THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING SPEAKER SERIES

PRESENTS

Robustness at Inference: Towards Explainability, Uncertainty, and Intervenability

Dr. Mohit Prabhushankar
Email: mohit.p@gatech.edu

Monday, September 16, 9:55 am – 11 am Central Time

Zoom: https://uh-edu-cougarnet.zoom.us/j/9762699678?pwd=RUp5ZmN3cHUyQ1FvUExVQjVsc1hVUT09
Meeting ID: 976 269 9678
Passcode: K91Bwy

LECTURE ABSTRACT

Neural networks provide generalizable and task independent representation spaces that have garnered widespread applicability in image understanding applications. The complicated semantics of feature interactions within image data has been broken down into a set of non-linear functions, convolution parameters, attention, as well as multi-modal inputs among others. The complexity of these operations has introduced multiple vulnerabilities within neural network architectures. These vulnerabilities include adversarial samples, confidence calibration issues, and catastrophic forgetting among others. Given that AI promises to herald the fourth industrial revolution, it is critical to understand and overcome these vulnerabilities. Doing so requires creating robust neural networks that drive the AI systems. Defining robustness, however, is not trivial. Simple measurements of invariance to noise and perturbations are not applicable in real life settings. In this tutorial, we provide a human-centric approach to understanding robustness in neural networks that allow AI to function in society. Doing so allows us to state the following: 1) All neural networks must provide contextual and relevant explanations to humans, 2) Neural networks must know when and what they don't know, 3) Neural Networks must be amenable to being intervened upon by humans at decision-making stage. These three statements call for robust neural networks to be explainable, equipped with uncertainty quantification, and be intervenable.

SPEAKER BIOSKETCH

Mohit Prabhushankar received his Ph.D. degree in electrical engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, Georgia, 30332, USA, in 2021. He is currently a Postdoctoral Research Fellow in the School of Electrical and Computer Engineering at the Georgia Institute of Technology in the Omni Lab for Intelligent Visual Engineering and Science (OLIVES). He is working in the fields of image processing, machine learning, active learning, healthcare, and robust and explainable AI. He has authored and co-authored more than 50 articles in international journals, conferences, and workshops. He has been issued five US patents. He is an IEEE Member.

Mohit Prabhushankar is the recipient of the Best Paper award at ICIP 2019 and Top Viewed Special Session Paper Award at ICIP 2020. He is the recipient of the ECE Outstanding Graduate Teaching Award, the CSIP Research award, and of the Roger P Webb ECE Graduate Research Assistant Excellence award, all in 2022.

Mohit Prabhushankar has delivered several short courses and tutorials at conferences including IEEE IV'23, ICIP'23, BigData'23, WACV'24 and AAAI'24. He has delivered an invited Educational short course for IEEE Signal Processing Society on Explainable AI. In Spring and Fall 2024, he served as the primary instructor for ECE 4252/8803/8803-QF Fundamentals and Applications of Machine Learning course at Georgia Tech. He has served ECE as a Teaching Fellow for ECE 2026 Signals and Systems course during Summer 2020, Fall 2020, Spring 2021, Summer 2021, and Fall 2021, playing a significant role in transitioning to online materials during the global pandemic.

Mohit Prabhushankar has been involved in several societal activities. He has served as a reviewer in prestigious conferences including NeurIPS, CVPR, and ICIP as well as journals including IEEE Transactions on Image Processing, IEEE Signal Processing, Elsevier Image Communications, Nature Scientific Data among others. He has served as a Session Chair at ICIP 2022, Big Data 2023, MIPR 2024, and as an Area Chair at ICIP 2024. He co-organized the IEEE Signal Processing Society's Video and Image Processing Cup at ICIP 2023 and served as one of the judges. He has published six large scale and open source datasets in diverse fields including Autonomous Driving, Healthcare, computer vision, and Human-Computer Interaction.