Jiajie Zhang

Mobile Autonomous Robotic Systems(MARS) Lab , ShanghaiTech University
Shanghai, China | (+86) 19858214897 | jerryzhangamen7@gmail.com | https://jiajiezhang7.github.io/ | Github

Experience

AI R&D Center, Central Research Institute, Wolong Electric — **July 2025 - Present**
Research Assistant — Shanghai, China

- Supervised by Alexander Kleiner

ShanghaiTech University — **Sept. 2024 - Jan. 2025**
Teaching Assistant — Shanghai, China
https://robotics.shanghaitech.edu.cn/teaching/moma2024

Mobile Manipulation

Education

ShanghaiTech University — **Sep 2023 - Present**
Computer Science and Technology — Master of Science
3.51/4.0 GPA

- Selected Courses: Robotics, Mobile Manipulation, Deep Learning
- Supervised by Professor Soeren Schwertfeger

Zhengzhou University — **Sep 2019 - Jun. 2023**
Automation — Bachelor of Engineering
3.66/4.0 GPA

Publications

WiFi-based Global Localization in Large-Scale Environments Leveraging Structural Priors from osmAG — **Aug. 2025**
Xu Ma*, Jiajie Zhang, Fujing Xie, Sören Schwertfeger
https://arxiv.org/abs/2508.10144

Submitted to ROBIO 2025

Automatic Generation of Open Street Map with Hierachical Topometrical Enhancement for Robot Navigation from CAD — **July 2025**
JiajieZhang*, ShenruiWu, Xu Ma and Sören Schwertfeger
https://arxiv.org/abs/2507.00552

Submitted to ICRA 2026

Intelligent LiDAR Navigation: Leveraging External Information and Semantic Maps with LLM as Copilot — **Mar. 2025**
Fujing Xie*, Jiajie Zhang and Sören Schwertfeger
https://arxiv.org/abs/2409.08493

Accepted by IROS 2025

Neural Surfel Reconstruction: Addressing Loop Closure Challenges in Large-Scale 3D Neural Scene Mapping — **Oct. 2024**
Jiadi Cui*, Jiajie Zhang*, Laurent Kneip and Sören Schwertfeger
https://www.mdpi.com/1424-8220/24/21/6919

Sensors (Basel, Switzerland)

Certifications

CET-6 (Score: 571)
Ministry of Education, China

Projects

VLM Pre-training with Latent Action from Human Videos **Aug. 2025 - Present**

- Designed and deployed an automated **YOLO-based** video pipeline to process industrial surveillance footage, converting unstructured video into segmented, trainable human action clips.
- Extracted discrete **Latent Action** representations from video by fine-tuning an LAQ model, enabling a Vision Language Model (VLM) to pre-train on general human behaviors without robot data.
- Validated a state-of-the-art **three-stage training framework** to prove the feasibility of improving robot policy generalization using third-person human video data.

PaliGemma2 for Motor Defect Detection **July 2025 - Present**

- Sourced and augmented a large-scale dataset by re-captioning images with **InstructBLIP** and synthesizing data from technical documents using **Qwen-Image-Edit**.
- Implemented a two-stage training process: conducted domain-specific continual pre-training on **PaliGemma2**, followed by Visual Question Answering (VQA) fine-tuning.
- Engineered a specialized VQA model capable of accurately identifying and classifying diverse defects on electric motor products in a factory setting.

AGLoc++: WiFi-Fused Global Localization and Monte Carlo Enhanced Tracking in Hierarchical Area Graph **Nov. 2024 - Jun. 2025**

https://jiajiezhang7.github.io/portfolio/AGLoc/

- **Led AGLoc++ development**: Ported Area Graph LiDAR localization system to ROS2, integrated with the Nav2 stack, implemented WiFi-aided kidnap recovery, developed odometry-fused Monte Carlo tracking with advanced re-localization.
- **Enhance Robustness:** implement techniques such as clutter filtering, architectural matching, weighted ICP, corridor-aware downsampling,ensuring robust, high-precision localization in dynamic indoor environments.

Campus Autonomy: Building and Navigating Autonomous Robots with Navigation2 **Sept. 2024 - Jan. 2025**

https://jiajiezhang7.github.io/portfolio/campus-autonomy-robot/

- Integrated advanced hardware, including the Agile X HUNTER SE **Ackermann** robot, **Hesai PandarQT64** Lidar, and **Insta360 Air** panoramic camera etc.
- Integrated navigation functionalities using the **ROS2** framework and **Navigation2** package, with testing performed in Gazebo simulator and on physical robots.
- Specialized in motion planning and control, applying and evaluating global planning algorithms (**Dijkstra, Hybrid A*, RRT***) for pathfinding within the Nav2 stack.

Skills

Python
Object-Oriented Programming (OOP), Data Structures, NumPy, Pandas, Scikit-learn, PyTorch, TensorFlow, Matplotlib

C++
Object-Oriented Design (OOD), Encapsulation, Inheritance, Polymorphism, Smart Pointers, Memory Management, STL.

Robot Learning
Proficient in algorithms (ACT, Diffusion Policy, RDT), VLA architectures, and model tuning (Pi-0, Gr00t).

Deep Learning
Proficient with Transformer/BERT principles, Vision Transformers (ViT, SigLiP, DINOv2), and Vision-Language Models (BLIP, PaliGemma). Experienced with Hugging Face libraries (Transformers, PEFT).

Tools
ROS/ROS2, Navigation2, PyTorch, BehaviorTree, Linux, Git, LaTex