

杨雪梦

Xuemeng Yang (Google Scholar)

上海市徐汇区 (地址)

13152147825 (电话)

xuemengyang96@gmail.com (邮件)

<https://jokester-zzz.github.io/> (个人主页)

个人经历

- 上海人工智能实验室 前沿探索中心 2022.05 - 至今
- 美团点评 无人车配送中心/感知组 2021.06 - 2021.09 实习
- 浙江大学 控制科学与工程 2019.09 - 2022.03 硕士学位
- 西北工业大学 自动化 2014.09 - 2018.06 学士学位

工作经历

通用智能体：记忆、搜索与自进化 上海人工智能实验室/前沿探索中心 2025.02-至今

- 构建闭环经验驱动的自进化智能体架构: 深入研究经验驱动的自进化 AI Agent 架构，将经验积累、自反思与能力更新解耦，实现长程任务场景下智能体在真实工作流程中“边干边学、持续自我改进”的生命周期闭环。成果: [Muse](ACL 2026 Findings)。
- 智能体持续学习能力评测基准: 构建 Trainee Bench 动态评测环境，模拟 Agent 面对未知业务规则、流式任务与不完整信息时的真实工作流程，系统评估其上下文感知调度、主动探索补全信息、从经验中持续学习的能力。成果: [Trainee Bench](ACL 2026 Findings)。
- 目标导向规划与鲁棒 Agent 搜索模型: 构建基于目标导向规划与自反思的鲁棒 Agentic Search 架构；攻克开放域开放式问答中检索、多源证据组织与可归因决策生成的协同难题。成果: [O²-Searcher](TMLR 2026), [RE-Searcher](arXiv 2025)。
- AI Agent Skill 版本管理与合并系统: 开发 skill-git 插件，将 Git 式版本控制、语义版本、安全回滚、冲突检测与相似 Skill 合并机制引入 Agent Skills 管理流程，支持 Claude、Codex 等多平台，提升智能体能力模块的可追踪、可复用和可维护性。成果: [skill-git]开源发布。

垂域应用：工具调用与工业智能体 上海人工智能实验室/前沿探索中心 2024.08-至今

- 代码生成驱动的 CAD 专业智能体: 构建以代码生成为核心的 CAD 建模智能体，通过大模型生成并迭代修改几何建模代码，结合模型渲染结果与交互式视觉反馈形成“代码生成—执行验证—视觉反馈—代码修正”的闭环迭代机制；进一步基于 Windows COM 接口实现对原生 CAD 软件的端到端自动化控制。成果: [IterCAD](arXiv 2026), [ComAct](arXiv 2026)。
- 书生大模型多模态工具调用能力建设与产品化: 深度参与书生大模型工具调用能力的方案设计与跨团队交付，成功上线高光视频剪辑、三维模型重建、图像修复等核心产品功能；并将工具链调用策略系统化总结。成果: 产品上线, [VisionLLM v3] 投稿 TPAMI 在审。
- 科研智能体与自动化内容生成引擎: 为 Migo 平台构建科研智能体核心能力，开发 QuestPaper 知识引导式文献阅读助手与 DeepResearch 自动化行业/科研报告生成引擎。通过自然交互的方式贡献高价值科研过程数据，形成数据-算法-用户体验的正向增长飞轮。成果: QuestPaper、DeepResearch 功能上线。

自动驾驶：全栈感知、认知智能体与生成式世界模型 浙江大学/上海人工智能实验室 2019-2024.08

- **闭环生成式仿真平台建设与世界模型研究**: 主导构建首个闭环生成式自动驾驶仿真平台 DriveArena，训练自回归多视角视频世界模型，打通“生成式世界模型-交通流仿真-驾驶智能体”的闭环链路，率先实现真实道路场景下的 agent-in-the-loop 闭环评测与时空规划闭环。**成果**: [\[DriveArena\]](#)(ICCV 2025), [\[DreamForge\]](#)(arXiv 2024)。
- **大模型认知驱动的驾驶智能体决策**: 系统评估多模态大模型 GPT-4V 在驾驶决策中的能力边界；进一步将显式交通知识、人类认知双过程和经验驱动学习引入驾驶控制模型，构建“慢思考分析-快决策执行-失败反思-长期记忆更新”的 LeapAD 闭环机制，在 CARLA 闭环环境中验证经验驱动的策略自我进化。**成果**: [\[GPT-4V AD\]](#)(ICLR 2024 Workshop), [\[LeapAD\]](#)(NeurIPS 2024), [\[LeapVAD\]](#)(IEEE TNNLS 2025), [\[Knowledge-driven AD\]](#)(arXiv 2023)。
- **多模态场景感知与三维环境理解**: 围绕复杂驾驶环境下的高精度场景感知，研发长时序点云离线检测框架 DetZero，通过轨迹级精修支撑大规模高质量数据自动生产；开展零样本全景感知、LiDAR 场景补全与跨模态占用预测研究，将语义先验与跨模态知识引入三维几何理解，提升模型在开放场景及极端天气下的感知能力与鲁棒性；并参与 BEV 融合、表示解耦与语义图闭环检测等研究。**成果**: [\[DetZero\]](#)(ICCV 2023), [\[ZOPP\]](#)(NeurIPS 2024), [\[LiCROcc\]](#)(IEEE RA-L 2024), [\[SSA-SC\]](#)(IROS 2021)。

代表论文与成果

- Hu T, Ai J, Wen L, et al. IterCAD: An iterative multimodal agent for visually-grounded CAD generation and editing[J]. arXiv preprint arXiv:2606.13368, 2026. **Yang X.** (通讯作者). [\[Link\]](#)
- Ai J, Hu T, **Yang X**, et al. ComAct: Reframing professional software manipulation via COM-as-Action paradigm[J]. arXiv preprint arXiv:2606.13239, 2026. [\[Link\]](#)
- Fu D*, Mei J*, Wu R*, **Yang X***, et al. The Agent's First Day: Benchmarking learning, exploration, and scheduling in the workplace scenarios[J]. arXiv preprint arXiv:2601.08173, 2026. [\[Link\]](#)
- Yang C*, **Yang X***, Wen L*, et al. Learning on the Job: An experience-driven self-evolving agent for long-horizon tasks[J]. arXiv preprint arXiv:2510.08002, 2025. [\[Link\]](#)
- Fu D*, Mei J*, Wen L, **Yang X**. RE-Searcher: Robust agentic search with goal-oriented planning and self-reflection[J]. arXiv preprint arXiv:2509.26048, 2025. [\[Link\]](#)
- Mei J, Hu T, Fu D, Wen L, **Yang X**, et al. O²-Searcher: A searching-based agent model for open-domain open-ended question answering[J]. arXiv preprint arXiv:2505.16582, 2025. [\[Link\]](#)
- Mei J, Hu T, **Yang X**, et al. DreamForge: Motion-aware autoregressive video generation for multi-view driving scenes[J]. arXiv preprint arXiv:2409.04003, 2024. [\[Link\]](#)
- **Yang X***, Wen L*, Wei T*, et al. DriveArena: A closed-loop generative simulation platform for autonomous driving[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2025: 26933-26943. [\[Link\]](#)
- Mei J, Ma Y, **Yang X**, et al. Continuously learning, adapting, and improving: A dual-process approach to autonomous driving[J]. arXiv preprint arXiv:2405.15324, 2024. [\[Link\]](#)

- Wen L*, **Yang X***, Fu D*, et al. On the road with GPT-4V(ision): Early explorations of visual-language model on autonomous driving[J]. arXiv preprint arXiv:2311.05332, 2023. [\[Link\]](#)
- Ma T, **Yang X**, Zhou H, et al. DetZero: Rethinking offboard 3d object detection with long-term sequential point clouds[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 6736-6747. [\[Link\]](#)
- **Yang X**, Zou H, Kong X, et al. Semantic segmentation-assisted scene completion for LiDAR point clouds[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 3555-3562. [\[Link\]](#)
- Kong X, **Yang X**, Zhai G, et al. Semantic graph based place recognition for 3d point clouds[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2020: 8216-8223. [\[Link\]](#)

技能专长

- 语言: Python, L^AT_EX
- 大模型与智能体: LLM Agent 架构设计, 工具调用, Benchmark 构建, 持续学习与自进化机制
- 工程能力: 熟练 Vibe Coding, Prompt/Workflow 设计, PyTorch, Git
- 全国大学英语: 六级 (590), 四级 (562)

荣誉奖项

• 西北工业大学校级奖学金 一等奖学金	2016/2017
• RoboCup 世界杯中国赛 RoboCup 家庭组技术挑战赛 冠军	2017
• 全国大学生物联网设计竞赛 (TI 杯) 西北赛区 特等奖	2017
• 全志杯微创客高校挑战赛全国总决赛 亚军	2017
• 恩德斯豪斯企业奖学金	2017
• 全国青少年信息学奥林匹克联赛 (NOIP) 吉林省一等奖	2012/2013