Wanpeng Zhang

Partner & Head of Post-Training @ BeingBeyond

wpzhang.edu@gmail.com https://wanpeng.me Google Scholar GitHub

Education

Peking University. *Ph.D. in Computer Science.* *Sep. 2022 – Jun. 2026*
Research Interest: Foundation Models / RL / Robotics *Beijing, China*

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

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

Experience

BeingBeyond. *Partner & Head of Post-Training.* *Apr. 2025 – Present*
Research Interest: VLA / World Model / RL / Robotics *Beijing, China*

Beijing Academy of Artificial Intelligence. *Research Intern.* *May. 2024 – Mar. 2025*
Research Interest: VLM / RL / Robotics *Beijing, China*

Tencent AI Lab (Rhino-bird Program). *Research Intern.* *Jun. 2020 – Jul. 2021*
Research Interest: Reinforcement Learning *Shenzhen, China*

Representative Work

Wanpeng Zhang*, BeingBeyond Team. *Being-H0.8: A Latent Tactile World-Action Model at Scale.* (**Tech Report 26.07**. **Lead of Post-Training, Core Contributor.*)

TLDR: `Being-H0.8` *extends latent world-action modeling to tactile-aware interaction, scaling visuo-tactile pretraining with 500,000+ hours of egocentric human video.*

Wanpeng Zhang*, BeingBeyond Team. *Being-H0.7: A Latent World-Action Model from Egocentric Videos.* (**NeurIPS 2026**. **Lead of Post-Training, Core Contributor.*)

TLDR: `Being-H0.7` *is a latent world-action model that scales egocentric video pretraining into future-aware robot control.*

Wanpeng Zhang, Hao Luo, Sipeng Zheng, Yicheng Feng, Haiweng Xu, Ziheng Xi, Chaoyi Xu, Haoqi Yuan, Zongqing Lu. *Conservative Offline Robot Policy Learning via Posterior-Transition Reweighting.* (**arXiv'26.03**.)

TLDR: PTR performs reward-free offline policy improvement by conservatively reweighting offline data according to whether an action leads to an identifiable downstream outcome.

Wanpeng Zhang*, BeingBeyond Team. *Being-H0.5: Scaling Human-Centric Robot Learning for Cross-Embodiment Generalization.* (**CoRL 2026**, **Oral**. **Lead of Post-Training, Core Contributor.*)

TLDR: `Being-H0.5` *is a foundational VLA model that scales human-centric learning with a unified action space to enable robust cross-embodiment robot control.*

Wanpeng Zhang, Ye Wang, Hao Luo, Haoqi Yuan, Yicheng Feng, Chaoyi Xu, Sipeng Zheng, Qin Jin, Zongqing Lu. *Transport Discrepancy as a Reliability Signal for Vision-Language-Action Models.* (**ECCV 2026**.)

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

Wanpeng Zhang*, BeingBeyond Team. *Being-H0: Vision-Language-Action Pretraining from Large-Scale Human Videos.* (**ICML 2026**. **Lead of Post-Training, Core Contributor.*)

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

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**)

TLDR: 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**)

TLDR: Proposed self-supervised visual BPE Tokenizer, enabling Transformers to learn and align multimodal 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**)

TLDR: 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**)

TLDR: 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**. **Co-first Author.*)

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

Other Publication (*Full List: Google Scholar.*)

- Haiweng Xu, Sipeng Zheng, Hao Luo, **Wanpeng Zhang**, Ziheng Xi, Zongqing Lu. *Unmasking the Illusion of Embodied Reasoning in Vision-Language-Action Models.* (**ECCV 2026**)
- 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.* (**CVPR 2026**)
- Hao Luo, Ye Wang, **Wanpeng Zhang**, Haoqi Yuan, Yicheng Feng, Haiweng Xu, Sipeng Zheng, Zongqing Lu. *Joint-Aligned Latent Action: Towards Scalable VLA Pretraining in the Wild.* (**CVPR 2026**)
- Chi Zhang, Penglin Cai, Ziheng Xi, Haoqi Yuan, Hao Luo, **Wanpeng Zhang**, Sipeng Zheng, Chaoyi Xu, Zongqing Lu. *Human-Centric Transferable Tactile Pre-Training for Dexterous Robotic Manipulation.* (**arXiv'26.07**)
- Chaoyi Xu, Yixuan Jiang, Jiahui Huan, Yuhui Fu, Haoyu Zhou, Weitian Yuan, Jiayi Yu, **Wanpeng Zhang**, Haoqi Yuan, Zongqing Lu. *RealDexUMI: A Wearable Universal Manipulation Interface for Dexterous Robot Learning.* (**CoRL 2026**)
- Ye Wang*, Sipeng Zheng*, Hao Luo*, **Wanpeng Zhang***, Haoqi Yuan, Chaoyi Xu, Haiweng Xu, Yicheng Feng, Mingyang Yu, Zhiyu Kang, Zongqing Lu, Qin Jin *Rethinking Visual-Language-Action Model Scaling: Alignment, Mixture, and Regularization* (**arXiv'26.02**. **Co-first Author.*)
- 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**)
- Zhicheng An, Xiaoyan Cao, Yao Yao, **Wanpeng Zhang**, Lanqing Li, Yue Wang, Shihui Guo, Dijun Luo. *A Simulator-based Planning Framework for Optimizing Autonomous Greenhouse Control Strategy.* (**ICAPS 2021**)
- Bowen Zhao, Xi Xiao, **Wanpeng Zhang**, Bin Zhang, Guojun Gan, Shutao Xia. *Self-Paced Probabilistic Principal Component Analysis for Data with Outliers.* (**ICASSP 2020**)

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 Talent Program of Tencent. (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 / BMVC
- **Journal Reviewer**: TNNLS / TIST / RAL / TMLR
- **Teaching Assistant**: Deep Reinforcement Learning, Peking University. *Spring, 2025.*