

Jiaxin Huang

Ph.D. Candidate, Machine Learning, MBZUAI
Project Leader, Embodied Navigation Team, AgiBot X-Lab

Last updated on Oct. 2026.

Machine Learning Department, MBZUAI
Abu Dhabi, United Arab Emirates
✉ jiaxin.huang@mbzuai.ac.ae
[GitHub](#), [LinkedIn](#)
[Google Scholar](#), [Personal Home Page](#)

Research Interests

Areas: 3D/4D World Models, seen from four perspectives: (1) Spatial Reasoning [[C7](#), [C6](#), [S14](#)]; (2) 3D/4D Reconstruction [[S1](#), [S2](#), [S9](#)]; (3) Embodied Navigation [[S16](#), [S15](#), [C9](#), [S5](#)]; (4) Safety of 3D/4D Generative Models [[S13](#), [S12](#), [S8](#), [S4](#)] (co-organizer of the [LifeGenIP Workshop at ECCV 2026](#)).

My research builds *3D/4D world models*: models that perceive, reason, act, and stay trustworthy in dynamic 3D environments, despite scarce 3D supervision and the gap between simulation scores and real robots. I study one world from four perspectives. (1) **Spatial Reasoning**: SURPRISE3D, a large-scale benchmark for spatial understanding and reasoning in complex 3D scenes, and MLLM-For3D, the first label-free framework that transfers 2D MLLM reasoning to 3D reasoning segmentation (both NeurIPS 2025). (2) **3D/4D Reconstruction**: semantic-guided Gaussian splatting for dynamic urban scenes and physically grounded Gaussian world models from video. (3) **Embodied Navigation**: as Project Leader of the Embodied Navigation Team at AgiBot X-Lab, I lead vision-and-language navigation for the mass-produced Lingxi X2 humanoid, from a simulation-to-real data engine to policies deployed on the robot. (4) **Safety of 3D/4D Generative Models**: protecting videos and 4D assets against unauthorized reconstruction and reuse (Safe4D, 4DGuard). Representative contributions include two (co-)first-author NeurIPS 2025 papers that define 3D spatial reasoning segmentation, an industrial team whose navigation policies run zero-shot on a production humanoid, and six ICLR 2027 submissions (four as first author); 220+ citations, h-index 9 (Oct. 2026). Before my Ph.D. I worked on cross-domain medical image reconstruction (AAAI, IEEE TMM, ICIP).

Academic Experience

- Sept. 2024 - **Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)**, Abu Dhabi, UAE.
Present Doctor of Philosophy (Ph.D.), Machine Learning.
Supervisors: [Prof. Tongliang Liu](#) and [Prof. Steve Liu](#).
- Sept. 2021 - **University of Electronic Science and Technology of China (UESTC)**, Chengdu, China.
July 2024 Master of Computer Science and Technology.
- Sept. 2017 - **University of Electronic Science and Technology of China (UESTC)**, Chengdu, China.
July 2021 Bachelor of Automation (B.Sc.). GPA: 3.89/4.00 (Top 10%).
National Scholarship (Ranking: 1/297); Outstanding Graduate of UESTC.
The University of Sydney, Sydney, Australia.
Visiting Student.
University of California, Berkeley, Berkeley, USA.
Summer Session.

Industrial Experience

Apr. 2026 - **AgiBot, X-Lab, Shenzhen, China.**

Oct. 2026 **Research Intern; Project Leader, Embodied Navigation Team.**

Mentors: [Dr. Shaoli Huang](#); [Zhihui Peng](#) (Co-Founder of AgiBot).

- Led the Embodied Navigation Team's vision-and-language navigation (VLN) work for AgiBot's mass-produced Lingxi X2 humanoid—model training, simulation and data, and robot deployment—with researchers from AgiBot, MBZUAI, the University of Melbourne, UESTC, and the University of Macau; the team produced two ICLR 2027 submissions in one internship term.
- Project leader of *GuidedNav* (arXiv:2610.13122), which shapes VLN representations with training-time subtask, landmark, and spatial guidance; among monocular-RGB policies it reports state-of-the-art success on R2R-CE and RxR-CE val-unseen (+8.1% / +13.1% over the strongest prior results) with faster inference, and released the Guided-R2R / Guided-RxR annotation pipelines.
- Project leader of *NavTTT*, an adaptive parametric memory built by hierarchical test-time training, which lets a VLN policy retain long-horizon history beyond the context window and follows instructions on a real robot (under review at ICLR 2027).
- Designed the AgiBot Nav technical report and its experiment system: a three-layer data engine (synthetic scenes, 3DGS-reconstructed real venues, and real-robot collection), a Real2Sim2Real pipeline from handheld scanning of a venue to a fine-tuned, deployable policy, and a unified real-robot evaluation protocol; deployed VLN policies zero-shot on the X2 humanoid through an agent runtime with a safety gateway.

Research Experience

3D/4D World Models & Spatial Reasoning (navigation at AgiBot: see Industrial Experience)

Jan. 2026 - **3D & 4D Spatial Reasoning, A Comprehensive Review.**

Present First Author, Project Leader

Designing and leading a comprehensive review on 3D and 4D spatial reasoning, covering scene understanding, world models, and embodied intelligence.

2026 **EgoAudit, Auditing Egocentric Spatial Reasoning in 3D Vision-Language Models.**

First Author (submitted to ICLR 2027)

An auditing framework that removes, perturbs, or replaces the viewpoint given to a 3D VLM, showing that egocentric scores measure viewpoint use rather than perspective taking.

2026 **Safe4D & 4DGuard, Proactive Protection of Videos and 4D Gaussian Assets.**

First Author (two submissions to ICLR 2027)

Safe4D is the first video protection against 4D reconstruction under causal, irrevocable 8-bit release: perturbations are transported along scene motion with a purification-aware objective and verified commits, and stay robust on four dynamic benchmarks. 4DGuard embeds the perturbation into the 4D Gaussian asset itself, so every rendered frame resists segmentation and reuse while looking unchanged.

Feb. 2025 - **SURPRISE3D, Dataset & Benchmark for 3D Spatial Reasoning Segmentation.**

June 2025 First Author, Project Leader

Designed and led a large-scale dataset and benchmark for 3D reasoning segmentation. Established the 3D Spatial Reasoning Segmentation (3D-SRS) benchmark for 3D-LLMs. Published at NeurIPS 2025.

Sept. 2024 - **MLLM-For3D, Adapting MLLM for 3D Reasoning Segmentation.**

June 2025 First Author

Proposed the first label-free framework transferring multimodal large language model (MLLM) reasoning abilities from 2D to 3D. Achieved SOTA performance on 3D reasoning segmentation benchmarks without using any labeled data. Published at NeurIPS 2025.

Dec. 2024 - **UrbanGS, Semantic-Guided 4D Gaussian Splatting.**

Mar. 2025 Co-First Author

Introduced a semantic-guided 4D Gaussian Splatting (4DGS) approach for dynamic urban scene reconstruction. Demonstrated superior reconstruction quality in large-scale urban scenes.

Previous Research: Medical Image Analysis & AI for Industry

2021 - 2024 **Cross-Domain Medical Image Reconstruction, UESTC.**

Led research on cross-domain low-dose CT image denoising using generative models and domain adaptation. Published first-authored papers at AAAI 2024, IEEE TMM 2024, and ICIP 2022 (Oral).

Publications

By October 2026 I have produced 18 peer-reviewed publications (10 as (co-)first author) at NeurIPS, CVPR, ECCV, ACL, EMNLP, AAAI, ACM MM, IEEE TMM, IEEE TIM and IEEE IoT-J, plus 16 preprints and papers under review, including six ICLR 2027 submissions (four as first author, two as project leader); 220+ citations (h-index 9) on [Google Scholar](#). **Detailed publication links are accessible via the item numbers listed below when available.**

* equal contribution; † project leader.

Preprints & Under Review

- S16** Jingyu Guo, **Jiaxin Huang**[†], Chengxing Lin, Xiaoyu Luo, Cheng Wen, Shanshan Ye, Tingjin Chu, Shaoli Huang, Mingming Gong. “*GuidedNav: Shaping Vision-Language Navigation Representations through Pre-Action Attention and Spatial Guidance*”. arXiv preprint arXiv:2610.13122, 2026. Under review at ICLR 2027. [\[Project page\]](#)
- S15** Xiaoyu Luo, Chengxing Lin, **Jiaxin Huang**[†], Jiale Liu, Jingyu Guo, Sirui Chen, Aolin Li, Shaoli Huang, Cheng-zhong Xu, Hui Kong. “*NavTTT: Adaptive Parametric Memory via Hierarchical Test-Time Training for Vision-and-Language Navigation*”. Under review at ICLR 2027.
- S14** **Jiaxin Huang**, Chengcheng Lin, Xiyang Huang, Yankai Chen, Min Peng, Qianqian Xie. “*EgoAudit: Auditing Egocentric Spatial Reasoning in 3D Vision-Language Models*”. Under review at ICLR 2027.
- S13** **Jiaxin Huang**, Ziming Hong, Mingming Gong, Tongliang Liu. “*Safe4D: Causal Video Protection against Unauthorized 4D Reconstruction*”. Under review at ICLR 2027.
- S12** **Jiaxin Huang**, Ziming Hong. “*4DGuard: Towards Proactive 4D Gaussian Assets Protection for Dynamic Novel-View Synthesis*”. Under review at ICLR 2027.
- S11** **Jiaxin Huang**, Hai Huang. “*Position: Interaction Understanding Requires More Than Object-Centric Representations*”. Under review at ICLR 2027.
- S10** Ye Yuan, Weien Li, Rui Song, Zeyu Li, Haochen Liu, Xiangyu Kong, Zixuan Dong, Linfeng Du, Zipeng Sun, Weixu Zhang, **Jiaxin Huang**, Changjiang Han, Yonghan Yang, Zichen Zhao, Xiuyuan Hu, Haolun Wu, Yankai Chen, Fengran Mo, Jikun Kang, Bowei He, Dawn Song, Philip S. Yu, Xue Liu. “*Discrete Diffusion Models: A Unified Framework from Tokenization to Generation*”. arXiv preprint arXiv:2607.13431, 2026.
- S9** Jiaxin Yan, Chaoning Zhang, Xudong Wang, Pengcheng Zheng, Ya Wen, Qigan Sun, **Jiaxin Huang**, Shuxu Chen, Yang Yang, Hyundong Shin. “*Predicting the World via Video Representation: A Comprehensive Survey on Video World Models*”. Preprints, 2026.
- S8** Jianwei Zhang, Sihan Cao, Chaoning Zhang, Ziming Hong, **Jiaxin Huang**, Pengcheng Zheng, Caiyan Qin, Wei Dong, Yang Yang, Tongliang Liu. “*Immunizing 3D Gaussian Generative Models Against Unauthorized Fine-Tuning via Attribute-Space Traps*”. arXiv preprint arXiv:2604.09688, 2026.

- S7 Jianwei Zhang, Chaoning Zhang, Sihan Cao, Wang Liu, Pengcheng Zheng, **Jiaxin Huang**, Caiyan Qin, arXiv Yalan Ye, Wei Dong, Yang Yang. “RCP: Representation Consistency Pruner for Mitigating Distribution Shift in Large Vision-Language Models”. arXiv preprint arXiv:2604.04972, 2026.
- S6 Yingxu Wang, Kunyu Zhang, **Jiaxin Huang**, Mengzhu Wang, Mingyan Xiao, Siyang Gao, Nan Yin. “DSBD: Dual-Aligned Structural Basis Distillation for Graph Domain Adaptation”. arXiv preprint arXiv:2604.03154, 2026.
- S5 Gaoge Han, Zhengqing Gao, Ziwen Li, **Jiaxin Huang**, Shaoli Huang, Fakhri Karray, Mingming Gong, arXiv Tongliang Liu. “ExecVLA: Following Fine-Grained Execution Constraints in Vision-Language-Action Models with Bi-Level Action Representation”. arXiv preprint arXiv:2603.17524, 2026.
- S4 Longjie Zhao, Ziming Hong, **Jiaxin Huang**, Runnan Chen, Mingming Gong, Tongliang Liu. “Intellectual arXiv Property Protection for 3D Gaussian Splatting Assets: A Survey”. arXiv preprint arXiv:2602.03878, 2026.
- S3 Pengcheng Zheng, Chaoning Zhang, Mingzheng Cui, Guo Chen, Qigan Sun, **Jiaxin Huang**, Jiaquan Zhang, TechRxiv Tae-Ho Kim, Caiyan Qin, Yazhou Ren, Guoqing Wang, Yang Yang. “Towards Visual Chain-of-Thought Reasoning: A Comprehensive Survey”. TechRxiv, 2026.
- S2 Zhengqing Gao, Ziwen Li, Xin Wang, **Jiaxin Huang**, Zhenyang Ren, Mingkai Shao, Hanlue Zhang, Tianyu arXiv Huang, Yongkang Cheng, Yandong Guo, Runqi Lin, Yuanyuan Wang, Tongliang Liu, Kun Zhang, Mingming Gong. “Mirage2Matter: A Physically Grounded Gaussian World Model from Video”. arXiv preprint arXiv:2602.00096, 2026.
- S1 Ziwen Li*, **Jiaxin Huang***, Runnan Chen, Yunlong Che, Yandong Guo, Tongliang Liu, Fakhri Karray, arXiv Mingming Gong. “UrbanGS: Semantic-Guided Gaussian Splatting for Urban Scene Reconstruction”. arXiv preprint arXiv:2412.03473, 2024.

Journal Articles

- J7 Pengcheng Zheng, Xiaorong Pu, Kecheng Chen, **Jiaxin Huang**, Meng Yang, Bai Feng, Yazhou Ren, Jianan IEEE TMM Jiang, Chaoning Zhang, Yang Yang, Heng Tao Shen. “Joint Lossless Compression and Steganography for Medical Images via Large Language Models”. IEEE Transactions on Multimedia, 2026.
- J6 Yingxu Wang, Kunyu Zhang, **Jiaxin Huang**, Nan Yin, Siwei Liu, Eran Segal. “ProtoMol: Enhancing Brief. Bioinf. Molecular Property Prediction via Prototype-Guided Multimodal Learning”. Briefings in Bioinformatics, 26(6), bbaf629, 2025.
- J5 Naixin Zhou*, **Jiaxin Huang***, Yijiu Zhao, Shibo Chen, Jiaming Zhao, Feiyu Long. “Data-Driven Soft Fault IEEE TIM Diagnosis for Analog Circuits Based on Contrastive Learning”. IEEE Transactions on Instrumentation and Measurement, 2025.
- J4 **Jiaxin Huang**, Kecheng Chen, Yazhou Ren, Jiayu Sun, Xiaorong Pu, Ce Zhu. “Cross-Domain Low-Dose CT IEEE TMM Image Denoising with Semantic Preservation and Noise Alignment”. IEEE Transactions on Multimedia, 26, pp. 8771–8782, 2024.
- J3 Siyan Liu*, Chang Wu*, **Jiaxin Huang***, Ying Zhang, Ming Ye, Yuhang Huang. “Blockchain-Based IEEE IoT-J Interpretable Electric Vehicle Battery Life Prediction in IoV”. IEEE Internet of Things Journal, 11(4), pp. 7214–7227, 2023.
- J2 **Jiaxin Huang**, Kecheng Chen, Yazhou Ren, Jiayu Sun, Yanmei Wang, Tao Tao, Xiaorong Pu. “CDDnet: CBM Cross-Domain Denoising Network for Low-Dose CT Image via Local and Global Information Alignment”. Computers in Biology and Medicine, 163, 107219, 2023.
- J1 Xiaorong Pu*, **Jiaxin Huang***, Junchi Liu, Jiayu Sun, Jixiang Luo, Yue Zhao, Kecheng Chen, Yazhou Ren. FDC “A Survey on Clinical Oriented CT Image Denoising”. Frontiers of Data and Computing, 3(6), pp. 35–49, 2022.

Conference Proceedings

- C11** Xiyang Huang, Jiawei Lin, Keying Wu, **Jiaxin Huang**, Kailai Yang, Renxiong Wei, Jiayi Xiang, Ziyang Kuang, Min Peng, Qianqian Xie, Sophia Ananiadou. “*Siming-Bench: Evaluating Procedural Correctness from Continuous Interactions in Clinical Skill Videos*”. In the 2026 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2026.
EMNLP Oral
- C10** Yingxu Wang, **Jiaxin Huang**, Mengzhu Wang, Nan Yin. “*TRACE: An Experiential Framework for Coherent Multi-hop Knowledge Graph Question Answering*”. In the 64th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 1830–1851, 2026.
ACL
- C9** Jingyu Guo, Ziyue Chen, Ziwen Li, Zhengqing Gao, **Jiaxin Huang**, Hanlue Zhang, Fengming Huang, Yu Yao, Tongliang Liu, Mingming Gong. “*HUGE-Bench: A Benchmark for High-Level UAV Vision-Language-Action Tasks*”. In the 19th European Conference on Computer Vision (ECCV), pp. 373–388, 2026.
ECCV
- C8** Zhenchen Wan, Ce Chen, Runqi Lin, **Jiaxin Huang**, Tianxi Chen, Yanwu Xu, Tongliang Liu, Mingming Gong. “*Mobile-VTON: High-Fidelity On-Device Virtual Try-On*”. In the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 38081–38090, 2026.
CVPR
- C7** **Jiaxin Huang***, Ziwen Li*, Hanlue Zhang*, Runnan Chen, Zhengqing Gao, Xiao He, Yandong Guo, Wenping Wang, Tongliang Liu, Mingming Gong. “*SURPRISE3D: A Dataset for Spatial Understanding and Reasoning in Complex 3D Scenes*”. In Advances in Neural Information Processing Systems 38 (NeurIPS), Datasets and Benchmarks Track, 2025.
NeurIPS D&B Track
- C6** **Jiaxin Huang**, Runnan Chen, Ziwen Li, Zhengqing Gao, Xiao He, Yandong Guo, Mingming Gong, Tongliang Liu. “*MLLM-For3D: Adapting Multimodal Large Language Model for 3D Reasoning Segmentation*”. In Advances in Neural Information Processing Systems 38 (NeurIPS), 2025.
NeurIPS
- C5** Kecheng Chen, Pingping Zhang, Hui Liu, Jie Liu, Yibing Liu, **Jiaxin Huang**, Shiqi Wang, Hong Yan, Haoliang Li. “*Large Language Models for Lossless Image Compression: Next-Pixel Prediction in Language Space is All You Need*”. In Advances in Neural Information Processing Systems 38 (NeurIPS), 2025.
NeurIPS
- C4** Pengcheng Zheng, Kecheng Chen, **Jiaxin Huang**, Bohao Chen, Ju Liu, Yazhou Ren, Xiaorong Pu. “*Lightweight Medical Image Restoration via Integrating Reliable Lesion-Semantic Driven Prior*”. In the 33rd ACM International Conference on Multimedia (ACM MM), pp. 2977–2986, 2025.
ACM MM
- C3** **Jiaxin Huang**, Qi Wu, Yazhou Ren, Fan Yang, Aodi Yang, Qianqian Yang, Xiaorong Pu. “*Sparse Bayesian Deep Learning for Cross Domain Medical Image Reconstruction*”. In the 38th AAAI Conference on Artificial Intelligence (AAAI), 38(3), pp. 2339–2347, 2024.
AAAI
- C2** **Jiaxin Huang**, Kecheng Chen, Jiayu Sun, Xiaorong Pu, Yazhou Ren. “*Cross Domain Low-Dose CT Image Denoising with Semantic Information Alignment*”. In the IEEE International Conference on Image Processing (ICIP), pp. 4228–4232, 2022.
ICIP Oral
- C1** Kecheng Chen*, **Jiaxin Huang***, Jiayu Sun, Yazhou Ren, Xiaorong Pu. “*Task-Driven Deep Learning for LDCT Image Denoising*”. In the 4th International Symposium on Image Computing and Digital Medicine (ISICDM), pp. 35–39, 2020.
ISICDM

Honors & Awards

- 2020 National Scholarship (Ranking: 1/297), *Ministry of Education, China*
- 2021 Outstanding Graduate Student of UESTC & First-Class Student Scholarship (Top 10%), *UESTC*
- 2020 The First Prize (only) in the Challenge of Optimization of Low-dose CT Image Quality, *ISICDM 2020*
- 2022 Rank 370/2992 in IEEEExtreme 16.0 Programming Competition, *IEEE*
- 2019 The Second Prize of the 15th “Challenge Cup” Competition (Team Leader)
- 2019 The Third Prize of the 3rd Sichuan Provincial Undergraduate Optoelectronics Design Competition
- 2019 The First Prize of “FLTRP Cup” English Public Speaking Contest

2018 The “Liren” Elite of UESTC (one of 5 students awarded in the entire university), *UESTC*

2018 The First-Class Undergraduate Student Scholarship (Ranking: 1/297), *UESTC*

Teaching & Mentorship

Teaching

- “*ML813 Dimensionality Reduction and Manifold Learning*” (Ph.D. course), Graduate TA, MBZUAI. Spring 2026.
- “*Linear Algebra and Calculus*”, Teaching Assistant for the first cohort of AI undergraduates at MBZUAI. 2025.
- “*Theory and Application of Neural Networks*”, Graduate Teaching Assistant, UESTC. 2022.

Mentorship

- Mentored the AgiBot X-Lab VLN interns (2026), resulting in two ICLR 2027 submissions (GuidedNav and NavTTT).
 - Xiaoyu Luo (Univ. of Macau)
 - Aolin Li (Wuhan University)
 - Sirui Chen (AgiBot)
 - Jingyu Guo (Univ. of Melbourne)
 - Chengxing Lin (UESTC)
- Mentored undergraduates and graduates in research projects, resulting in 2 CCF-A papers (incl. ACM MM 2025) and 3 IEEE Transactions papers (TMM 2026, TIM 2025, IoT-J 2023). **Highlight: mentored Xiyang Huang (Ph.D. student, Wuhan University) to a first-author EMNLP 2026 Oral paper, Siming-Bench [C11].**
 - Pengcheng Zheng (UESTC)
 - Jianwei Zhang (UESTC)
 - Xiyang Huang (Wuhan University)
 - Naixin Zhou (UESTC)
 - Yintong Wang (MBZUAI)
 - Siyan Liu (UESTC)
 - Xize Duan (McGill University)

Professional & Community Service

- **Workshop Organization:** Co-organizer, [LifeGenIP: Life-Cycle Intellectual Property Governance of Visual Generative Models](#), ECCV 2026 Workshop, Malmö, Sweden.
- **Conference Reviewer**
 - CVPR / ECCV / ICCV 2024, 2025, 2026
 - NeurIPS 2025, 2026
 - ICLR / ICML 2025, 2026
 - AAAI 2025

Skills & Activities

- **Technical:** Python, PyTorch, multi-GPU training; VLM post-training (Qwen-VL, LoRA, DAgger); Isaac Sim, Habitat/VLN-CE; 3DGS and 4D reconstruction; real-robot deployment (ROS, RGB-D calibration, safety gating).
- **Languages:** Mandarin Chinese (native); English (TOEFL 103); German (A2, studied in Munich for half a year).
- **Interests:** Violin, CrossFit, Public (Impromptu) Speech.

References

- **Tongliang Liu (Ph.D. Co-Advisor)** [\[Homepage\]](#)
Director, Sydney AI Centre; Director, Trustworthy ML Lab;
Deputy Associate Head of School (Research)
Professor, *School of Computer Science, Faculty of Engineering, The University of Sydney, Australia*
- **Steve Liu (Ph.D. Co-Advisor)** [\[Homepage\]](#)
Fellow of the Canadian Academy of Engineering (FCAE);
Fellow of the IEEE (FIEEE)
Associate VP of Research, Professor of Computer Science & Machine Learning, *MBZUAI, Abu Dhabi, UAE*
Professor & William Dawson Scholar, *School of Computer Science, McGill University, Montreal, Canada*
- **Mingming Gong (Research Collaborator)** [\[Homepage\]](#)
Associate Professor in Data Science, *School of Mathematics and Statistics, University of Melbourne, Australia*
Associate Professor in Machine Learning (Affiliated), *Department of Machine Learning, MBZUAI, Abu Dhabi, UAE*
- **Shaoli Huang (Industrial Mentor)** [\[Homepage\]](#)
AgiBot X-Lab, Shenzhen, China; [Google Scholar](#)