

· 姓名: 吴志鹏 | 年龄: 23岁 | 电话: 13641896821 | 微信: WNPPP0114 | 邮箱: 13641896821@163.com

· 目标岗位: 嵌入式 AI/模型部署/边缘计算工程师

个人优势

- 软硬件全栈工程能力: 具备贯穿 MCU/RTOS 固件、Linux BSP 构建到上层 AI 模型部署的全栈能力; 熟悉在资源受限平台进行系统构建、异构计算与性能优化。
- 商业级产品量产经验: 拥有大模型应用、智能机器人、工业物联网领域的完整工程落地经验, 具备从硬件选型、固件设计、原型验证到批量生产的技术产品化能力。



教育背景

2026.09-2027.06	香港科技大学	信息与通信工程（电信学）	硕士
2021.09-2025.06	上海大学	电子信息工程	本科

专业技能

- AI 与部署: 熟悉 PyTorch 至 ONNX/TensorRT/RKNN 的转换链路与算子适配; 理解 YOLO/LLM/VLM 架构; 具备 Jetson/Rockchip 平台移植与环境构建经验; 熟悉微调/剪枝/量化/Zero-Copy/流水线并行等调优技术。
- 嵌入式与硬件架构: 熟悉 Rockchip 平台的 Linux BSP 构建 (U-Boot/Kernel/设备树/文件系统) 与驱动/QT 开发; 熟悉 STM32 开发和 FreeRTOS 系统移植; 拥有相关 PCB 设计的商业产品量产经验。
- 通信协议: 熟悉 I2C/SPI/UART/RS485 等总线, 以及 Zigbee/NB-IoT/MQTT/Modbus-TCP 等物联网通信协议。
- 开发语言与工具: C/C++, Python, Shell; PyCharm, VS Code, Keil, CubeMX, Ubuntu, Cursor/Trae

荣誉与证书

- 竞赛与奖项: 2024 年 TI 杯上海市大学生电子设计竞赛三等奖; 2024 年上海大学大学生电子设计竞赛二等奖
- 证书与技能: 国家新型实用专利 (共同发明人, 202521718502.8); 雅思 IELTS 7.0; 大学英语六级 CET-6

实习经历

2025.07-至今	上海晨晖智能科技有限公司	嵌入式 AI 工程师
· 自研微调框架与 VLM 工业落地 (Agent/LoRA/Qwen3-VL): 针对 YOLO+CV 在工业环境下识别目标困难的痛点, 使用 Agent 工具从零设计 LLM/VLM 通用的 LoRA 微调框架; 完成 Qwen3-VL 的工业适配与微调, 结合提示词工程, 实现宝钢厂区除尘器管道泄漏检测准确率超 99%。(开源链接: https://github.com/WNPPP0114/Lora-Struct)		
· 机器人视觉推理优化 (Jetson/TensorRT/YOLO11): 针对高清视频流实时检测需求, 设计多线程异步流水线架构, 实现了模型推理与后处理的并行执行, 并通过模型剪枝与微调, 保障检测质量的同时实现系统吞吐量额外提升 80%。		
· 巡检机器人固件重构 (FreeRTOS): 独立负责机器人固件从裸机到 FreeRTOS 架构的迁移; 解耦数据采集/逻辑判断/网络通信等模块, 运用信号量/事件标志组/互斥锁, 有效解决裸机轮询的低效与竞态风险, 提升工作稳定性。		
· 低功耗无线系统研发 (IoT/STM32/PCB): 负责多款低功耗无线传感器的 PCB 与固件设计; 实现终端设备整体功耗降低最多 60% (工作/休眠电流降至 25mA/9uA), 预期续航超 8 年; 自研 Zigbee-TCP 网关, 将单设备 BOM 成本降低 50%, 并实现生产周期自主可控。产品获宝钢等客户批量采购。		

2025.03-2025.05	中冶集团・上海二十冶建设有限公司	物联网开发实习生
· 边缘侧数据清洗与自动化 (Python/Modbus): 负责工业传感器数据的采集链路验证, 编写 Python 脚本批量解析私有 Modbus 协议数据; 设计异常值过滤算法清洗现场噪声数据, 为上层业务平台提供高质量的数据集支撑。		

2024.12-2025.02	上海中治横天智能科技股份有限公司	算法工程师
· 专利贡献: 作为共同发明人研发以 STM32 为核心的用于工业环境的控制柜, 获得国家新型实用专利。		

2023.07-2023.08	宝武集团・上海金艺检测技术有限公司	视觉算法实习生
· 视觉增强算法预研: 针对金属反光, 部署 CLAHE 作为深度学习模型的前置增强模块, 显著提升微小缺陷检出率。		

项目经历

2025.11-至今	基于 RK3588 的 Linux 系统构建与端侧异构视觉/LLM 部署	个人开源项目
· 项目链接: https://github.com/WNPPP0114/RK-Linux-Hetero-Fusion		

· Linux 系统构建 (BSP): 基于 Rockchip SDK 定制 Linux 系统 (U-Boot/Kernel/RootFS), 集成 PREEMPT_RT 补丁实现硬实时; 基于 Qt6+EGL/DMA-BUF 实现 NPU/RGA 显存直通 GPU 渲染, 达成 2K UI 60FPS 流畅交互。
· 异构视觉流水线 (Zero-Copy): 构建 RGA+NPU 全硬件加速链路 (YOLO11/26), 利用 DRM Zero-Copy 消除数据拷贝; 设计多线程异步调度引擎实现 3 核 NPU 满载并行, 视频吞吐量稳定 60FPS+。
· 端侧多模态推理 (Token Injection): 部署 Qwen3-1.7B (W8A8), 设计“语义 Token 直通”机制, 建立 YOLO 类别 ID 与 LLM 词表 ID 直接映射, 跳过 Tokenizer 与 JSON 开销, 推理速度达 14 Tokens/s。