Graph Based Representation and Classification of Whole Slide Tissue Images

Marion Schoenenweid

Master thesis in Computer Science

Cancer diagnosis depend on a precise and rapid analysis of tissue images by expert pathologists in order to increase the survival outcome and provide adequate treatment to patients. The analysis of histological images is the gold standard for the elaboration of the diagnosis. Whole Slide Images, the digital version of those images, since they have been validated as non inferior to classical glass slides for visual analysis, have lead to increase researches in the Computer Aided Diagnosis domain, and in particular in the area of artificial intelligence.

To alleviate the technical challenges associated with Whole Slide Images, such as their great size and the difficulty to obtain large labelled datasets, the present work proposes a fully unsupervised process for the representation of the images and their classification by leveraging graphs techniques.

This exploratory project is divided into two parts. Where researches usually address the problem of the size of Whole Slide Images by using patches or manually selected Region Of Interest to used with deep learning techniques, this project first proposes to represent the entire images as graphs. These structures allow to keep the topological information of tissues, which is particularly important for colorectal cancer diagnosis, and are memory efficient. This first step leverage superpixels for the segmentation of the images and Region Adjacency Graph in order to create the graphs.

The second step regards the classification of those graphs. Encouraged by recent studies showing promising results with deep learning approaches adapted to graphs, we tested a graph classification problem on the graphs previously created, through a Graph Neural Network. As proof of concept, the proposed classification focuses on the four T stages of the Tumour-Node-Metastasis system. Even if the presented results here did not reach the expected outcome, similar state-of-the-art results confirm that the hypothesis tested is a valid approach, that is worth to be further explored.

Prof. Rolf Ingold