# A Deep Learning Algorithm for the Diagnosis and Gleason Grading of Whole Slide Images of Prostate Cancer Core Biopsies

Ohad Kott[1], Sizhe Li[2], Drew Linsley[3], Ali Amin[4], Bora Golijanin[1], Dragan Golijanin[1], Thomas Serre[3], Boris Gershman[5]

[1]Minimally Invasive Urology Institute, The Miriam Hospital; Warren Alpert Medical School of Brown University, Providence, RI
[2]University of Rochester, Rochester, NY
[3] Carney Institute for Brain Science, Department of Cognitive Linguistic & Psychological Sciences, Brown University, Providence, RI
[4]Department of Pathology, The Miriam Hospital; Warren Alpert Medical School of Brown University, Providence, RI
[5]Division of Urologic Surgery, Beth Israel Deaconess Medical Center, Boston, MA

# 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

# Disclosures

- Dr. Serre serves on the scientific advisory board for Vium, Inc.
- The remaining investigators have nothing to disclose

# Introduction

- The pathologic diagnosis and grading of prostate cancer is time-consuming, error-prone, and subject to inter-observer variability
- 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

# Objective

- To develop a weakly supervised, deep learning approach for the diagnosis and Gleason grading of whole-slide images of prostate core biopsies

# Methods

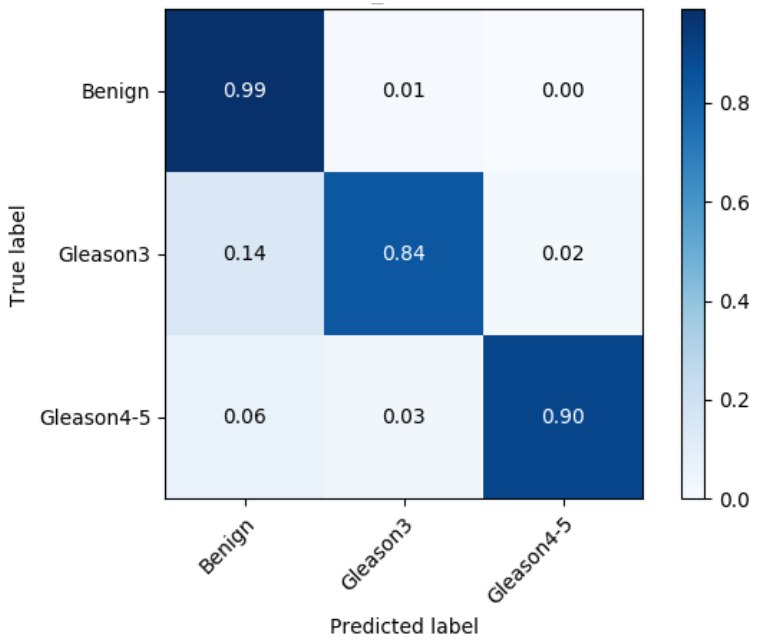
- 3,680 prostate core biopsy specimens were digitized from 291 patients as whole slide images at 20x magnification
- Two-stage classification pipeline for whole slide image classification:
  - Encoder network trained using multiple instance learning to extract features from tiles of a given core image
  - Second stage classifier provided a classification for each tile
- Two classification tasks:
  - (1) Benign vs. malignant
  - (2) Primary Gleason score: benign vs. 3 vs. 4-5
- Heatmaps 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
  - Sensitivity: 95.7%
  - Specificity: 93.9%
  - Average precision: 94.7%

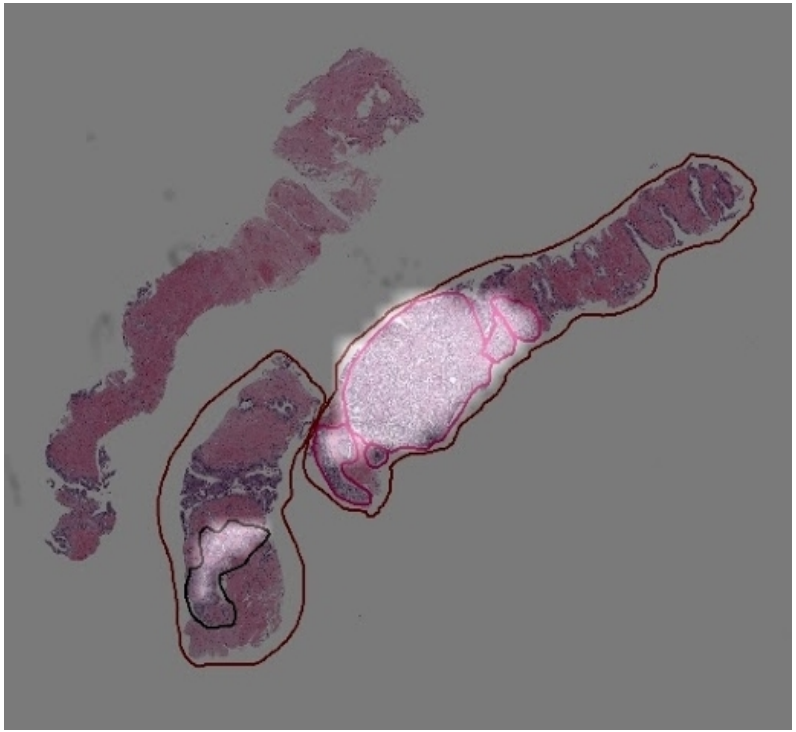
# Results

- The model achieved 93.0% accuracy for the classification of biopsy cores as benign vs Gleason 3 vs Gleason 4-5
  - Sensitivity: 87.3%
  - Specificity: 98.9%
  - Average precision: 93.3%



# Results - Heat map generated using Grad–CAM

- Heat maps confirmed network sensitivity to malignant image regions as confirmed by a trained pathologist



# Conclusion

- 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

- Additional studies are planned to externally validate the algorithm and to improve model performance