

Xu, Huijuan

Email: xuhuijuan6@gmail.com

·Education

2016.6 -	Boston University, Computer Science PhD program of Computer Science	Boston, USA
2013.9 - 2016.6	University of Massachusetts-Lowell, Computer Science PhD program of Computer Science	Lowell, USA
2012.9 - 2013.5	University of Massachusetts-Lowell, Computer Engineering PhD program of Computer Engineering	Lowell, USA
2009.9 - 2012.7	Graduate University of Chinese Academy of Sciences (GUCAS) Master of Computer Application Technology	Beijing, China
2005.9 - 2009.7	Hefei University of Technology Bachelor of Computer Science	Hefei, China

·Activities

2016.6 -	Research Assistant at Boston University CS Department Research area: deep learning, vision and language, captioning, VQA, temporal action detection
2012.9 - 2016.6	Teaching Assistant / Research Assistant at UMass Lowell CE/CS Department Research area: deep learning, vision and language, video language description, captioning, VQA
2009.9 - 2012.6	Teaching Assistant / Research Assistant at GUCAS Research area: Bayesian network and probability inference

·Academic Honors

·National Scholarship of P. R. China (2007)	1%
·Outstanding Graduate of Anhui Province (2009)	1%

·Publications

- Huijuan Xu, Abir Das and Kate Saenko. R-C3D: Region Convolutional 3D Network for Temporal Activity Detection. In International Conference on Computer Vision (ICCV), 2017.
- Huijuan Xu and Kate Saenko. Dual Attention Network for Visual Question Answering. ECCV2016 2nd Workshop on Storytelling with Images and Videos (VisStory), 2016.
- Huijuan Xu and Kate Saenko. Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering. Proceedings of the 14th European Conference on Computer Vision (ECCV 2016), Amsterdam, Netherlands, 2016.
- Huijuan Xu, Subhashini Venugopalan, Vasili Ramanishka, Marcus Rohrbach and Kate Saenko. A Multi-scale Multiple Instance Video Description Network. ICCV15 workshop on Closing the Loop Between Vision and Language, 2015.
- Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney and Kate Saenko. Translating Videos to Natural Language Using Deep Recurrent Neural Networks. Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics -- Human Language Technologies (NAACL HLT 2015), Denver, Colorado, June 2015.
- Jinke Jiang, Juyun Wang, Hua Yu, Huijuan Xu. Poison Identification Based on Bayesian Network: A Novel Improvement on K2 Algorithm via Markov Blanket. Proceedings of the 4th International Conference of Advances in Swarm Intelligence, Harbin, China, June 2013.
- Huijuan Xu, Hua Yu and Juyun Wang. Poison Identification Based on Bayesian Method in Biochemical Terrorism Attacks. *Advanced Science Letters* (SCIE, ISSN: 1936-6612), Vol. 5, 1-5, 2012.
- Huijuan Xu, Hua Yu, Juyun Wang and Jinke Jiang. Exploration of Greedy Hill-climbing Search in Markov Equivalent Class Space. Proceedings of the 14th International Conference on Artificial Intelligence (ICAI 2012), Las Vegas, Nevada, USA, July 2012.

·Research Experience

- 2016.9 - present **Temporal Action Detection**
- Apply the image object detection pipelines (Faster RCNN) to solve the temporal activity detection in videos, use C3D as video feature extractor, design the RPN and ROI pooling for video data.

2015.6 - 2016.12 **Visual Question Answering (VQA)**

- Design one deep neural network model called SMem-VQA based on spatial attention to solve the VQA problem, show the inference process of SMem-VQA model by visualizing the attention weights, and achieve good result on VQA challenge.
- Propose Dual Attention Network (DAN) to investigate the idea of attention on both question and image for the VQA problem, and get improved result on VQA challenge.

2014.5 - 2105.6 **Natural Language Generation for youtube videos**

- Run LSDA object detection (Large Scale Detection Through Adaptation detection algorithms) on youtube clipped frames, extract LSDA features for detected bounding boxes, and form video descriptor using the LSDA features of all the bounding boxes in one video. Based on video descriptor, use SVM to classify subject, verb and object, and generate sentence using template methods.
- Design a multi-scale fully convolutional architecture, use multiple instance learning mechanism (MIL) to fine-tune the network on youtube data, and integrate this multi-scale framework with LSTM to realize end-to-end video description generation.

2009.9 - 2012.7 **Bioterrorism Emergency Decision Support Program**

- Supported by National Natural Science Foundation of P. R. China, mainly studies the key technologies for dealing with biochemical terrorism attack emergency, including poison identification in biochemical terrorism attacks, hospital rescue capacity analysis, the wounded transportation, and evacuation simulation. I am Responsible for poison identification, namely using the preliminary symptoms to diagnose poisoning person.
- Investigate several Bayesian network structure learning algorithms and data processing methods, do comparison experiments, improve some algorithms, and apply the algorithms into practical problems. Gibbs sampling, K2/Hill-Climbing, Independent Component Analysis(ICA), Markov Equivalent Class space, Greedy Equivalent Search algorithm(GES), etc.