

# UNIVERSITY OF SOUTH FLORIDA

## Defense of a Doctoral Dissertation

Knowledge Extraction and Inference Based on Visual Understanding of Cooking Contents

by

**Ahmad Babaeian Jelodar**

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

This dissertation focuses on analyzing cooking content for the ultimate goal of automatic robotic manipulation. For a robot to perform a cooking task, it will need to both have an understanding of the scene and utilize prior knowledge. We will explore two main sub-problems: knowledge extraction and inference, and visual understanding of the scene in this dissertation. Visual understanding of a scene, requires computer vision based algorithms which can visually infer information from a single image or video. Although great advances has been achieved by the emergence of deep learning, state-of-the art algorithms in this area have limitations. To attempt to overcome this lack of performance, we propose to use structured knowledge representations combined with state of the art deep learning techniques for visual understanding of cooking videos. Besides objects, and motions, we recognize that states of objects are important in interpreting the scene and introduce the state identification challenge in cooking applications and collect a dataset for research in the area of ingredient state analysis. We look into the problem of simultaneous knowledge extraction from a single image and extracting information about ingredients, their states, the inter-connection between different objects in the scene and the motion-object interconnections. This problem requires an algorithm that can model the correlation of various concepts in a single image simultaneously. We propose to incorporate auto-regressive self-attention based mechanisms to extract knowledge from a single image. We show that the knowledge acquired from a single image can be used for calorie estimation and suggest that it can be utilized for robotic manipulation in the future.

### Examining Committee

Lingling Fan, Ph.D., Chairperson

Yu Sun, Ph.D., Major Professor

Shaun Canavan, Ph.D.

Heather Culbertson, Ph.D.

John Licato, Ph.D.

Kyle Reed, Ph.D.

Wednesday, June 30, 2021

11:00 AM

Online (Microsoft Teams)

Please email for more information

[ajelodar@usf.edu](mailto:ajelodar@usf.edu)

**THE PUBLIC IS INVITED**

### Publications

- 1) **A. B. Jelodar**, D. Paulius, and Y. Sun, "Long Activity Video Understanding using Functional Object-Oriented Network", IEEE Transactions on Multimedia, Dec. 2018.
- 2) David Paulius, **A. B. Jelodar**, and Y. Sun. "Functional Object-Oriented Network: Construction & Expansion", International Conference on Robotics and Automation (ICRA) 2018.
- 3) **A. B. Jelodar**, and Y. Sun, "Joint Object and State Recognition using Language Knowledge", International Conference on Image Processing (ICIP) 2019.

*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.