

陈浩锟

多模态大模型 / Agent 系统 / 后训练与强化学习 — 算法工程师

杭州 imchenhaokun@foxmail.com +86 186 6808 9675 github.com/chenhk-chn

— SUMMARY

简介

专注多模态大模型、Agent 系统与后训练方向。目前在字节跳动 Seed 视觉应用主导豆包 / 即梦创作 Agent 的架构演进与强化学习训练，此前负责 DreamPoster、ETTA 的数据工程与开放式输入建模；更早在腾讯自研 QQ 业务大语言模型并落地生成式图像编辑。国家发明专利 18 项，共同一作论文发表于 arXiv 与 WACV。

— FOCUS

核心技能

AGENT 系统

Planner + Function Calling /
ReAct Loop 架构、工具调用、任务规划、多模态智能体

后训练与强化学习

SFT / RM / DPO / RL 全流程，
rule-based reward、trace
verifier、带图 verifier

多模态大模型与数据

VLM / T5 / LLaMA / CLIP 训练；自
动化标注、协议化合成、百万级数
据构建与评测

— EXPERIENCE

工作经历

Seed · 视觉应用（历经创意营销 → 智能创作 → Seed 视觉应用）

创作 Agent 架构与强化学习 · 算法 Owner / RL 负责人 · 2025.3 – 至今

主导创作 Agent 架构演进：Planner + Function Calling（豆包「超能创意」1.0 / 2.0）→ 三合一 / 三合二 PlanList → ReAct Loop，把线上串联 VLM 工程链路升级为可训练、可评测、可 AB 的 Agent 模型体系。负责 Agent 强化学习方向，自研工具参数 + 工具选择的 rule-based RL，并沉淀 trace verifier、带图 verifier，使 RL 能约束中间推理与工具调用过程。

工具参数错误兜底率 3%→0.08% · 工具选择准确率 95.14% · Resource ID 幻觉率 0.71% · ReAct 复杂任务 Avg 1.186→1.511

豆包「超能创意」Web/客户端 LT30 +0.70%/+0.80% · 生图 DAU +169.4% · 渗透率 6.21%→12.07% · ReAct 升级后生图/生视频 UV 采纳率 +4.2%/+5% · App 创作失败请求 3.5%→0.5% · 复用至 Pippit / 剪映 / 豆包 Mobile

即梦创作 Agent · 算法 Owner · 2025.7 – 2025.11

完成即梦 Agent 架构四期迭代，走通面向 VLM 的后训练全流程（人工标注 → SFT / RM / RL），并落地创意顾问、记忆工具、上下文治理、灵感搜索等能力模块。

文生图可用率 31%→42% · 文生视频 6%→17% · 创意度 Best 率 21%（豆包 3% / Agent-Claude 1% / GPT-5 0%）

国内 Agent DAU 0→34.4 万 · 活跃渗透 0→45.48% · 生成→保存 PV 转化 31.3% · 大盘 LT +1.27% · 会员 ARPU +9%

DreamPoster / ETTA 图文生成 · 算法 Owner · 2024.3 – 2025.6

数据侧构建渲染 + Inpainting + 裁剪的成对数据链路（500 万+ 图片对、万级高质量海报），并把 OCR + 多模型组合链路收敛为单一多模态标注模型；输入侧基于 Doubao-1.5-Vision-Pro / Qwen2.5-VL 开发海内外 PE 模块，用 persona 模拟 + DPO 把开放式输入转成模型偏好的结构化 Prompt。

DreamPoster 可用率 88.55%（GPT-4o 47.56% / SeedEdit3.0 25.96%）· ETTA 中英文海报 GSB +40%/+35% · 文字标注编辑距离 4.94→3.40 · SDXL+ETTA 文字准确率 18%→82%

即梦首发带动图生图 UV 渗透 +1.8%、保存渗透 +9.7%、图生成 LT7 +1% · ETTA 带动 Web 文生图保存渗透 +1.4%，Web DAU 破 10 万 · 上线即梦 / Pippit / CapCut / 剪映 / 豆包 / 抖音

早期项目 · 2023.10 – 2024.8

AI 商品图与抖音相机的开放式输入（PE）；剪映 / CapCut 营销智能文案与 CCPush 多语言推送文案模型。

商品图 PE 对比美图 / PhotoRoom CMOS 109%/114% · CommercePro DAU 导出率 41.46% · Push 文案可用率 50%→88%，美区 LT7 +0.365%

其他：公司级技术分享 veTalk《创作 Agent 构建之道》；2025 Q3 Spot Bonus（个人）、2025 Q1 Spot Bonus（优秀团队）。

Hu, X.⁺, Chen, H.⁺, Qi, Z.⁺, et al. *DreamPoster: A Unified Framework for Image-Conditioned Generative Poster Design*, arXiv:2507.04218, 2025.（⁺ 共同一作）

大语言模型 · 算法 Owner

为 QQ 业务自研系列语言大模型（1B / 13B / 65B），基于 DeepSpeed、PyTorch-Lightning、Megatron-LM 完成 200B+ 中文 token 预训练与 SFT，负责训练框架搭建、Tokenizer 训练与幻觉问题优化。

MMLU / HumanEval 表现优异，业务场景效果超越 GPT-3.5 · 落地 QQ 小世界毕业季活动、魔法画室等场景

生成式图像编辑：发型替换 & AI 换脸 · 一期 Owner / 二期核心参与者

基于 StyleGAN 特征劫持 + GFPGAN 构造图片对，实现任意发型替换与染发能力；基于 DeepFake + FaceShifter 实现通用变脸模型，并通过 OMGD 蒸馏完成移动端小型化。此外参与优化 QQ-AIGC 文本特征提取器（mT5）与 CLIP 跨模态对齐（13B 语言塔）。

换脸 ID 保留指标达 78% · 上线 QQ 小世界发型屋、Anyu/黄昏发型、国庆及万圣节变脸活动 · 产出发明专利 5 项 · 相同 clip score 下 FID 降低约 2 点

阿里巴巴达摩院（杭州）科技有限公司 · 实习算法工程师

2020.03 – 2020.12

监控图像人脸老化编辑 · 算法 Owner

基于 cGAN 框架实现人脸年龄编辑，模型落地于杭州市城市大脑项目。

— EDUCATION

教育经历

浙江大学 · 电子与通信工程 · 硕士

2018.09 – 2021.03

机器视觉与自主导航实验室 / 阿里巴巴-浙大前沿技术联合研究中心

浙江大学 · 信息工程 · 学士

2014.09 – 2018.06

Wang, M., Lai, B., Chen, H., et al. *Towards Precise Intra-camera Supervised Person Re-identification*, WACV 2021.

— PATENTS & PUBLICATIONS

专利与学术成果

- 国家发明专利 18 项（授权 8 项，在审 10 项），涵盖图像生成优化、多模态模型架构设计与跨模态对齐算法 — 实习阶段授权 3 项，腾讯阶段授权 5 项，字节阶段在审 10 项
- 代表性论文：DreamPoster（arXiv 2025，共同一作）、Towards Precise Intra-camera Supervised Person Re-identification（WACV 2021）