

YE ZHU

Google Scholar
yzhu96@hawk.iit.edu
Homepage

RESEARCH INTERESTS

Computer Vision; Multi-Modal Learning; Vision, Audio and Language

EDUCATION

| | |
|-----------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Illinois Institute of Technology, USA Ph.D. candidate in Computer Science, advised by Prof. Yan Yan. GPA: 4.0/4.0 | January 2021 - Present |
| Texas State University, USA (<i>Transferred</i>) Ph.D. in Computer Science, advised by Prof. Yan Yan. GPA: 4.0/4.0 | September 2019 - December 2020 |
| Shanghai Jiao Tong University(SJTU), China M.S. in Mechanical Engineering. GPA: 3.72/4.0 | September 2016 - March 2019 |
| Ecole Polytechnique(X), France Exchange Master Student in Engineering. | September 2016 - March 2017 |
| Shanghai Jiao Tong University(SJTU), China B.S. in Mechanical and Automation. GPA: 3.63/4.3 | September 2012 - August 2016 |

SELECTED PUBLICATIONS

- [1] Ye Zhu, Kyle Olszewski, Yu Wu, Panos Achlioptas, Menglei Chai, Yan Yan, and Sergey Tulyakov. Quantized GAN for Complex Music Generation from Dance Videos. (arXiv preprint: arXiv:2204.00604), 2022.
- [2] Ye Zhu, Xiaoguang Zhu, Yuzhang Shang, Zhenhao Zhao, Yan Yan. Supplementing Missing Visions via Dialog for Scene Graph Generations. (arXiv preprint: arXiv:2204.11143), 2022.
- [3] Ye Zhu, Yu Wu, Yi Yang, and Yan Yan. Saying the Unseen: Video Descriptions via Dialog Agents, in IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**), 2021.
- [4] Ye Zhu, Yu Wu, Hugo Latapie, Yi Yang, Yan Yan. Learning Audio-Visual Correlations From Variational Cross-Modal Generations, in IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP**), 2021.
- [5] Xiaoguang Zhu, Ye Zhu, Haoyu Wang, Honglin Wen, Yan Yan, Peilin Liu. Skeleton Sequence and RGB Frame Based Multi-Modality Feature Fusion Network for Action Recognition, in Transactions on Multimedia Computing Communications and Applications (**TOMM**), 2021.
- [6] Ye Zhu, Yu Wu, Yi Yang, and Yan Yan. Describing Unseen Videos via Multi-Modal Cooperative Dialog Agents, in European Conference on Computer Vision (**ECCV**), 2020.
- [7] Ye Zhu, Yan Yan, and Oleg Komogortsev. Hierarchical HMM for Eye Movement Classification, in European Conference on Computer Vision Workshop (**ECCVW**), 2020.
- [8] Y. Zhu, S. Shepstone, P. Martínez-Nuevo, et al. Multiview Based 3D Scene Understanding On Partial Point Sets. (arXiv:1812.01712)

EXPERIENCE

Snap Inc., Remote, USA

May 2021 - August 2021

Research intern in Computer Vision, advised by Dr. Kyle Olszewski

- Project: Music generation conditioned on dance videos.

Bang & Olufsen, Struer, Denmark

July 2018 - December 2018

Research intern in Computer Vision, advised by Dr. Sven Ewan Shepstone and Dr. Pablo Martinez-Nuevo

- Project: 3D indoor scene understanding via point clouds.

Robotics Institute, Shanghai Jiao Tong University, China

March 2017 - July 2018

*Graduate research assistant***SELECTED HONORS AND AWARDS**

| | |
|---------------------------------------------------------------------------------|-----------|
| Merrick Merit Fellowship, Texas State University | 2019 |
| First Class Academic Excellence Scholarship for Graduate Students of SJTU | 2017&2018 |
| Meritorious Winne in Mathematical Contest in Modeling (MCM) | 2015 |
| Second Class Academic Excellence Scholarship for Undergraduate Students of SJTU | 2015 |

PROFESSIONAL ACTIVITIES

| | |
|------------------------------------------------------------------------------------------------|-----------|
| Research Assistant at the Computer Vision and Multimedia Lab, Illinois Institute of Technology | 2021-Now |
| Teaching Assistant for CS4310 Computer Networks, Texas State University | 2019&2020 |
| Teaching Assistant for CS4328 Operating Systems, Texas State University | 2020 |
| Reviewer for CVPR, ECCV, TMM, ACMM, ICASSP and etc. | |

PROGRAMMING

Python (Pytorch), Matlab, C/C++

LINGUISTIC SKILLS AND OTHERS

Chinese(Native)
English (TOEFL iBT 106/120, GRE 322+3/340+6)
French (Professional, DALF C1)
Chinese-French-English Translator for Science Magazine *Science&Vie*