards
Department of Computer Science, University of Houston, Houston, TX, USA.
randyrichards99@uh.edu

Abstract

Histopathological image analysis presents a distinct set of representational and infrastructural challenges because diagnostically meaningful patterns extend across widely separated regions of gigapixel whole slide images. Convolutional architectures are efficient but local by design, while attention-based models enable global interactions at substantial computational cost. State space models have recently emerged as an architectural middle ground that supports long-range sequence modeling with a bounded state representation. This paper provides a systems-oriented examination of state space models for histopathological image analysis, moving beyond accuracy comparisons to analyze structural trade-offs in deployment, data governance, fairness, clinical accountability, and regulatory compliance. The discussion situates state space models within the larger computational pathology ecosystem, including whole slide imaging pipelines, laboratory information systems, federated learning infrastructure, and postmarket surveillance. The paper argues that the value of state space models depends not only on their capacity to capture long-range dependencies but also on how their architectural properties intersect with institutional workflows, equity safeguards, sustainability requirements, and evolving policy frameworks. The analysis integrates perspectives from machine learning, digital pathology, health informatics, and socio-technical systems research.

Keywords

state space models, histopathology, long-range dependencies, computational pathology, digital pathology, system governance.

1. Introduction

The digitization of tissue specimens through whole slide imaging has transformed histopathology from a predominantly optical discipline into a hybrid visual and computational practice [1]. Whole slide images now routinely contain billions of pixels and encode complex spatial arrangements of cells, glands, stroma, and immune infiltrates. The diagnostic interpretation of these images depends on recognizing architectural patterns that can span very large spatial extents, from subcellular details to tissue-level structures. Systems that support automated analysis must therefore address a central representational problem: how to model dependencies between distant regions without sacrificing computational tractability or clinical interpretability.

Computational pathology has produced a substantial body of work on deep learning architectures for segmentation, classification, and prognostic prediction [2]. Early successes

often relied on convolutional neural networks applied to small image tiles, with weak labels aggregated across entire slides. Large-scale studies have demonstrated that such approaches can approach clinical-grade performance for selected diagnostic tasks, including cancer detection in prostate biopsies and skin specimens [3]. Nevertheless, the patch-based processing paradigm can fragment the tissue landscape, and many clinically meaningful features emerge only when interactions across broad spatial ranges are preserved.

The need for long-range dependency modeling has motivated interest in architectures beyond convolutional networks. Transformers and attention mechanisms have shown promise in general visual recognition, but their adoption in histopathology raises concerns about memory growth, data appetite, and the instability of attention over very long sequences. State space models have recently emerged as a distinct family of sequence models that combine linear or near-linear scaling with a recurrent state representation. This paper examines state space models for histopathological image analysis from a systems perspective, focusing not only on accuracy but also on deployment, governance, fairness, maintenance, and policy. The aim is to provide a critical architectural and socio-technical assessment of how state space models may fit into the broader computational pathology ecosystem.

2. Long-Range Dependency Challenges in Histopathological Images

Histopathological images differ from natural images in several important ways. The objects of interest are not discrete entities with fixed scales; they include nuclei, tubules, ducts, inflammatory clusters, necrosis, fibrosis, and vascular structures. Diagnostic criteria often require evaluating the spatial relationship between these components across millimeters of tissue. In whole slide imaging, a single tissue section may be represented as a pyramid of tens of thousands of tiles, and the relevant contextual field can extend far beyond a conventional convolutional receptive field. The system must therefore handle scale variation, tissue heterogeneity, and long-range architectural cues simultaneously.

Multiple instance learning has been widely adopted for weakly supervised whole slide classification because it allows models to operate on bags of patches while leaving the underlying graph or spatial arrangement partially implicit [4]. This approach reduces memory pressure, but it can also discard positional relationships unless explicitly augmented. Context-aware convolutional approaches have attempted to restore local and regional context by incorporating larger fields of view or multi-resolution inputs [5]. These methods improve performance for grading tasks such as colorectal cancer assessment, yet they remain bounded by the inductive biases of local convolution and fixed tile sizes.

Attention mechanisms provide a more flexible route to long-range dependency modeling. Transformers replace the sequential assumption of recurrent networks with a content-addressed retrieval process over token representations [6]. Vision transformers adapted this principle to image patches, showing that pure attention can compete with convolutional architectures when sufficient data are available [7]. In histopathology, vision transformers allow tokens from distant tissue areas to interact directly. However, the computational and memory costs of global self-attention grow quadratically with sequence length, which is particularly problematic for gigapixel histopathological images. Although approximations such as windowed attention and hierarchical tokenization reduce this burden, they reintroduce locality constraints and can weaken the very long-range interactions that motivated the transformer in the first place.

The practical consequence is a structural trade-off. Convolutional stacks are efficient but local by design. Attention mechanisms are expressive but expensive and data-hungry. Multiple instance learning scales well but often underuses spatial topology. A histopathology system must balance these properties while remaining robust across staining variations, scanner differences, and tissue types. This trade-off defines the design space into which state space models enter.

3. State Space Models as Architectural Middle Ground

State space models offer an architectural alternative that is neither strictly convolutional nor attention-based. Recent work has extended this family to multi-view medical image classification through cross-fusion architectures that integrate complementary views and modalities [8]. This line of work illustrates how state space representations can be used to combine evidence across different observational channels, a requirement that has increasing relevance in computational pathology when multiple stains, magnifications, or imaging modalities are available. However, the broader significance of state space models for histopathology lies in their capacity to model long sequences without incurring the quadratic cost of full attention.

Foundational structured state space models formalize a sequence-to-sequence mapping through a latent state that evolves according to a continuous-time dynamical system [9]. The recurrent form of this computation allows efficient sequential processing, while the convolutional form supports parallel training over long inputs. Subsequent selective state space models introduced input-dependent gating, allowing the dynamics to adapt to the content of the sequence [10]. This selectivity is important for histopathology because tissue is heterogeneous; a fixed state transition may be insufficient to capture differences between dense tumor, inflamed stroma, and necrotic regions. Selective models can, in principle, modulate their memory based on the diagnostic content of each image region.

Adaptations of state space models to visual data have demonstrated that bidirectional scanning can preserve spatial context while maintaining linear complexity [11]. In histopathology, this suggests an appealing compromise: the model can traverse image tokens in multiple directions to encode neighborhood orientation, while retaining the ability to propagate information across long distances. Unlike local convolutional networks, the state transition can carry a summarized representation of earlier tissue regions forward. Unlike global attention, the state has a fixed dimensionality, which decouples memory usage from the number of image patches. Residual connections remain important for optimization and gradient flow, continuing a design principle established in deep convolutional networks [12].

The system-level advantage of state space models is not merely algorithmic speed. Because their memory footprint is controlled by the state size rather than the square of sequence length, they enable different deployment configurations. A hospital might process entire tissue sections as long patch sequences on a single accelerator, rather than relying on complex hierarchical sampling or patch-wise ensembling. The recurrent state also aligns naturally with streaming inference, in which tiles are processed sequentially as they are extracted from a whole slide image. This property may reduce the need to store all intermediate patch features, which is a significant operational concern in digital pathology.

However, state space models also introduce new trade-offs. The fixed-dimensional state can become a bottleneck if it is too small relative to the diversity of tissue patterns. Selectivity helps, but it also increases the complexity of the state transition and may reintroduce stability

concerns during training. The design of scanning paths across two-dimensional tissue images is nontrivial, and poor ordering can introduce artificial discontinuities. Multi-view and multi-omic fusion creates further governance questions about which modalities are allowed to dominate the state update and whether the resulting representation remains auditable. These are not merely technical details; they shape how such systems can be validated and deployed.

4. System-Level Trade-offs in Deployment and Infrastructure

Deploying histopathological image analysis systems requires careful attention to preanalytical, analytical, and postanalytical variability. Stain color normalization and data augmentation have been shown to affect convolutional network robustness, and similar considerations will apply to state space models that ingest large image sequences [13]. A state space model that has learned to rely on long-range chromatic gradients may fail when scanner calibration or staining protocols change. Therefore, infrastructure must include quality control pipelines that monitor color distributions, tissue processing artifacts, and slide preparation variation before model inference.

Open source platforms such as QuPath have become important components of digital pathology workflows because they provide interoperable tools for image annotation, tile export, and review [14]. State space model deployments should be compatible with such platforms rather than existing as isolated engines. Interoperability requirements include standardized file formats, metadata exchange, and the ability to map model outputs back onto whole slide coordinates. The architectural choice of a model is only one element in a larger system that comprises slide scanners, image management systems, laboratory information systems, and clinical viewers [15]. Long-range sequence models must therefore be embedded within a governance layer that tracks model version, input provenance, preprocessing parameters, and output uncertainty.

Cloud and on-premise trade-offs are particularly acute in pathology because whole slide images are large and patient-sensitive. State space models may reduce peak memory relative to attention-based models, but large-scale training still requires substantial storage and network bandwidth. Federated or distributed training arrangements can preserve data locality, but they introduce synchronization overhead, model drift, and auditing complexity. The selected architecture affects these system choices because model size, sequence length, and optimizer state determine the minimum hardware envelope for training. An architecture that enables linear scaling may make incremental training on regional hospital clusters more feasible, but it does not eliminate the need for robust data governance and secure multi-site coordination.

Another infrastructure concern is resilience to distribution shift. Histopathology models trained in one institution often degrade when applied elsewhere due to scanner differences, population variation, and changing diagnostic criteria. The state space formulation may help by maintaining a compact running summary that can be updated with limited calibration data, but this also raises questions about local adaptation without centralized oversight. A balanced system would combine federated fine-tuning with continuous performance monitoring, alerting pathologists when the model encounters tissue types that are underrepresented in its training distribution.

5. Data Governance, Fairness, and Clinical Accountability

Histopathological machine learning systems operate on data that are simultaneously clinically sensitive and socially stratified. Large-scale studies have shown that deep learning can

classify histopathological images and even predict genetic alterations from routine tissue slides [16]. Diagnostic assessment studies have further demonstrated that algorithm performance can approach human experts for defined tasks such as detecting lymph node metastases [17]. These capabilities strengthen the case for state space models as scalable representation learners, but they also intensify the need for governance. If long-range models make predictions based on subtle architectural features distributed across a slide, the rationale for a specific decision may be more difficult to summarize for a pathologist or a patient.

The convergence of high-performance algorithms and clinical practice raises broader questions about human oversight and accountability [18]. Long-range state representations may be less interpretable than attention maps because the state summarizes information over time rather than assigning pairwise token weights. Although attention maps are themselves imperfect explanations, they offer a familiar visual heuristic. State space models will require new explanation strategies that surface which regions most influenced the evolving state, perhaps through influence functions, perturbation analysis, or trajectory visualization. Without such tools, clinicians may be hesitant to adopt models that appear as opaque memory systems.

Fairness is a system-level requirement, not an afterthought. Algorithms used in health care have been shown to exhibit racial and social biases despite being trained on seemingly objective clinical data [19]. A survey of bias and fairness in machine learning highlights the many ways bias can enter through data collection, label definition, model design, and evaluation [20]. Histopathology images carry demographic and environmental information that may correlate with ancestry, geography, or socioeconomic conditions. State space models that are trained on large tissue cohorts may learn spurious long-range associations if the training data are not carefully curated. Governance must include demographic documentation, subgroup performance evaluation, and transparent limits on model use.

Model reporting practices provide a mechanism for making these risks explicit. Model cards and related documentation frameworks encourage developers to state intended use, training data composition, evaluation metrics, and known limitations [21]. For state space models in histopathology, a model card should also document the scanning order, state size, sequence length, pretraining corpus, and the modalities used for fusion. Such transparency supports institutional review boards, regulatory reviewers, and clinical end users in assessing whether a model is appropriate for a particular population and diagnostic question.

6. Integration with Hospital and Laboratory Information Systems

The integration of artificial intelligence into clinical pathology is not only a modeling challenge but also an information systems challenge. Health care systems already manage complex electronic health record data at scale, and successful deployment requires harmonization with existing clinical workflows [22]. State space models that consume image data must be connected to laboratory information systems, picture archiving and communication systems, and order entry workflows. The model output must be delivered in a format that pathologists can review, annotate, and either accept or contest. This demands careful interface design, standardized messaging protocols, and audit trails that preserve the chain of evidence from slide preparation to model inference.

Digital pathology platforms are increasingly positioned as clinical decision support tools embedded within broader precision oncology ecosystems [23]. A state space model that identifies long-range architectural features associated with tumor progression may be used

alongside molecular tests, radiology findings, and clinical history. The system must therefore represent uncertainty in a way that supports clinical reasoning. A single scalar risk score is rarely sufficient; clinicians need to know the confidence interval, the basis for the prediction, and the conditions under which the model should not be used. Long-range sequence models can be particularly useful for encoding spatial context, but their outputs still need to be translated into actionable clinical reports.

Regulatory and quality management considerations shape how such systems are built and maintained. The United States Food and Drug Administration has articulated a risk-based framework for artificial intelligence and machine learning software as a medical device, emphasizing lifecycle management and iterative improvement [24]. State space models that are updated continuously or that adapt their state transition during fine-tuning fall squarely within this lifecycle-oriented regulatory view. Manufacturers and health systems must establish procedures for change control, retraining, validation, and postmarket surveillance. These procedures are not secondary to model design; they are constitutive of safe deployment.

7. Sustainability, Maintenance, and Lifecycle Governance

Sustainability in histopathological image analysis refers both to computational resource use and to organizational maintenance. Large transformer models may require substantial energy and hardware budgets, which can limit adoption in resource-constrained health systems. State space models have the potential to reduce inference memory and enable longer sequence processing on more modest accelerators, but their training cost depends heavily on the scale of the whole slide corpus and the chosen state expansion. Institutions must evaluate whether the marginal accuracy gain justifies the carbon, financial, and staffing costs. A linear-complexity model that requires fewer accelerators for deployment may still demand significant data engineering and quality assurance labor.

Maintenance is a long-term governance activity. Pathological definitions evolve, new biomarkers emerge, and scanner technologies change. A deployed model cannot remain static. State space models, like other deep learning systems, exhibit performance decay when the input distribution shifts. Continuous monitoring should track concordance with pathologist reviews, flag out-of-distribution slide characteristics, and trigger recalibration or retraining when needed. Because state space models can process streaming sequences, they may be more amenable to online adaptation, but this must be balanced against the risk of unplanned model drift in clinical practice. Lifecycle governance therefore requires versioned model registries, reproducibility protocols, and clear ownership of model updates.

Cross-domain comparisons are instructive. In natural language processing, state space models have been adopted to handle long documents and streaming text. In medical imaging, they have been explored for segmentation and classification tasks. Histopathology differs because the spatial organization of tissue is not a linear narrative; the ordering of patches is an analytical convenience rather than an intrinsic property. This distinction has governance implications. The scanning path used to serialize a two-dimensional slide must be justified and held constant across training and inference. If the path changes, the state sequence changes, and model behavior may shift. Documentation of this serialization strategy is therefore part of the model specification.

8. Policy Implications and Regulatory Trajectories

The adoption of state space models in histopathology will occur within evolving regulatory and policy environments. Regulators increasingly expect evidence that a model is robust,

generalizable, and safe across intended populations [24]. For state space models, this evidence must include evaluation of long-range dependency behavior. It is not enough to report patch-level area under the curve; validation studies should demonstrate that the model captures relevant architectural context without propagating bias across distant regions. This requires benchmark suites that represent diverse tissue types, stain protocols, scanner vendors, and patient populations.

Fairness and accountability policies must go beyond aggregate performance. Health systems adopting histopathology AI should conduct subgroup analyses and community consultation, particularly when the model may influence cancer diagnosis or treatment selection. Model cards can serve as public-facing accountability artifacts, but they must be linked to enforceable institutional policies [21]. If a state space model fails disproportionately for a particular demographic group, there must be a clear process for disclosure, correction, and review. Algorithmic bias in health care is a documented risk [19], and long-range sequence models are not immune simply because they use a different mathematical mechanism.

The policy landscape also includes data sharing and privacy. Whole slide images are difficult to anonymize fully because tissue morphology may contain identifying information in rare cases. Federated learning and distributed governance may help, but they require agreements on data use, model ownership, and liability. State space models with compact state representations may facilitate some privacy-preserving workflows by reducing the need to transmit raw feature maps across institutions, yet this advantage must not be overstated. The state itself can encode sensitive patient information, and access to model parameters or cached states may still constitute a privacy risk.

International harmonization is another forward-looking concern. Regulatory frameworks differ across the United States, the European Union, and other jurisdictions. A model validated in one setting may require additional evidence elsewhere. State space model developers should work with regulatory science communities to define common evaluation protocols for long-range visual models in pathology. The goal should be to create a shared evidentiary standard that supports innovation while protecting patients. This will require input from pathologists, computer scientists, bioinformaticians, ethicists, and health system administrators.

9. Conclusion

State space models represent a promising architectural direction for long-range dependency modeling in histopathological image analysis. Their ability to maintain a bounded state while traversing long sequences addresses a central limitation of both local convolutional processing and global attention. From a systems perspective, this property may reduce memory pressure, enable streaming inference, and support more flexible deployment across hospital infrastructure. However, architectural elegance does not automatically produce clinical readiness. The real value of state space models depends on how they are integrated into the broader computational pathology ecosystem, including data quality pipelines, laboratory information systems, regulatory review, and long-term maintenance.

The trade-offs are structural and organizational as much as technical. Fixed-dimensional states may act as bottlenecks if they are too small, while selective dynamics may increase training complexity. Scanning order and multi-view fusion introduce new documentation obligations. Fairness, accountability, and regulatory compliance demand evidence that goes well beyond aggregate accuracy. Institutions must consider the lifecycle costs of deployment,

from initial training to postmarket surveillance. A model that is efficient at inference but difficult to explain or monitor may impose hidden burdens on clinical teams.

Future research should move beyond benchmark comparisons and examine how state space models behave under real-world shifts in staining, scanning, and patient population. Studies should assess whether the recurrent state can be made auditable and whether long-range representations support or undermine clinical interpretability. Multi-institutional governance models are needed to test federated training and local adaptation without compromising safety. If these challenges are addressed, state space models may become a durable component of computational pathology. If they are not, their architectural advantages will remain confined to controlled academic settings rather than translating into trustworthy clinical systems. The path forward is therefore not only about modeling long-range dependencies more efficiently; it is about building the socio-technical conditions under which such models can be used responsibly.

References

1. Van der Laak, J., Litjens, G., & Ciompi, F. (2021). Deep learning in histopathology: the path to the clinic. Nature Medicine, 27(5), 775–784.

2. Srinidhi, C. L., Ciga, O., & Martel, A. L. (2021). Deep neural network models for computational histopathology: A survey. Medical Image Analysis, 67, 101813.

3. Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Werneck Krauss Silva, V., Busam, K. J., & Fuchs, T. J. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nature Medicine, 25(8), 1301–1309.

4. Ilse, M., Tomczak, J. M., & Welling, M. (2018). Attention-based deep multiple instance learning. Proceedings of the 35th International Conference on Machine Learning, 2127–2136.

5. Shaban, M., Awan, R., Fraz, M. M., Azam, A., Tsang, Y.-W., Snead, D., & Rajpoot, N. M. (2020). Context-aware convolutional neural network for grading of colorectal cancer histology images. IEEE Transactions on Medical Imaging, 39(7), 2395–2405.

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

7. 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.

8. 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.

9. Gu, A., Goel, K., & Ré, C. (2022). Efficiently modeling long sequences with structured state spaces. Proceedings of the International Conference on Learning Representations.

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

11. Zhu, L., Liao, B., Zhang, Q., Wang, X., Liu, W., & Wang, X. (2024). Vision Mamba: Efficient visual representation learning with bidirectional state space model. Proceedings of the 41st International Conference on Machine Learning, 26795–26812.

12. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.

13. Tellez, D., Litjens, G., Bándi, P., Bulten, W., Bokhorst, J.-M., Ciompi, F., & van der Laak, J. (2019). Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. Medical Image Analysis, 58, 101544.

14. Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., ... & Hamilton, P. W. (2017). QuPath: Open source software for digital pathology image analysis. Scientific Reports, 7(1), 16878.

15. Clunie, D. A., Hosseinzadeh, D., Wintell, M., De Mena, D., Lajara, N., Garcia-Rojo, M., ... & Hanna, M. G. (2018). Digital imaging and communications in medicine (DICOM) supplement 145: Whole slide imaging. Journal of Pathology Informatics, 9, 14.

16. Kather, J. N., Heij, L. R., Grabsch, H. I., Loeffler, C., Echle, A., Muti, H. S., ... & Luedde, T. (2020). Pan-cancer image-based detection of clinically actionable genetic alterations. Nature Cancer, 1(8), 789–802.

17. Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., ... & Venâncio, R. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA, 318(22), 2199–2210.

18. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

19. 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.

20. 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.

21. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019). Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency, 220–229.

22. Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., ... & Dean, J. (2018). Scalable and accurate deep learning with electronic health records. npj Digital Medicine, 1(1), 18.

23. Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. Nature Reviews Clinical Oncology, 16(11), 703–715.

24. U.S. Food and Drug Administration. (2021). Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. FDA.