Proposal Defense

Doctor of Philosophy in Intelligent Systems

“Trustworthy Deep Learning on Medical Images” by Degan Hao

Date: February 9, 2024
Time: 1:00 p.m. – 2:00 p.m.
Place: https://pitt.co1.qualtrics.com/jfe/form/SV_268kEVcaDi2jQ6a

Committee:
- Shandong Wu, Ph.D., Intelligent Systems Program, School of Computing and Information, University of Pittsburgh
- Pradeep Reddy Raamana, Ph.D., Intelligent Systems Program, School of Computing and Information, University of Pittsburgh
- Ahmad Pahlavan Tafti, Ph.D., Intelligent Systems Program, School of Computing and Information, University of Pittsburgh
- Marius George Lingurar, DPhil MA MSc, Department of Radiology, School of Medicine and Health Sciences, George Washington University

Abstract:
This PhD research proposal focuses on advancing trustworthy artificial intelligence (AI) in medical imaging, a field where deep learning has shown immense promise but faces challenges in reliability and safety. The initial phase of the research addressed crucial issues such as the accuracy of AI models with limited or inaccurate labels, the importance of model explainability, and the integration of clinical knowledge to improve disease prognosis and lesion localization. A significant achievement was enhancing the adversarial robustness of these models against white-box attacks.

The proposed research aims to build upon these foundations by concentrating on the defense against black-box adversarial attacks, a critical threat to AI in medical diagnostics. A novel approach involving label-independent data augmentation is proposed, ensuring stable model performance on both unaltered and adversarially modified data. Additionally, the research will develop a bias-resilient feature learning method, aimed at reducing the impact of biased synthetic data on model training. This approach is vital for creating fair deep learning models, contributing significantly to safe, equitable and trustworthy healthcare.