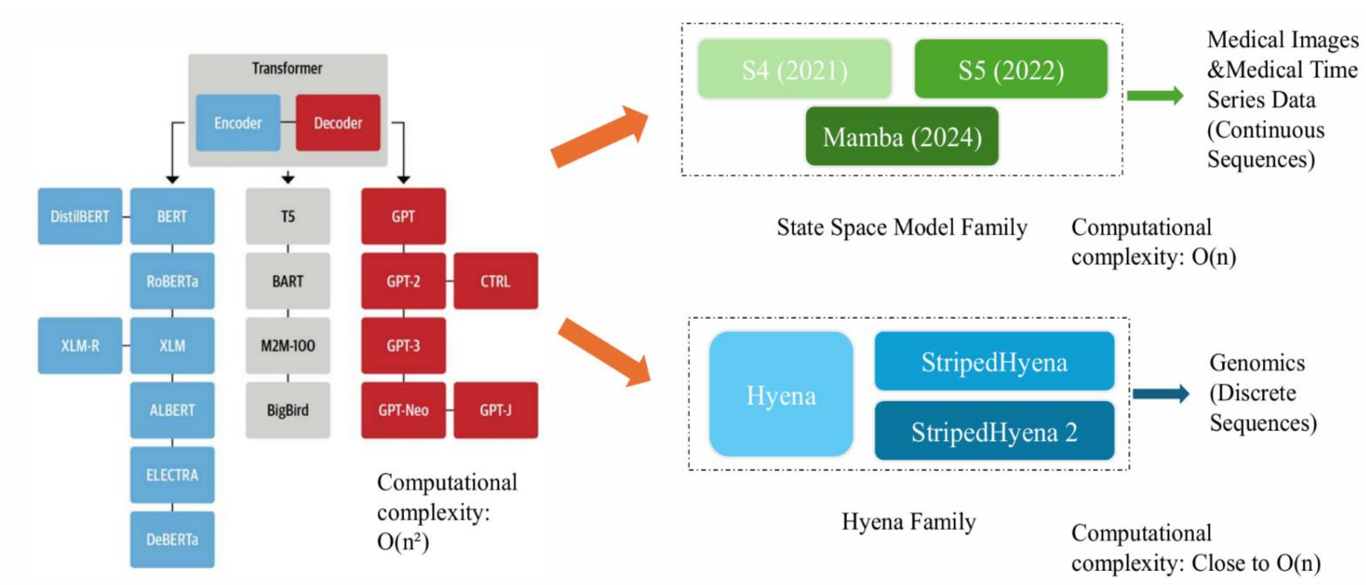
ASEC 2025 Conference
The 6th International Electronic Conference on Applied Sciences
09–11 December 2025 | Online

Mamba in Medical Imaging: A Comprehensive Survey of State Space Models

Jinglin Liang[1], Sijia Zhu[2], Zhe Liu[3,4,*]
1 School of Biological Science, Nanyang Technological University, 637551, Singapore
2 Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore 21218, USA
3 School of Computer Sciences, Universiti Sains Malaysia, Penang 11800, Malaysia
4 College of Mathematics and Computer, Xinyu University, Xinyu 338004, China

INTRODUCTION

The Mamba architecture, based on the Selective State Space Model (SSM), represents a new generation of vision frameworks characterized by linear computational complexity and the capability to efficiently capture long-range dependencies. Owing to these advantages, Mamba and related SSM-based architecture has garnered significant attention in the computer vision community and are increasingly being investigated for applications in medical image analysis.



METHOD

This study provides a comprehensive review of Mamba and Selective State Space Model (SSM) architectures and their emerging applications in medical imaging from 2023 to 2025. We systematically examined representative frameworks—VM-UNet, Mamba-UNet, and 2D-Mamba—that integrate SSM-based modules into medical vision pipelines. The review encompasses a broad spectrum of tasks, including 2D and 3D medical image segmentation, whole-slide pathology classification, and surgical or endoscopic video understanding, highlighting how Mamba's linear-complexity design facilitates efficient long-range dependency modeling in diverse imaging contexts. Furthermore, we conducted a comparative evaluation of these architectures against conventional Transformer and CNN-based baselines, focusing on computational efficiency, scalability to high-resolution data, and training stability. The findings underscore Mamba's potential as a next-generation vision backbone for medical image analysis, capable of bridging the gap between expressive global modeling and practical computational feasibility.

RESULTS & DISCUSSION

State Space Models (SSMs) have recently demonstrated the ability to match or surpass Transformer-based architectures in a range of multimodal medical imaging tasks. Owing to their linear computational complexity and memory-efficient design, SSMs substantially reduce GPU memory usage and computational overhead, thereby facilitating effective training and inference on high-resolution medical images. These properties make SSMs particularly advantageous for high-resolution and long-sequence applications, such as whole-slide pathology analysis and surgical video understanding, where conventional architectures struggle to balance performance and scalability. Nevertheless, challenges persist—training stability remains fragile, and the development of standardized, large-scale pre-training pipelines for medical domains is still at an early stage, underscoring the need for further methodological and infrastructural advancements.

CONCLUSION

The Mamba architecture establishes a new performance–efficiency trade-off well aligned with the computational and practical needs of clinical imaging workflows. By modeling long-range dependencies with linear complexity, it offers strong scalability without sacrificing accuracy. However, key challenges persist, including optimization stability, effective pre-training strategies, and improved interpretability. Future progress will likely emerge from hybrid CNN–Transformer–SSM architectures, enhanced cross-modality generalization, and robust 3D temporal extensions. Ultimately, successful clinical translation will depend on building transparent, reliable, and scalable deployment pipelines suitable for real-world medical use.

Keywords: Mamba Architecture; Selective State Space Model (SSM); Medical Image Analysis; Semantic Segmentation; Computational Efficiency; Long-Range Dependency Modeling; Clinical AI Systems