Artificial Intelligence Initiative (Aii)

https://ai.ucf.edu/



Invited Speaker

Dr. Zhenyi Wang, University of Maryland

"Building Adaptive and Resilient AI: Bridging Continual Learning and Model Extraction Defense"

Wednesday, April 16, 2025 • 2:00PM • HEC 101A • Zoom

Abstract:
Deep learning has driven remarkable progress across diverse domains. However, most methods assume stationary data distributions, which rarely hold in real-world applications. In practice, data distributions evolve over time, creating significant challenges such as catastrophic forgetting and the need to balance performance across both past and newly acquired knowledge. These challenges are particularly critical in two key settings: (1) Resource -constrained environments (e.g., edge devices), where memory and computational resources are limited. (2) Large language models (LLMs), where initial pre-training data is often unavailable after deployment, making adaptation difficult.

In this talk, I will present strategies to address these challenges from two complementary perspectives: (1) Data-centric approaches, which focus on optimizing memory buffer evolution for continual learning; (2) Model-centric approaches, which reduces parameter sensitivity to retain past knowledge while efficiently incorporating new information. I will then establish a novel connection between continual learning and model extraction defense, demonstrating how continual learning principles can enhance the efficiency of protecting pre-trained models behind black-box APIs, safeguarding both functionality and architecture from model extraction attacks. Finally, I will discuss future directions for enabling continual learning in dynamic, real-world settings.

Bio:
Zhenyi Wang is a Postdoctoral Associate at the University of Maryland, College Park. He earned his Ph.D. in Computer Science from the State University of New York at Buffalo. His research focuses on continual learning, trustworthy AI, and data-efficient learning. He has published more than 30 papers in top-tier conferences and journals, including ICML, ICLR, NeurIPS, CVPR, ICCV, ECCV, ACL, and TPAMI.