

Semi-Supervised Learning for Postoperative Early Recurrence Prediction of Hepatocellular Carcinoma Using Histopathological Image

Gan Zhan¹, Fang Wang², Yinhao Li¹, Qingqing Chen², Hongjie Hu², and Yen-Wei Chen¹

¹Graduate School of Engineering, Ritsumeikan University, Japan

² Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, China
E-mail: gr0502vs@ed.ritsumei.ac.jp

Hepatocellular Carcinoma, which is also known as HCC, is one of the primary liver cancers around the world, and it accounts for a high mortality rate in clinical. Surgical resection is the first-line treatment choice for patients with HCC, but the problem is patients will still have a great chance of recurrence after surgery, and patients with early recurrence tends to have a shorter survival time than patients with non-early recurrence. Therefore, postoperative early recurrence prediction of HCC patients is necessary, which can guide physicians to manage the individualized follow-up so that we could increase the survival time of patients.

In the realm of postoperative whole-slide image (WSI) analysis, annotation-free methodologies, including multi-instance learning and self-supervised learning, have showcased impressive performance. Nevertheless, such annotation-free approaches typically necessitate abundant data and substantial computing power to compensate for the absence of existing pathology priors during training. On the other hand, annotation-based methods involve a two-step process for WSI analysis. Initially, pathologists provide annotations to generate tissue segmentation maps of WSIs, which are subsequently utilized to perform desired tasks. While annotation-based methods offer the advantage of incorporating pathology priors, the process of fully annotating WSIs proves to be time-consuming and labor-intensive due to complex boundaries and large scales. Moreover, tissue boundaries often encompass other tissues that are difficult to exclude via annotation, thereby introducing potential noise and erroneous label data during WSI tissue segmentation learning.

To address these challenges, we propose a semi-supervised learning approach for early recurrence prediction of HCC using WSIs. Our proposed method adopts an annotation-based framework, which also involves two steps. However, we depart from the conventional practice of fully annotating WSIs and instead, pathologists only annotate a few tiny areas for each tissue category within gigapixel WSIs. Subsequently, we could still generate high quality tissue segmentation heatmaps for each WSI based on these provided minor annotations. Thereafter, we construct our prediction model employing these tissue segmentation heatmaps. In comparison to both annotation-free and fully annotation-based methods, our proposed approach leverages the pathology priors provided by the minor annotations, thereby mitigating the demand for extensive data and computing power. Furthermore, the minor annotations significantly reduce the time and labor required, while concurrently yielding promising results. In addition to early recurrence prediction, we performed a comprehensive tissue analysis based on the generated tissue segmentation heatmaps. Our investigation revealed the pivotal role of stromal tissue in the early recurrence of HCC.