UNIVERSITY OF CENTRAL FLORIDA

CENTER FOR RESEARCH IN COMPUTER VISION

Dr. Du Tran

Facebook AI Research

"Video Understanding: from Spatio-Temporal Representation Learning to Open-World Long-Term Reasoning"

Monday, November 21, 2022 · 2:00PM · Zoom



ABSTRACT

Video understanding is one of the fundamental problems in computer vision with various applications, including autonomous vehicles, robot learning, and visual perception. Compared with traditional image understanding, video understanding: (i) has higher model complexity and requires learning from a much larger amount of data; (ii) requires more expensive annotations; (iii) and sometimes demands long-term video modeling and flexible open-world tasks. In this talk, I will present some of our approaches addressing these challenges, such as efficient and scalable spatiotemporal learning, cross-modal self-supervised learning of video and audio representations, and open-world instance segmentation. Finally, I will outline several potential future research directions in this area.

BIOGRAPHY

Du Tran is a staff research scientist at Facebook AI Research. He graduated with a Ph.D. in computer science from Dartmouth College and an M.S. in computer science from the University of Illinois at Urbana-Champaign, receiving the Dartmouth Presidential Fellowship and the Vietnam Education Fellowship. His research interests are in computer vision, machine learning, and computer graphics, with specific interests in video understanding, representation learning, and multimodal modeling. His work on C3D was instrumental in steering the field towards the widespread adoption of 3D CNNs as the model of choice for video analysis. His video understanding architectures have been deployed in production at Facebook, to process hundreds of millions of videos daily for various tasks, including video classification, violence prediction, and advertisement ranking.