Prostate Cancer: Markers II

Podium 52

Sunday, May 17, 2020 **3:30 PM-5:30 PM**

PD52-01

A DEEP LEARNING ALGORITHM FOR THE DIAGNOSIS AND GLEASON GRADING OF WHOLE SLIDE IMAGES OF PROSTATE CANCER CORE BIOPSIES

Ohad Kott, Miami Beach, FL; Sizhe Li, Rochester, NY; Drew Linsley, Ali Amin, Bora Golijanin, Dragan Golijanin, Thomas Serre, Providence, RI; Boris Gershman, Boston, MA*

INTRODUCTION AND OBJECTIVE: Deep learning algorithms have shown promising early results in the automated diagnosis and grading of prostate cancer. However, training such algorithms typically requires a large amount of manually annotated training data. Herein, we developed a weakly supervised, deep learning approach for the diagnosis and Gleason grading of whole-slide images of prostate core biopsies.

METHODS: We digitized 3,680 prostate biopsy cores as whole-slide images from 291 patients at 20x magnification. The main challenge for a whole-slide classification approach, as opposed to a tile-based approach, arises because of hardware memory limitations. We thus trained and tested a two-stage classification pipeline. First, an encoder network was trained using a multiple-instance learning setting to extract features from every tile of a given core image. We then trained a second-stage classifier for (1) cancer diagnosis (benign vs. malignant) and (2) primary Gleason scoring (benign vs 3 vs 4-5). Heatmaps were generated using Grad-CAM to produce a localization map of the class-discriminative regions in the image.

RESULTS: The model demonstrated an accuracy of 94.4% for the classification of prostate biopsy cores as benign vs. malignant (95.7% sensitivity, 93.9% specificity, and 94.7% average precision). The model achieved 93.0% accuracy for the classification of biopsy cores as benign vs Gleason 3 vs Gleason 4-5 (87.3% sensitivity, 98.9% specificity, and 93.3% average precision) (Figure 1). Heat maps confirmed network sensitivity to malignant image regions as confirmed by a trained pathologist (Figure 2).

CONCLUSIONS: In this study, a weakly supervised deep learning algorithm demonstrated excellent performance for the diagnosis of prostate cancer and the classification of primary Gleason score for whole slide images of prostate core biopsies.

Figure 1: Confusion matrix for primary Gleason score classification as benign vs Gleason 3 vs Gleason 4-5.

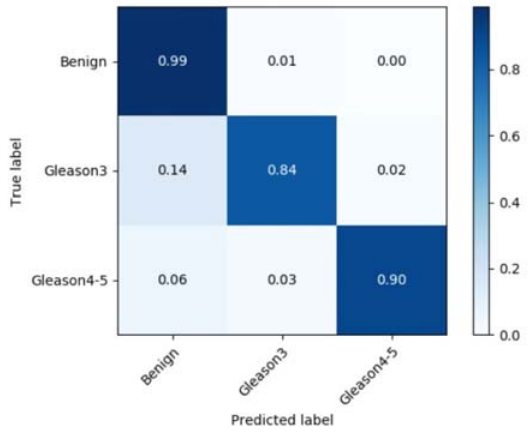
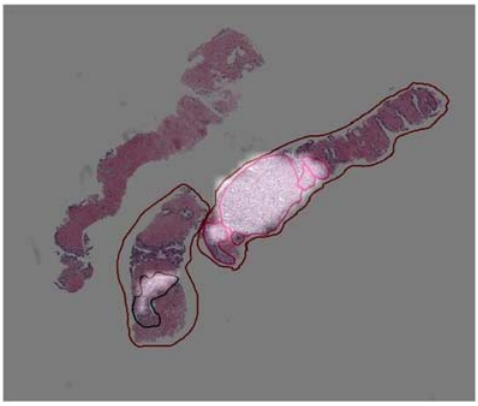


Figure 2: Heat map generated using Grad-CAM to identify class-discriminative regions in the image.



Source of Funding: This study was supported by NIGMS/Advance-CTR through the IDeA-CTR grant NIGMS/Advance-CTR (U54GM115677). Additional support provided by ONR (N00014-19-1-2029), NSF (IIS-1912280 and EAR-1925481), DARPA (D19AC00015), the ANR-3IA Artificial and Natural Intelligence Toulouse Institute (ANITI), the Carney Institute for Brain Science and the Center for Computation and Visualization. We also acknowledge the Cloud TPU hardware resources that Google made available via the TensorFlow Research Cloud (TFRC) program.

PD52-02

COMPUTER-EXTRACTED FEATURES OF GLAND MORPHOLOGY FROM DIGITAL TISSUE IMAGES IS COMPARABLE TO DECIPHER FOR PROGNOSIS OF BIOCHEMICAL RECURRENCE RISK POST-SURGERY

Patrick Leo, Robin Elliott, Andrew Janowczyk, Cleveland, OH; Nafiseh Janaki, Boston, MA; Kaustav Bera, Rakesh Shiradkar, Cleveland, OH; Ayah El-Fahmawi, Jessica Kim, Mohammed Shahait, Abhishek Shah, Philadelphia, PA; Hari Thulasidass, Ashutosh Tewari, New York City, NY; Sanjay Gupta, Cleveland, OH; Natalie Shih, Michael Feldman, Priti Lal, David Lee, Philadelphia, PA; Anant Madabhushi, Cleveland, OH*

INTRODUCTION AND OBJECTIVE: It is currently clinically challenging to stratify patients' risk of biochemical recurrence (BCR) of localized prostate cancer following radical prostatectomy (RP). Decipher, a 22-gene genomic classifier, is currently a part of NCCN guidelines to determine risk of metastasis and BCR following RP with improved accuracy beyond existing clinicopathologic factor based nomograms. Recently, computer extracted quantitative histomorphometric (QH) analysis of hematoxylin and eosin (H&E) images alone has shown increasing value in predicting risk of BCR following RP. In this work, we sought to compare the prognostic ability of QH against Decipher in BCR prognosis post-RP. As compared to Decipher, QH is non tissue-destructive, less time consuming, and cheaper.

METHODS: A single diagnostic slide was collected from N=388 patients from three institutions and compose. Patients were split into training (N=215) and validation (N=173) sets. One institution was split across the training and testing set according to availability of Decipher results. Each slide was annotated for a single large, representative cancer region, from which 26 texture features were extracted. 41 training set patients were used to train a deep learning model for lumen segmentation, which was applied to all the cancerous regions. From the lumen segmentations, 216 features of lumen arrangement, shape, and disorder were extracted. Training set patients were used to