WANPENG ZHANG

Ph.D. Candidate. School of Computer Science, Peking University

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

Education

Peking University. *Ph.D. Candidate in Computer Science.* *Sep. 2022 – Jun. 2026 (Expected)*
Research Interest: Foundation Models / Embodied AI *Beijing, China*

Tsinghua University. *M.S. in Computer Science.* *Sep. 2019 – Jun. 2022*
Research Interest: Reinforcement Learning *Beijing, China*

Nankai University. *B.S. in Applied Mathematics.* *Sep. 2015 – Jun. 2019*
Research Interest: Applied Mathematics / Machine Learning *Tianjin, China*

Experience

BeingBeyond. *Startup Team Member.* *Apr. 2025 – Present*
Research Interest: Foundation Models / VLA / Embodied AI *Beijing, China*

Beijing Academy of Artificial Intelligence. *Research Scientist Intern.* *May. 2024 – Mar. 2025*
Research Interest: Foundation Models / VLM / Embodied AI *Beijing, China*

Tencent AI Lab. *Research Scientist Intern.* *Jun. 2020 – Jul. 2021*
Research Interest: Reinforcement Learning *Shenzhen, China*

Representative Work

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**)

> <u>TLDR:</u> *Building upon the self-supervised visual BPE Tokenizer proposed in the previous work, we further designed a complete training framework and our* `Being-VL-0.5` *model.*

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**)

> <u>TLDR:</u> *Proposed self-supervised visual BPE Tokenizer, enabling Transformers to learn and align multi-modal information more effectively, providing a new learning paradigm for Unified MLLMs.*

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

> <u>TLDR:</u> *By adaptively learning the causal relationship joint graph in the environment and providing representations with causal relationships, RL algorithms can effectively tackle non-stationarities.*

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

> <u>TLDR:</u> *Proposed* `AdaRefiner` *to achieve the co-learning of LLMs and RL agents by enabling them to provide feedback to each other, optimizing both perception and decision-making capabilities.*

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

> <u>TLDR:</u> *Proposed* `EnDi` *framework, achieving agent goal division and collaboration enhancement in multi-agent systems through language and entity binding.*

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**.)

> <u>TLDR:</u> `DiG-Flow` *is a plug-and-play module for flow-matching based VLAs that rebalances control between the autoregressive foundation model and the flow expert.*

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**. ***Equal Contribution**.)

> <u>TLDR:</u> *We introduce* `Being-H0`, *the first dexterous Vision-Language-Action model pretrained from large-scale human videos via explicit hand motion modeling.*

Other Publication (*Full List: [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**)

Patent

- Zongqing Lu, **Wanpeng Zhang**. *Multimodal data processing method, device, storage medium, and electronic equipment.* (CN119226992B)

- **Wanpeng Zhang**, Dijun Luo, Xi Xiao. *Method, device and equipment for determining parameters and storage medium.* (CN112527104A)

Award

- **National Scholarship**. (2025)

- **Top 10 Students at the National Engineering Research Center of Visual Technology**. (2025)

- Merit Student of Peking University. (2025)

- **Presidential Scholarship of Peking University**. (2024)

- Award for Scientific Research of Peking University. (2024)

- Rhino-bird Elite Training Program of Tencent AI Lab. (2021)

- Mathematical Contest in Modeling (MCM/ICM), Meritorious Winner (First Prize). (2017)

- China Undergraduate Mathematical Contest in Modeling (CUMCM), Second Prize. (2016)

- National High School Mathematics Competition, Second Prize. (2014)

Service

- **Conference Reviewer**: ICML / NeurIPS / ICLR / CVPR / ICCV / AAAI / ICRA / AISTATS

- **Journal Reviewer**: TNNLS / TIST / RAL

- **Teaching Assistant**: Deep Reinforcement Learning, Peking University. *Spring, 2025.*