

# Huijuan Xu

Department of Computer Science and Engineering  
Pennsylvania State University–University Park

☎ +1 (510) 660-9108 • ✉ hqx5063@psu.edu • 🌐 cs-people.bu.edu/hxu/

## Research Interests

---

My research focuses on deep learning, computer vision and natural language processing, particularly in the area of action understanding. The goal of my research is building robust action understanding algorithms with human level structural knowledge and multi-modal complementary ability. To achieve this goal, my research touches the following two aspects: compositional action recognition with knowledge reasoning, and multimodal action detection with less supervision.

## Research Appointments

---

<b>Pennsylvania State University–University Park</b> Department of Computer Science and Engineering <i>Assistant Professor</i>	Aug 2021 - Present <i>State College, PA</i>
<b>University of California, Berkeley</b> <i>Postdoc Scholar</i> Advisor: Prof. Trevor Darrell	Oct 2018 - Aug 2021 <i>Berkeley, CA</i>
<b>Disney Research Pittsburgh</b> <i>Research Intern</i> Advisor: Prof. Leonid Sigal	Jun 2017 - Sep 2017 <i>Pittsburgh, PA</i>

## Education

---

<b>Boston University</b> <i>Ph.D., Computer Science</i> Dissertation title: Vision and language understanding with localized evidence Advisor: Prof. Kate Saenko Committee: Prof. Emily Whiting, Prof. Margrit Betke, Prof. Stan Sclaroff, Prof. Leonid Sigal	Sep 2018 <i>Boston, MA</i>
<b>Graduate University of Chinese Academy of Sciences</b> <i>M.S., Computer Science</i> Thesis title: Bayesian network and probability inference Advisor: Prof. Hua Yu	Jul 2012 <i>Beijing, China</i>
<b>Hefei University of Technology</b> <i>B.S., Computer Science</i> Thesis title: Adaptive online exam system	Jul 2009 <i>Hefei, China</i>

## Peer-Reviewed Conference Publications

---

- [1]. Reuben Tan, **Huijuan Xu**, Kate Saenko, and Bryan A. Plummer. LoGAN: Latent Graph Co-Attention Network for Weakly-Supervised Video Moment Retrieval. IEEE Winter Conference on Applications of Computer Vision (WACV), 2021.
- [2]. Baifeng Shi, Judy Hoffman, Kate Saenko, Trevor Darrell, and **Huijuan Xu**. Auxiliary Task Reweighting for Minimum-Data Learning. Neural Information Processing Systems (NeurIPS), 2020. (Acceptance rate 20.1%)
- [3]. Zhekun Luo, Devin Guillory, Baifeng Shi, Wei Ke, Fang Wan, Trevor Darrell, and **Huijuan Xu**. Weakly-Supervised Action Localization with Expectation-Maximization Multi-Instance Learning. European Conference on Computer Vision (ECCV), 2020. (Acceptance rate 26%)
- [4]. Roei Herzig\*, Amir Bar\*, **Huijuan Xu**, Gal Chechik, Trevor Darrell, and Amir Globerson. Learning Canonical Representations for Scene Graph to Image Generation. European Conference on Computer Vision (ECCV), 2020. (Acceptance rate 26%)
- [5]. Joanna Materzynska, Tete Xiao, Roei Herzig, **Huijuan Xu**\*, Xiaolong Wang\*, and Trevor Darrell\*. Something-Else: Compositional Action Recognition with Spatial-Temporal Interaction Networks. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. (Acceptance rate 22.1%)

- [6]. Ximeng Sun, **Huijuan Xu**, and Kate Saenko. TwoStreamVAN: Improving Motion Modeling in Video Generation. IEEE Winter Conference on Applications of Computer Vision (WACV), 2020. (Acceptance rate 34.5%)
- [7]. Roei Herzig\*, Elad Levi\*, **Huijuan Xu\***, Eli Brosh, Amir Globerson, Trevor Darrell. Classifying Collisions with Spatio-Temporal Action Graph Networks. IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019.
- [8]. Yi Zhu, Yanzhao Zhou, **Huijuan Xu**, Qixiang Ye, David Doermann, and Jianbin Jiao. Learning Instance Activation Maps for Weakly Supervised Instance Segmentation. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019. (Acceptance rate 25.2%)
- [9]. **Huijuan Xu**, Kun He, Bryan A. Plummer, Leonid Sigal, Stan Sclaroff and Kate Saenko. Multilevel Language and Vision Integration for Text-to-Clip Retrieval. Thirty-Third AAAI Conference on Artificial Intelligence (AAAI), 2019. (Acceptance rate 16.2%)
- [10]. **Huijuan Xu**, Boyang Li, Vasili Ramanishka, Leonid Sigal and Kate Saenko. Joint Event Detection and Description in Continuous Video Streams. IEEE Winter Conference on Applications of Computer Vision (WACV), 2019. (Acceptance rate 39%)
- [11]. Kia Xu, Guorong Li, **Huijuan Xu**, Weigang Zhang, and Qingming Huang. Edge Guided Generation Network for Video Prediction. IEEE International Conference on Multimedia and Expo (ICME), 2018. (Acceptance rate 30%)
- [12]. **Huijuan Xu**, Abir Das, and Kate Saenko. R-C3D: Region Convolutional 3D Network for Temporal Activity Detection. IEEE/CVF International Conference on Computer Vision (ICCV), 2017. (Acceptance rate 29%)
- [13]. **Huijuan Xu** and Kate Saenko. Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering. European Conference on Computer Vision (ECCV), 2016. (Acceptance rate 26.6%)
- [14]. Subhashini Venugopalan, **Huijuan Xu**, Jeff Donahue, Marcus Rohrbach, Raymond Mooney and Kate Saenko. Translating Videos to Natural Language Using Deep Recurrent Neural Networks. North American Chapter of the Association for Computational Linguistics (NAACL), 2015. (Acceptance rate 29.1%)
- [15]. Jinke Jiang, Juyun Wang, Hua Yu, **Huijuan Xu**. Poison Identification Based on Bayesian Network: A Novel Improvement on K2 Algorithm via Markov Blanket. International Conference of Advances in Swarm Intelligence (ICSI), 2013.
- [16]. **Huijuan Xu**, Hua Yu, Juyun Wang and Jinke Jiang. Exploration of Greedy Hill-Climbing Search in Markov Equivalent Class Space. International Conference on Artificial Intelligence (ICAI), 2012.
- [17]. **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.

## Journals

---

- [1]. **Huijuan Xu**, Abir Das, and Kate Saenko. Two-Stream Region Convolutional 3D Network for Temporal Activity Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2019.
- [2]. Tianliang Zhang, Zhenjun Han, **Huijuan Xu**, Baochang Zhang, Qixiang Ye. CircleNet: Reciprocating Feature Adaptation for Robust Pedestrian Detection. IEEE Transactions on Intelligent Transportation Systems (ITS), 2019.

## Workshop Short Papers

---

- [1]. **Huijuan Xu**, Lizhi Yang, Stan Sclaroff, Kate Saenko, and Trevor Darrell. Spatio-Temporal Action Detection with Multi-Object Interaction. The Seventh International Workshop on Egocentric Perception, Interaction and Computing (EPIC) at ECCV, 2020.
- [2]. **Huijuan Xu**, Ximeng Sun, Eric Tzeng, Abir Das, Kate Saenko, and Trevor Darrell. Revisiting Few-Shot Activity Detection with Class Similarity Control. Visual Learning with Limited Labels (VL3) Workshop at CVPR, 2020.
- [3]. **Huijuan Xu**, Boyang Li, Vasili Ramanishka, Leonid Sigal and Kate Saenko. Joint Event Detection and Description in Continuous Video Streams. Human Activity Detection in Multi-Camera, Continuous, Long-Duration Video (HADCV) Workshop at WACV, 2019.
- [4]. **Huijuan Xu**, Abir Das, and Kate Saenko. R-C3D: Region Convolutional 3D Network for Temporal Activity Detection. ActivityNet Large Scale Activity Recognition Challenge at CVPR, 2017.
- [5]. **Huijuan Xu**, Abir Das, and Kate Saenko. R-C3D: Region Convolutional 3D Network for Temporal Activity Detection. Charades Activity Challenge at CVPR, 2017.

- [6]. **Huijuan Xu** and Kate Saenko. Dual Attention Network for Visual Question Answering. 2nd Workshop on Storytelling with Images and Videos (VisStory) at ECCV, 2016.
- [7]. **Huijuan Xu** and Kate Saenko. Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering. Visual Question Answering (VQA) Challenge at CVPR, 2016.
- [8]. **Huijuan Xu**, Subhashini Venugopalan, Vasili Ramanishka, Marcus Rohrbach and Kate Saenko. A Multi-Scale Multiple Instance Video Description Network. First workshop on Closing the Loop Between Vision and Language (CLVL) at ICCV, 2015.

## Arxiv Preprints & Working Papers

---

- [1]. Baifeng Shi, Judy Hoffman, Kate Saenko, Qi Dai, Trevor Darrell, and **Huijuan Xu**. A Flexible Model for Action Localization with Multi-level supervision. (in progress)
- [2]. Zhekun Luo, Shalini Ghosh, Devin Guillory, Keizo Kato, Trevor Darrell, and **Huijuan Xu**. Factorized Zero-shot Compositional Action Recognition using Multiple Knowledge Bases. (in progress)
- [3]. **Huijuan Xu**, Lizhi Yang, Stan Sclaroff, Kate Saenko, and Trevor Darrell. Spatio-Temporal Action Detection with Multi-Object Interaction. arXiv preprint arXiv:2004.00180, 2020.
- [4]. **Huijuan Xu**, Ximeng Sun, Eric Tzeng, Abir Das, Kate Saenko, and Trevor Darrell. Revisiting Few-Shot Activity Detection with Class Similarity Control. arXiv preprint arXiv:2004.00137, 2020.
- [5]. Yinbo Chen, Xiaolong Wang, Zhuang Liu, **Huijuan Xu**, and Trevor Darrell. A New Meta-Baseline for Few-Shot Learning." arXiv preprint arXiv:2003.04390, 2020.
- [6]. Yancheng Bai, **Huijuan Xu**, Kate Saenko and Bernard Ghanem. Contextual Multi-Scale Region Convolutional 3D Network for Activity Detection. arXiv preprint arXiv:1801.09184, 2017.

## Professional Activities

---

### Journal Reviewer

- ◇ Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- ◇ International Journal of Computer Vision (IJCV)
- ◇ IEEE Transactions on Image Processing (TIP)
- ◇ ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)
- ◇ Pattern Recognition (PR)

### Conference Reviewer/Program Committee Member

- ◇ Conference on Computer Vision and Pattern Recognition (CVPR) 2018-2021
- ◇ International Conference on Computer Vision (ICCV) 2017, 2019
- ◇ European Conference on Computer Vision (ECCV) 2018, 2020
- ◇ Asia Conference on Computer Vision (ACCV) 2018
- ◇ Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2019
- ◇ International Joint Conferences on Artificial Intelligence (IJCAI) 2019, 2020
- ◇ Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence 2020
- ◇ International Conference on Machine Learning (ICML) 2020

### Workshop Reviewer/Program Committee Member

- ◇ Women in Computer Vision (WiCV) Workshop 2018, 2019
- ◇ Human Activity Detection in Multi-Camera, Continuous, Long-Duration Video (HADCV) Workshop 2020
- ◇ International Workshop on Deep Learning for Human-Centric Activity Understanding (DL-HAU) 2021

## Teaching Experiences

---

### Machine Learning

Spring 2015

*Teaching Assistant*

- ◇ Designed assignments and examinations; grading; office hours.
- ◇ Class of approx. 40 graduate students and 40 undergraduate students.

### Computing I Lab (C Programming)

Fall 2013, Spring 2014

*Lab Instructor*

- ◇ Lectures; tutorials; grading; office hours.
- ◇ Class of approx. 40 undergraduate students.

## Advising & Mentoring

---

- ◇ Zhekun Luo, Master, UC Berkeley (now at Nvidia)
- ◇ Lizhi Yang, Undergraduate, UC Berkeley
- ◇ Baifeng Shi, Undergraduate, Peking University (now PhD at UC Berkeley)
- ◇ Joanna Materzynska, Master, Oxford University (now PhD at MIT)
- ◇ Roei Herzig, PhD, Tel Aviv University
- ◇ Yinbo Chen, Undergraduate, Tsinghua University (now PhD at UC San Diego)
- ◇ Reuben Tan, PhD, Boston University
- ◇ Ximeng Sun, PhD, Boston University

## Other Service

---

- ◇ BAIR Undergraduate Mentoring Program, Mentor Fall 2019
  - Anjali Kantharuban, Undergraduate, UC Berkeley
  - Grace Luo, Undergraduate, UC Berkeley
- ◇ Graduate Admission Committee at UC Berkeley, Member 2019
- ◇ Boston University Computer Vision and Deep Learning Reading Group, Organizer Fall 2016

## Honors & Awards

---

- ◇ UC Berkeley Rising Star in EECS 2020
- ◇ Distinguished Program Committee Member at IJCAI 2019
- ◇ Doctoral Consortium Award at WACV 2019
- ◇ MIT Rising Star in EECS 2018
- ◇ Travel Grant for Women in Computer Vision (WiCV) Workshop at CVPR 2017-2019
- ◇ Most Innovative Award in ActivityNet Challenge at CVPR 2017
- ◇ Outstanding Graduate of Anhui Province 2009
- ◇ China National Scholarship 2007
- ◇ Hefei University of Technology Undergraduate Scholarship 2006-2008

## Invited talks & Presentations

---

- ◇ Spatio-Temporal Action Detection with Multi-Object Interaction  
Seventh International Workshop on Egocentric Perception, Interaction and Computing (EPIC) at ECCV 2020
- ◇ Revisiting Few-shot Activity Detection with Class Similarity Control  
Visual Learning with Limited Labels (VL3) Workshop at CVPR 2020

- ◇ Learning Activities from Few Examples  
BAIR/BDD Computer Vision Fall 2019 Workshop, University of California, Berkeley 2019
- ◇ Vision and Language Understanding with Localized Evidence  
University of California, Berkeley 2018
- ◇ Joint Event Detection and Description in Continuous Video Streams  
Women in Computer Vision (WiCV) Workshop at CVPR 2018
- ◇ Bridge Image and Video Visual Features for Video Captioning  
Disney Research, Pittsburgh 2017
- ◇ R-C3D: Region Convolutional 3D Network for Temporal Activity Detection  
The ActivityNet Large Scale Activity Recognition Challenge Workshop at CVPR 2017