A Vision-Transformer based pipeline for investigating the association between cancer pathology images, T-cell receptors and immuno-characteristics

Meiling Liu[1], Yang Liu[2], Amy Peng[3], Qianchuan He[1]*.

[1] Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, United States

[2] Department of Mathematics and Statistics, Wright State University, Dayton, OH, United States

[3] Department of Biological Sciences, University of the Pacific, Stockton, CA, United States

Correspondence: Qianchuan He, Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, United States.

Email: qhe@fredhutch.org

Abstract:

Pathology images, such as the hematoxylin and eosin stain (H&E) images, provide a comprehensive characterization of the cellular and tissue structures of tumor samples, and have the potential to serve as an important predictor for tumor progression and prognosis. T-cell receptor (TCR) repertoire, a complex set of proteins that are expressed on the surface of T-cells and responsible for recognizing tumor antigens, play crucial roles in immune responses against cancer and can significantly influence the progression of cancer. It is worthwhile to study if pathology images and T-cell receptor (TCR) repertoire are associated with cancer immuno-characteristics, because these characteristics represent the infiltration of immune cells in tumor tissue and can serve as a surrogate for immunotherapy efficacy. Here, we develop a Vision-Transformer-based AI pipeline to investigate whether and how well the immune score could be predicted from histological slides and TCR using deep learning approaches. We apply our pipeline to the immune score which is derived from a large set of immune gene expressions, and show that pathology image and TCR can make meaningful contribution to the prediction of the immune score. We further show that our pipeline can be used to predict binary clinical outcomes of cancer patients. Overall, our proposed AI pipeline provides a timely tool for incorporating pathology images and TCR repertoire for clinical outcome prediction.