

欧南涵 | 大模型算法实习生

预训练 | 后训练 | 强化学习 | 具身智能

教育背景

北京大学软件与微电子学院
软件工程（2026 级硕士研究生）
浙江大学光电学院
光电信息科学与工程（本科）

中国，北京
2026.09–2029.06
中国，杭州
2022.09–2026.06

科研成果

本科阶段发表一篇 SCI 独立一作 | *Reliability-Controlled Stitching for Large-Area White-Light Interferometric Surface Topography Reconstruction under Weak Overlap* **2025–2026**

- **Optics Communications** (DOI: 10.1016/j.optcom.2026.133668)。独立完成弱重叠白光干涉形貌拼接的算法设计、实现、实验评估与英文论文写作；融合 SuperPoint+LightGlue、guarded ICP、可靠性筛选、桥接边和加权位姿图优化。
- 在 WLI39 基准完成 **39/39** 拼接， $RMSE_{XY}$ 从 $27.17\ \mu m$ 降至 $0.247\ \mu m$ 。

实习经历

数坤科技 | 多模态大模型算法实习生（医疗 ASR 后训练优化）

2026.07–至今

- **长对话 ASR 后训练**：在 **5335** 条通用对话、超声、TTS 与医疗门诊语音上，以 CER 构造奖励完成 GRPO 后训练；结合错误位置加权与监督交叉熵约束，兼顾定向纠错与识别稳定性。
- **模型评测与选型**：完成 **10** 个模型版本的医疗长对话评测并选出最优模型；在内部测试集上，Mean/Micro CER 较基础模型分别相对降低 **9.9%/10.9%**。
- **语音切分优化**：以非平稳降噪和能量回补优化语音切分，按最终时间边界从原始音频切分片段输入 ASR；在内部测试集上，Mean/Micro CER 较基础模型分别相对降低 **12.9%/13.2%**。
- **医学对话 SFT 数据构建**：针对“一诉七史”直接生成难以兼顾问诊结构与口语风格的问题，设计病历结构种子与真实语音口语锚点结合的双种子流程；先生成主诉/现病史核心问诊，再按未覆盖事实续写其余病史，累计输出 **2154** 条多轮医患对话及质检报告。

项目经历

【预训练 / 后训练】Mini-LLaVA 两阶段训练与视觉语言对齐

2025

- 面向图像 caption/VQA，基于 CLIP ViT-B/16 与 Qwen 搭建 Mini-LLaVA；以两层 MLP Projection 将视觉 token 映射至语言模型空间，打通图像编码、视觉投影与文本生成链路。
- 两阶段训练：第一阶段冻结 CLIP/Qwen，取 SA1B-Dense-Caption 前 **10** 万图文对训练 Projection，在约 **1** 万条固定验证样本上将 PPL 从 **10.2** 降至 **4.1**；第二阶段冻结 CLIP，以 CogVLM-SFT-311K 联合微调 Projection 与 LLM，并通过 VQA/caption 样例及错误案例评估。

【具身智能 / 后训练】OpenVLA-OFT 监督微调与动作建模

2025–2026

- 面向双臂连续操控，将 RoboTwin 轨迹由 HDF5 转为 ALOHA/RLDS，并对齐图像、语言指令与关节状态；训练 OpenVLA 根据当前观测一次输出 **25 步 × 14 维** 双臂绝对关节角动作。
- 使用约 **500** 条双臂轨迹，在 **4×RTX 4090** 上采用 FSDP+LoRA 训练约 **40k steps**；固定单任务成功率 **76%**。

【具身智能 / 强化学习】VLA PPO 在线强化学习与评测

2025–2026

- 参与 VLA PPO 在线训练；在 **128** 个并行环境中持续采样，模型根据图像和任务指令生成 **25 步** 动作块，并记录动作、奖励、终止信号与动作对数概率。
- 基于采样轨迹进行优势估计与 PPO 更新，监控 reward/KL/value loss 并回放失败轨迹；固定任务成功率 **72%**。

专业技能

- **编程与深度学习**：Python、PyTorch，熟悉 Linux 开发环境。
- **模型训练与推理**：SFT、LoRA、FSDP、PPO、GRPO、vLLM；熟悉 Qwen、LLaVA、OpenVLA。

奖项荣誉

- 2026 年美国大学生数学建模竞赛（MCM/ICM）一等奖（Meritorious Winner）