Munich Health Foundation Model Symposium



Contribution ID: 32

Type: Poster

Efficient Vision Transformer for End-to-End Whole Slide Image Representation Learning in Cancer Subtyping Classification

Histopathology plays a central role in clinical medicine for tissue-based diagnostics and in biomedical research as a basis for understanding cellular and molecular mechanisms of disease. The microscopic evaluation of morphological information is a key component of computational pathology. The end-to-end whole slide image representation learning imposes a great challenge to the field of computational pathology with enormous sizes (i.e., ranging from 100M pixels to 10G pixels) for both computation and memory. In this work, we train an efficient Vision Transformer for gigapixel whole slide images in an end-to-end manner for cancer subtyping classification task. The model directly takes the whole slide image as a sequence of millions of patches and generates the representations to capture both short-range and long-range dependencies. Experimental results demonstrate that the proposed model effectively encodes gigapixel images and gives better results on cancer subtyping task with previous methods.

Primary author: UMER, Rao Muhammad (Helmholtz Munich)

Co-authors: Mr MADNI, Hussain Ahmad (University of Udine, Italy); Dr MARR, Carsten (Helmholtz Munich)

Session Classification: Break + Posters session

Track Classification: Poster