Talk Title: Towards Open World Video Event Understanding - Flexible Representations, Commonsense Priors, and Self-Supervised Learning

Sudeep Sarkar, Professor and Chair
Computer Science and Engineering
University of South Florida, Tampa

Abstract
Events are central to the content of human experience. From the constant stream of the sensory onslaught, the brain segments, extracts, represents aspects related to events, and stores in memory for future comparison, retrieval, and re-storage. Contents of events consist of objects/people (who), location (where), time (when), actions (what), activities (how), and intent (why). Many deep learning-based approaches extract this information from videos. However, most methods cannot adapt much beyond what they were trained and are incapable of recognizing new events beyond what they were explicitly programmed or trained. The main limitation of current event analysis approaches is the implicit closed world assumption. The ability to support open world inference is limited by three main aspects: the underlying representation, the source of semantics, and the ability to continuously learn or adapt.

In this talk, I will focus on flexible representations, amenable for open-world, and self-supervised learning that is not dependent on the existence of a large amount of training data. We will see how Grenander's pattern theory-based canonical representation offers an elegant, flexible, compositional mechanism. It naturally models semantic connections between what is observed directly in the image and prior knowledge in large-scale commonsense knowledge bases, such as ConceptNet. The use of knowledge bases such as ConceptNet allows expanding the set of primary objects and actions to a very large (not infinite) set without the need for massive annotated training sets. And finally, if we have time, how predictive learning can be used to continuously learn how to segment a video into elementary event segments, again without training annotations.

Biography:

Dr. Sudeep Sarkar is a Professor and Chair of Computer Science and Engineering and the Associate Vice President for Special Programs at the University of South Florida in Tampa. He received his B. Tech. degree from the Indian Institute of Technology, Kanpur, M.S. and Ph.D. degrees in Electrical Engineering, on a University Presidential Fellowship, from The Ohio State University, Columbus. He has 25-year expertise in computer vision and pattern recognition algorithms and systems, holds nine U.S. patents, licensed technologies, and has published high-impact journal and conference papers. He is the recipient of the National Science Foundation CAREER award in 1994, the USF Teaching Incentive Program Award for Undergraduate Teaching Excellence in 1997, the Outstanding Undergraduate Teaching Award in 1998, and the Ashford Distinguished Scholar Award in 2004. He is a Fellow of the American Association for the Advancement of Science (AAAS), Institute of Electrical and Electronics Engineers (IEEE) and International Association for Pattern Recognition (IAPR), American Institute for Medical and Biological Engineering (AIMBE), and a Fellow and member of the Board of Directors of the National Academy of Inventors (NAI). He has served on many journal boards and is currently the Editor-in-Chief for Pattern Recognition Letters.