

YE ZHU (SHE/HER)

[Homepage](#)

[Google Scholar](#)

yezhu@princeton.edu

RESEARCH INTERESTS

Main Research: Generative Models, Multimodal Learning, Computer Vision

Interdisciplinary: Machine Learning for Astrophysics

PROFESSIONAL APPOINTMENT

Princeton University, USA

September 2023 - Now

Postdoctoral Researcher 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

March 2017

French Engineering-track Student.

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)

Peer-Reviewed Conference and Journal Publications

[1] Yongqi Yang, Zhihao Qian, **Ye Zhu**, Olga Russakovsky, and Yu Wu. D³: Scaling Up Deepfake Detection by Learning from Discrepancy, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. [[Paper](#)] [[Code](#)]

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

- [3] **Ye Zhu**, 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](#)]
- [4] William Yang, **Ye Zhu**, Zhiwei Deng, Olga Russakovsky. What is Dataset Distillation Learning?, in *International Conference on Machine Learning (ICML)*, 2024. [[Paper](#)] [[Code](#)]
- [5] Ruoyu Wang*, Yongqi Yang*, Zhihao Qian, **Ye Zhu**, 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](#)]
- [6] Duo Xu, **Ye Zhu**. Surveying Image Segmentation Approaches in Astronomy. in *Astronomy and Computing*, 2024. [[Paper](#)] (Invited Paper)
- [7] **Ye Zhu***, 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](#)]
- [8] Bin Duan, Hao Tang, Changchang Sun, **Ye Zhu**, 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](#)]
- [9] **Ye Zhu**, 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](#)]
- [10] 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](#)]
- [11] **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](#)]
- [12] **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](#)]
- [13] **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](#)]
- [14] **Ye Zhu**, 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](#)]
- [15] Xiaoguang Zhu, **Ye Zhu**, 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](#)]
- [16] **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](#)]

Peer-Reviewed Workshop Publications

- [1] Sai Wang*, **Ye Zhu***, Ruoyu Wang, Amaya Dharmasiri, Olga Russakovsky, and Yu Wu. DETER: Detecting Edited Regions for Deterring Generative Manipulations, in *CVPR Responsible Data Workshop (CVPR Workshop Spotlight)*, 2024. [[Paper](#)] [[Project](#)]
- [2] 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](#)]
- [3] 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. [[Paper](#)]

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

Recent 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] **Ye Zhu**, 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](#)]

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.

SELECTED PUBLIC TALKS

DAML Seminar, École Polytechnique, Paris, France (Remote)

February, 2025

- Talk on *A Sustainable Vision for GenAI through the Lens of Dynamic Generative Models*

AI Lab, Princeton, USA

February, 2025

- Spotlight talk on *Generative AI beyond Scaling*

NYC Computer Vision Day, New York City, USA

February, 2025

- Lightning talk on *generative dynamics for image controlling and astrophysical modeling*

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*

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*

Shanghai Jiao Tong University (SJTU), Shanghai, China

May 2023

- Topic on diffusion generative models, ML4Astrophysics.

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 and the computer vision project.	
COS429 Computer Vision, Princeton University, USA	Fall 2023
- Guest Lecturer on the Generative Models.	
AI4ALL Program, Princeton University, USA	July 2023
- Instructor for general lectures and 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 workshop](#) (Workshops on Computer Vision for Science),

CVPR 2024&2025 [ReGenAI Workshop](#) (Workshops on **R**esponsible **G**enerative **AI**)

Conference Reviewer

In Machine Learning and Statistics: NeurIPS 2023-2024, ICLR 2024-2025, ICML 2023-2025,

AISTATS 2025, AAAI 2023-2024

In Computer Vision: CVPR 2022-2025, ECCV 2022-2024, ICCV 2023-2025,

WACV 2023-2024, ACM-MM 2021-2022

In Computer Graphics: SIGGRAPH 2024

In Signal Processing: ICASSP 2022

Journal Reviewer

Transactions on Machine Learning Research (TMLR), IEEE Transactions on Image Processing (TIP),

IEEE Transactions on Multimedia (TMM), Scientific Reports (Springer Nature),

Neurocomputing, Knowledge-Based Systems

STUDENT PROJECT MENTORSHIP

Amaya Dharmasiri, DETER dataset for Detecting Generative Manipulations	PhD, Princeton University
William Yang, Dataset Disillation Analysis	PhD, Princeton University
Xinran Liang, Dataset Debias via Generated Synthetic Data	PhD, Princeton University
Nobline Yoo, Numeracy correction for text-to-image generative models	MS, Princeton University
Matthew Coleman, Discrete Diffusion Reward Guidance for Offline RL	BS, Princeton University
Zhenhao Zhao, Scene Graph Generation	PhD, Illinois Tech
Sai Wang, DETER dataset for Detecting Generative Manipulations	PhD, Wuhan University
Chi Zuo, Noise Masking via Music Generations	PhD, Wuhan University
Ruoyu Wang, Cyclic Diffusion for Text-Visual Conditioned Generation	BS, Wuhan University
Yongqi Yang, Cyclic Diffusion for Text-Visual Conditioned Generation	MS, Wuhan University

LINGUISTIC SKILLS AND OTHERS

Chinese (Native Proficiency)

French (Professional Proficiency, DALF & TCF C1 Diploma)

English (Professional Proficiency)

French-Chinese Translator for the European Science Magazine *Science&Vie*