Jianke Zhang

zhangjianke53@gmail.com · (+86) 158-0101-3822 · Github.com/CladernyJorn

Education

Tsinghua University Beijing, China

Institute for Interdisciplinary Information Science, PhD Student Sep 2024 - Jun 2029 (Expected)

- **交叉信息研究院**
- Major: Computer Science and Technology
- Avdisor: Assistant Professor Jianyu Chen
- Research Area: Embodied AI, Computer Vision, Robot Learning, Vision-Language Model

Beijing Institute of Technology Beijing, China

Computer Science and Technology (Elite Program), Undergraduate Sep 2020 - Jun 2024

- **Overall GPA**: **94.92/100 (3.91/4.00)**
- Major: Computer Science and Technology (**Rank 1/223**)
- **Rank 1/968** in College of Information Sciences and Technology

Experience

Seed Robotics, ByteDance Group Oct 2025 - May 2026

Research Intern

- **Large Unified MLLM in VLA**: Using Bagel to train a large unified VLA for long horizon manipulation tasks, which can perform both understanding, world modeling, robotic control.
- **UAM: Exploring Mechanisms in Unified Models and World Model in VLA**: Try explore and explain how unified models and world models interact with actions in VLA architecture, in order instruct how to design next generation of VLA.

Qwen Team, Alibaba Group Mar 2025 - Oct 2025

Research Intern, supervised by Shuai Bai, Junyang Lin *from Qwen Team*

- **Contributor in Qwen3-VL**: Enhancing Embodied Understanding in VLMs. Assisted in collecting and processing datas with spatial position annotations derived from embodied tasks, improving the performance of benchmarks related to spatial reasoning and perception.
- **Evaluation Framework VLM4VLA**: Designed a framework for fair evaluation of different vision-language models applied to robot operation tasks (VLA), validating VLMs effectiveness in robotic scenarios.

ISR LAB, Tsinghua University Oct 2023 - Present

PhD student, supervised by Prof. Jianyu Chen *from Tsinghua University*

- **End2End Embodied Models (VLA)**: Exploring end-to-end approach to embodied large models based on visual language models. Small models are used to accelerate the reasoning speed of large models. Supervised reinforcement learning is also explored for fine-tuning embodied large models for fast task adaptation. Research includes fast-slow system, unified models for VLA, online reinforcement learning for VLA, and mechnism design for VLA.

KEG LAB, Tsinghua University Feb 2023 - Aug 2023

Intern, supervised by Prof. Yuxiao Dong & Prof. Jie Tang *from Tsinghua University*

- **Research on Pre-training Algorithms for Graph Neural Networks**: Developed a codebase that integrates different algorithms, datasets, and graph sampling methods to enable fair and convenient comparison of graph self-supervised algorithms. Additionally, I had been working on building a benchmark for SSL-Graph using this codebase.

LEAP LAB, Tsinghua University Sep 2022 - Mar 2023

Intern, supervised by Prof. Gao Huang *from Tsinghua University*

- **Cross-Modal Adapter for Text-Video Retrieval**: Explored the efficient transfer of large models in the multi-modal domain. To improve the early fusion of image and text encoding features, introduced a multi-modal weight sharing mechanism to the *Adapter* module.

Publications

** represents equal contribution.*

1. UAM: A Dual-Stream Perspective on Forgetting in VLA Training
   Jianke Zhang*, Luoyuan Fei*, Yucheng Hu*, Xiaoyu Chen, Yanjiang Guo, Ziyang Liu, Hongbin Xu, Lan Tian, Jianyu Chen
   Under Review [Paper] [Project Page]
2. VLM4VLA: Revisiting Vision-Language-Models in Vision-Language-Action Models
   Jianke Zhang, Xiaoyu Chen, Qiuyue Wang, Mingsheng Li, Yanjiang Guo, Yucheng Hu, Jiajun Zhang, Shuai Bai, Junyang Lin, Jianyu Chen
   ICLR 2026 (High Score) [Paper] [Project Page]
3. BagelVLA: Enhancing Long-Horizon Manipulation via Interleaved Vision-Language-Action Generation
   Yucheng Hu*, **Jianke Zhang***, Yuanfei Luo, Yanjiang Guo, Xiaoyu Chen, Xinshu Sun, Kun Feng, Qingzhou Lu, Sheng Chen, Yangang Zhang, Wei Li, Jianyu Chen
   RSS 2026 [Paper] [Project Page]
4. UniJEPA: Enhancing Robot Policy via Unified Continuous and Discrete Representation Learning
   Jianke Zhang*, Yucheng Hu*, Yanjiang Guo, Xiaoyu Chen, Yichen Liu, Chaochao Lu, Jianyu Chen
   ICML 2026 [Paper] [Project Page]
5. PlotCraft: Pushing the Limits of LLMs for Complex and Interactive Data Visualization
   Jiajun Zhang*, **Jianke Zhang***, Zeyu Cui, Jiaxi Yang, Lei Zhang, Binyuan Hui, Qiang Liu, Zilei Wang, Liang Wang, Junyang Lin
   ICML 2026 [Paper] [Project Page]
6. UP-VLA: A Unified Understanding and Prediction Model for Embodied Agent
   Jianke Zhang*, Yanjiang Guo*, Yucheng Hu*, Xiaoyu Chen, Xiang Zhu, Jianyu Chen
   ICML 2025 [Paper]
7. Video Prediction Policy: A Generalist Robot Policy with Predictive Visual Representations
   Yucheng Hu*, Yanjiang Guo*, Pengchao Wang, Xiaoyu Chen, Yen-Jen Wang, **Jianke Zhang**, Koushil

Sreenath, Chaochao Lu, Jianyu Chen
ICML 2025 Spotlight [Paper] [Project Page]

8. Improving Vision-Language-Action Model with Online Reinforcement Learning
Yanjiang Guo*, **Jianke Zhang***, Xiaoyu Chen*, Xiang Ji, Yen-Jen Wang, Yucheng Hu, Jianyu Chen
ICRA 2025 [Paper]
9. HiRT: Enhancing Robotic Control with Hierarchical Robot Transformers
Jianke Zhang*, Yanjiang Guo*, Xiaoyu Chen, Yen-Jen Wang, Yucheng Hu, Chengming Shi, Jianyu Chen
CoRL 2024 [Paper]
10. Prediction with action: Visual policy learning via joint denoising process
Yanjiang Guo*, Yucheng Hu*, **Jianke Zhang**, Yen-Jen Wang, Xiaoyu Chen, Chaochao Lu, Jianyu Chen
NeurIPS 2024 [Paper]
11. Cross-modal Adapter for Vision-Language Retrieval
Haojun Jiang, **Jianke Zhang**, Rui Huang, Chunjiang Ge, Zanlin Ni, Jiwen Lu, Shiji Song, Gao Huang
Pattern Recognition [Paper]

Honors and Awards

- **National Scholarship × 2** 2020-2021, 2021-2022
- **First Prize × 2**, 13th and 14th National College Mathematics Competitions Oct 2021, Oct 2022
- **First Prize**, 38th National College Physics Competition May 2021
- **Third Prize**, International Genetic Engineering Machine Design Competition (IGEM) Oct 2021
- **First Prize × 2**, 35th and 36th Chinese Physics Olympiad (CPHO) semi-finals, Beijing Division Oct 2018, Oct 2019

Professional Experience

- Student Staff, CollegeAI Summer Recruitment Camp, Tsinghua University May 2024
- OBS Broadcast Training Intern (BTP), Beijing Winter Olympics Nov 2021 - Mar 2022
- Peer Advisor, Class Learning Committee Member, Beijing Institute of Technology Aug 2021 - Dec 2023

Technical Skills

- Programming Languages: Python, C/C++, LaTeX, Linux, Shell (Bash/Zsh), SQL, Java, HTML/CSS
- English: CET4: 612, CET6: 562, TOEFL: 99