RESEARCH & ENGINEERING

祝仰坤 Yangkun Zhu

具身智能、多模态大模型

上海
zyk_seu_job@163.com
https://tenstep.top/

个人概况

7 年 AI 算法研发与落地经验，5 年带队经验。聚焦 VLA、力控后训练与推理部署，兼具前沿模型研究、真机系统集成和车载量产经验；参与 InternW0、InternVLA 系列工作，论文成果包括 ICLR 与 CoRL。

工作经历

上海人工智能实验室 2025.03 - 至今

具身智能 / 物理智能中心
参与具身智能预训练、后训练及部署基础设施研发；获年度员工荣誉。

地平线 / 红西瓜半导体（GUA） 2021.02 - 2025.03

HALO 座舱部门 · 行为手势组 TL
兼任项目负责人、技术架构与算法职责，主导 VLM 预研、3D 人手重建与空间交互，服务十数款量产车型。

电科云（中国电科集团） 2020.02 - 2021.02

算法工程师
从零搭建边缘智能平台算法框架，参与总体方案设计与重点项目推进。

地平线 2019.07 - 2020.02

研究院助理研究员 / DMS 算法工程师
负责疲劳检测算法，支撑地平线首款量产车载芯片 J2 发版；担任研究院 A 班班长。

教育背景

东南大学 · 工程热物理 · 学术硕士 2016.09 - 2019.07

能源与环境学院 · 能源信息自动化系
研究方向：吸收光谱、光场及温度场 / 浓度场重建。

南京林业大学 · 热能与动力工程 · 学士 2012.09 - 2016.07

材料科学与工程学院
专业 Top 2；国家励志奖学金、优秀毕业生。

技术能力

模型与算法	VLA / VLM、空间 Grounding、力控后训练、多模态时序建模、长程任务拆解
系统与部署	PyTorch、vLLM-Omni、异步推理、模型量化、机器人接口与数据采集
研发与协作	Python、C++、Git；项目带队、跨团队推进、真机验证与车规芯片部署

近期项目

InternW0 / 浦江创新论坛湿实验力控原型 2026.08 - 2026.09

W0 力控算法 Owner · 团队成员 1 人

- 牵头力控后训练，约 1.5 个月内协同完成双臂双手平台、采集链路与湿实验移液 Demo；完成 π0.5、InternVLA-A1.5 力控模块接入。
- 开展力输入量化抗 OOD、标定噪声优化与位控表征消融；支撑 InternW0 定量移液任务进度率 65.3%，较 π0.5 提升 18.6 个百分点。

祝仰坤

代表项目与论文成果

https://tenstep.top/

代表项目

WAIC 2026 电催化具身实验原型验证

2026.04 - 2026.07

旋蒸场景负责人 · 团队成员 1 人

- 采用力控算法与结构件辅助提升现场工况适应能力，旋蒸机场景执行成功率达到 ≥ 99%。
- 开发一键回零、一键启动与远程调度，形成冗余保障与应急恢复能力，完成面向科技部的现场展示。

VLA 推理部署基础设施（Rdeploy / vLLM-Omni）

2025.11 - 2026.03

Infra 部署负责人 · 团队成员 3 人

- 自主搭建插件化 Rdeploy，统一内部 6 台机器人接口；完成 vLLM-Omni 框架首个 VLA 模型 InternVLA-A1 开源部署，推理速度较原生 PyTorch 提升 100%。
- 落地实验室首套异步推理方案，支撑衣物展叠和高速传送带动态夹取；在 acone、lift2 上完成拖动 / VR 示教等 DAgger 方案验证。

InternVLA-M1 空间引导 VLA 框架

2025.03 - 2025.09

Sys2（操作大脑）Owner / 下游任务算法工程师

- 负责 M1-VLM 预训练，增强 freespace、point、bbox、traj 等空间表征，在多个 Grounding 榜单取得 SOTA；构建 190 万条多模态数据并验证 Scaling 效果。
- 建设覆盖 8 个数据集、15 个维度的评测框架；长程操作任务成功率达到 90%，较 GPT baseline 提升 30%+。

车载多模态模型与空间交互

2023.07 - 2025.03

行为手势组 TL / 交警指挥算法模块负责人

- 主导交警场景 VLM 方案与数据闭环，将关键任务 Precision 从 GPT-4o 基线约 30% 提升至 95%+，完成公司内部首次 VLM 自动驾驶场景量产级上线。
- 负责 Vision Pro 手眼交互等展示项目，牵头 15 名核心开发、协同 43 人跨团队推进，完成车载 15.6 寸屏首发落地与实车验证。

论文成果

以下均为共同作者，完整作者列表与资源见个人主页 Works。

1. FlowVLA: Learning Flow-Aware Physical Completion for Quantitative Robotic Pouring
   arXiv · 2026
2. VOMMI: Collecting and Leveraging Portable Demonstrations for Mobile Manipulation
   arXiv · 2026
3. InternW0: A Foundational Physical World Model for Efficient Real-World Interactions
   Technical Report · 2026
4. RoboSnap: One-Shot Real-to-Sim Scene Generation for Generalizable Robot Learning and Evaluation
   CoRL 2026 · Accepted
5. InternVLA-A1.5: Unifying Understanding, Latent Foresight, and Action for Compositional Generalization
   Technical Report · 2026
6. ST4VLA: Spatially Guided Training for Vision-Language-Action Models
   ICLR 2026 · Accepted
7. InternVLA-M1: A Spatially Guided Vision-Language-Action Framework for Generalist Robot Policy
   Technical Report · 2025
8. InternVLA-A1: Unifying Understanding, Generation and Action for Robotic Manipulation
   Technical Report · 2026