

教育背景

- 北京大学，物理学院

物理学学士（在读） 卓越计划（物理竞赛金牌保送） 预期毕业：2027 年 6 月
- 2023 年 9 月 – 2027 年 6 月

• 学业: GPA 3.74/4.00，第一学年物理学院排名前 10%，获国家奖学金（北大前 1%）。大二即几乎完成全部培养方案（141/149 学分），并修读 3 门研究生课程。

研究与实习经历

- 研究实习生，Seed VLM 后训练团队

VLM 多阶段训练和交付
- 字节跳动

2025 年 7 月 – 至今
- HiPhO 交付: 持续负责 RL、SFT、Mid-training 等多阶段的交付工作，其中 HiPhO 从 Seed 2.0 Lite 0428 版本分数达到 83.8 (2.0 Lite 72.5, Pro 74.1)

◦ 参与多阶段训练和交付交付: 参与 Mid-training 早期探索，负责数据清洗。近 90 天日均占用 1156 H100 卡天。

◦ 数据质量控制: 使用对齐技术，研究出了一种使用低价模型对齐外场模型，或者较少的标注拟合大量 rollout 结果的框架，节省数倍算力。

◦ 尝试数据合成、新的训练范式: 搭建了从书籍中大规模提取习题的 pipeline（滑动窗口裁剪 + prompt 提取），生产亿级 QA pair。做了 CoT compression (summary + 滑动窗口摘取原文关键步骤)，然后蒸馏回模型。目前也在探索 TransferRL（将 SFT 和 RL 结合）。

◦ Agent 与工作提效: 搭建了个人的 autoresearch agent，通过 loop-adversary pair 作为单位 agent，总控 agent 监督和介入，人类可以与总控对话，实现可控的 research 优化。可以连续执行数天，研究上百个 setting。

◦ 正在研究 Agent 相关:
- 联合发起人 & 第一作者之一

PHYBench: 面向大语言模型物理推理能力的综合性评测基准
- Eureka Lab（合作者：仇是等）

2025 年 2 月 – 2025 年 9 月
- 项目发起: 发现现有 LLM 评测在物理推理领域的系统性空白，与合作者发起 PHYBench 项目，从 POC 验证到数据流设计全程主导。

◦ 数据集构建: 组织 178 名北大学生，两周内完成 500 道高质量原创物理题的数据集构建；设计评测标准、质量控制流程与失败模式分析框架。

◦ 实验与分析: 完成论文中主要实验与结果分析，系统评估多个前沿 LLM 在物理子领域的推理能力边界。

◦ 论文: Shi Qiu, Shaoyang Guo, et al. “PHYBench: Holistic Evaluation of Physical Perception and Reasoning in Large Language Models.” arXiv:2504.16074 (已投稿 NeurIPS 2025)
- 研究助理

视觉-语言-动作（VLA）模型综述
- 北京大学 PsiRobot 实验室（导师：杨耀东助理教授）

2025 年 3 月 – 2025 年 8 月
- 综述撰写: 参与撰写 VLA 模型综述，聚焦动作表征（Action Tokenization）视角。负责 Raw Action 章节，梳理了 end-to-end VLA 架构，综述 30 余篇关键文献。

◦ 论文: “A Survey on Vision-Language-Action Models: An Action Tokenization Perspective.” arXiv:2507.01925

荣誉奖项

- 国家奖学金: 教育部，北京大学前 1% (2024)
- 全国中学生物理竞赛（CPhO）金牌: 全国第 57 名 (2022)，经卓越计划保送北大物理学院
- 全国信息学奥林匹克联赛（NOIP）一等奖: (2020)

技能

- 编程: Shell, Python, PyTorch, Spark, C/C++, MATLAB
- ML/AI: RL for LLM (RLHF, BoN), SFT, mid-training, 评测方法论, 搭建 multi-agent 系统
- 物理: 量子力学, 量子场论, 统计物理, 实验物理中的统计方法, 计算物理（数值方法）
- 语言: 中文（母语），英文（四六级 657 分，GRE 150+170）

研究兴趣

LLM/VLM 推理与后训练（RL/SFT/数据合成）、Agent、Benchmarking、具身智能、Physics of AI(大模型的定量解释)