



# 周伯涵

- 政治面貌：正式党员
- 联系电话：15797895657
- 联系邮箱：zhoubh@stu.pku.edu.cn
- 通讯地址：北京市海淀区颐和园路 5 号北京大学

## 教育背景

2019.9-2023.6	南开大学-人工智能学院	智能科学与技术
◆ GPA: 3.95/4, 学绩分: 93.644/100, 专业排名: 2/98。		
◆ 荣誉奖项: 优秀毕业生, 获得天津市人民政府奖学金(市级)、公能奖学金(校级)以及企业奖学金等七项。		
2023.9-2026.6	北京大学-视频与视觉技术国家工程研究中心-硕士	计算机应用技术
◆ GPA: 3.88/4。		
◆ 荣誉奖项: 国家奖学金, 北京大学三好学生。		

## 论文发表 (蓝色为第一作者)

- Bohan Zhou, et al. Learning from Visual Observation via Offline Pretrained State-to-Go Transformer. [NeurIPS 2023](#).
- Bohan Zhou, et al. MEgoHand: Multimodal Egocentric Hand-Object Interaction Motion Generation. [NeurIPS 2025](#).
- Bohan Zhou, Jiangxing Wang, and Zongqing Lu. NOLO: Navigate Only Look Once. [IROS 2025 Oral](#).
- Haoqi Yuan, Bohan Zhou, Yuhui Fu, Zongqing Lu. Cross-Embodiment Dexterous Grasping with Reinforcement Learning. [ICLR 2025](#).
- Hao Luo, Bohan Zhou, Zongqing Lu. Pre-trained Visual Dynamics Representations for Efficient Policy Learning. [ECCV 2024](#).
- Sipeng Zheng, Bohan Zhou, et al. UniCode: Learning a Unified Codebook for Multimodal Large Language Models. [ECCV 2024](#).
- ..., Bohan Zhou (7<sup>th</sup>/28), et al. Cradle: Empowering Foundation Agents Towards General Computer Control. [ICML 2025](#).
- ..., Bohan Zhou (5<sup>th</sup>), et al. GFIE: A Dataset and Baseline for Gaze-Following from 2D to 3D in Indoor Environments. [CVPR 2023](#).
- ..., Bohan Zhou (3<sup>th</sup>), et al. Gaze Target Estimation Inspired by Interactive Attention. [TCSVT 2022](#).
- Bohan Zhou, et al. Learning diverse bimanual dexterous manipulation skills from human demonstrations. ([under review](#))
- ..., Bohan Zhou (4<sup>th</sup>), et al. Being-0: A Humanoid Robotic Agent with Vision-Language Models and Modular Skills. ([under review](#))

## 科研实习

2022.12-2025.5	北京智源人工智能研究院(BAAI)	实习
----------------	-------------------	----

研究核心：从离线视觉观测序列中学习智能体策略

- STG [NeurIPS 2023]: 提出一种通用视频预训练获取先验知识-在线强化学习的策略学习框架, 在 Atari 和 Minecraft 中表现突出。
- PVDR [ECCV 2024]: 提出一种从广泛互联网无标注视频数据预训练视觉转移表示克服领域差距从而适应下游任务强化学习的方法。

研究核心：机器人-导航&五指灵巧手操作

- NOLO [IROS 2025]: 提出首个无需全局地图、深度相机、位姿传感器仅根据 RGB 视频离线训练 3D 视觉室内导航策略方法。
- BiDexHD [Arxiv 2024]: 提出首个通用、可扩展的从人类演示数据集中构建丰富双手灵巧操作任务并学习多样技能的三阶段框架。

研究核心：多模态大模型

- MEgoHand [NeurIPS 2025]: 提出首个以任务指令、第一人称图像以及初始手部位姿为输入的多模态手物交互位姿序列生成器。
- UniCode [ECCV 2024]: 提出一种新的高效学习统一 CodeBook 的多模态模型训练范式, 解放纯文本码本对图像生成能力的限制。
- Being-0 [Arxiv 2025]: 提出三层机器人智能体系统, 上层 GPT4o 任务拆解+中层自研 VLM 导航/工具链规划+底层技能库实现十余种双手灵巧操作, 能够部署至宇树 H1 机器人上完成路线长度数十米的从导航到接咖啡递给顾客的长程任务。
- Cradle v2.0 [ICML 2025] / Cradle v1.0 [ICLR+NeurIPS Workshop 2024]: 首次提出 Cradle 框架赋能 foundation models 以统一的屏幕截图为输入键鼠动作为输出执行复杂的计算机控制任务, 该框架被应用于 3A 大作“荒野大镖客”和“黑神话悟空”, 成为 Claude 3.5 在 OSWorld 基准上认可并比较的最强大的 baseline。

2021.3-2022.9	天津市大学生创新训练计划	队长
---------------	--------------	----

项目主题：室内场景下人的意图分析及服务机器人的应用 (市创三等奖)

- 项目①构筑了一套不改变实验者外貌、不引入主观性偏差 GFIE 数据集[CVPR 2023]; ②研究交互式注意力启发下的注视目标估计, 提出一套同时估计 2D&3D 场景下注视目标热区的方法[TCSVT 2022]; ③提出一种复杂室内环境下激光点检测算法[CCDC 2022]。

## 爱好特长

- 熟悉 Python/C++/MATLAB 语言、Pytorch 深度学习&DeepSpeed 并行训练框架、各种大语言模型、生成模型、强化学习算法。
- 爱好: 国际象棋 (国家二级运动员、棋协大师)、钢琴、吉他、乒乓球。