

Cong Wan (万聪)

17711081241 | wancong@stu.xjtu.edu.cn | [Homepage](#)

 [Github](#) |  [Google-Scholar](#) |  [import_torch-as_tf](#)



Education

Xi'an Jiaotong University	2024.09-2027.06
Master of Computer Science and Technology, Advisor: Yihong Gong	Xi'an
Xi'an Jiaotong University	2020.09-2024.06
Bachelor of Mathematics	Xi'an

Experience

Seed, ByteDance	2025.11- Now
Research Intern, Lead by Chang Zhou	Beijing, China
Foundation Model, Embodied AI	
ML Lab, MSRA	2025.06-2025.11
Research Intern, Lead by Li Zhao	Beijing, China
World Model	
DAMO, Alibaba	2024.12-2025.06
Research Intern, Lead by Hao Luo	Hangzhou, China
AIGC, Unified model	
MIV Lab, Xi'an Jiaotong University	2024.09-2027.06
Master, Advised by Yihong Gong	Xi'an, China
Foundation Model, RL, World Model	

Publications

- [C.12] **DataClaw0: Agentic Tailoring Multimodal Data from Raw Streams**
Cong Wan, Zeyu Guo, Zijian Cai, Jiangyang Li, SongLin Dong, Lin Peng, Xiangyang Luo, Zhiheng Ma, Yihong Gong
NIPS, 2026 Under review
- [C.11] **Retrieve-then-Steer: Online Success Memory for Test-Time Adaptation of Generative VLAs**
Jianchao Zhao, Huoren Yang, Hu Yusong, Yuyang Gao, Qiguan Ou, Cong Wan, SongLin Dong, Zhiheng Ma, Yihong Gong
NIPS, 2026 Under review
- [C.10] **ProSR: Process-Shaped Spatial Reasoning for Reliable Chain-of-Thought in VLMs**
Jiangyang Li, Cong Wan, Changjie Wu, SongLin Dong, Lingjun Zhang, Linzhe Shi, Xu Wang, Zhiheng Ma, Hang Zhang, Mu Xu, Yihong Gong
NIPS, 2026 Under review
- [C.09] **WorldCanvas: Context-Guided Embodied Video Synthesis with a Unified Visual Canvas**
Cong Wan, Jingzhou You, SongLin Dong, Kailai Chen, Zhiheng Ma, Yihong Gong
NIPS, 2026 Under review
- [C.08] **Grounding Visual Chain-of-Thought in Evidence**
Lin Peng, Cong Wan, Zijian Cai, Jiangyang Li, Zeyu Guo, Xiangyang Luo, Dian Yang, SongLin Dong, Yihong Gong
AAAI, 2026 Under review
- [C.07] **CoRe: A Comprehensive Framework for Cross-Image Comparative Reasoning in VLM**

Lin Peng, **Cong Wan**, Zeyu Guo, Songlin Dong, Yihong Gong
MM, 2026

- [C.06] **ReMoT: Reinforcement Learning with Motion Contrast Triplets.**
Cong Wan, Zeyu Guo, Jiangyang Li, SongLin Dong, Yifan Bai, Lin Peng, Zhiheng Ma, Yihong Gong
CVPR, 2026 **Highlight(Top 3.8%)**
- [C.05] **Trajectory-Diversity-Driven: Robust Vision-and-Language Navigation**
Jiangyang Li, **Cong Wan**, SongLin Dong, Chenhao Ding, Qiang Wang, Zhiheng Ma, Yihong Gong
CVPR, 2026 *Finding*
- [C.04] **Canonswap: High-fidelity and consistent video face swapping via canonical space modulation**
Xiangyang Luo, Ye Zhu, Yunfei Liu, Lijian Lin, **Cong Wan**, Zijian Cai, Yu Li, Shao-Lun Huang
ICCV, 2025
- [C.03] **CIA: Class-and Instance-aware Adaptation for Vision-Language Models**
Lin Peng, **Cong Wan**, Shaokun Wang, Xiang Song, Yuhang He, Yihong Gong
MM, 2025
- [C.02] **Grid: Omni Visual Generation**
Cong Wan, Xiangyang Luo, Hao Luo, Zijian Cai, Yiren Song, Yunlong Zhao, Yifan Bai, Fan Wang,
Yuhang He, Yihong Gong
Technical Report, 2024
- [C.01] **Prompt-Agnostic Adversarial Perturbation: for Personalized Diffusion Models**
Cong Wan, Yuhang He, Xiang Song, Yihong Gong
NIPS, 2024
- [J.01] **SDA: Structure-aware Distribution Alignment for Vision-Language Models**
Lin Peng, **Cong Wan***, Shaokun Wang, Yuhang He, Yihong Gong
TMM, 2026

Academic Service

Conference Reviewer: NeurIPS
Journal Reviewer: TMM

Honors and Awards

Outstanding Undergraduate
National Scholarship
Special-Class Scholarship
National Postgraduate Mathematical Modeling Competition, Second Prize
National Undergraduate Mathematics Competition, Second Prize
Mathematical Contest in Modeling (MCM), Meritorious Winner
National High School Mathematics League, Second Prize