Jue Wang

4 Hutton Street, Acton, Canberra, ACT
(+61)414191375
Jue.Wang@anu.edu.au

Objective

With the astronomical increase of video data on the web, the Ph.D project of mine aims at improving video summarization by modeling the semantics of video data. Given the feature of videos, video descriptors are automatically generated using proposed methods, which is favorable for state-of-the-art features and end-to-end trainable. The video descriptor will be used for the further tasks and they should embed the necessary discriminative information of the video, such as temporal order, and action grammar. Based on this, the performance of video-based applications is improved, such as video dialog and video captioning.

To this end, several projects have been proposed:

1. Dynamic Optical Flow image (WACV, 2017) that summarize multiple optical flow into one single image embedding the temporal order.
2. SVM Pooling (CVPR, 2018) that generate a decision boundary for separating the discriminative frames out of the background/noise in one video and use it as the video descriptor.
3. Discriminative pooling with ordered constraint (TPAMI, 2019), that improve the SVM pooling by introducing temporal constraint and enable the end-to-end training.
4. Discriminative Subspace Pooling (ECCV, 2018) that generate multiple hyperplanes in a subspace to distinguish the discriminative feature in the video by using adversarial perturbations.
5. Audio-Visual Scene-Aware Dialog (CVPR, 2019) that generate a complete and natural response to a question about a scene, given video and audio of the scene and the history of previous turns in the dialog.
6. Generalized One-class Discriminative Subspaces (ICCV, 2019) that propose a novel objective for one-class learning by using a pair of orthonormal frames -- as subspaces -- to ``sandwich'' the labeled data. To study the effectiveness of the scheme, a new dataset, namely Dash-Cam-Pose, is generated.

In the near future, the emphasis of my research interest are: 1). Creating an aesthetic and interesting summarization, based on an existing temporal pooling scheme and optimization algorithms. 2) Action grammar learning using discriminative clustering. 3) Investigating the problem of video generation by using GAN.

Professional Summary

Self-directed and motivated PhD Student who works in the area of Computer Vision, especially in Action recognition, Video semantic learning and Action detection. Work effectively in a dynamic environment and fluent in programming languages and machine learning software packages .

Education

Ph.D   2016 - Now
College of Engineering Australian National University
PhD committee member: Prof. Richard Hartley

Prof. Fatih Porikli

Dr.Anoop Cherian

Dr. Xuming He

Dr. Mehrtash Harandi

B.Eng   2013 - 2014
College of Engineering Australian National University

First Class Honors

B.Eng   2010 - 2012
School of Information Beijing Institute of Technology
First Class Honors

Awards:

- Grant an oral presentation in the ECCV 2018
- Charades challenge 2017--Rank 6 in the Action Recognition Task, 2017
- Australian Government Research Training Program Scholarship, Australian National University, 2017 - 2019
- National ICT Australia Ltd Research Project Award, NICTA, 2016 – 2019
- Australian Nationcal University Supplementary Scholarship, Australian National University, 2016
- Australian Postgraduate Award, Australian National University, 2016 - 2017
- Double degree scholarship, Beijing Institute of Technology, 2012 - 2014
- Competition Award, ISCC (Information Security Counterwork Competition), NSFOCUS, 2012
- The National scholarship, Beijing Institute of Technology, 2010 - 2012
- Award of Squad Leader, Beijing Institute of Technology, 2010 - 2012

Work experience

Intern   2018 - 2018
Mitsubishi Electric Research Laboratories

Investigating the algorithm of video captioning, visual dialog and VQA problems.

Visitor Researcher   2014 - 2015
National ICT Australia Ltd

Participating a research group to generate the algorithm for self-calibration cameras

Strengths, Expertise, Skills and Experience

- Solid skills in programming the algorithm using Matlab & Python
- Sufficient experience with machine learning software packages (e.g., Caffe, TensorFlow, Matconvnet, Keras, PyTorch, Chianer)
- Sufficient experience with different Computer Vision tasks, such as Action recognition, Action detection, Video mining, Semantic modeling.
- Keeping the state-of-the-art results in many Computer Vision benchmarks.
- Co-organize AVSD DSTC7 workshop in AAAI 2019.
- Solid English skills. Overall ban Score 7.5 in IELTS with each task over Score 7.

Publications and Papers

- Honor Thesis: Dense Flow Estimation with Application to Motion Segmentation
- Technical Report: Self-calibration of cameras mounted on ARRB Surveying Platform in a real-world context
- Wang, J., Cherian, A., & Porikli, F. Ordered pooling of optical flow sequences for action recognition. In Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on (pp. 168-176). IEEE.

- Wang, J, Cherian, A, Porikli F, et al. Video representation learning using discriminative pooling. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (CVPR) 2018: 1149-1158.
- Wang, J., Cherian, A. Learning Discriminative Video Representations Using Adversarial Perturbations. In European Conference on Computer Vision. (ECCV) 2018 (oral presentation).
- Hori, C., Hori, T., Wichern, G., et al. Multimodal Attention for Fusion of Audio and Spatiotemporal Features for Video Description. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2018.
- Alamri, H., Cartillier, V., Lopes R, G., et al. Audio Visual Scene-Aware Dialog (AVSD) Challenge at DSTC7. In Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2018.
- Hori, C., Alamri, H., Wang, J., et al. End-to-End Audio Visual Scene-Aware Dialog using Multimodal Attention-Based Video Features. arXiv preprint arXiv:1806.08409, 2018.(AVSD DSTC7 workshop)
- Alamri, H., Cartillier, V., Das, A., Wang, J., et al. Audio-Visual Scene-Aware Dialog. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (CVPR) 2019
- Wang, J., Cherian, A. Discriminative Video Representation Learning Using Support Vector Classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence. (TPAMI) 2019
- Wang, J., Cherian, A., GODS: Generalized One-class Discriminative Subspaces for Anomaly Detection. IEEE International Conference on Computer Vision. (ICCV) 2019
- Cherian, A., Wang, J., Hori, C., Marks, T., Hori, T., STaTS: A Spatio-Temporal and Temporo-Spatial Attention Model for Video Captioning. (Under review)
- Hori, C., Hori, T., Cherian, A., Wang, J., et al. Dynamic Scene-Aware Dialog using Multimodal Attention-Based Video Description Features. (Under review)

Reference

References will be available upon request