JIN CAO

Homepage ◇ Github ◇ Google Scholar ◇
◇ Email: xjincaox@gmail.com

EDUCATION

Xi'an Jiaotong University (XJTU) *Sep 2022 - June 2026*
Undergraduate in Artificial Intelligence (Advanced Class)
GPA: 4.13/4.3 (*Professional Courses*), 3.85/4.3 (*All Courses*)
Score: 94.71/100 (*Professional Courses*), 90.83/100 (*All Courses*)
TOFEL: 101(S24)

Xi'an Jiaotong University (XJTU) *Sep 2021 - June 2022*
Special Class for the Gifted Young

RESEARCH INTEREST & EXPERIENCE

I am interested in developing **4D world models**—generative systems that learn to create physically grounded yet imaginative environments evolving consistently over space and time. My ultimate goal is to model worlds where not only geometry and physics emerge naturally, but also interactive agents and NPCs can act, react, and cohabit coherently within the generated universe. While such models have broad applications in robotics and simulation, I am most drawn to their potential for open-ended virtual worlds and creative **entertainment**.

Currently, I am a research intern at the Stanford Vision and Learning Lab (SVL) and the Computer Science Department at Stanford University, working with Koven Yu and Prof. Jiajun Wu on cross-scale 3D scene generation. Previously, I was a research intern at ZJU3DV, collaborating with Prof. Sida Peng on robust radiance field reconstruction.

SELECTED PUBLICATIONS

[1] **J. Cao***, H. Yu*, and J. Wu, "Wonderzoom: Multi-scale 3d world generation," in *Submitted to The Fourteenth International Conference on Learning Representations*, under review, 2025.
[2] **J. Cao***, H. Wu*, Z. Feng, H. Bao, X. Zhou, and S. Peng, "Universe: Unleashing the scene prior of video diffusion models for robust radiance field reconstruction," *ICCV*, 2025.
[3] **J. Cao**, Y. Cao, L. Pang, D. Meng, and X. Cao, *Hair: Hypernetworks-based all-in-one image restoration*, 2024. arXiv: 2408.08091 [cs.CV].
[4] **J. Cao***, X. Rui*, L. Pang, D. Meng, and X. Cao, "Latenthsi: Restore hyperspectral images in a latent space," *Information Fusion*, vol. 117, p. 102 848, 2025.

SELECTED RESEARCH PROJECTS

WonderZoom: Multi-Scale 3D World Generation [1]
[Website] [Paper] May 2025 - Oct 2025
Supervisors: Prof. Jiajun Wu and Hong-Xing (Koven) Yu

- We propose WonderZoom, the first approach to enable multi-scale 3D world generation from a single image, supporting seamless transitions from macro to micro scales.
- We introduce scale-adaptive Gaussian surfels, a dynamically updatable representation that grows incrementally with newly generated finer-scale content, while maintaining real-time rendering performance.
- We demonstrate and evaluate multi-scale 3D generation across diverse scenarios including natural environments, villages, and urban scenes, achieving consistent quality across scale transitions while significantly outperforming state-of-the-art video and 3D generation models in both perceptual quality and prompt alignment.
- **Under review at *ICLR 2026*.**

UniVerse: Unleashing the Scene Prior of Video Diffusion Models for Robust Radiance Field Reconstruction [2] [Website] Dec 2024 - Mar 2025

Supervisors: Prof. Sida Peng

- We propose UniVerse, a video generative model for robust 3D reconstruction from inconsistent multi-view images. Specifically, given a set of unstructured multi-view images, we first sort them to obtain a camera trajectory and insert blank images along this trajectory to transform images into a video, finally restore the video and extract the corresponding frames for following reconstruction tasks
- This work highlight the potential of decoupling robust reconstruction into restoration & reconstruction, instead of directly performing reconstruction on inconsistent images, bringing new insights to the community.
- **Accepted by *ICCV 2025*.**

HAIR: Hypernetworks-based All-in-One Image Restoration [3] [Paper] [Code] Jul 2024 - Sep 2024

Supervisors: Prof. Xiangyong Cao and Prof. Deyu Meng.

- We propose HAIR, a novel Hypernetworks-based All-in-One image restoration method that is capable of dynamically generating parameters based on the degradation information of input image. Extensive experiments demonstrate that HAIR can significantly improve the performance of existing image restoration models in a plug-and-play manner, both in single-task and All-in-One settings.
- We theoretically prove that, for a given small enough error threshold ϵ in image restoration tasks, HAIR requires fewer parameters compared to mainstream embedding-based All-in-One methods
- I'm responsible for most of the idea, code, experiments and writing of this work.
- **Submitted to *TPAMI*.**

LatentHSI: Restore Hyperspectral Images in a Latent Space [4] [Paper] Feb 2024 - Apr 2024

Supervisors: Dr. Xiangyu Rui and Prof. Xiangyong Cao.

- We propose LatentHSI, an unsupervised method for HSI restoration that utilizes a VAE to construct a latent space for HSIs. Sampling within this latent space is enabled by a diffusion model and the guidance of the observed images.
- We have developed a unified approach applicable to HSI restoration tasks within this latent space, making it general for many HSI applications including pansharpening, denoising, super-resolution, etc.
- I'm responsible for most of the idea, code, experiments and writing of this work.
- **Accepted by *Information Fusion.* (IF 18.6)**

LANGUAGE

English	TOEFL 101(S24)
Chinese	Native Speaker
Japanese	Beginner, actively learning Gojūon

ACHIEVEMENTS

National Third Prize in the C++ A Group of the 15th Langiao Cup National Software and Information Technology Professional Talent Competition (top 6%) *June 2024*

Provincial First Prize in the C++ A Group of the 15th Langiao Cup National Software and Information Technology Professional Talent Competition (rank 18^{th} top 2.7%) *Apr 2024*

Provincial Second Prize of China's 2024 National Mathematics Competition for College Students (top 9%) *Nov 2023*

Provincial First Prize in the C++ A Group of the 14th Langiao Cup National Software and Information Technology Professional Talent Competition (rank 8^{th} top 1.9%) *Apr 2023*

SKILLS

Programming Languages	Python, C/C++, MATLAB, Latex
Machine Learning Tools	Pytorch, Sklearn, Matplotlib, Pandas, Numpy