UNIVERSITY OF SOUTH FLORIDA

Defense of a Doctoral Dissertation

Beyond Labels and Captions: Contextualizing Grounded Semantics for Explainable Visual Interpretation

by

Sathyanarayan Aakur Narasimhan

For the Ph.D. degree in Computer Science and Engineering

While significant milestones have been achieved in the field of computer vision, majority of the work has been concentrated on *supervised* visual recognition. This implies a closed world where the underlying assumption is that all environments contain the same objects and events, which are in one-to-one correspondence with the ground evidence in the image. Hence, the learned knowledge is limited to the annotated training set. Increasingly complex models require massive amounts of training data and offer little to no explainability due to the lack of transparency in the decision-making process. In this dissertation, we develop an inherently explainable approach for generating rich interpretations of visual scenes. We move towards an *open world, open-domain* visual understanding by decoupling the ideas of recognition and reasoning. We show that the proposed approach is able to advance the state-of-the-art results in complex benchmarks to handle data imbalance, complex semantics and complex visual scenes without the need for vast amounts of domain-specific training data. Extensive experiments on several publicly available datasets show the efficacy of the proposed approaches.

Examining Committee
Anuj Srivastava, Ph.D., Chairperson
Sudeep Sarkar, Ph.D., Major Professor
Yu Sun, Ph.D.
John Licato, Ph.D.
Rajiv Dubey, Ph.D.
Kenneth Malmberg, Ph.D.

Friday, May 31, 2019
10:00 AM
ENB 109

THE PUBLIC IS INVITED

Publications

1. **Sathyanarayanan Aakur**, Sudeep Sarkar. A Perceptual Prediction Framework for Self-Supervised Event Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.*
2. **Sathyanarayanan Aakur***, Daniel Sawyer*, Sudeep Sarkar. Fine-grained Action Detection in Untrimmed Surveillance Videos. *IEEE Winter Conference on Applications of Computer Vision Workshops, 2019.*
3. **Sathyanarayanan Aakur**, Fillipe DM de Souza, Sudeep Sarkar. Going Deeper with Semantics: Exploiting Semantic Contextualization for Interpretation of Human Activity in Videos. *Winter Conference on Applications of Computer Vision, 2019.*
4. **Sathyanarayanan Aakur**, Fillipe DM de Souza, Sudeep Sarkar. Generating Open World Descriptions of Video using Commonsense Knowledge in a Pattern Theory Framework. *Quarterly of Applied Mathematics.*
5. Gilbert Rotich*, **Sathyanarayanan Aakur***, Rodrigo Minetto, Mauricio Segundo, Sudeep Sarkar. Using semantic relationships among objects for geospatial land use classification. *IEEE Applied Imagery Pattern Recognition Workshop, 2018.*
6. **Sathyanarayanan Aakur**, Fillipe DM de Souza, Sudeep Sarkar. On the Inherent Explainability of Pattern Theory-based Video Event Interpretations. *Explainable and Interpretable Models in Computer Vision and Machine Learning*, *Springer Series on Challenges in Machine Learning*.
7. **Sathyanarayanan Aakur**, Fillipe DM de Souza, Sudeep Sarkar. Inherently Explainable Model for Video Activity Recognition. *AAAI Workshop On Reasoning and Learning for Human-Machine Dialogues, 2018.*
8. **Sathyanarayanan Aakur**, Fillipe DM de Souza, Sudeep Sarkar. Towards a Knowledge-based Approach to Video Comprehension. *In Conference on Computer and Robot Vision (CRV), 2017.*

Robert Bishop, Ph.D.
Dean, College of Engineering

Dwayne Smith, Ph.D.
Dean, Office of Graduate Studies

Disability Accommodations:
If you require a reasonable accommodation to participate, please contact the
Office of Diversity & Equal Opportunity at 813-974-4373 at least five (5) working days prior to the event.