Deep and Model-Merging-Based Continual Learning for Multi-Task Whole Slide Image Analysis

Name BUI CAO DOANH

Laboratory Computing Architecture Laboratory

Supervisor Prof. Yasuhiko Nakashima

Abstract

Digital computational pathology has rapidly evolved to support cancer diagnosis and prognosis. However, Whole Slide Images (WSIs) are extremely large, gigapixel-scale data that present challenges for transfer and storage. With the continuous emergence of new cancer-related tasks, retraining a separate model for each task becomes inefficient and unsustainable in terms of scalability. To address this, continual learning methods have been introduced in computational pathology, enabling models to learn sequentially from new data without forgetting previously acquired knowledge. Nonetheless, existing rehearsal-based and regularization-based continual learning approaches face several limitations. First, rehearsal methods require storing old samples, which conflict with data privacy and storage constraints. Second, most methods assume a fixed number of classes per task, which contradicts the open-ended nature of clinical workflows where classes (e.g., cancer types) are defined by experts or pathologists. Third, they typically assume that all training data are co-located on a single node, which restricts multi-institutional collaboration due to privacy and data transfer limitations. With the emergence of vision-language pathology foundation models (VLMs), information from textual annotations can be leveraged to enhance WSI representations. This thesis develops an efficient continual learning framework based on pathology VLMs, focusing on three key aspects: (1) leveraging VLMs to enable more effective lifelong learning on WSIs; (2) comparing zero-shot VLM performance with training-based continual learning methods; and (3) introducing a buffer-free, distributed approach that reduces inter-institutional data transfer, protects privacy, and facilitates the creation of a unified model for WSI analysis.