焦鹏昆

CURRICULUM VITAE

28岁 | 18217186385 | 303705335@qq.com | pengkun-jiao.github.io/

教育经历 EDUCATION

复旦大学 · 人工智能（硕博连续） 2021.09 – 2026.06

研究方向：大语言模型、多模态学习、高效学习 · **论文发表**：一作论文 6A1B · **荣誉**：启明星辰企业冠名奖学金

华东师范大学 · 计算机科学与技术（学士） 2017.09 – 2021.06

工作经历 EXPERIENCE

阿里巴巴 · 通义实验室 · Researcher · 模型架构 | 分布式训练 2026.07 – 至今

- **Prompt-level MoE 架构**：针对高频专家切换与通信开销，负责设计指令级 MoE 架构，利用小型 LLM 与逐层预测头为大型 MoE 生成专家掩码，将受控中间层的专家选择由 Token 级调整为请求级，固定单次请求内各受控层的激活专家集合，从而为专家预加载与执行调度规划提供依据。
- **分布式训练实现：**负责基于 Megatron 设计并行训练策略，利用 NCCL 实现专家掩码的跨流水线阶段分发与梯度回传，支持 Prompt-level MoE 架构的大规模训练；采用直通估计器（STE）实现离散专家掩码的近似梯度传播，通过自定义 backward buffer 汇总各流水线阶段的梯度贡献；结合分层激活检查点与重计算降低峰值显存占用，实现千卡规模高效训练。

阿里巴巴 · 通义实验室 · 实习 · 多教师在线蒸馏 | Agentic RL 训推优化 2026.04 – 2026.07

- **多教师在线蒸馏**：负责搭建 MOPD 训练框架，支持 397B 模型、10 教师与 256K 上下文的稳定训练；结合重要性采样与 ratio 裁剪处理采样分布差异，支持通用 Agent、代码等领域教师能力融合。**MOPD 算法优化**，针对教师跨领域短板引起的负迁移，设计 advantage 裁剪与负 advantage 影响衰减策略。保留教师性能95%+，其中部分指标甚至超越教师。
- **Agentic RL 训推优化**：负责 397B 模型的异步训练与推理优化链路，结合 early return 缓解 rollout 耗时长尾，通过重要性采样修正策略分布差异，在项目训练设置下实现约 2× 训练效率，并在 Terminal Bench 等基准上取得性能提升。

近期研究 RESEARCH

LU-KV：面向长期信息保留的 KV Cache 压缩 · 百度松果高效合作 · ICML 2026 2025.09 – 2026.01

- 针对现有 KV Cache 压缩方法侧重单步 Token 重要性、难以刻画长期信息价值及不同 Attention Head 重要性估计误差差异的问题，基于长期边际效用构建全局 KV Cache 预算优化目标，结合 PAVA 凸松弛与贪心策略求解预算分配；通过离线 Profiling 生成预算表，在线结合头内排序执行缓存压缩，兼容 SnapKV、KeyDiff 等头内打分策略。
- 在 LongBench 上，基于 Mistral-v0.3 和 Qwen-32B，将 KV Cache 占用减少 80%，同时保留 Full KV Cache 基线平均得分的 96.8%～99.5%；**NVIDIA KVPress leadboard 单 Forward 方法第一。**

Dual-LoRA：双低秩空间协同适配 · ICCV 2025 2024.11 – 2025.03

- 双空间条件化适配：针对多任务微调中的知识冲突与 LoRA-MoE 多专家计算开销，设计共享技能空间与指令条件任务空间；通过任务分支的 ReLU 门控，对归一化技能特征进行逐元素调制，实现低秩维度上的细粒度知识选择，在单个适配器内兼顾跨任务知识共享与任务特定响应。
- 针对高层视觉特征对局部细节表达不足的问题，设计 VCE 模块，利用可变形注意力从视觉编码器多个中间层采样并聚合局部信息，经跨层融合与残差注入增强视觉投影特征。

RoDE：线性整流路由与异构低秩专家 · ICMR 2026 2024.01 – 2024.04

- 针对多任务微调中的监督冲突与任务复杂度差异，将 LoRA 专家建模为可跨任务复用的技能模块；设计ReLU 线性整流路由，将非正响应置零、保留正响应加权，实现无需预设 Top-k 的动态稀疏专家组合，支持单任务调用多个专家及不同任务共享专家。

稀疏奖励下的步骤级信用分配与引导探索 · 淘宝AI搜索 · WWW 2025 2025.5 – 2025.9

- 长轨迹 RL 优化： 针对长轨迹任务中终局奖励稀疏、信用分配困难及难样本探索低效的问题，设计 Step-level PRM + Step-level GAE 的细粒度奖励与信用分配机制，将终局反馈细化至步骤级；结合 Ground-truth 引导探索提升有效轨迹发现率，并采用难度感知采样与动态数据平衡抑制训练中的策略塌缩，提升复杂任务上的探索与学习效率。

一作论文

PUBLICATIONS

[1] From Holistic to Localized: Local Enhanced Adapters for Efficient Visual Instruction Fine-Tuning · ICCV 2025

***Pengkun Jiao**, Bin Zhu, Jingjing Chen, Chong-Wah Ngo, Yu-Gang Jiang*

[2] Rode: Linear Rectified Mixture of Diverse Experts for Food Large Multi-modal Models · ICMR 2026

***Pengkun Jiao**, Xinlan Wu, Bin Zhu, Jingjing Chen, Chong-Wah Ngo, Yu-Gang Jiang*

[3] Predicting Future Utility: Global Combinatorial Optimization for Task-Agnostic KV Cache Eviction · ICML 2026

Z Tang, **Pengkun Jiao***, X Chen, W Liu, S Li, J Chen*

[4] Disentangling Semantic Attention from Structural Bias in the Attention Manifold · ACM MM 2026

***Pengkun Jiao**, Bin Zhu, Jingjing Chen, Yu-Gang Jiang*

[5] Spatiotemporal Sycophancy: Negation-Based Gaslighting in Video Large Language Models · ACL 2026 Findings

Z Tang, **Pengkun Jiao***, B Zhu, H Qi, J Chen, YG Jiang*

[6] Open-Vocabulary 3D Object Detection Enhanced by Comprehensive Guidance from Text and Image · ECCV 2024

***Pengkun Jiao**, Na Zhao, Jingjing Chen, Yu-Gang Jiang*

[7] Domain Expansion and Boundary Growth for Open-Set Single-Source Domain Generalization · TMM 2024

***Pengkun Jiao**, Na Zhao, Jingjing Chen, Yu-Gang Jiang*