Chengxuan Li

cxli04@link.cuhk.edu.hk • +86 15257935582

Research Interests: Embodied AI, Robotics, World Model

EDUCATION

MPhil student, The Chinese University of Hong Kong (Aug 2026 – now) **Advisor: Hongliang REN**

Bachelor of Engineering in Automation, Beijing Forestry University (Sep 2022 – Jun 2026)

• GPA: 3.78/4.0 Average score: 90.9/100

• Awards: The 2nd Prize in the 15th National College Students Mathematics Competition, the 2nd Prize in the 34th Beijing College Students Mathematics Competition, the Merit Student Scholarship

PUBLICATIONS

• **AAAI 2026 (CCF-A)**[1] DiTEA: Mixture-of-Experts for Vision-Language-Action Model in Robotic Manipulation

Li, C., Wang, X. **(First Author)** **(Oral)**

• **AAAI 2026 (CCF-A)**[2] TTF-VLA: Temporal Token Fusion via Pixel-Attention Integration for Vision-Language-Action Models Liu, C*., Zhang, J*., **Li, C***., et al. **(Co-first Author)(Poster)**

• **ICML 2026 (CCF-A)**[3] Mask World Model: Predicting What Matters for Robust Robot Policy Learning

Lou, Y., Chi, X., Zhang, X., Qian, Z., **Li, C**., et al. **(Fifth Author)(Poster)**

RESEARCH EXPERIENCE

Peking University **Advisor:** Songfang Huang **Jun 2025—Aug 2025**

VLA-driven robotic manipulation, Pi0 and other VLA models deployed on Franka Research 3

- A novel feature reuse mechanism for neighbouring pixels during VLA inference has been designed, achieving improvements when compared against baselines such as OpenVLA and VLA-Cache
- Led all real-world experiments and deployed several key VLA models on Franka Research 3

the Institute of Automation, Chinese Academy of Sciences (CASIA) **Advisor:** Haoran Li **Feb 2025—Jun 2025**

Application of the Mixture of Experts(MoE) in VLA model for Imitation Learning

- For the first time, the MoE architecture has been incorporated into the Action Head of VLA, resulting in a novel VLA model designed to enhance the multi-task learning capabilities and instruction-following performance of Diffusion-based VLA models. As the project lead, responsible for all experiments and algorithm design
- Responsible for conducting real-world experiments for part of the ConRFT (RSS 2025 accepted) project, utilizing robotic arms such as ARX and AgileX for data collection and model deployment

the Institute of Automation, Chinese Academy of Sciences (CASIA) **Advisor:** Shichao Chen **Aug 2024—Jan 2025**

A Cloud-Edge Collaboration and Multimodal Large Language Model System Design

- As a core member, responsible for designing cloud-edge collaborative multimodal large language model systems and participating in the design of cloud-edge collaborative inspection solutions for unmanned aerial vehicles.

TECHNICAL SKILLS

Software

Programming Language: Python, C++

Tools and Libraries: PyTorch, Tensorflow, Jax, ROS, Gym, Git, OpenCV, Mujoco, Pybullet

VLA Benchmarks: LIBERO, SIMPLER, RLBENCH, CALVIN

Hardware

Franka, ARX, AgileX, Unitree G1 robotic control, Robotic data collection (Teleoperation)