YE ZHU (SHE/HER)

Homepage
Google Scholar
yezhu@princeton.com

RESEARCH INTERESTS

Main Research: Generative Models, Multimodal Learning, Computer Vision
Interdisciplinary: Machine Learning for Astrophysics

PROFESSIONAL APPOINTMENT

Princeton University, USA   September 2023 - Now
Postdoctoral Research Associate in Computer Science.
Advisor: Prof. Olga Russakovsky.

EDUCATION

Illinois Institute of Technology, USA   August 2023
Ph.D. in Computer Science.
Thesis: Multimodal Learning and Generation - Toward a Multisensory and Creative AI System. (Award for Excellence in Dissertation)
Advisor: Prof. Yan Yan.

Princeton University, USA   September 2022 - June 2023
Visiting Ph.D. in Computer Science.
Advisor: Prof. Olga Russakovsky.

Shanghai Jiao Tong University (SJTU), China   March 2019
M.S. in Mechanical Engineering.
French Engineering Diploma.

École Polytechnique (l'X), France   September 2016 - March 2017
Master Student in Engineering.

Shanghai Jiao Tong University (SJTU), China   August 2016
B.S. in Mechanical and Automation.
Pre-enrollment before Chinese College Entrance Examination.

PUBLICATIONS

(* for equal contributions)

Preprints

[1] Ruoyu Wang, Huayang Huang, **Ye Zhu**, Olga Russakovsky, Yu Wu. The Silent Prompt: Initial Noise as Implicit Guidance for Goal-Driven Image Generation. (arXiv preprint, arXiv: arXiv:2412.05101), 2024. [*Paper*]

[2] Yuhan Pei*, Ruoyu Wang*, Yongqi Yang, **Ye Zhu**, Olga Russakovsky, Yu Wu. SOWing Information: Cultivating Contextual Coherence with MLLMs in Image Generation. (arXiv preprint, arXiv:2411.19182), 2024. [*Paper*] [*Project*]

[3] Yongqi Yang*, Zhihao Qian*, <u>**Ye Zhu**</u>, and Yu Wu. D^3: Scaling Up Deepfake Detection by Learning from Discrepancy. (arXiv preprint, arXiv:2404.04584), 2024. *[Paper]*

[4] <u>**Ye Zhu**</u>, Yu Wu, Duo Xu, Zhiwei Deng, Yan Yan, and Olga Russakovsky. Discovery and Expansion of New Domains within Diffusion Models. (arXiv preprint, arXiv:2310.09213), 2024. *[Paper] [Code]*

[5] Sai Wang*, <u>**Ye Zhu***</u>, Ruoyu Wang, Amaya Dharmasiri, Olga Russakovsky, and Yu Wu. DETER: Detecting Edited Regions for Deterring Generative Manipulations. (arXiv preprint, arXiv:2312.10539), 2023. *[Paper] [Project]*

Computer Science Conference and Journal Publications, 2020 - Now

[1] <u>**Ye Zhu**</u>, Yu Wu, Nicu Sebe, and Yan Yan. Vision + X: A Survey on Multimodal Learning in the Light of Data. in *IEEE Transactions on Pattern Analysis and Machine Intelligence* (**TPAMI**), 2024. *[Paper]*

[2] William Yang, <u>**Ye Zhu**</u>, Zhiwei Deng, Olga Russakovsky. What is Dataset Distillation Learning?, in *International Conference on Machine Learning* (**ICML**), 2024. *[Paper] [Code]*

[3] Ruoyu Wang*, Yongqi Yang*, Zhihao Qian, <u>**Ye Zhu**</u>, and Yu Wu. Diffusion in Diffusion: Cyclic One-Way Diffusion for Text-Vision-Conditioned Generation, in *International Conference on Learning Representations* (**ICLR**), 2024. *[Paper] [Code] [Project]*

[4] <u>**Ye Zhu***</u>, Zhenhao Zhao*, Xiaoguang Zhu, Yuzhang Shang, and Yan Yan. Supplementing Missing Visions via Dialog for Scene Graph Generations, in *IEEE International Conference on Acoustics, Speech and Signal Processing* (**ICASSP**), 2024. *[Paper] [Code]*

[5] Bin Duan, Hao Tang, Changchang Sun, <u>**Ye Zhu**</u>, and Yan Yan. Mining and Unifying Heterogeneous Contrastive Relations for Weakly-Supervised Actor-Action Segmentation, in *Winter Conference on Applications of Computer Vision* (**WACV**), 2024. *[Paper]*

[6] <u>**Ye Zhu**</u>, Yu Wu, Zhiwei Deng, Olga Russakovsky, and Yan Yan. Boundary Guided Learning-Free Semantic Control with Diffusion Models, in *Conference on Neural Information Processing Systems* (**NeurIPS**), 2023. *[Paper] [Code] [Project] [Hugging face]*

[7] <u>**Ye Zhu**</u>, 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]*

[8] <u>**Ye Zhu**</u>, 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]*

[9] <u>**Ye Zhu**</u>, 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]*

[10] <u>**Ye Zhu**</u>, Yu Wu, Hugo Latapie, Yi Yang, and 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]*

[11] Xiaoguang Zhu, <u>**Ye Zhu**</u>, Haoyu Wang, Honglin Wen, Yan Yan, and 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]*

[12] <u>**Ye Zhu**</u>, 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]*

Astrophysics Journal Publications, 2023 - Now

[1] Duo Xu, Jenna Karcheski, Chi-Yan Law, <u>**Ye Zhu**</u>, Chia-Jung Hsu, and Jonathan Tan. Exploring Magnetic Fields in Molecular Clouds through Denoising Diffusion Probabilistic Models, in *The Astrophysics Journal* (**APJ**), 2025. *[Paper]*

[2] Duo Xu, **Ye Zhu**. Surveying Image Segmentation Approaches in Astronomy. in *Astronomy and Computing*, 2024. [*Paper*] (Invited Paper)

[3] Duo Xu, Jonathan Tan, Chia-Jung Hsu, and **Ye Zhu**. Denoising Diffusion Probabilistic Models to Predict the Density of Molecular Clouds, in *The Astrophysics Journal* (**APJ**), 2023. [*Paper*]

Workshop **Publications**

[1] Matthew Coleman, Olga Russakovsky, Christine Allen-Blanchette, and **Ye Zhu**. Discrete Diffusion Reward Guidance Methods for Offline Reinforcement Learning, in *International Conference on Machine Learning, Sampling and Optimization in Discrete Space (SODS) Workshop* (**ICML Workshop**), 2023. [*Paper*]

[2] Duo Xu, Jonathan Ta, 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. [*Paper*]

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

WORK 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.

PUBLIC TALKS

Vision Lab, Yale University, New Haven, USA   November, 2024
- Invited talk on *A Sustainable Vision for Generative AI*

TTIC Summer Workshop on Multimodal Artificial Intelligence, Chicago, USA   August, 2024
- Invited talk on *Taming Multimodal Generations via Fundamental Inspirations from Mathematics and Physics*

CVPR Responsible Data Workshop, Seattle, USA   June, 2024
- *DETER: Detecting Edited Regions for Deterring Generative Manipulations*

Talking to Machines Workshop, Riga, Latvia *(Remote)*   May, 2024
- *GenAI as Content Creators and Beyond*

Immersive Computing Lab, New York University, New York, USA   April, 2024
- *Mining the Latent: A Tuning-Free Paradigm for Versatile Applications with Diffusion Models*

ByteDance *(Remote)*   March, 2024
- *Mining the Latent: A Tuning-Free Paradigm for Versatile Applications with Diffusion Models*

MaVi Group, University of Bristol, UK *(Remote)*   February, 2024
- *Mining the Latent: A Tuning-Free Paradigm for Versatile Applications with Diffusion Models*

AI forum, Bang & Olufsen, Copenhagen, Denmark   October, 2023
- *Multimodal Learning and Generation*

Joint talk with Prof. Olga Russakovsky at ICML Workshop, Hawaii, USA   July 2023
- *Art, Science and Challenges of Generative AI*

Wuhan University, Wuhan, China   May 2023
- Topics on diffusion generative models, ML4Astrophysics.

Shanghai Jiao Tong University (SJTU), Shanghai, China   May 2023
- Topic on diffusion generative models, ML4Astrophysics.

Guest course lecture, Princeton University, USA   April 2023
- Guest lecture for the COS429 Computer Vision, topic on diffusion generative models.

PIXL talk, Princeton University, USA   April 2023
- Topic on diffusion generative models, ML4Astrophysics.

FEATURED HONORS AND AWARDS

Top Reviewer Award for NeurIPS	2024
EECS Rising Stars, MIT, Cambridge, USA	2024
NeurIPS 2023 Scholar Award, New Orleans, USA	2023
ICCV 2023 DEI Grant, Paris, France	2023
ICLR 2023 Financial Assistance Award, Kigali, Rwanda	2023
ACM-Women Scholarship [Coverage]	2023
Award for Excellence in Dissertation Research for the College of Computing, IIT, USA	2022
CVPR 2022 Travel Grant Award, New Orleans, USA	2022
Merrick Merit Fellowship, Texas, USA	2019
First Class Academic Excellence Scholarship for Graduate Students of SJTU, China	2018
First Class Academic Excellence Scholarship for Graduate Students of SJTU, China	2017
Meritorious Winner in Mathematical Contest in Modeling (MCM)	2015

TEACHING

AI4ALL Program, Princeton University, USA   July 2024
- Instructor for general lectures.

AI4ALL Program, Princeton University, USA   July 2023
- Instructor for the NLP project.
- Mentor for high school students to create an emotionally supportive Chatbot.
- Media Coverage on AI4ALL Princeton

PROFESSIONAL SERVICE

Workshop Organizer
CVPR 2025 CV4Science (Computer Vision for Science) workshop,
CVPR 2025 ReGenAI (Responsible Generative AI) workshop,
CVPR 2024 ReGenAI Workshop (First Workshop on **Re**sponsible **Gen**erative **AI**)
Conference Reviewer
In Machine Learning: NeurIPS 2023-2024, ICLR 2024-2025, ICML 2023-2025, AISTATS 2025, AAAI 2023-2024
In Computer Vision: CVPR 2022-2025, ECCV 2022-2024, ICCV 2023, WACV 2023-2024, ACMMM 2021-2022
In Computer Graphics: SIGGRAPH 2024
In Signal Processing: ICASSP 2022
Journal Reviewer
IEEE Transactions on Image Processing (TIP), IEEE Transactions on Multimedia (TMM),
Neurocomputimg, Knowledge-Based Systems

RESEARCH MENTORSHIP

Amaya Dharmasiri, DETER dataset for Detecting Generative Manipulations PhD student, Princeton University
William Yang, Dataset Disllation Analysis   PhD student, Princeton University
Xinran Liang, Dataset Debias via Generated Synthetic Data   PhD student, Princeton University
Matthew Coleman, Discrete Diffusion Reward Guidance for Offline RL Undergrad, Princeton University
Zhenhao Zhao, Scene Graph Generation   PhD student, Illinois Tech
Sai Wang, DETER dataset for Detecting Generative Manipulations   PhD student, Wuhan University
Chi Zuo, Noise Masking via Music Generations   PhD student, Wuhan University
Ruoyu Wang, Cyclic Diffusion for Text-Visual Conditioned Generation Undergrad, Wuhan University
Yongqi Yang, Cyclic Diffusion for Text-Visual Conditioned Generation MS student, Wuhan University

LINGUISTIC SKILLS AND OTHERS

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