University of Pittsburgh | School of Computing and Information

Dissertation Defense
Doctor of Philosophy in Computer Science

"**Deep Learning for Medical Imaging from Diagnosis Prediction to its Explanation**"
by **Sumedha Singla**

Date: July 12, 2022
Time: 10:00AM – 12:00PM
Place: Room 5317, Sennott Square, 3810 Forbes Ave., Pittsburgh, PA 15213
https://pitt.co1.qualtrics.com/jfe/form/SV_77dBuycfy4bQptI

Committee:

- Dr. Kayhan Batmanghelich, Assistant Professor, Department of Biomedical Informatics, University of Pittsburgh School of Medicine
- Dr. Milos Hauskrecht, Professor, Department of Computer Science, School of Computing and Information
- Dr. Adriana Kovashka, Associate Professor, Department of Computer Science, School of Computing and Information
- Dr. Jia Xiaowei, Assistant Professor, Department of Computer Science, School of Computing and Information
- Dr. Sofia Triantafyllou, Assistant Professor, Department of Mathematics and Applied Mathematics, University of Crete

Abstract:
Deep neural networks (DNN) have achieved unprecedented performance in computer-vision tasks almost ubiquitously in business, technology, and science. While substantial efforts are made to engineer highly accurate architectures and provide usable model explanations, most state-of-the-art approaches are first designed for natural vision and then translated to the medical domain. This dissertation seeks to address this gap by proposing novel architectures that integrate the domain-specific constraints of medical imaging into the DNN model and explanation design.
Prior work on DNN design commonly performs lossy data manipulation to make volumetric data compatible with 2D or low-resolution 3D architectures. We proposed a novel DNN architecture that transforms volumetric medical imaging data of any resolution into a robust representation that is highly predictive of disease. For DNN model explanation, current explanation methods primarily focus on highlighting the essential regions (where) for the classification decisions. The location information alone is insufficient for applications in medical imaging. We designed counterfactual explanations to visually demonstrate how adding or removing image-features changes the DNN decision to be positive or negative for a diagnosis.
Further, we reinforced the explanations by quantifying the causal relationship between neurons in DNN and relevant clinical concepts. These clinical concepts are derived from radiology reports and are corroborated by the clinicians to be useful in identifying the underlying diagnosis. In the medical domain, multiple conditions may have a similar visual appearance, and it's common to have images with conditions that are novel for the pre-trained DNN. DNN should refrain from making over-confident predictions on such data and mark them for a second reading. Our final work proposed a novel strategy to make any off-the-shelf DNN classifier adhere to this clinical requirement.