



Qinghongbing Xie
Tsinghua University
M.Eng. in Intelligent Manufacturing
Beijing, China

+86-180-7992-9847
binicey@outlook.com
davidxie.online
David Xie

Education

Tsinghua University 2023.09 – Present
M.Eng. in Mechanical Engineering (Intelligent Manufacturing) – Recommended Admission GPA: 3.97/4.0, Top 1%
Advisor: Prof. Long Zeng

Beijing Institute of Technology 2019.09 – 2023.06
B.Eng. in Intelligent Manufacturing Engineering / AI (Dual Degree) GPA: 92/100, Rank: 2
Advisors: Prof. Yaoguang Hu (Intelligent Manufacturing), Prof. Wenyao Zhang (AI)

Research Experience

Semantic Map Construction for Embodied Intelligence 2023.09 – Present
Master's Thesis Research, Tsinghua University Advisor: Prof. Long Zeng

- Built geometric-semantic maps using Open-Vocabulary detection and segmentation for open-vocabulary scene understanding
- Constructed 3D scene graphs via multimodal LLMs, enhancing robot spatial intelligence and scene reasoning
- Achieved indoor robot scene perception, reasoning, and 3D Visual Grounding; 1st-authored paper accepted at IEEE RA-L 2026

SenseTime – Algorithm Research Intern 2025.04 – Present
Spatiotemporal Reasoning Agent & Multimodal LLM Shenzhen
Advisors: Rui Zhao, Feng Zhu, Ziyue Li

- Proposed GTR-Bench, a novel benchmark for evaluating geographic spatial-temporal reasoning of VLMs across maps and multi-camera video networks with non-overlapping views
- Designed 7 geo-temporal task types spanning basic reasoning (geo-location, time-interval, motion-state) and combinatorial tasks (causal reordering, trajectory forecasting, multi-target tracking)
- Evaluated 10+ VLMs, revealing a critical gap between best model (Gemini-2.5-Pro, 34.9%) and human performance (78.61%) on spatial-temporal intelligence
- 1st-authored paper accepted at ICLR 2026 (CCF-A)

Tencent (Collaborative) – Algorithm Research Intern 2024.08 – 2025.03
Complex 3D Scene Layout Generation Remote
Advisors: Xiaoming Zhu, Zhi Deng

- Curated high-quality expert 3D layout dataset; fine-tuned Flux with LoRA for enhanced layout and asset memory
- Parsed layout guidance images via multimodal LLM and 3D geometric algorithms, generating scene graphs and pre-placement poses
- Optimized full-scene layout through semantic space optimization and physics simulation; 2nd-authored paper at SIGGRAPH Asia (TOG)

Experience

DJI – Algorithm Intern 2024.04 – 2024.07
Video Segmentation Pre-annotation & Temporal Stability Shenzhen

- Unlocked SAM2 multi-object segmentation with SOTA detection models, improving segmentation pass rate by 20%
- Deployed full pipeline to automated production line, increasing pre-annotation automation rate by 40%
- Surveyed video segmentation and temporal stability methods; findings adopted by the department

Pudu Robotics – Embodied Algorithm Intern 2024.04 – 2024.07
Vision-Language Grasping for Service Robots Shenzhen

- Implemented 2D real-time detection with YOLO-World and MobileSAM for front-end perception
- Built 3D scene reconstruction with point-cloud object-level reconstruction and CLIP semantic fusion
- Generated grasp points via GraspNet; deployed to table-top grasping and elevator operation for Pudu D7

Publications

- [1] Xie Q^{*}, Liang Z^{*}, Li F, Zeng L. DSM: Constructing a Diverse Semantic Map for 3D Visual Grounding[J]. IEEE RA-L, 2026. (Accepted, SCI Q1, IF 5.3)
Proposes DSM, a diverse semantic map construction framework for open-vocabulary 3D visual grounding in embodied intelligence.
- [2] Xie Q^{*}, Xia Z^{*}, Zhu F, Gong L, Li Z, Zhao R, Zeng L^{*}. GTR-Bench: Evaluating Geo-Temporal Reasoning in Vision-Language Models[C]. ICLR, 2026. (CCF-A)
Introduces the first comprehensive benchmark for evaluating geo-temporal reasoning capabilities of vision-language models.
- [3] Zhu X^{*}, Huang X^{*}, Xie Q, Deng Z, Yu J, Guan Y, Liu Z, Zhu L, Zhao Q, Liu L, et al. Imaginarium: Vision-guided High-Quality 3D Scene Layout Generation[J]. ACM TOG, 44(6): 1–24, 2025. (SCI Q1 Top, IF 7.8)
Presents a vision-guided framework for high-quality 3D scene layout generation using expert data and semantic optimization.
- [4] Zheng Q^{*}, Zhang C^{*}, Liang Z^{*}, Lin ET, Cui S, Xie Q, Xu Z, Zeng L. AssemMate: Graph-Based LLM for Robotic Assembly Assistance[C]. ICRA, IEEE, 2026.
Proposes AssemMate, a graph-based LLM framework for robotic assembly task planning and assistance.
- [5] Zheng Z^{*}, Li Z^{*}, Wang Y^{*}, Xie Q, Zeng L. Demonstrating DVS: Dynamic Virtual-Real Simulation Platform for Mobile Robotic Tasks[C]. RSS, 2025.
Builds a dynamic virtual-real simulation platform bridging sim-to-real gap for mobile robotic task evaluation.
- [6] Ni Z^{*}, Deng X^{*}, Tai C^{*}, Zhu X, Xie Q, Huang W, Wu X, Zeng L. Grid: Scene-graph-based instruction-driven robotic task planning[C]. IROS, IEEE, 2024: 13765–13772.
Develops a scene-graph-based approach for instruction-driven robotic task planning in unstructured environments.
- [7] Zhang C^{*}, Zhang C^{*}, Xu Z, Xie Q, Hou J, Feng P, Zeng L. Embodied intelligent industrial robotics: Concepts and techniques[J]. arXiv:2505.09305, 2025. (Under Review at JMS, SCI Q1 Top, IF 12.2)
Provides a comprehensive survey on concepts and techniques for embodied intelligence in industrial robotics.
^{*}Equal contribution.

Technical Skills

Programming: Python, C/C++, CUDA, Bash

Deep Learning: PyTorch, Large-scale Data Processing, Multi-GPU Training, LoRA Fine-tuning

Robotics: ROS, IsaacSim, STM32+CubeMX

Tools: Linux (Ubuntu), Git, Docker, LaTeX, Office, PR

Language: English (CET-4: 590, CET-6: 510), Chinese (Native)

Awards & Honors

National Scholarship Top 0.2% nationwide

Beijing Outstanding Graduate Beijing Municipality

National Encouragement Scholarship, University Outstanding Student Multiple University-level Scholarships