



黎家辉

求职意向：ai 应用开发

22 岁 | 苏州市 | 15262073926 | 3314794663@qq.com

实习经历

AI 视觉行为监测系统 · 苏州鲲因科技有限公司 · 系统开发

2025.11 – 2026.06

面向生物制药生产车间 B/C/D 级洁净区实时视频 AI 行为分析平台，覆盖 PPE 合规检测、关键操作行为监控与 GMP 审计追溯，现场接入海康威视 NVR 16 路生产/检验通道

- 参与 FastAPI 后端开发，实现摄像头管理、分级报警、RBAC 权限、审计日志、报表导出等 REST API，支撑 GMP 合规要求
- 参与本地 YOLO 推理服务开发与优化，对接 16 路 HLS/RTSP 视频流，实现口罩、护目镜、手套、防护服等 PPE 检测，单路响应 ≤ 3 秒、准确率 $\geq 95\%$
- 基于 FFmpeg+ZLMediaKit 搭建视频流子系统，实现海康 NVR 多码流转 HLS 及多路并发预览与 AI 推理取帧
- 参与海康 AIOP 报警接入与本地推理方案切换，实现报警去重、截图存档、洁净等级映射及 WebSocket 实时通知
- 参与 Vue 3 + Electron 前端开发（监控大屏、实时报警推送、HLS.js 多路播放）及安全合规模块（JWT 鉴权、AES-256 加密、面部脱敏、操作审计追踪）
- 完成 PyInstaller 打包部署与现场联调，编写 Windows 生产环境运维文档

技术栈：Python、FastAPI、SQLAlchemy、YOLOv8、OpenCV、PyTorch、HLS/RTSP、海康 SDK、Vue 3、Pinia、Electron、WebSocket、SQLite、ffmpeg

项目经验

RAG 知识库智能问答系统

从零设计并实现企业级 RAG 系统，支持 PDF/DOCX/CSV 等 6 种格式文档的智能问答。自研四阶段检索管线，每阶段独立可拔插、可评估对比。

- 混合检索：FAISS (HNSW) 向量检索 + BM25 (jieba 分词) $\rightarrow$ RRF 融合，Recall@5 提升 18% vs 纯向量检索
- Cross-Encoder 精排：粗筛 20 候选 $\rightarrow$ bge-reranker-v2-m3 联合编码精排 Top-5，MRR 提升 15%，检索失败率降低 67%
- Query Rewriting：LLM 将口语问题改写为书面检索词（"去年赚多少" $\rightarrow$ "2025 年净利润数据"），Recall 提升 23%
- 多轮对话：追问检测 + 上下文指代消解，追问回答准确率提升 40%
- ReAct Agent：LLM 自主调用搜索/计算工具，思考过程可视化
- 评估体系：自建 MRR/NDCG/Recall@k 多维评估框架，支持 A/B 对比

技术栈：BGE Embedding、FAISS、BM25、Streamlit

教育背景

江苏师范大学科文学院 · 人工智能专业（本科）

2022.09 – 2026.06

主修课程：人工智能基础、机器学习、深度学习、计算机视觉、数据结构、计算机网络、计算机操作系统、计算机组成原理、MySQL、Python 程序设计

个人荣誉

- 论文《基于深度学习涉诈 App 智能识别分析系统》第一作者，发表于《互联网周刊》
- 计算机软件著作权证书《涉诈 App 智能识别分析系统 V1.0》
- 全国大学生英语六级证书 (CET-6)
- 江苏省大学生创新创业训练计划项目

个人评价

热衷人工智能与软件工程交叉领域开发，具备扎实的 Python 后端与计算机视觉工程能力：熟练 Flask/FastAPI、MySQL、SQLAlchemy 等后端技术栈，及 OpenCV、YOLO、PyTorch 等图像算法与深度学习框架；拥有从 0 到 1 的 AI 项目全流程落地经验，涵盖数据与特征工程、模型训练调优（关键指标提升 5%+）及高性能部署（响应延迟<500ms）；善于快速学习新技术并解决复杂技术问题（如大一自学 Spark 处理医疗保险欺诈数据，获“优秀奖”），沟通协作能力强，能将业务需求转化为可落地的技术方案。

其他经历

软件著作权技术顾问

2024.03 – 至今

- 独立完成 7 个以上软件项目的需求分析、开发与技术文档撰写，客户一次性通过率超 85%；灵活运用 Python/Java/JavaScript、MySQL、Vue/React 等技术栈按需交付