UCF

Artificial Intelligence Initiative (Aii)

https://ai.ucf.edu/



Invited Speaker

Dr. Jaehong Yoon, University of North Carolina at Chapel Hill

"Trustworthy and Continually Adaptable Multimodal AI Systems"

Thursday, April 10, 2025 • 11:00AM • HEC 101A • Zoom

Abstract:
As our world advances, AI systems must evolve beyond static models to become more adaptive and continuously improving. These systems should seamlessly integrate new knowledge and capabilities throughout their operational lifespan while ensuring safe and robust interactions in dynamic, multimodal environments. Achieving this vision requires addressing key challenges across multiple fields, with a particular focus on memory, integration, adaptation, and reliability in multimodal AI systems. In this talk, I will introduce these challenges and present approaches that enable AI to continuously evolve while maintaining robustness, reliability, and safety in response to the growing complexity of the real world and user demands. First, I will explore multiple essential aspects to develop multimodal models that continuously accumulate and integrate knowledge from new tasks, datasets, or modalities. I will then introduce approaches that promote long-term model evolution and generalization while effectively mitigating catastrophic forgetting. Second, I will present methods to enhance a model's ability to acquire new capabilities and further refine its existing skills through the generation of targeted and controllable synthetic data. These techniques empower models to expand their knowledge beyond initial training, enabling more precise and efficient adaptation to complex and long-horizon challenges. Finally, I will highlight the importance of ensuring model reliability and safety by enhancing reasoning, mitigating unintended biases, and preventing harmful outputs. These advancements are essential for developing trustworthiness AI systems that can evolve without compromising their integrity and effectiveness. I will conclude by outlining future directions for developing a scalable and trustworthy embodied continual learning agent and its applications in high-stakes real-world domains.

Bio:
Jaehong Yoon is a postdoctoral research associate at the University of North Carolina at Chapel Hill, working with Prof. Mohit Bansal. He completed his Ph.D. in the School of Computing at KAIST advised by Prof. Sung Ju Hwang. His primary research focuses on developing continually adaptable, trustworthy, interactive AI systems addressing a wide range of problems in a dynamic and rapidly changing multimodal world. He is the recipient of multiple awards, including the Best Ph.D. Dissertation Award by both the KAIST College of Engineering and the School of Computing, the best student paper award at the ICML '20 Federated Learning Workshop, and a PaliGemma Academic Program Award from Google (2024). He also serves on the program committee of leading NLP and ML conferences and was an area chair for EMNLP 24', NAACL '24, and NeurIPS '24 Workshop on Scalable Continual Learning for Lifelong Foundation Models, and a senior reviewer for CoLLAs '25.