This dissertation addresses the problem of action understanding in videos, which includes action recognition in trimmed video, temporal action localization in untrimmed video, spatial-temporal action detection and video object/action segmentation. For video action recognition, we propose a category level feature learning method. Our proposed method automatically identifies such pairs of categories using a criterion of mutual pairwise proximity in the (kernelized) feature space, and a category-level similarity matrix where each entry corresponds to the one-vs-one SVM margin for pairs of categories. For temporal action localization, we propose to exploit the temporal structure of actions by modeling an action as a sequence of sub-actions and present a computationally efficient approach. For video action detection, we propose 3D Tube Convolutional Neural Network (TCNN) based pipeline. The proposed architecture is a unified deep network that is able to recognize and localize action based on 3D convolution features. It generalizes the popular faster R-CNN framework from images to videos. For video object and action segmentation, an end-to-end encoder-decoder based 3D convolutional neural network pipeline is proposed, which is able to segment out the foreground objects from the background. Moreover, the action label can be obtained as well by passing the foreground object into an action classifier. Extensive experiments on several video datasets demonstrate the superior performance of the proposed approach for video understanding compared to the state-of-the-art.

Major: Computer Science

Educational Career:
 Bachelor's of Telecommunication Engineering, BS, 2010, Beijing University of Posts and Telecommunications
 Master's of Electrical Engineering, MS, 2012, Arizona State University

Committee in Charge:
Mubarak Shah, Chair, Computer Science
Abhijit Mahalanobis, Computer Science
Kien Hua, Computer Science
Rahul Sukthankar, Google Research

Approved for distribution by Mubarak Shah, Committee Chair, on June 11, 2019.

The public is welcome to attend.