

XUPENG ZHANG

✉ daoshi1593@gmail.com  github.com/daoshi1593

Education

Harbin Institute of Technology

Sep 2023 – Jun 2027 (expected)

Faculty of Computing

Undergraduate

Computer Science and Technology GPA: 87.984/100 Rank: 24/141

Research Interests

Embodied intelligence and robotic systems, with a focus on compositional generalization and long-horizon execution in vision-language-action (VLA) models, reliable robot runtimes, performance engineering, and systems infrastructure.

Employment History

IDEA

Sep 2026 – Present

Intern, Vision-Language-Action Models Group

Publications & Manuscripts

- **Exploring the Limits of Combinatorial Generalization in Embodied Manipulation**
Xupeng Zhang, Yantai Yang, Chang Guo, Zhaokai Yin, Zhipeng Zhang
NeurIPS 2026 · Evaluations and Datasets Track (E&D Track) · **Accepted (Poster)**
- **InfiniteVGGT: Visual Geometry Grounded Transformer for Endless Streams**
Shuai Yuan, Yantai Yang, Xiaotian Yang, **Xupeng Zhang**, Zhonghao Zhao, Lingming Zhang, Zhipeng Zhang
Submitted (under review)
- **LatentRisk-VLM: Failure-Guided Representation Editing for Frozen Vision-Language Models**
Ke Lu, **Xupeng Zhang**, Jiawei Gu, Chaomeng Chen
Submitted (under review)

Research & Engineering Experience

VLA Skill Composition and Handoff Robustness

2025 - 2026

AutoLab-SAI, SJTU

As first author, led problem formulation, experimental design, and paper writing for research on atomic skill recombination and handoff robustness in VLA policies.

- **Evaluation and diagnosis:** Built an atomic-only recombination benchmark with **19 held-out skill compositions**. Proposed HSE to decompose subtask execution and diagnose start-state gaps, handoff failures, and unstable downstream skill initiation.
- **Method and validation:** Developed PRTI using handoff-like starts, achieving **18.2% Ordered SR** versus **2.6%** for random perturbations. Across **13 real-robot tasks with 2-5 stages**, Mean Progress improved from **0.752 (HSE) to 0.802**, with gains on **8/13 tasks**.
- **Simulation infrastructure:** Built an end-to-end embodied-policy evaluation pipeline with Spine, covering environment setup, task logic, policy integration, batch execution, metric aggregation, and reproducible analysis.

Research & Engineering Experience (continued)

Doogle: Distributed CyberDog2 Runtime

Aug 2026

Core Developer

Designed, implemented, and open-sourced a heterogeneous PC-robot runtime integrating perception, planning, and motion control. GitHub: Doogle-Release

- **Coordination and recovery:** Unified sensor and control interfaces; decomposed six-stage tasks into independent processes with readiness synchronization, state management, and cross-machine tracing for independent deployment, restart, and recovery. Adopted by **3 competition teams**, which earned **two first prizes and one second prize** in the 2026 Xiaomi Cup intelligent-systems competition.
- **Image throughput:** Profiled camera acquisition and vision processing, benchmarked PC/ARM placement, and improved critical image-processing throughput from approximately **4.3 FPS to 15.1-30.0 FPS**.
- **IPC efficiency:** Replaced the C++ - ZMQ - Python image pipeline with a C++ Camera API and seqlock shared memory; real-robot stress tests reduced **peak CPU by 37.1%** and **peak RSS by 42.1%** relative to ZMQ.

Tactile Data Acquisition for Human-Robot Alignment

Sep 2026 - Present

Undergraduate Thesis

Ongoing design of unified human-hand and robot-gripper acquisition interfaces and tactile alignment using a bare-hand device with a **9-axis tactile sensor**. Completed acquisition-software interface verification documentation and a review of relevant papers and code.

- **Acquisition design:** Planning unified fields, dual timestamps, dropped-frame checks, and calibration versioning; organizing RGB, proprioception, tactile data, and expert actions by episode in HDF5 for traceability and quality checks.
- **Planned validation:** Pair human and robot samples by load, contact location, and relative motion; learn tactile encoders with representation consistency and physical-quantity prediction, integrate with ACT, and evaluate transfer benefits under equal robot-demonstration budgets.

Honors & Awards

- National Encouragement Scholarship (2024, 2025); university scholarships: one second-class and two third-class awards
- First Prize, Shensi Cup Finals, 7th Global Campus AI Algorithm Elite Competition 2025
- Third Prize, Finals, 18th National College Student Information Security Contest (Works Track) 2025

Technical Skills

Languages: Python, C++, Bash, Ascend C

ML / Robotics: PyTorch, PyG, VLA, Robot Learning

Engineering: Linux, Git, Docker, CMake, network programming, interprocess communication, performance analysis