

# 赖彦涛

电话: 13033081901(微信同号) | 邮箱: yantaolai@nuaa.edu.cn | 中共党员  
个人主页: yantaolai.github.io



## 教育经历

南京航空航天大学 (211, 推免)	硕士	2023 年 9 月-2026 年 3 月
计算机技术   计算机科学与技术学院   模式分析与机器智能 (工信部) 重点实验室		
安徽大学 (211)	本科	2019 年 9 月-2023 年 6 月
智能科学与技术   互联网学院		

## 实习经历

北京百度在线网络技术有限公司	自动驾驶感知模型算法工程师	2025 年 5 月-2025 年 9 月
• 实习内容: 百度智能驾驶 apollo (L4) 离线大模型训练相关工作。主要包含 1、开城数据挖掘: 对原始数据提取关键 clips、质量检测, 送给数据产线标注 2、大模型数据质检: 对数据产线标注结果使用离线大模型推理、查找 diff、可视化质检 3、AIHUB 工具链: 设计前端网页, 封装模型评测、指标获取、数据质检、case 回归、corner case 挖掘等功能 4、SFT 阶段训练: 使用倒地行人、拉树车等特定精标数据, 通过训练来提升模型整体能力。		
北京小米移动有限公司	AIGC 与大模型算法工程师	2024 年 12 月-2025 年 3 月
• 实习内容: 小米文生图大模型 “梵高” 项目。参与了 1、文生图大模型 FLUX 的 LoRA 微调与全参微调 2、FLUX 的超分模型训练; 主要负责 3、高质量图像-文本对收集 4、图像编辑功能 (如 Inpainting, Outpainting) 实现。通用文生图效果超过了市面大部分闭源/开源模型, 并在内部投入使用。 • 相关知识: FLUX, Stable Diffusion, ControlNet, IP-Adapter, SUPIR, LORA, DiT, UNet, Flow Match, DDPM, DDIM, Diffusers, Accelerate, PEFT		
合肥科大讯飞教育发展有限公司	数据资源开发工程师	2023 年 3 月-2023 年 6 月
• 实习内容: 科大讯飞 “讯飞星火认知大语言模型” 数据项目。主要包含 1、大语言模型对话数据清理、标注 (如: CoT、润色) 2、数据质量/安全审查。		

## 科研经历

ScanLLM: Object Referring-Guided Scanpath Prediction using MLLM (第一作者)	AAAI 2026 (在投)
• 任务描述: 针对于日常中普遍使用的指代文本表达, 预测人类在逐步听到或者看见其中每一个单词后, 视线注视点坐标的即时转移轨迹。 • 相关知识: 指代表达、增量预测、视觉语言模型 VLM、Qwen2.5-VL、InternVL 2.5、LoRA、Sa2VA、LISA、SAM、DeepSpeed、ZeRO、Xtuner	
CLIPGaze: Zero-Shot Goal-Directed Scanpath Prediction Using CLIP (第一作者)	ICASSP 2025 (oral, 10%)
• 任务描述: 针对在零样本/目标存在/目标缺失等三种情况下, 预测人类在给定完整目标文本/视觉 prompt 后, 视线注视点坐标的转移轨迹。 • 相关知识: 意图预测、眼动、目标导向、多模态、零样本、预训练、指代分割、视觉-语义融合、CLIP、DETR、CLIPSeg、Transformer	
Pathformer3D: A 3D Scanpath Transformer for 360° Images (与导师共一)	ECCV 2024
• 任务描述: 针对在全景图像的场景下, 预测人类在自由观看时视线注视点的坐标转移轨迹。 • 相关知识: 全景图像/360°图像、意图预测、眼动、时序预测、三维混合密度网络、SphereCNN/球面卷积、Transformer	
一种目标导向的扫视路径预测方法 (导师一作, 本人二作) 专利号: ZL 202411720546.4	专利已授权
一种面向全景图像的人眼扫视轨迹预测方法 (导师一作, 本人二作) 专利号: ZL 202410512599.0	专利已授权

## 项目经历:

基于三维点云的牙齿修复研究
• 项目描述: 针对目前医疗患者缺失牙齿修复过程依赖医生手动设计的问题, 我们提出了一种适用于牙齿点云的自动修复方法。为了适用于牙齿点云数据, 我们对点云补全方法 CRA-PCN 进行了全参微调, 然后采用了点云分割方法 PTv2 和点云上采样方法 RepKPU 来生成密度均匀且光滑的目标点云 • 相关知识: 点云补全, 点云上采样, 点云分割
Mini Qwen
• 项目描述: Mini Qwen 包含了预训练(PT)、微调(SFT)和直接偏好优化(DPO)3 个阶段, 从头开始训练了一个 1B 参数量的大型语言模型 • 相关知识: Qwen, PreTrain, SFT, DPO, Transformers, Trainer, Accelerate, TRL(SFTTrainer, DPOTrainer)

## 技术细节 & 荣誉

- 研究领域: 计算机视觉, 意图预测 (眼动), 视觉语言模型 VLM, 文生图模型 AIGC, 自动驾驶感知算法
- 主要课程: 计算机视觉, 机器学习, 模式识别, 图像处理与分析, 智能控制理论, 智能软件开发与测试技术, 数据库, 自然语言处理
- 编程语言 & 框架: Python, C++, C, PyTorch, Linux Shell, Hugging face, Git, Docker, LaTeX, OpenCV, Xtuner, OpenMMLab, VeRL, TRL
- 本硕期间荣誉: 科研创新先进个人, 三好学生, 优秀共青团员, 同峻奖学金, 一等奖学金, 研究生国家奖学金
- 其他: CET-4/6