BOWEN ZHANG

E-mail: zbwglory@gmail.com, Homepage: http://zbwglory.github.io

MCB 407, 1002 West Childs Way

University of Southern California

EDUCATION

University of Southern California *Sept. 2017 - Aug. 2022*
Ph.D in Computer Science, Advisor: Prof. Fei Sha and Prof. Leana Golubchik.

Tongji University *Sept. 2014 - Mar. 2017*
M.Eng. in Computer Engineering, Advisor: Prof. Hanli Wang.

Tongji University *Sept. 2010 - Jul. 2014*
B.Eng. in Computer Science and Technology, minor in Mathematics.
Best Undergraduate Thesis of Tongji University.

RESEARCH EXPERIENCE

Google Brain *May 2021 - Jan. 2022*
Hosts: Jiahui Yu and Andrew M. Dai *Student Researcher & Research Intern*

· Action recognition with co-training on videos and images. [Arxiv]

Google Research Language *Feb. 2020 - May 2021*
Hosts: Eugene Ie, Vihan Jain, and Pete Shaw *Student Researcher & Research Intern*

· Visually grounded compositional generalization. [Findings of EMNLP'21]
· Language guided clip localization from a large video corpus. [ArXiv]
· Learning from structured image and language data. [EMNLP'20 Oral]

Amazon AWS Rekognition *May 2019 - Dec. 2019*
Hosts: Yuanjun Xiong, Hao Chen, and Meng Wang *Intern*

· Online video understanding system for streaming videos. [ArXiv]

Tencent AI Lab *Jun. 2018 - Aug. 2018*
Hosts: Boqing Gong *Intern*

· Model-based reinforcement learning.

University of Southern California *Sept. 2017 - Aug. 2022*
Advisor: Prof. Fei Sha *Research Assistant & PhD student*

· Vision + Language

    Learning visual and linguistic representation from video and language. [ECCV'18]

    Learning word embedding from image and text. [EMNLP'18]

    Visual Storytelling generation. [ICCV'19 Workshop]

    Topic based text summarization. [BayLearn'19]

SIAT, Chinese Academy of Sciences *Jul. 2015 - Jul. 2017*
Research Center for Integrate Multimedia, Advisor: Prof. Yu Qiao *Visiting Student*

· Large Scale Action Recognition

Real-time action recognition. [CVPR'16, TIP]

ActivityNet Challenge [CVPR'16 Workshop, CVPR'17 Workshop]

PREPRINT

1. **Bowen Zhang**, Jiahui Yu, Christopher Fifty, Wei Han, Andrew M. Dai, Ruoming Pang, and Fei Sha, Co-training Transformer with Videos and Images Improves Action Recognition, Arxiv, 2021

2. **Bowen Zhang**⋆, Hexiang Hu⋆, Joonseok Lee, Ming Zhao, Sheide Chammas, Vihan Jain, Eugene Ie, and Fei Sha, A Hierarchical Multi-Modal Encoder for Moment Localization in Video Corpus, Arxiv, 2020

3. **Bowen Zhang**, Hao Chen, Meng Wang, and Yuanjun Xiong, Online Action Detection in Streaming Videos with Time Buffers, Arxiv, 2020

4. Melissa Ailem, **Bowen Zhang**, and Fei Sha, Topic Augmented Generator for Abstractive Summarization, Arxiv, 2019 (Short version is on BayLearn 2019)

PUBLICATIONS

1. **Bowen Zhang**, Hexiang Hu, Linlu Qiu, Pete Shaw, and Fei Sha, Visually Grounded Concept Composition, *Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), 2021, Findings, Long Paper

2. Linlu Qiu, Hexiang Hu, **Bowen Zhang**, Pete Shaw, and Fei Sha, Systematic Generalization on gSCAN: What is Nearly Solved and What is Next?, *Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), 2021. Short paper, *Oral Presentation*

3. **Bowen Zhang**⋆, Hexiang Hu⋆, Vihan Jain, Eugene Ie, and Fei Sha, Learning to Represent Images and Texts with Denotation Graphs, *Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), 2020, Long paper, *Oral Presentation*

4. Melissa Ailem, **Bowen Zhang**, Aurélien Bellet, Pascal Denis, and Fei Sha, A Probabilistic Model for Joint Learning of Word Embeddings from Texts and Images, *Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), 2018, Long paper

5. **Bowen Zhang**⋆, Hexiang Hu⋆, and Fei Sha, Cross-Modal and Hierarchical Modeling of Video and Text, *European Conference on Computer Vision* (*ECCV*), 2018

6. **Bowen Zhang**, Limin Wang, Zhe Wang, Yu Qiao, and Hanli Wang, Real-time Action Recognition with Enhanced Motion Vector CNNs, *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, USA, 2016.

7. **Bowen Zhang**, Limin Wang, Zhe Wang, Yu Qiao, and Hanli Wang, Real-Time Action Recognition with Deeply-Transferred Motion Vector CNNs, *IEEE Transactions on Image Processing* (*TIP*), 2018

8. Zhe Wang, Limin Wang, Yali Wang, **Bowen Zhang**, and Yu Qiao, Weakly supervised patchnets: Describing and aggregating local patches for scene recognition, *IEEE Transactions on Image Processing* (*TIP*), 2017

9. **Bowen Zhang**, and Hanli Wang, Encoding Scale into Fisher Vector for Human Action Recognition, *IEEE Conference on Visual Communications and Image Processing* (*VCIP*), Singapore, 2015, *Oral Presentation*

WORKSHOPS

1. **Bowen Zhang**, Hexiang Hu, and Fei Sha, Visual Storytelling via Predicting Anchor Word Embeddings in the Stories, *ICCV 2019 Workshop on Closing the Loop Between Vision and Language*, 2019

2. Yue Zhao, **Bowen Zhang**, Zhirong Wu, Shuo Yang, Lei Zhou, Sijie Yan, Limin Wang, Yuanjun Xiong, Dahua Lin, Yu Qiao, and Xiaoou Tang, CUHK & ETHZ & SIAT Submission to ActivityNet Challenge 2017, *ActivityNet Large Scale Activity Recognition Challenge, CVPR'17 Workshop*, Hawaii, USA, 2017.

3. Yuanjun Xiong, Limin Wang, Zhe Wang, **Bowen Zhang**, Hang Song, Wei Li, Dahua Lin, Yu Qiao, Luc Van Gool, and Xiaoou Tang, CUHK & ETHZ & SIAT Submission to ActivityNet Challenge 2016, *ActivityNet Large Scale Activity Recognition Challenge, CVPR'16 Workshop*, Las Vegas, USA, 2016.

4. Yun Yi, Hanli Wang, **Bowen Zhang**, and Jian Yu, MIC-TJU in MediaEval 2015 Affective Impact of Movies Task, *MediaEval 2015 Workshop*, Wurzen, Germany, 2015.

5. **Bowen Zhang**, Yun Yi, Hanli Wang, and Jian Yu, MIC-TJU at MediaEval Violent Scenes Detection (VSD) 2014, *MediaEval 2014 Workshop*, Barcelona, Spain, 2014.

CONTESTS

The ActivityNet Challenge 2017 on CVPR'17: Untrimmed Classificaiton. *Rank* $2/37$.	*Jul. 2017*
The ActivityNet Challenge 2017 on CVPR'17: Trimmed Classification. *Rank* $2/27$.	*Jul. 2017*
The ActivityNet Challenge 2017 on CVPR'17: Action Localization. *Rank* $2/15$.	*Jul. 2017*
The ActivityNet Challenge 2016 on CVPR'16: Untrimmed Video Classification. *Rank* $1/25$.	*Jul. 2016*
LSUN Challenge 2016 on CVPR'16: Scene Classification. *Rank* 1st.	*Jul. 2016*
MediaEval'15 Affective Impact. Violence: *Rank:* 2nd; Valence: *Rank:* 2nd; Arousal: *Rank:* 1st.	*Sept. 2015*
MediaEval'14 Affect in Multimedia. Main: *Rank:* 4th; Generalization: *Rank:* 4th.	*Sept. 2014*

HONORS AND AWARDS

Shanghai Outstanding Graduates	*2017*
National Graduate Scholarship of China	*2016*
Best Undergraduate Thesis of Tongji University (2 recipients in CS department)	*2014*
SAIC Volkswagen Automotive Scholarship	*2013*
Outstanding Student of Tongji University	*2012*
Scholarship for Academic Excellence of Tongji University	*2011, 2012, 2013*
China Undergraduate Mathematical Contest in Modeling (CUMCM) Province Third Prize	*2012*
The Mathematical Contest in Modeling (MCM) Honorable Mention.	*2012*

ACADEMIC SERVICE

Conference Reviewers:

Machine Learning: NeurIPS, ICML, ICLR

Computer Vision: CVPR, ECCV, ICCV, BMVC, WACV

Artificial Intelligence: IJCAI, AAAI

Natural Language Processing: NAACL, ACL, EMNLP

Journal Reviewers: IJCV, T-PAMI, T-IP