

张万鹏

博士研究生，北京大学计算机学院

✉ wpzhang.edu@gmail.com 🔗 <https://zhangwp.com> 🎓 Google Scholar 🐙 GitHub



教育背景

北京大学. 计算机学院, 博士研究生.

研究方向: 语言模型 / 具身智能 / 强化学习

2022 年 9 月 - 2026 年 6 月 (预期)

北京, 中国

清华大学. 计算机科学与技术系, 硕士学位.

研究方向: 强化学习

2019 年 9 月 - 2022 年 6 月

北京, 中国

南开大学. 数学科学学院, 学士学位.

研究方向: 应用数学 / 机器学习

2015 年 9 月 - 2019 年 6 月

天津, 中国

工作经历

智在无界 (BeingBeyond). 初创团队.

研究方向: 语言模型 / VLA / 具身智能

2025 年 4 月 - 至今

北京, 中国

北京智源人工智能研究院. 科研实习.

研究方向: 语言模型 / VLM / 具身智能

2024 年 5 月 - 2025 年 3 月

北京, 中国

腾讯 AI Lab. 科研实习.

研究方向: 强化学习

2020 年 6 月 - 2021 年 7 月

深圳, 中国

代表工作

Wanpeng Zhang, Yicheng Feng, Hao Luo, Yijiang Li, Zihao Yue, Sipeng Zheng, Zongqing Lu. *Unified Multimodal Understanding via Byte-Pair Visual Encoding*. (ICCV 2025, **Highlight**)

简介: 基于前作提出的自监督视觉 BPE Tokenizer 范式, 进一步设计了完整的统一多模态大模型训练框架, 并推出我们的 *Being-VL-0.5* 模型。

Wanpeng Zhang, Zilong Xie, Yicheng Feng, Yijiang Li, Xingrun Xing, Sipeng Zheng, Zongqing Lu. *From Pixels to Tokens: Byte-Pair Encoding on Quantized Visual Modalities*. (ICLR 2025)

简介: 提出了适用于图像的自监督视觉 BPE Tokenizer, 使 Transformer 能更有效地学习和对齐多模态信息, 为统一多模态大模型提供新的学习范式。

Wanpeng Zhang, Yilin Li, Boyu Yang, Zongqing Lu. *Tackling Non-Stationarity in Reinforcement Learning via Causal-Origin Representation*. (ICML 2024)

简介: 通过自适应地学习环境中的因果关系联合图, 并提供带有因果关系的表征, 使得强化学习算法能够有效应对非平稳性问题。

Wanpeng Zhang, Zongqing Lu. *AdaRefiner: Refining Decisions of Language Models with Adaptive Feedback*. (NAACL 2024)

简介: 提出了 *AdaRefiner* 框架, 通过让语言模型与强化学习智能体互相提供反馈信息, 实现两者的共同学习, 最终实现感知与决策能力的同时优化。

Ziluo Ding*, Wanpeng Zhang*, Junpeng Yue, Xiangjun Wang, Tiejun Huang, Zongqing Lu. *Entity Divider with Language Grounding in Multi-Agent Reinforcement Learning*. (ICML 2023. 共同一作.)

简介: 提出 *EnDi* 框架, 在多智能体系统中通过语言与实体的绑定, 实现了智能体的目标划分与协作增强。

Wanpeng Zhang, Ye Wang, Hao Luo, Haoqi Yuan, Yicheng Feng, Sipeng Zheng, Qin Jin, Zongqing Lu. *DiG-Flow: Discrepancy-Guided Flow Matching for Robust VLA Models*. (arXiv'25.12.)

简介: *DiG-Flow* 是适配流匹配 VLA 模型的即插即用模块, 可重新平衡基座模型与动作专家之间的控制权。

Hao Luo*, Yicheng Feng*, Wanpeng Zhang*, Sipeng Zheng*, Ye Wang, Haoqi Yuan, Jiazheng Liu, Chaoyi Xu, Qin Jin, Zongqing Lu. *Being-H0: Vision-Language-Action Pretraining from Large-Scale Human Videos*. (arXiv'25.07. 共同一作.)

简介: 我们推出 *Being-H0*, 首个通过显式手部动作建模、基于大规模人类视频预训练的 VLA 模型。

其他论文 (完整列表: [Google Scholar](#).)

- Yicheng Feng, **Wanpeng Zhang**, Ye Wang, Hao Luo, Haoqi Yuan, Sipeng Zheng, Zongqing Lu. *Spatial-Aware VLA Pretraining through Visual-Physical Alignment from Human Videos*. ([arXiv'25.12](#))
- Hao Luo, Zihao Yue, **Wanpeng Zhang**, Yicheng Feng, Sipeng Zheng, Deheng Ye, Zongqing Lu. *OpenMMego: Enhancing Egocentric Understanding for LMMs with Open Weights and Data*. ([NeurIPS 2025](#))
- Yicheng Feng, Yijiang Li, **Wanpeng Zhang**, Hao Luo, Zihao Yue, Sipeng Zheng, Zongqing Lu. *VideoOrion: Tokenizing Object Dynamics in Videos*. ([ICCV 2025](#))
- Xiaopeng Yu, **Wanpeng Zhang**, Zongqing Lu. *LLM-Based Explicit Models of Opponents for Multi-Agent Games*. ([NAACL 2025](#))
- Xingrun Xing, Boyan Gao, David A. Clifton, Zheng Liu, Shitao Xiao, **Wanpeng Zhang**, Li Du, Zheng Zhang, Guoqi Li, Jiajun Zhang. *SpikeLLM: Scaling up Spiking Neural Network to Large Language Models via Saliency-based Spiking*. ([ICLR 2025](#))
- Xingrun Xing, Zheng Liu, Shitao Xiao, Boyan Gao, Yiming Liang, **Wanpeng Zhang**, Haokun Lin, Guoqi Li, Jiajun Zhang. *EfficientLLM: Scalable Pruning-Aware Pretraining for Architecture-Agnostic Edge Language Models*. ([arXiv'25.02](#))
- Xiaopeng Yu, Jiechuan Jiang, **Wanpeng Zhang**, Haobin Jiang, Zongqing Lu. *Model-Based Opponent Modeling*. ([NeurIPS 2022](#))
- Xiaoyan Cao, Yao Yao, Lanqing Li, **Wanpeng Zhang**, Zhicheng An, Zhong Zhang, Li Xiao, Shihui Guo, Xiaoyu Cao, Meihong Wu, Dijun Luo. *iGrow: A Smart Agriculture Solution to Autonomous Greenhouse Control*. ([AAAI 2022](#))
- Mingzhe Chen, Xi Xiao, **Wanpeng Zhang**, Xiaotian Gao. *Efficient and Stable Information Directed Exploration for Continuous Reinforcement Learning*. ([ICASSP 2022](#))
- **Wanpeng Zhang**, Xiaoyan Cao, Yao Yao, Zhicheng An, Dijun Luo, Xi Xiao. *Robust Model-based Reinforcement Learning for Autonomous Greenhouse Control*. ([ACML 2021](#))
- **Wanpeng Zhang**, Xi Xiao, Yao Yao, Mingzhe Chen, Dijun Luo. *MBDP: A Model-based Approach to Achieve both Robustness and Sample Efficiency via Double Dropout Planning*. ([arXiv'21.08](#))
- Yao Yao, Li Xiao, Zhicheng An, **Wanpeng Zhang**, Dijun Luo. *Sample Efficient Reinforcement Learning via Model-Ensemble Exploration and Exploitation*. ([ICRA 2021](#))

专利

- 卢宗青, 张万鹏. 多模态数据处理方法、装置、存储介质及电子设备. (CN119226992B)
- 张万鹏, 罗迪君, 肖喜. 确定参数的方法、装置、设备及存储介质. (CN112527104A)

获奖情况

- 国家奖学金. (2025)
- 北京大学视觉技术国家工程研究中心十佳学生. (2025)
- 北京大学三好学生. (2025)
- 北京大学校长奖学金. (2024)
- 北京大学优秀科研奖. (2024)
- 腾讯 AI Lab 犀牛鸟精英计划. (2021)
- 美国大学生数学建模竞赛, 一等奖. (2017)
- 全国大学生数学建模竞赛, 二等奖. (2016)
- 全国高中数学联赛, 二等奖. (2014)

学术服务

- 会议审稿: ICML / NeurIPS / ICLR / CVPR / ICCV / AAAI / ICRA / AISTATS
- 期刊审稿: TNNLS / TIST / RAL
- 课程助教: 深度强化学习, 北京大学. 2025 年春.