

基本信息

姓名：吴瀚

年龄：21

电话：+86 19355693187

邮箱：wuhan9918@gmail.com

学校：上海交通大学

专业：自动化（IEEE 试点班）

预计毕业时间：2027 年 7 月

GitHub：github.com/HanWu9918



教育背景

上海交通大学（985 高校，全日制）

2023 年 8 月 – 2027 年 7 月（预计）

自动化（IEEE 试点班）工学学士；辅修：法学（知识产权方向）

上海

- 电子信息与电气工程学院
- 核心课程：人工智能原理与应用、机器学习、计算机视觉、机器人学
- 英语能力：大学英语四级 630，大学英语六级 609，雅思 7.0
- 荣誉：上海交通大学致远荣誉奖学金

基础能力

- 编程语言：Python（熟练），C++（中等），熟悉 JavaScript、HTML/CSS、MATLAB、LaTeX。
- AI 与工程：熟悉 PyTorch、OpenCV、FastAPI、LMDeploy、OpenCompass；熟练使用 Git 与 Linux；具备大模型多卡集群推理部署与自动化评测流水线搭建经验。
- 评测方法：熟悉 LLM-as-a-Judge（pointwise / pairwise）评测范式、rubric 设计与人工抽检校准流程，覆盖大语言模型与多模态模型。
- AI 辅助开发：熟练运用 Codex、GitHub Copilot、Trae 等 AI 编程工具，能准确评估大模型能力边界并合理规划人机协作流程。

论文发表

Jize Wang, Han Wu, et al.. *RouteMoA: Dynamic Routing without Pre-Inference Boosts Efficient Mixture-of-Agents*.
录用于 ACL 2026 主会（长文，CCF-A）；代码：github.com/Jize-W/RouteMoA。

Teamwork at Alibaba Group. *FilmBench: A Film-Grade Benchmark for Cinematic Video Generation*. arXiv 预印
本；AAAI 2027 在投；代码：github.com/Neo-yk/FilmOps。

实习经历

阿里巴巴控股集团

2026 年 6 月 – 至今

大模型评测实习生

杭州

项目一 – FilmBench

- 参与 FilmBench（电影级 T2V/R2V 评测基准，AAAI 2027 在投），评测 HappyHorse、Seedance 2.0、Kling 3.0/Omni、Veo 3.1、Vidu、Hailuo 等主流视频生成模型。
- 从镜头语言、多镜头连续性、指令遵循和动态美学等维度评估生成结果，并将电影制作概念转化为一致、可执行的标注规范。

项目二 – QwenEditBench（通义实验室合作）

- 参与 QwenEditBench（与通义实验室合作的开源图像编辑评测基准），评测 Seedream 5.0 Lite、GPT Image 2 等模型，并完善编辑意图、指令遵循、视觉质量和失败案例等评测分类体系。
- 参与基于 AI Agent 的评测集构建与数据生成，负责 prompt 设计、样例筛选和质量审核；项目计划投稿 ICLR 2027。

项目三 – Qwen-Image 训练数据支持

- 根据 Qwen-Image 训练团队的需求，整理、筛选并质检部分图像生成训练数据，按照统一规范完成交付。

RouteMoA：面向高效混合智能体的动态路由方法	2025 年 5 月 – 2025 年 9 月
核心贡献者	
- 提出并验证动态路由 MoA 框架：推理前筛选高潜力候选智能体，推理后通过 LLM 自评与交叉评估进行后验校正。 - 基于真值匹配分数与 LLM-as-a-Judge 结果构建路由器监督信号；为 86M 参数的 mDeBERTaV3-base 评分器生成、清洗并审核训练数据。 - 使用 LMDeploy、FastAPI 与 OpenCompass 搭建评测栈，覆盖 iFEval、GSM8K、AIME、GPQA、AlpacaEval、HumanEval、BBH 等基准，完成超过 2 万条问答样本的评测。 - 完成与 MoA、SMoA 的模型池对比及消融实验，在保持更优效果的同时将大模型池推理成本降低 89.8%、延迟降低 63.6%。 - 参与实验方案设计、结果分析、实验记录整理与论文撰写，成果以长文发表于 ACL 2026 主会。	
飞机客舱异常声音识别与分类	2024 年 6 月 – 2024 年 9 月
上海交通大学—中国商飞合作项目；核心成员	上海
- 面向客舱质量保障任务，采集并清洗巡航、起飞和降落场景下的音频，构建 STFT/Mel 频谱特征以识别 12 类异常声音。 - 使用 SpecAugment 与时域片段替换扩充数据，并针对真实飞行噪声和机载硬件约束优化训练与验证流程。 - 设计轻量级 MobileNetV3 分类器，引入 ECA、深度可分离卷积和 h-swish，并结合 RMSprop warmup、cosine annealing、EMA 与 early stopping 优化训练。 - 通过混淆矩阵、逐类指标和强噪声测试验证模型；平均准确率达 98.01%，强噪声准确率提升 10.66%，0.4–1 秒片段的 CPU 推理耗时控制在 20ms 以内。	
代表性工程项目	
多智能体生产规划系统 <i>CoppeliaSim</i> 、 <i>LLM Agents</i>	2026 年 5 月 – 2026 年 6 月
- 搭建双线柔性制造仿真器，将自然语言订单解析为产品类型和数量，通过 OpenAI 兼容的本地 LLM 接口生成生产计划，并保留规则基线方案。 - 设计三智能体规划架构，覆盖汽车与手机装配、零件转运、带碰撞感知的滑块调度、多批次生产和质量检查。 - 实现基于 AABB 的逐帧碰撞检测和零件映射逻辑，支持连续跨线转运、协同装配与完整生产流程仿真。 - 对比 multi-agent、single-agent 和 rule-based 三种模式，分析不同规划方式在执行稳定性和任务完成度上的差异。	
面向策略型台球的 AI 智能体 <i>MCTS</i> 、贝叶斯优化、 <i>LLM Agents</i>	2026 年 2 月 – 2026 年 7 月
- 为美式八球设计 physics-grounded Agent 框架，将球桌几何、ghost-ball 瞄准、射线-圆碰撞预测和击球难度转化为可计算的决策特征。 - 构建 LLM iterative replanning 流程，根据仿真反馈更新击球方案，并结合 MCTS、Bayesian Optimization 和 ensemble voting 提高决策稳定性。 - 对比几何启发式、MCTS、Bayesian Optimization 和 LLM Agent 等方案，最佳专用 Agent 对基线取得 86.75% 胜率。	
车牌识别系统 <i>OpenCV</i> 、 <i>PyTorch</i> 、 <i>LPRNet</i> 、 <i>PyQt5</i>	2024 年 11 月 – 2025 年 1 月
- 构建完整识别流水线，覆盖高斯模糊、Canny 边缘检测、轮廓筛选、HSV 颜色过滤、车牌定位、裁剪与字符分割。 - 在 CBLPRD330k 上训练轻量级 LPRNet，结合数据增强、CTC loss 和 greedy decoding 完成端到端字符识别。 - 开发支持本地图片上传和实时摄像头输入的 PyQt5 GUI，并使用 PyInstaller 打包模型、资源与应用程序。 - 完成从图像预处理、模型推理到结果可视化的可部署系统，单字符识别准确率达 86%。	