

Jiayao Li

Email: 2351405@tongji.edu.cn | Tel: +86 132 9198 9727 | [HomePage](#) | [GitHub](#)

Education

Tongji University

Bachelor of Computer Science and Technology (Elite Class) GPA: 4.66/5 (93.2/100) CET-6: 582 Shanghai, CN 09/2023 - 06/2027

Research interests: Embodied AI & Robotics; Physical-World AI & Simulation; 3D Perception & Spatial Understanding

Selected Competitions

Outstanding Winner & INFORMS Award (Top Prize, 0.12%) , Mathematical Contest in Modeling	May 2026
Second Prize , National Artificial Intelligence Application Scenario Innovation Challenge	Jan. 2026
Second Prize , Final Round of the 7th Global Campus Artificial Intelligence Algorithm Elite Competition	Dec. 2025
Third Prize , Lanqiao Cup National Software and Information Technology Professionals Competition	May 2025

Publications

SparseVLN: Action-aware Lossless Pruning for Vision-Language Navigation <i>Jiayao Li, Kaiwen Tuo, Chengju Liu, Qijun Chen</i> Work in Progress · Targeting ICLR 2027	Sept. 2026
Two-Stream Interactive Joint Learning of Scene Parsing and Geometric Vision Tasks <i>Guanfeng Tang, Hongbo Zhao, Ziwei Long, Jiayao Li, Bohong Xiao, Wei Ye, Hanli Wang, Rui Fan</i> Under review at IEEE Transactions on Intelligent Vehicles . arXiv preprint . [arXiv] [PDF]	Feb. 2026
Are LLMs Reliable for Recommending Libraries and Their Usage? An Empirical Study <i>Yuying Wang, Yilian Yu, Jiayao Li, Maojun Huang, et al.</i> Under review at Internetware 2026 (CCF C). [PDF]	May 2026
Can video generation replace cinematographers? Research on the cinematic language of generated video <i>Xiaoze Li, Kai Wu, Siyi Yang, YiZhan Qu, Guohua Zhang, Zhiyu Chen, Jiayao Li et al.</i> arXiv preprint . [arXiv] [PDF]	Dec. 2024

Research Experiences

Vision-and-Language Navigation, VLN — Inference Acceleration Study Mar.2026 – Present
Advisor: [Prof. Qijun Chen, Tongji University](#)

- Exploring long-horizon memory, spatial grounding, and closed-loop correction for real-time VLN, with an in-depth study of representative works such as StreamVLN; currently **leading a first-author project**, targeting ICLR 2027.

Scene Parsing and Geometric Vision — Two-Stream Interactive Joint Learning Framework Sep.2025 – Dec.2025
Advisor: [Prof. Rui Fan, Tongji University](#)

- Project Description:** Developed a two-stream interactive **joint learning framework for scene parsing and geometric vision tasks**. The framework enhances segmentation performance by incorporating cross-view geometric information from the geometric vision stream, improving semantic segmentation accuracy on *vKITTI2*, *Cityscapes*, and *KITTI 2015*.
- Personal Contributions:** Conducted **benchmark evaluation and visualization analysis** across multiple models, including *DFormerv2*. Experimentally validated the effectiveness of the *Cross-Task Adapter* for feature fusion. *TwInS* achieved up to a 7.23% improvement in semantic segmentation *mIoU* and up to a 13.64% improvement in instance segmentation *mAP*.
- Outcome:** **Fourth author of the manuscript** *Two-Stream Interactive Joint Learning of Scene Parsing and Geometric Vision Tasks*, currently under review at *IEEE Transactions on Intelligent Vehicles*.

Single-View 3D Occupancy Prediction and Spatial Perception — ViPOcc Framework Sep.2025 – Dec.2025
Advisor: [Prof. Rui Fan, Tongji University](#)

- Project Description:** Developed a **single-view 3D occupancy prediction framework**, leveraging depth and semantic-instance priors from *visual foundation models*. Designed *inverse-depth alignment* and *Semantic-guided Non-overlapping Gaussian Mixture Sampling (SNOG)* to improve 3D spatial perception from monocular RGB images.
- Personal Contributions:** Conducted **functional validation** of the *inverse-depth alignment module* and the *SNOG sampler*. Quantitatively improved performance, including a 2.9% increase in invisible-region accuracy, a 6.7% increase in object-level invisible-region recall, and a 6.2% reduction in Abs Rel for depth estimation. Led the **organization of the technical solution for the national competition** and optimized the presentation of the project methodology.
- Outcome:** Second Prize, 7th Global Artificial Intelligence Algorithm Elite Competition, 2025.

Dynamic Robot Navigation — Optimization of 2D LiDAR-Based Pedestrian Detection Sep.2025 – Jan.2026
Advisor: Mohan Jiang, Shanghai Innovation Institute

- Project Description:** Developed a **dynamic navigation system** integrating perception and decision-making for mobile robot navigation in dynamic environments. The perception module was optimized based on the *DR-SPAAM*, while the decision-making module adopted *TD3 reinforcement learning* with a *four-stage curriculum learning strategy*.
- Personal Contributions:** **Optimized the perception module** by replacing *Mask R-CNN* with *YOLOv8* to generate higher-quality visual pseudo-labels for training the *DR-SPAAM* detector. Improved the pedestrian detection metric *AP@0.5m* by 35.7%, from 0.42 to 0.57. Developed a systematic understanding of the perception–decision–control closed loop.

Selected Honors

Qidi Scholarship	2024-2025	School-wide, First-Class, top 1%
Guo Xie Birong Scholarship	2024-2025	School-wide, First-Class, top 1%
Social Activity Scholarship & Outstanding Student Cadre	2023-2024 / 2024-2025	School-wide

Technical Skills

- Tech Stack (selected):** Vibe Coding (CLI Agent, IDE Agent, et al.); PyTorch; Django+Vue 3 (Full-Stack); OpenCV; Git.
- Social Activities (selected):** Volunteer for the 8th China International Import Expo (CIIE); Class Monitor.