

# Songyang Zhang

Homepage: <http://sy-zhang.github.io/>  
Email : [szhang83@ur.rochester.edu](mailto:szhang83@ur.rochester.edu)

## Education

---

### University of Rochester

*Ph.D. Student in Computer Science, Advisor: Jiebo Luo*

Rochester, NY

*Aug. 2018 – Present*

### Zhejiang University

*M.S. in Computer Science, Advisor: Jun Xiao*

Hangzhou, China

*Sep. 2015 – Mar. 2018*

### Southeast University

*B.E. in Computer Science*

Nanjing, China

*Aug. 2011 – Jun. 2015*

## Professional Experience

---

### Facebook AI Research

*Research Intern, Mentor: Xi Yin*

Menlo Park, CA

*Jan 2022 - Aug 2022*

- Text-to-Video Generation (ECCV 2022, Two Submissions)

### Tencent AI Lab

*Research Intern, Mentor: Linfeng Song, Lifeng Jin, Kun Xu*

Bellevue, WA

*Sep 2021 - Dec 2021*

- Unsupervised Grammar Induction (EMNLP 2022)

### Microsoft

*Research Intern, Mentor: Jiang Wang, Quanzeng You*

Redmond, WA

*May 2021 - Aug 2021*

- Moment Localization with Natural Language

### Tencent AI Lab

*Research Intern, Mentor: Linfeng Song, Lifeng Jin, Kun Xu*

Bellevue, WA

*May 2020 - Aug 2020*

- Unsupervised Grammar Induction (NAACL 2021)
- Best Long Paper Award in NAACL 2021 (1/1797)

### Microsoft Research Asia

*Research Intern, Mentor: Houwen Peng, Jianlong Fu*

Beijing, China

*May 2019 - Aug 2019*

- Moment Localization with Natural Language (AAAI 2020, TPAMI 2021)
- Rank 1 in HACS Temporal Action Localization Challenge at ICCV 2019

### Tencent AI Lab

*Research Intern, Mentor: Lin Ma, Linchao Bao*

Shenzhen, China

*Mar 2018 - Aug 2018*

- Temporal Action Localization

## Publications

---

*\* Equal contribution, ordered alphabetically*

Jie An\*, **Songyang Zhang**\*, Harry Yang, Sonal Gupta, Jia-Bin Huang, Jiebo Luo, Xi Yin, “Latent Diffusion with Temporal Shift for Efficient Text-to-Video Generation”, *Submitted, 2023*

**Songyang Zhang**, Linfeng Song, Lifeng Jin, Kun Xu, Dong Yu and Jiebo Luo, “Learning a Grammar Inducer from Massive Uncurated Instructional Videos”, *EMNLP 2022* [Oral]

Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, **Songyang Zhang**, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, Devi Parikh, Sonal Gupta, Yaniv Taigman, “Make-A-Video: Text-to-Video Generation without Text-Video Data”, *In Submission to ICLR 2023, Under Review*

Xingchen Li, Long Chen, Jian Shao, Shaoning Xiao, **Songyang Zhang**, Jun Xiao, “Rethinking the Evaluation of Unbiased Scene Graph Generation”, *BMVC 2022* [Oral]

Bolin Ni, Houwen Peng, Minghao Chen, **Songyang Zhang**, Gaofeng Meng, Jianlong Fu, Shiming Xiang, Haibin Ling, “Expanding Language-Image Pretrained Models for General Video Recognition”, *ECCV 2022* [Oral]

Thomas Hayes\*, **Songyang Zhang**\*, Xi Yin, Guan Pang, Sasha Sheng, Harry Yang, Songwei Ge, Isabelle Hu, Devi Parikh, “MUGEN: A Playground for Video-Audio-Text Multimodal Understanding and GEneration”, *ECCV 2022*

Lin Li, Long Chen, Yifeng Huang, Zhimeng Zhang, **Songyang Zhang**, Jun Xiao, “The Devil is in the Labels: Noisy Label Correction for Robust Scene Graph Generation”, *CVPR 2022* [Oral]

Jiahui Li, Kun Kuang, Lin Li, Long Chen, **Songyang Zhang**, Jian Shao, Jun Xiao, “Instance-wise or Class-wise? A Tale of Neighbor Shapley for Concept-based Explanation”, *ACM MM 2021*

Zhengyuan Yang, **Songyang Zhang**, Liwei Wang, Jiebo Luo, “SAT: 2D Semantics Assisted Training for 3D Visual Grounding”, *ICCV 2021* [Oral]

**Songyang Zhang**, Linfeng Song, Lifeng Jin, Dong Yu, Jiebo Luo, “Video-aided Unsupervised Grammar Induction”, *NAACL 2021* [Best Long Paper Award]

**Songyang Zhang**, Houwen Peng, Jianlong Fu, Yijuan Lu, Jiebo Luo, “Multi-Scale 2D Temporal Adjacency Networks for Moment Localization with Natural Language”, *TPAMI 2021*

Shaoning Xiao, Long Chen, **Songyang Zhang**, Wei Ji, Jian Shao, Lu Ye, Jun Xiao, “Boundary Proposal Network for Two-Stage Natural Language Video Localization”, *AAAI 2021*

**Songyang Zhang**, Tolga Aktas, Jiebo Luo, “Mi YouTube es Su YouTube? Analyzing the Cultures using YouTube Thumbnails of Popular Videos”, *IEEE Big Data 2021*

Li Sun\*, Haoqi Zhang\*, **Songyang Zhang**, Jiebo Luo, “Content-based Analysis of the Cultural Differences between TikTok and Douyin”, *IEEE Big Data 2020*

Jie An, Tianlang Chen, **Songyang Zhang**, Jiebo Luo, “Global Image Sentiment Transfer”, *ICPR 2020*

**Songyang Zhang**, Houwen Peng, Le Yang, Jianlong Fu, Jiebo Luo, “Learning 2D Temporal Adjacent Networks for Temporal Action Localization”, *HACS Temporal Action Localization Challenge at ICCV 2019* [Winner]

**Songyang Zhang**, Houwen Peng, Jianlong Fu, Jiebo Luo, “Learning 2D Temporal Adjacent Networks for Moment Localization with Natural Language”, *AAAI 2020*

**Songyang Zhang**, Jinsong Su, Jiebo Luo, “Exploiting Temporal Relationships in Video Moment Localization with Natural Language”, *ACMMM 2019*

Jiageng Feng, **Songyang Zhang**, Jun Xiao, “Explorations of skeleton features for LSTM-based action recognition”, *MTA 2017*

**Songyang Zhang**, Jun Xiao, Xiaoming Liu, Yang Yi, Yueting Zhuang, “Fusing Geometric Features for Skeleton-Based Action Recognition using Multilayer LSTM Networks”, *TMM 2017*

**Songyang Zhang**, Xiaoming Liu, Jun Xiao, “On Geometric Features for Skeleton-Based Action Recognition Using Multilayer LSTM Networks”, *WACV 2017*

## Honors & Awards

---

Best Long Paper Award at NAACL 2021	Jun 2021
Winner, HACS Temporal Action Localization Challenge at ICCV 2019	Oct 2019
Student Travel Award at ACM Multimedia 2019	Oct 2019
Large Scale 3D Human Activity Analysis Challenge in Depth Videos	
<i>Rank 4 in Recognition Track, 6 in Detection Track, granted by ICME 2017 Workshop</i>	Apr 2017
<i>Rank 6, granted by ACCV 2016 Workshop</i>	Oct 2016
Robocup 3D Simulation	
<i>Winner, RoboCup China Open 2014 Competition</i>	Oct 2014
<i>Rank 5-8, Robocup 2014 Competition, Brazil</i>	Jul 2014

## Services

---

### Reviewer:

- Conference: CVPR, ICCV, ECCV, ACL, EMNLP, NAACL, AAAI, ACCV, WACV, ACMMM
- Journal: TPAMI, IJCV, TMM, TIP, TCSVT, Pattern Recognition, Neural Computing