

# YE ZHU (SHE/HER)

[Homepage](#)

[Google Scholar](#)

[ye.zhu@polytechnique.edu](mailto:ye.zhu@polytechnique.edu)

## RESEARCH AREAS

---

**Main Research:** Generative Models, Computer Vision, Multimodal Learning

**Interdisciplinary:** Machine Learning for Astrophysics

## PROFESSIONAL APPOINTMENT

---

**École Polytechnique, Institut Polytechnique de Paris (IPP), France** September 2025 - Now  
Monge Tenure-Track Assistant Professor in Computer Science (Laboratoire d'Informatique, LIX).

**Princeton University, USA** September 2023 - August 2025  
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, 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.

## RESEARCH PUBLICATIONS

---

**Peer-Reviewed Conference and Journal Publications** (\* for equal contributions)

[1] **Ye Zhu**, Duo Xu, Zhiwei Deng, Jonathan Tan, Olga Russakovsky. Dynamic Diffusion Schrödinger Bridge in Astrophysical Observational Inversions, in *Conference on Neural Information Processing Systems (NeurIPS)*, 2025. [[Paper](#)] [[Code](#)]

[2] Chi Zuo, Martin B. Møller, Pablo Martínez-Nuevo, Huayang Huang, Yu Wu, **Ye Zhu**. BNMUSIC: Blending Environmental Noises into Personalized Music, in *Conference on Neural Information Processing Systems (NeurIPS)*, 2025. [[Paper](#)] [[Code](#)]

- [3] Ruoyu Wang, Huayang Huang, **Ye Zhu**, Olga Russakovsky, Yu Wu. The Silent Assistant: NoiseQuery as Implicit Guidance for Goal-Driven Image Generation, in *International Conference on Computer Vision (ICCV Highlight)*, 2025. [[Paper](#)] [[Code](#)]
- [4] Yongqi Yang\*, Zhihao Qian\*, **Ye Zhu**, Olga Russakovsky, and Yu Wu. D<sup>3</sup>: Scaling Up Deepfake Detection by Learning from Discrepancy, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. [[Paper](#)] [[Code](#)]
- [5] 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](#)]
- [6] **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](#)]
- [7] William Yang, **Ye Zhu**, Zhiwei Deng, Olga Russakovsky. What is Dataset Distillation Learning?, in *International Conference on Machine Learning (ICML)*, 2024. [[Paper](#)] [[Code](#)]
- [8] 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](#)]
- [9] Duo Xu, **Ye Zhu**. Surveying Image Segmentation Approaches in Astronomy. in *Astronomy and Computing*, 2024. [[Paper](#)] (Invited Paper)
- [10] **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](#)]
- [11] 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](#)]
- [12] **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](#)]
- [13] 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](#)]
- [14] **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](#)]
- [15] **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](#)]
- [16] **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](#)]
- [17] **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](#)]
- [18] 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](#)]
- [19] **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] **Ye Zhu**, Yu Wu, Duo Xu, Zhiwei Deng, Yan Yan, and Olga Russakovsky. Discovery and Expansion of New Domains within Diffusion Models, in *Conference on Neural Information Processing Systems, Frontiers in Probabilistic Inference Workshop (NeurIPS Workshop)*, 2025. [[Paper](#)] [[Code](#)]
- [2] Sai Wang\*, **Ye Zhu\***, Ruoyu Wang, Amaya Dharmasiri, Olga Russakovsky, and Yu Wu. DETER: Detecting Edited Regions for Detering Generative Manipulations, in *CVPR Responsible Data Workshop (CVPR Workshop Spotlight)*, 2024. [[Paper](#)] [[Project](#)]
- [3] 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](#)]
- [4] 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](#)]
- [5] **Ye Zhu**, Yan Yan, and Oleg Komogortsev. Hierarchical HMM for Eye Movement Classification, in *European Conference on Computer Vision Workshop (ECCV Workshop)*, 2020. [[Paper](#)]

## Other Preprints

- [1] Nobline Yoo, Olga Russakovsky, **Ye Zhu**. D2D: Detector-to-Differentiable Critic for Improved Numeracy in Text-to-Image Generation. (arXiv preprint, 2025). [[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](#)]

## TEACHING

---

<b>AI4ALL Program, Princeton University, USA</b> - Instructor for general lectures and the computer vision project.	July 2024
<b>COS429 Computer Vision, Princeton University, USA</b> - Guest Lecturer on the Generative Models.	Fall 2023
<b>AI4ALL Program, Princeton University, USA</b> - Instructor for general lectures and the NLP project. - <a href="#">Media Coverage on AI4ALL Princeton</a>	July 2023

## SELECTED HONORS AND AWARDS

---

Top Reviewer Award for NeurIPS	2024
<a href="#">EECS Rising Stars</a> , 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 [ <a href="#">Coverage</a> ]	2023
Award for Excellence in Dissertation Research for the College of Computing, Illinois Tech, USA	2022

## SELECTED PUBLIC TALKS

---

<b>Columbia University, New York City, USA</b> - Talk on <i>Controllability of Dynamic Generative Models with Theoretical Groundings</i>	July, 2025
<b>DAML Seminar, École Polytechnique, Paris, France (Remote)</b> - Talk on <i>A Sustainable Vision for GenAI through the Lens of Dynamic Generative Models</i>	February, 2025
<b>AI Lab, Princeton, USA</b> - Spotlight talk on <i>Generative AI beyond Scaling</i>	February, 2025

<b>NYC Computer Vision Day, New York City, USA</b>	February, 2025
- Lightning talk on <i>Generative Dynamics for Image Controlling and Astrophysical Modeling</i>	
<b>Vision Lab, Yale University, New Haven, USA</b>	November, 2024
- Invited talk on <i>A Sustainable Vision for Generative AI</i>	
<b>TTIC Summer Workshop on Multimodal Artificial Intelligence, Chicago, USA</b>	August, 2024
- Invited talk on <i>Taming Multimodal Generations via Fundamental Inspirations from Mathematics and Physics</i>	
<b>CVPR Responsible Data Workshop, Seattle, USA</b>	June, 2024
- <i>DETER: Detecting Edited Regions for Deterring Generative Manipulations</i>	
<b>Talking to Machines Workshop, Riga, Latvia (Remote)</b>	May, 2024
- <i>GenAI as Content Creators and Beyond</i>	
<b>Immersive Computing Lab, New York University, New York, USA</b>	April, 2024
- <i>Mining the Latent: A Tuning-Free Paradigm for Versatile Applications with Diffusion Models</i>	
<b>MaVi Group, University of Bristol, UK (Remote)</b>	February, 2024
- <i>Mining the Latent: A Tuning-Free Paradigm for Versatile Applications with Diffusion Models</i>	
<b>PIXL talk, Princeton University, USA</b>	April 2023
- Topic on diffusion generative models, ML4Astrophysics.	

## PROFESSIONAL SERVICE

---

### Workshop Organizers

CVPR 2025 [CV4Science workshop](#) (Workshops on Computer Vision for Science),  
CVPR 2024-2025 [ReGenAI Workshop](#) (Workshops on **R**esponsible **G**enerative **A**I)

### Conference Reviewers

*Machine Learning and Statistics*: NeurIPS 2023-2025, ICLR 2024-2026, ICML 2023-2025, AISTATS 2025-2026, AAAI 2023-2024

*Computer Vision and Graphics*: CVPR 2022-2025, ECCV 2022-2024, ICCV 2023-2025, WACV 2023-2024, ACM-MM 2021-2022, SIGGRAPH 2024

*Signal Processing*: ICASSP 2022

### Journal Reviewers

Transactions on Machine Learning Research (TMLR), IEEE Transactions on Image Processing (TIP),  
IEEE Transactions on Multimedia (TMM), Computer Vision and Image Understanding (CUIV)

## EARLY INDUSTRIAL EXPERIENCE

---

<b>Snap Inc., CA, USA</b>	May 2021 - August 2021
Research intern in Computer Vision, advised by Dr. Kyle Olszewski.	
<b>Bang &amp; Olufsen, Struer, Denmark</b>	July 2018 - December 2018
Research intern in Computer Vision, advised by Dr. Sven Ewan Shepstone and Dr. Pablo Martinez-Nuevo.	

## LINGUISTIC SKILLS AND OTHERS

---

**Chinese** (Native), **French** (Professional, DALF & TCF C1 Diploma), **English** (Professional)