YE ZHU

Homepage
Google Scholar
yzhu96@hawk.iit.edu

RESEARCH INTERESTS

Main Research: Generative Models; Multimodal Learning (Vision, Audio and Language); Computer Vision

EDUCATION

Illinois Institute of Technology, USA January 2021 - Present
Ph.D. candidate in Computer Science, advised by Prof. Yan Yan.
GPA: 4.0/4.0

Princeton University, USA September 2022 - Present
Visiting Ph.D. at VisualAI lab, advised by Prof. Olga Russakovsky.

Texas State University, USA *(Transferred)* September 2019 - December 2020
Ph.D. in Computer Science, advised by Prof. Yan Yan.
GPA: 4.0/4.0

Shanghai Jiao Tong University(SJTU), China September 2016 - March 2019
M.S. in Mechanical Engineering.
GPA: 3.72/4.0

Ecole Polytechnique(X), France September 2016 - March 2017
French Engineering Diploma.
Exchange Master Student in Engineering

Shanghai Jiao Tong University(SJTU), China September 2012 - August 2016
Pre-enrollment before Chinese College Entrance Examination
B.S. in Mechanical and Automation (in Sino-French Program).
GPA: 3.63/4.3

FEATURED PUBLICATIONS AND PREPRINTS

Preprints

[1] Ye Zhu, Yu Wu, Zhiwei Deng, Olga Russakovsky, and Yan Yan. Boundary Guided Mixing Trajectory for Semantic Control with Diffusion Models. (arXiv preprint: arXiv:2302.08357), 2023. [*Paper*] [*Code*]

[2] Ye Zhu, Yu Wu, Nicu Sebe, and Yan Yan. Vision+X: A Survey on Multimodal Learning in the Light of Data. (arXiv preprint: arXiv:2210.02884), 2022. [*Survey Paper*]

Computer Science Conference and Journal Publications, 2020 - Now

[1] Ye Zhu, Yu Wu, Kyle Olszewski, Jian Ren, Sergey Tulyakov, and Yan Yan. Discrete Contrastive Diffusion for Cross-Modal Music and Image Generation, in *International Conference on Learning Representations* (**ICLR**), 2023. [*Paper*] [*Code*] [*Project*]

[2] Duo Xu, Jonathan Tan*, Chia-Jung Hsu, and Ye Zhu. Denoising Diffusion Probabilistic Models to Predict the Number Density of Molecular Clouds in Astronomy, in *International Conference on Learning Representations Physics4ML Workshop* (**ICLR Workshop**), 2023. [1] [*Paper*]

[1] The author list for interdisciplinary collaborations with astrophysics follows the culture of the astronomy community, the last author does not signify the corresponding authorship.

[3] Ye Zhu, Kyle Olszewski, Yu Wu, Panos Achlioptas, Menglei Chai, Yan Yan, and Sergey Tulyakov. Quantized GAN for Complex Music Generation from Dance Videos, in *European Conference on Computer Vision* (**ECCV**), 2022. [*Paper*] [*Code*] [*Project*]

[4] 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**), 2022. [*Paper*] [*Code*]

[5] 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. [*Paper*] [*Code*]

[6] 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 *ACM Transactions on Multimedia Computing Communications and Applications* (**TOMM**), 2021. [*Paper*]

[7] 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. [*Paper*] [*Code*]

[8] Ye Zhu, Yan Yan, and Oleg Komogortsev. Hierarchical HMM for Eye Movement Classification, in *European Conference on Computer Vision Workshop* (**ECCV Workshop**), 2020. [*Paper*]

INDUSTRIAL 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.

TALKS

AI on the Big Stage, Bang & Olufsen, Denmark *(Remote)* November 2022
- Guest speaker for Bang & Olufsen AI club
- Multimodal Learning and Generation for Audio and Music

FEATURED HONORS AND AWARDS

ICLR Financial Assistance Award	2023
ACM-Women Scholarship for ICLR2023 Traveling	2023
Award for Excellence in Dissertation Research for the College of Computing, IIT	2022
CVPR Travel Grant Award	2022
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

TEACHING

Teaching Assistant for CS4310 Computer Networks, Texas State University	2019&2020
Teaching Assistant for CS4328 Operating Systems, Texas State University	2020

PROFESSIONAL SERVICE

Conference Reviewer
CVPR 2022-2023, ECCV 2022, ICCV 2023, ICML 2023
AAAI 2023, ACMMM 2021-2022, WACV 2023, ICASSP 2022
Journal Reviewer
IEEE Transactions on Multimedia, Neurocomputimg, Knowledge-Based Systems

LINGUISTIC SKILLS AND OTHERS

Chinese (Native Proficiency)
English (Professional Proficiency)
French (Professional Proficiency, DALF & TCF C1 Diploma)
French-Chinese Translator (part-time) for the European Science Magazine *Science&Vie*