# Active Learning for Scalable Histopathological Image Analysis

Yan Zhu

Dept. of Computer Science

9/22/2016 at 03:00 pm
CBIM 22

## Abstract

Training an effective and scalable system for medical image analysis usually requires a large amount of labeled data, which incurs a tremendous annotation burden for pathologists. Recent progress in active learning can alleviate this issue, leading to a great reduction on the labeling cost without sacrificing the predicting accuracy too much.

In this talk, I will talk about a novel batch-mode active learning method which explores and leverages such structured information in annotations of medical images to enforce diversity among the selected data, therefore maximizing the information gain. We formulate the active learning problem as an adaptive submodular function maximization problem subject to a partition matroid constraint, and further present an efficient greedy algorithm to achieve a good solution with a theoretically proven bound. We demonstrate the efficacy of our algorithm on thousands of histopathological images of breast microscopic tissues. Well conclude the talk and indicate ongoing work in the end.

Examination Committee: Prof. Dimitris N. Metaxas (chair), Prof. Vladimir Pavlovic, Prof. Kostas Berkris and Prof. Casimir Kulikowski