Chaojun Ni

Xin'ao Engineering Building, Peking University, 5 Yiheyuan Road, Haidian District, Beijing, China
chaojun.ni@pku.stu.edu.cn, +86 18668328472, https://nichaojun.github.io/

EDUCATION

2020.09–2024.06	**B.Eng.**, School of Information Management, Jiangxi University of Finance and Economics (JUFE). Supervisor: Yuming Fang.
2024.09–2027.06	**M.Eng.**, College of Engineering, Peking University (PKU). Supervisor: Wenjun Mei.

WORK EXPERIENCE

2024.05–2025.11	**Algorithm Intern**, Foundation Model Research Division, GigaAI. Supervisor: Zheng Zhu (Co-founder & Chief Scientist, GigaAI)

AWARDS

- **2025** – Recipient of the "Tongzhi Star" Award at the 1st Young Scholars Forum on Artificial General Intelligence
  - See: https://news.pku.edu.cn/xwzh/de4b012cafe840c496cabfeb53009bc4.htm
- **2024** – Top Ten Outstanding Graduation Theses of Jiangxi University of Finance and Economics
  - See: https://jwc.jxufe.edu.cn/news-show-1746.html
- **2024** – Outstanding Thesis of the School of Information Management, Jiangxi University of Finance and Economics
- **2022** – Jiangxi University of Finance and Economics Academic Scholarship
- **2021** – Jiangxi University of Finance and Economics Alumni Scholarship
- **2021** – Silver Medal at the 2021 ACM/ICPC National Invitational Contest for College Students

RESEARCH

My research aims to leverage **world models** to advance the application and development of autonomous driving and embodied intelligence, with a particular focus on the following two complementary research directions:

- **World-model-driven construction of explorable worlds.** I build explorable 3D/4D worlds for autonomous driving and embodied agents, transforming videos or single images into high-fidelity, controllable environments that support closed-loop interaction. These environments enable realistic simulation and large-scale evaluation, providing a practical foundation for reinforcement learning and policy optimization under diverse, safety-critical conditions.

  - **4D scene reconstruction for autonomous driving.** By integrating the generative capabilities of video world models with scene reconstruction, *DriveDreamer4D* [5], *ReconDreamer* [2], and *ReconDreamer++* [4] generate realistic and controllable 4D scene representations, enabling closed-loop driving simulations and large-scale evaluations. Moreover, *ReconDreamer-RL* [8] leverages reinforcement learning for end-to-end driving policy learning through direct interaction with these high-fidelity reconstructed environments, leading to significant performance improvements.

  - **Single-image to interactive 3D worlds.** I propose *WonderTurbo* [1] and *WonderFree* [9] to synthesize rich, explorable 3D worlds from a single input image. In particular, *WonderTurbo* [1] can construct a fully interactive 3D scene in just 0.72s, enabling real-time exploration, rapid design iteration, and an overall smoother creative workflow.

- **World-model-driven VLA data synthesis and training paradigms.** I explore how world models can generate high-quality, controllable, and scalable synthetic data to alleviate the high cost and limited scale of embodied interaction data. I further investigate how to leverage the generated data to improve the training efficiency and cross-environment generalization of VLA models.

  - **Data synthesis using world models.** I participate in the *GigaWorld* [10] project, which explores how to achieve large-scale VLA data synthesis and expansion. In this work, we focus on three key challenges: (i) how to extract and structure supervisory signals from Internet-scale data for VLA training; (ii) how to augment existing VLA datasets to improve diversity; and (iii) how to obtain high-quality VLA data from simulations while minimizing the sim-to-real gap. Specifically, *MimicDreamer* [13] transfers motion from human manipulation videos to synthesize high-fidelity robot demonstration videos,

enabling the low-cost collection of diverse demonstrations for policy learning. *EMMA* [12] performs instruction-guided, controllable editing of robot videos via text prompts, flexibly modifying the foreground, background, and lighting while preserving 3D structure and geometric plausibility. Additionally, *EmbodieDreamer* [14] leverages a world model to align the physics and visual domains between simulation and the real world, enabling imitation and reinforcement learning policies trained in simulation to transfer more effectively to real-world deployment.

- **More effective and generalizable training paradigms for VLA.** I contributed to *GigaBrain* [16], which explore training VLA models with large-scale synthetic data and achieve state-of-the-art performance by using 80% synthetic data and only 20% real data. Moreover, in *SwiftVLA* [7], I investigate VLA architectures and training paradigms, emphasizing the effective integration of 4D information extracted from existing pretrained models without additional sensors, thereby improving spatial understanding, temporal reasoning, and cross-environment generalization.

PUBLICATIONS

[1] **Chaojun Ni**, Xiaofeng Wang, Zheng Zhu, Weijie Wang, Haoyun Li, Guosheng Zhao, Jie Li, Wenkang Qin, Guan Huang, Wenjun Mei, "WonderTurbo: Generating Interactive 3D World in 0.72 Seconds", IEEE International Conference on Computer Vision (***ICCV***), 2025.

[2] **Chaojun Ni**, Guosheng Zhao, Xiaofeng Wang, Zheng Zhu, Wenkang Qin, Guan Huang, Chen Liu, Yuyin Chen, Yida Wang, Xueyang Zhang, et al., "Recondreamer: Crafting World Models for Driving Scene Reconstruction via Online Restoration", IEEE Conference on Computer Vision and Pattern Recognition (***CVPR***), 2025.

[3] **Chaojun Ni**, Wenhui Jiang, Chao Cai, Qishou Zhu, Yuming Fang, "Feature Adaptive YOLO for Remote Sensing Detection in Adverse Weather Conditions", IEEE International Conference on Visual Communications and Image Processing (***VCIP***), 2023.

[4] Guosheng Zhao, Xiaofeng Wang, **Chaojun Ni**, Zheng Zhu, Wenkang Qin, Guan Huang, Xingang Wang, "Recondreamer++: Harmonizing Generative and Reconstructive Models for Driving Scene Representation", IEEE International Conference on Computer Vision (***ICCV***), 2025.

[5] Guosheng Zhao, **Chaojun Ni**, Xiaofeng Wang, Zheng Zhu, Xueyang Zhang, Yida Wang, Guan Huang, Xinze Chen, Boyuan Wang, Youyi Zhang, et al., "Drivedreamer4d: World Models are Effective Data Machines for 4D Driving Scene Representation", IEEE Conference on Computer Vision and Pattern Recognition (***CVPR***), 2025.

[6] Boyuan Wang, Xiaofeng Wang, **Chaojun Ni**, Guosheng Zhao, Zhiqin Yang, Zheng Zhu, Muyang Zhang, Yukun Zhou, Xinze Chen, Guan Huang, et al., "Humandreamer: Generating Controllable Human-Motion Videos via Decoupled Generation", IEEE Conference on Computer Vision and Pattern Recognition (***CVPR***), 2025.

[7] **Chaojun Ni**, Cheng Chen, Xiaofeng Wang, Zheng Zhu, Wenzhao Zheng, Boyuan Wang, Tianrun Chen, Guosheng Zhao, Haoyun Li, Zhehao Dong, et al., "SwiftVLA: Unlocking Spatiotemporal Dynamics for Lightweight VLA Models at Minimal Overhead", arXiv preprint arXiv:2512.00903, 2025.

[8] **Chaojun Ni**, Guosheng Zhao, Xiaofeng Wang, Zheng Zhu, Wenkang Qin, Xinze Chen, Guanghong Jia, Guan Huang, Wenjun Mei, "Recondreamer-RL: Enhancing Reinforcement Learning via Diffusion-Based Scene Reconstruction", arXiv preprint arXiv:2508.08170, 2025.

[9] **Chaojun Ni**, Jie Li, Haoyun Li, Hengyu Liu, Xiaofeng Wang, Zheng Zhu, Guosheng Zhao, Boyuan Wang, Chenxin Li, Guan Huang, et al., "Wonderfree: Enhancing Novel View Quality and Cross-View Consistency for 3D Scene Exploration", arXiv preprint arXiv:2506.20590, 2025.

[10] GigaWorld Team, Angen Ye, Boyuan Wang, **Chaojun Ni**, Guan Huang, Guosheng Zhao, Haoyun Li, Jiagang Zhu, Kerui Li, Mengyuan Xu, et al., "Gigaworld-0: World Models as Data Engine to Empower Embodied AI", arXiv preprint arXiv:2511.19861, 2025.

[11] Weijie Wang, Jiagang Zhu, Zeyu Zhang, Xiaofeng Wang, Zheng Zhu, Guosheng Zhao, **Chaojun Ni**, Haoxiao Wang, Guan Huang, Xinze Chen, et al., "Drivegen3d: Boosting Feed-Forward Driving Scene Generation with Efficient Video Diffusion", arXiv preprint arXiv:2510.15264, 2025.

[12] Zhehao Dong, Xiaofeng Wang, Zheng Zhu, Yirui Wang, Yang Wang, Yukun Zhou, Boyuan Wang, **Chaojun Ni**, Runqi Ouyang, Wenkang Qin, et al., "Emma: Generalizing Real-World Robot Manipulation via Generative Visual Transfer", arXiv preprint arXiv:2509.22407, 2025.

[13] Haoyun Li, Ivan Zhang, Runqi Ouyang, Xiaofeng Wang, Zheng Zhu, Zhiqin Yang, Zhentao Zhang, Boyuan Wang, **Chaojun Ni**, Wenkang Qin, et al., "Mimicdreamer: Aligning Human and Robot Demonstrations for Scalable VLA Training", arXiv preprint arXiv:2509.22199, 2025.

[14] Boyuan Wang, Xinpan Meng, Xiaofeng Wang, Zheng Zhu, Angen Ye, Yang Wang, Zhiqin Yang, **Chaojun Ni**, Guan Huang, Xingang Wang, "Embodiedreamer: Advancing Real2Sim2Real Transfer for Policy Training via Embodied World Modeling", arXiv preprint arXiv:2507.05198, 2025.

[15] Boyuan Wang, Runqi Ouyang, Xiaofeng Wang, Zheng Zhu, Guosheng Zhao, **Chaojun Ni**, Guan Huang, Lihong Liu, Xingang Wang, "Humandreamer-X: Photorealistic Single-Image Human Avatars Reconstruction via Gaussian Restoration", arXiv preprint arXiv:2504.03536, 2025.

[16] GigaBrain Team, Angen Ye, Boyuan Wang, **Chaojun Ni**, Guan Huang, Guosheng Zhao, Haoyun Li, Jie Li, Jiagang Zhu, Lv Feng, et al., "Gigabrain-0: A World Model-Powered Vision-Language-Action Model", arXiv e-prints, 2025.

[17] Xiaolei Lang, Yang Wang, Yukun Zhou, **Chaojun Ni**, Kerui Li, Jiagang Zhu, Tianze Liu, Jiajun Lv, Xingxing Zuo, Yun Ye, Guan Huang, Xiaofeng Wang, Zheng Zhu, "Vag: Dual-Stream Video-Action Generation for Embodied Data Synthesis".

[18] **Chaojun Ni**, Xiaofeng Wang, Zheng Zhu, Weijie Wang, Haoyun Li, Guosheng Zhao, Jie Li, Wenkang Qin, Guan Huang, Wenjun Mei, "*Wonderturbo++*: Generating Interactive 3D World in 0.58 Seconds".

[19] Chenxin Li, Hengyu Liu, Panwang Pan, Yifan Liu, **Chaojun Ni**, Zhiqin Yang, Yunlong Lin, Kairun Wen, Zheng Zhu, Guan Huang, Yixuan Yuan, "Worldagent: Empowering 3D Scene Interaction via an Intelligent Versatile Field-Driven Agent".

PROFESSIONAL SERVICE

- Reviewer for Conferences
  - AAAI Conference on Artificial Intelligence (AAAI)
  - Computer Vision and Pattern Recognition (CVPR)

REFEREE CONTACTS

- Wenjun Mei, mei@pku.edu.cn, College of Engineering, Peking University

- Zheng Zhu, zhengzhu@ieee.org, GigaAI