

杨一君

电话：15910626500 | 邮箱：yangyijun94@163.com | [个人主页](#) [谷歌学术主页](#)

🎓 教育背景

- **博士，人工智能、强化学习 (2020.01 - 2024.07)**
 - 悉尼科技大学 (US News Top 100)，澳大利亚人工智能研究院；
 - 导师：张成奇教授 (IJCAI 2024大会主席，澳洲人工智能理事会会长，IEEE Fellow)、史玉回教授 (IEEE Fellow, 群体智能算法先驱)、周天翼博士 (马里兰大学帕克分校助理教授)；
 - **研究方向**：自适应决策智能体构建，聚焦如何利用先验知识 (离线数据、历史数据、互联网数据) 提升智能体 (Agent) 的自适应决策能力。
 - **核心工作 (强化学习算法研究)**：
 - **研究基于模型的离线强化学习算法**，创新性地利用Pareto优化/多目标SGD解决回报-不确定性权衡问题，提升离线RL算法的效率；
 - **研究免模型的持续强化学习算法**，首次从神经网络优化动力学角度分析可塑-稳定性权衡，设计基于任务相似性的权重划分与优化算法，融合预训练LLM、在线字典学习、最大熵策略优化，实现当时 (截止2024年6月) 最强的持续强化学习算法基准。
- **科研助理 (2017.09 - 2019.10)**
 - 南方科技大学 (双一流大学)，计算机科学与工程系
 - 研究方向：演化策略、梯度估计、随机优化、黑箱优化 (部分成果发表于PPSN, JMLR)。
- **本科，自动化 (2013.09 - 2017.06)**
 - 北京航空航天大学 (985, 双一流大学)

👛 工作经历

- **腾讯 - AI Lab - 绝悟 AI | RL Agent 算法研究员 (2024.08 - 至今)**
 - Leader: [叶德珩博士](#)
 - **项目：FPS游戏中LLM/VLM Agent应用**
 - 负责应用大型语言模型 (LLM) 和视觉语言模型 (VLM) 技术，为FPS游戏研发具备高水平协作和决策能力的全自动AI队友 (人机交互智能体方向)；
 - 探索LLM/VLM在复杂游戏环境中的理解、决策、规划与协作机制。
- **腾讯 - 数据平台部 - 混元大模型 | AIGC 算法研究员 (2024.02 - 2024.08)**
 - Leader: [刘威博士](#)
 - **项目：多模态后训练、Discrete Visual Tokenizers与视频生成模型研发**
 - 参与Hunyuan-Vision模型的后训练；
 - 负责多模态模型视觉推理 (Visual Reasoning) 能力的探索与测评工作；
 - 参与VQ-GAN模型训练流程设计与实现，研究统一的多模态自回归建模技术。
 - 独立负责并训练人脸特征解耦的VAE模型，成功复现并优化MSRA VASA-1方案，实现高质量视频/音频驱动的Talking Head生成。

- 京东 - 探索研究院 | 算法研究实习生 (2023.06 - 2023.12)
 - Mentor: [沈力博士](#)
 - 项目: **大语言模型赋能的具身决策系统**
 - 研究如何利用互联网尺度数据作为先验, 赋予基于**视觉语言模型 (InstructBLIP)** 的具身智能体媲美**大语言模型**的推理与规划能力;
 - 独立负责核心算法设计, 采用**在线模仿学习**技术对具身智能体与LLM的推理规划能力。

代表性论文成果 (完整列表请见谷歌学术主页)

1. **Yang, Y.**, Jiang, J., Zhou, T., Ma, J., & Shi, Y. (2022). Pareto Policy Pool for Model-based Offline Reinforcement Learning. **ICLR**.
2. **Yang, Y.**, Zhou, T., Li, K., Tao, D., Li, L., Shen, L., He, X., Jiang, J., & Shi, Y. (2024). Embodied Multi-Modal Agent trained by an LLM from a Parallel TextWorld. **CVPR**.
3. **Yang, Y.**, Zhou, T., Jiang, J., Long, G., & Shi, Y. (2023). Continual Task Allocation in Meta-Policy Network via Sparse Prompting. **ICML**.
4. Wen, X., **Yang, Y.**, Gao, C., Xiao, Y., Ye, D. (2025) Boosting Vulnerability Detection of LLMs via Curriculum Preference Optimization with Synthetic Reasoning Data. **ACL Long Paper**.
5. Li, K., Yu, B., Zheng, Q., Zhan, Y., Zhang, Y., Zhang, T., **Yang, Y.** ... & He, X. (2024). MuEP: A Multimodal Benchmark for Embodied Planning with Foundation Models. **IJCAI**.
6. Duan, Q., Zhou, G., Shao, C., Wang, Z., Feng, M., **Yang, Y.**, Zhao, Q., & Shi, Y. (2024). PyPop7: A Pure-Python Library for Population-Based Black-Box Optimization. **JMLR**.

获奖情况

- 春晖杯中国留学人员创新创业大赛优胜奖 (教育部, 2022)
- 澳大利亚人工智能研究院最佳学生论文奖 (算法与理论创新方向) (2022)
- Parallel Problem Solving from Nature (PPSN) 大会最佳论文奖提名 (2021)
- **悉尼科技大学院长奖 (Top 3/120) (2025)**
- 多次获得北京航空航天大学、悉尼科技大学校级奖学金 (2013-2025)