



教育背景

北京邮电大学 211 本科 信息工程 双一流 2022.09-2026.06

- 曾获奖项：“青聚AI”人工智能+专项赛 北京市三等奖 | “互联网+”大学生创新创业大赛 国家级项目 & 北京市级项目 | 中国大学生服务外包创新创业大赛北部赛区 三等奖
- CET-4 (535分) / CET-6 (530分)

北京邮电大学 211 硕士 人工智能 双一流 2026.09-2029.06

项目经历

天翼AI DataEvolver: 目标驱动的可控视觉数据自演化系统 一作+项目负责人 2026.03-至今

项目背景：针对可控视觉生成与图像编辑中合成数据高度依赖人工Blender调参及固定规则质检的痛点，设计并实现首个基于VLM + AI Agent驱动的自演化数据构建框架，实现从图像生成、3D重建到自动质检修复的全链路数据闭环。
Github 仓库：PRIS-CV/DataEvolver

- 多阶段合成数据流水线：构建从文本 Prompt 扩展、RGB/SAM3 前景分割、Hunyuan3D 2.1 拓扑重建到 Blender 自动化渲染的端到端数据流水线，实现图像、mask、mesh 与相机/几何元数据资产链的自动化生成与版本化管理。
- VLM/CV 双门控与有界动作修复：引入 VLM 自由文本诊断（诊断悬浮/光照/构图/语义漂移）与 CV 质量门控（清晰度/曝光/遮罩），设计含 24 个原子动作的有界修复控制器（覆盖光照/材质/场景参数），实现无人工干预的样本质检与精准修复。
- 内外双层自演化与防震荡控制：构建样本级内循环（门控过滤/修复/淘汰）与数据集级外循环（根据下游弱角度评测定向补全数据）；设计方向翻转追踪、参数冻结与步长衰减调度（100%→70%→50%），成功解决 Agent 调参震荡难题。
- 落地成果与学术验证：基于 Qwen-Edit 统一 Probe 及 SpatialEdit 评测，Ours+DualGate 最终集达到 PSNR 16.79 dB（较 Public LoRA 提升 +1.03 dB）、CLIP-I 0.9525 (+0.0625)、DINO 0.8918 (+0.0474)。论文（arXiv:2605.01789）入选 KDD 2026 Workshop (AIDataSci) Oral 报告与 2026 BAAI Agent for Science Competition 现场海报展示，代码与项目主页完全开源。

天翼AI 多任务图像编辑 LoRA 微调系统 AI算法实习生 2025.09-2026.07

项目背景：基于 Qwen-Image-Edit-2511 大规模图像编辑基座模型，构建多任务图像编辑 SFT 微调系统，支持物体旋转、移动、插入、移除四类核心编辑任务。

- 分角度 SFT 训练：针对多角度统一训练导致的旋转失真与不精准痛点，设计分角度独立训练策略，完成 11 个独立 LoRA（顺/逆时针 30°-180°），大幅提升工业级旋转编辑精度。
- 课程学习策略：设计二阶段 Preset LoRA 机制：白底高容量 LoRA（Rank=128）显式学习几何变换 → 冻结后 Rank=64 残差 LoRA 学习复杂背景融合，训练效率提升 50%。
- 统一分角度 LoRA 模型：设计 6 阶段增量式数据重放（Additive Rehearsal）课程学习，按历史角度远近递减追加重放数据防止灾难性遗忘，成功训练出支持 30°-180° 全角度的统一 LoRA 模型。
- 分布式调度与上线：调度 3 台服务器（4+4+8 卡 H100）分布式训练与并行推理，分角度模型上线 TeleStudio 平台服务工业应用（精度提升 15%+），统一模型具备高度学术投稿价值。

“邮爱” - 自进化多模态UI智能体 项目负责人 2024.06-2025.07

项目背景：针对移动端自动化测试中长程复杂任务规划困难、智能体自我纠错能力弱的核心痛点，主导研发自进化 UI Agent 框架。

Github 仓库：Kamisato520/MobileAgent

- **分层多代理协作架构**：作为项目负责人，设计基于 Mobile-Agent-E 的分层协作框架，将长程移动端任务解耦为 Manager 高级规划与 Operator 低级执行，显著提升长程规划与动作执行鲁棒性。
- **双轨记忆自进化与 SOTA 性能**：引入 Tips（经验）与 Shortcuts（技能）双轨持久化记忆机制，使智能体具备历史经验沉淀与错误恢复能力。在 Mobile-Eval-E 评测基准上满意度得分达 86.9%（较 SOTA 提升 22%），动作准确率 90.4%，终止错误率降至 12.0%。

科研与项目经历

DataEvolver: Let Your Data Build and Improve Itself via Goal-Driven Loop Agents

AI4Data Goal-Driven Loop Agents Synthesized Data Evolution

第一作者+项目负责人：KDD (CCF-A) 2026 Workshop (AI4DataSci) Oral | BAAI 2026 Agent for Science 海报展示

核心贡献：提出首个目标驱动的可控视觉自演化数据构建框架。设计 Prompt-RGB-3D-Blender 自动化流水线、VLM/CV 双门控及 24 有界动作修复控制器，配合内外双层反馈闭环与防震荡调度，实现数据质量自主演化与下游模型性能显著提升（SpatialEdit PSNR 16.79 dB / CLIP-I 0.9525）。

Traj-MCPO: Trajectory-wise Marginal Consistent Policy Optimization

第二作者：ACM MM 2026 已接收 (CCF-A) | Kamisato520/Traj-MCPO

Rectified Flow Alignment RLHF / Post-training Text-to-Image Policy Optimization

核心贡献：针对整流流 (Rectified Flow) 扩散模型 RLHF 后训练中 ODE-SDE 边际不匹配与奖励粒度不匹配难题，提出轨迹级边际一致策略优化框架。设计长度归一化轨迹似然比与轨迹级裁剪目标，显著提升优化稳定性与推理一致性。在 Qwen-Image / FLUX 基座上图像质量偏好率达 75.0%、指令对齐偏好率达 83.2%，全面超越 DanceGRPO / MixGRPO 等基线。

Panel2Page: Structured Visual Narrative Generation with Omni-Panel Alignment

第二作者 | AAAI 2027 (CCF-A) 在投

Structured Visual Generation Hierarchical Semantic Alignment Omni-Panel Alignment

核心贡献：针对结构化视觉叙事生成（漫画页面生成）中全局版面构图与跨分镜局部语义一致性的协同难题，提出单阶段双过程生成框架 Panel2Page。采用 Page-Level SFT 显式学习全局结构先验；提出 Omni-Panel Alignment (OPA) 训练策略，融合文本感知、分镜事件及跨分镜视图的局部特征增强对齐（不增加推理开销）。构建 20,000 组高质量 Prompt-Page 数据集与 320 复杂结构 Benchmark，各项指标达 SOTA。

GeoEdit++: Learning Residual-Conditioned Rotation References for Geometric Image Editing

第二作者 | IEEE Transactions on Multimedia (TMM) 在投 (CCF-A)

Geometric Image Editing Rotation Control Residual Condition Multi-task LoRA

- **核心贡献**：针对复杂场景下物体几何旋转编辑导致的拓扑失真与背景一致性破坏问题，提出基于残差条件与分阶段/统一 LoRA 课程学习的精准几何图像编辑框架，提升多角度旋转控制精度与背景融合稳定性。

专业技能

编程与工程基础

- Python，具备大模型训练、推理与系统开发经验；熟悉 C++ / Java。
- 具备丰富的 AI-Native Coding 经验，精通 Claude Code、Codex 等主流 AI Coding 工具。

生成模型与图像编辑

- 熟悉 Diffusion / DiT 架构，具备大规模图像编辑模型（Qwen、FLUX）微调与工程化落地经验。
- 具备 T2I → 3D 重建 → Blender 渲染 → 质量闭环的合成数据构建经验。

强化学习与后训练

- 理解并实践 SFT + RL（GRPO / PPO / DPO）的完整后训练流程，能设计基于 VL 的多维奖励模型并实现 LoRA-Only 高效强化学习。

多模态大模型与智能体

- 具备 VLM + MultiAgent 自动化决策闭环的设计经验，熟悉结构化动作空间、防震荡控制与步长调度策略。