Automatic Prostate Cancer Detection and Localisation on Digital Histopathology Images

Wenchao Han
Supervisor: Aaron Ward

**Purpose:** The successful development of a system for computer-assisted pathologic prostate cancer quantification on digital histology images of the post-surgery specimen will provide accurate and repeatable decision support for guidance of potentially life-saving adjuvant therapy. To achieve that goal, our objectives are: (1) to develop an automatic system labeling tissue components on digital histology images, (2) to use labeled tissue component distributions to detect cancer regions, and (3) to grade cancer on digital histology images. The work presented in this abstract focuses on steps (1) and (2).

**Methods:** Each haematoxylin and eosin (H&E) stained digital histology image of surgically removed prostate tissue sections was automatically labeled by our system with three tissue components (lumen, stroma, nuclei) to generate tissue component maps. Nuclei were segmented by adaptively selecting a threshold to optimize the number of connected components in a hematoxylin image obtained through colour deconvolution. Lumen regions were detected by setting thresholds within the red-green-blue colour space. All remaining tissue was labeled as stroma. We automatically calculated the proportions of those tissue components, as well as 22 additional texture features for each 480×480 μm subregion of the image, and used machine learning methods to classify it as cancerous or non-cancerous. The experiment was conducted using 20 different mid-gland histology images from 20 different patients using leave-one-patient-out cross validation.

**Results:** Our tissue component labeling algorithms generated qualitatively accurate tissue component maps, overcoming staining differences among different digital histology images. Quantitative evaluation for cancerous vs. noncancerous tissue classification showed an overall error rate of 17.5% ± 10.1% and an area under the receiver operating characteristic curve (AUC) of 90.8% ± 9.1%.

**Conclusion:** The combination of tissue component proportions and texture features yields an accurate machine learning-based classification of cancer vs. non-cancer within small subregions of prostate digital histopathology images. The next step is to refine the classification for automated cancer grading, leading to full quantification of the positions and sizes of high-grade lesions. This will ultimately lead to more quantitative and repeatable clinical pathology reporting, providing improved decision support for adjuvant therapy post-prostatectomy.