이름: Thanaporn Viriyasaranon
직위: Postdoctoral Researcher
소속: Ewha Womans University
기타소속:

강연제목: 자기지도 트랜스포머와 적응형 스펙트럼-공간 모델: 의료 영상의 데이터 부족·다중 모달리티 복잡성 극복

From Self-Supervised Transformers to Adaptive Spectral-Spatial Models: Bridging Data Scarcity and Multi-Modal Complexity in Medical Imaging

Abstract:

Medical image analysis faces two critical challenges: (1) annotation scarcity , where deep learning models typically require large-scale labeled datasets to achieve optimal accuracy, yet such datasets are costly and time-consuming to curate in medical imaging due to the necessity for expert annotations; and (2) spectral-spatial heterogeneity , wherein conventional deep learning frameworks (e.g., CNNs, transformers, state space models) struggle to adaptively address multi-scale anatomical variations, modality-specific artifacts, and dynamic frequency signatures.

To address annotation scarcity , we propose a Convolutional Pyramid Vision Transformer (CPT) trained via self-supervised geometric shape segmentation on unlabeled CT scans. By integrating multi-kernel convolutional patch embeddings and local spatial reduction, the CPT learns robust semantic features without reliance on annotations, thereby mitigating overfitting in small datasets (e.g., pancreatic cancer CT, breast MRI). This approach achieves state-of-the-art performance in liver and pancreatic cancer classification while significantly reducing dependence on costly labeled data. To tackle spectral-spatial heterogeneity , we introduce FREXMamba , a novel architecture combining Convolutional State Space Modules (CSSM) and Frequency-Aware Dynamic Expert Routing (FADER-MoE) . FREXMamba dynamically routes modality-specific frequency bands (e.g., low-frequency tumors in CT, high-frequency textures in dermatoscopy) to specialized experts via hybrid gating mechanisms, preserving phase information through FFT-based decomposition. This framework demonstrates superior accuracy across five modalities (X-ray, CT, Ultrasound, MRI, Dermatoscopy) and achieves a 68.3% reduction in computational cost compared to baseline Mamba-based models.

Collectively, these contributions bridge self-supervised feature learning and adaptive spectral-spatial modeling, enabling scalable, annotation-efficient diagnosis across diverse clinical workflows. By synergizing data-driven and architecture-driven innovations, this work advances unified AI solutions for resource-constrained healthcare settings, emphasizing both technical rigor and clinical translatability.

Brief Biosketch:

Dr. Thanaporn Viriyasaranon is a postdoctoral researcher at Ewha Womans University, specializing in advancing deep learning methodologies for medical imaging and biomedical engineering. She earned her PhD in System Health and Engineering (2024), where her doctoral research pioneered innovative neural architectures for medical image analysis, addressing critical tasks such as cancer classification and segmentation, anatomical landmark detection, and prohibited object detection in security imaging. Furthermore, she developed a framework to reduce motion artifacts in cone-beam CT (CBCT) via unsupervised landmark detection and engineered annotation-efficient deep learning models for pancreatic cancer diagnosis using CT imaging. In her current role, Dr. Viriyasaranon focuses on three key areas: (1) developing a multimodal framework that integrates genomic biomarkers with whole-slide imaging (WSI) data to enhance diagnostic precision in bladder tumor analysis; (2)designing robust deep learning models for abnormality diagnosis across multimodal medical imaging, including X-ray, ultrasound, dermatoscopic imaging, CT, and MRI. Her work has been published in high-impact journals such as IEEE Transactions on Biomedical Engineering and Expert Systems with Applications and presented at premier conferences, including MICCAI and ACCV. Recognized for her contributions, she received the Best Paper Award at IPIU 2022, Best Oral Presentation Award at KSIIM 2024, and the Graduate School Award for Dissertation with Highest Honors (2024).