JIA CHEN

Ph: +1 (412) 218-6159 ⋄ Email: jiac@cs.cmu.edu
Homepage: http://www.cs.cmu.edu/~jiac

RESEARCH INTERESTS

My research area is in the interdisciplinary field of multimedia, machine learning and computer vision, which specifically includes image/video captioning, event detection and event localization.

EDUCATION

Shanghai Jiao Tong University, China *2008 - 2015*
Doctor, Computer Science
Advisor: Prof. Yong Yu
Thesis: On Exploitation and Exploration in Image Collection

Shanghai Jiao Tong University, China *2004 - 2008*
Bachelor, Mathematics (Major Degree) and Computer Science (Minor Degree)

HONORS & AWARDS

Best performer at video caption task in TRECVID *2017, 2018*
Best paper runner up at ICMR *2018*
Best performer at MSR video to language challenge *2016, 2017*
Best grand challenge paper award at ACM MM 2017 *2017*
2nd place in surveillence event detection task at TRECVID *2017*
IBM PhD Fellowship *2012*
Morgan Stanley Fellowship *2009*

RESEARCH EXPERIENCE

Post Doctoral Associate at Carnegie Mellon University *2016 - present*
Supervisor: Prof. Alexander Hauptmann
Research area: video event detection, video captioning, video event reconstruction

Fellowship Research Intern at IBM China Research Laboratory *2012-2013, 2014-2015*
Supervisor: Dr. Shenghao Bao, Dr. Zhong Su
Research area: product recommendation

RESEARCH PROJECTS

IARPA DIVA: Deep Intermodal Video Analytics *2018 - present*
The key contributor and system designer in CMU team of IARPA DIVA project phase one. The project aims to detect events such as loading, entering car and carrying bags in surveillance videos. Coordinated more than 10 students and 5 researchers from different groups. Built a modularized pipeline composed of event proposal, event classification and event localization. Proposed joint event classification and localization, which boosts the performance significantly by around 7 absolute points in the final evaluation metric.

CMU-RUC Video Captioning *2016 - 2018*
The key contributor in CMU-RUC team of MSR video to language challenge (2016, 2017) and video caption task at TRECVID (2017, 2018). Coordinated more than 5 students from 2 universities in the team. Best Performer in these challenges, 2 years in a row. Proposed topic-guided caption model to generate more accurate and detailed video descriptions. Implemented and conducted comprehensive study of state-of-the-art captioning algorithms.

Video Event Reconstruction *2016 - 2017*

The key contributor to the 2013 Boston Marathon dataset, an event reconstruction dataset for synchronization and localization. Proposed joint saliency estimation and matching using robust building related parts to localize online event videos, which improves more than 10 percent on mean average precision compared to methods without saliency estimation.

History Event Image Profiling *2015 - 2016*

The project aims to automatically construct an image/multimedia profile given a one sentence description of the historic event which contains where, when, who and what elements. Proposed to generate pseudo positive in-domain labels from out-of-domain labeled datasets. Proposed a multi-modal joint embedding algorithm to robustly learn the model from the noisy pseudo positive labels. The automatically generated data in this way achieves 66.1 on average precision, which saves a lot of human effort in constructing history event dataset.

Product Recommendation *2014-2015*

Worked on product recommendation in cold-start setting of unexplored product category. Studied using price to help improve commodity recommendation in unexplored categories. Proposed joint matrix factorization to embed the price information into the classical user-item matrix factorization, which boosts the performance by 3 percent on area under curve metric.

SELECTED PUBLICATIONS

Shizhe Chen, Qin Jin, **Jia Chen**, Alexander Hauptmann. Generating Video Descriptions with Latent Topic Guidance. Transaction on Multimedia (TMM), 2019.

Shizhe Chen, **Jia Chen**, Qin Jin, Alexander Hauptmann. Class-aware Self-Attention for Audio Event Recognition. ACM International Conference on Multimedia Retrieval (ICMR), 2018 (**Best paper runner up**).

Jia Chen, Shizhe Chen, Qin Jin. Video Captioning with Guidance of Multimodal Latent Topics. ACM Multimedia, 2017.

Haoyue Shi, **Jia Chen**, Alexander Hauptmann. Joint Saliency Estimation and Matching using Image Regions for Geo-Localization of Online Video. ACM International Conference on Multimedia Retrieval (ICMR), 2017.

Qin Jin, Shizhe Chen, **Jia Chen**, Alexander Hauptmann. Knowing Yourself: Improving Video Caption via In-depth Recap. ACM Multimedia, 2017 (**Best Grand Challenge Paper**).

Jia Chen, Qin Jin, Yifan Xiong. Semantic Image Profiling for Historic Events: Linking Images to Phrases. ACM Multimedia, 2016.

Jia Chen, Qin Jin, Shiwan Zhao, Shenghua Bao, Li Zhang, Zhong Su, Yong Yu. Boosting Recommendation in Unexplored Categories by User Price Preference. ACM Transactions on Information Systems (TOIS), 2016.

Jia Chen, Qin Jin, Yong Yu, Alexander G Hauptmann. Image Profiling for History Events on the Fly. ACM Multimedia, 2015.

Jia Chen, Qin Jin, Shenghua Bao, Zhong Su, Shimin Chen, Yong Yu: Exploitation and Exploration Balanced Hierarchical Summary for Landmark Images. Transactions on Multimedia (TMM), 2015.

Jia Chen, Qin Jin, Shiwan Zhao, Shenghua Bao, Li Zhang, Zhong Su, Yong Yu. Does Product Recommendation Meet its Waterloo in Unexplored Categories? No, Price Comes to Help. SIGIR, 2014.

Jie Shen, Guangcan Liu, **Jia Chen**, Yuqiang Fang, Jianbin Xie, Yong Yu, Shuicheng Yan. Unified Structured Learning for Simultaneous Human Pose Estimation and Garment Attribute Classification. Transactions on Image Processing (TIP), 2014.

PROFESSIONAL SKILLS

Python, Java, MATLAB, C/C++

Docker, PyTorch, Tensorflow, Caffe, MatConvNet, MxNet