Computer-aided histological analysis for prostate cancer diagnosis

Prostate cancer is the most frequently diagnosed non-skin cancer and the second leading cause of cancer death among men in the United States. Conventionally, pathologists review prostatic tissue visually under a microscope and a diagnosis of prostate cancer is rendered basedon subjective perceptionsofa numberof histological features. Thehypothesisof this dissertation is that computer histological image-processing techniques can be developed to detect prostate cancer automatically at a level of accuracy comparable to an expert pathologist. The long-term goal of this research is to develop a computer-aided diagnosis (CAD) system that can be used to help pathologists diagnose prostate cancer with improved accuracy, reproducibility, and efficiency. In this dissertation research, I developed computer image-processing techniques for classifying prostate cancer and non-malignant tissue in digital histological images acquired from tissue sections with immunohistochemistry (IHC) and hematoxylin and eosin (H&E) stains.

Digital color histological images were acquired using light microscopes and charge-coupled device (CCD) cameras from regions of interest (ROIs) marked by pathologists on prostatic tissue sections collected from Northwestern University and the University of Chicago. I used an image-artifact correction method to remove image artifacts caused by image acquisition. The consistencyofthe artifact correctionwas testedintwoexperiments.

I developed a CAD technique for images of tissue sections processed with a triple-antibody-cocktail immunostaining technique, combining alpha-methylacyl-CoA racemase (AMACR, a prostate cancer marker), high-molecular-weight cytokeratin (HMWCK, a negative prostate cancer marker), and p63 (another negative prostate cancer marker), for prostate cancer classification. The CAD system, based on color

recognition and a rule-based classifier, was developed using 35 (17 malignant) images. Blinded validation on an independent database of 299 (114 malignant) images achieved the sensitivity and specificity of 85%

and 89%, respectively; and 88% and 97%, when we excluded images that initially lacked a consensus diagnosis (n = 27).

For digital histological imagesof tissue sections processed with routine H&E staining,I developed image segmentation methods to segment tissue components, glands, and epithelial cell nucleoli. Accuracyof the segmentation of glands and cell nucleoli was evaluated by comparing the segmentation results with manual outlines and/or by subjective assessment. Image features were then calculated from the segmented structures to measure qualitative histological features of prostate cancer quantitatively. I used receiver operating characteristic (ROC) analysis to evaluate the performance of image features in classifying prostate cancer and non-malignant tissue. The area under the curve (AUC) values were equal to or greater than 0.95 on three independent datasets (n=62, 142, and 148, respectively) when promising features were combined by linear discriminant analysis (LDA) classifiers using the leave-one-out (LOO) cross-validation method.

This dissertation provides preliminaryevidence that supports thehypothesis that appropriate CAD techniques can be developed for classifying prostate cancer and non-malignant tissue accurately in digital histological images of tissue sections with both IHC and H&E stains. With further development in image processing, image feature extraction, and classification, it is likely that CAD systems could be developed to classify prostate cancer and benign tissue accurately.

Yahui Peng
Medical Physics Program
University of Chicago