中国科学技术大学 University of Science and Technology of China

# 王运韬



199-9326-3836 | wangyuntao@mail.ustc.edu.cn | **中共党员** | 个人主页

## 教育经历

**中国科学技术大学** **合肥** 计算机科学与技术（**博士**，2023.09–至今）

**中国科学技术大学** **合肥** 数学与应用数学（**本硕连读**，2017.09–2023.06，辅修 国际组织与项目管理 学士）

**荣誉奖项：** 唐仲英奖学金、优秀学生奖学金（铜奖）、一等学业奖学金、全国大学生数学竞赛（数学类）一等奖（2019）

**研究方向**：**扩散模型与流匹配、视觉语言模型、最优传输理论与应用、统一多模态模型、数值优化**

## 科研成果

**Hierarchical Cross-modal Representation Alignment for Text-to-image Diffusion Models.** **（一作）ICLR 2026 在投**

- 针对文生图中的属性错绑、目标遗漏与空间关系错误，提出**层级跨模态表征对齐算法 HIP**，在微调阶段将 LLM 语义信息蒸馏至生成过程，无需外部教师模型或区域级训练标注，不增加推理参数和计算开销。
- 设计共享 Poincaré 双曲空间，联合优化短语—区域 InfoNCE 对比损失与模态内整体—局部约束，在跨模态蒸馏的同时，保持生成侧表征的语义稳定，且增强整体场景中不同对象之间的区分度。
- 基于约 3.2 万对 Flickr30k-Entities 图文数据使用本方法进行轻量微调，在 Stable Diffusion 3.5 与 FLUX.1-dev 上均取得最优结果，组合生成准确率相对标准微调基线提升 **8.2%–14.5%**，且不影响生图质量。

**MIRROR: Aligning Semantic Relations from Language to Image via Gromov-Wasserstein.** **（共一）ECCV 2026**

- 指出多模态大模型 SFT 仅监督答案生成，缺少跨模态关系结构约束，导致空间、计数、细粒度关系理解不足；
- **提出 MIRROR 结构正则项并设计高效算法**，显式约束文本-视觉关系一致性，提升视觉 token 关系建模；
- 零额外推理开销，在 LLaVA-1.5/NeXT 7B/13B 上 GQA / BLINK / VQAv2 较常规 STF 全面提升，并保持通用能力。

**Finding Wasserstein Ball Center: Efficient Algorithm and The Applications in Fairness.** **（一作）ICML 2025**

- 提出 Wasserstein 球心模型，改善了多分布聚合传统方法中存在的公平性问题，并将其建模为大规模线性规划；
- 发现该问题的特殊矩阵结构，提出基于预测-矫正内点法的时空高效优化算法，并在真实图像数据验证效果和效率。

## 项目经历

**文生图扩散模型的跨模态层级表征对齐**

- ✧ 基于 PyTorch、Diffusers 与 Accelerate 搭建 LoRA 微调流程，实现混合精度、梯度累积及多卡训练。
- ✧ 对单/双流的不同 MMDiT 架构，实现 spaCy 短语抽取、T5 Token 跨度对齐、联合注意力提取与软区域特征池化，贯通文本短语到视觉区域的可微监督链路，并通过跨卡特征聚合构建对比学习负样本池。
- ✧ 搭建 GenEval、T2I-CompBench 与 COCO-10k 评测流程，完成组合生成、图像质量及模块消融评估。

**多模态大模型评测偏差与视觉依赖诊断**

- ✧ 发现 MLLM 内部表示的**“对齐幻觉”**：使用噪声替代视觉 token 后，任务准确率显著下降，但多种常用标量对齐指标如 CKA、MIR 等仍保持较高水平，说明内部表征几何相似不能证明跨模态交互的有效性。
- ✧ 揭示共享 MLP 权重诱导的对齐幻觉机制，并提出主角余弦间隙，可以诊断模型跨模态对齐的实现情况。

**Wasserstein 球心数值优化求解器**

- ✧ 基于 MATLAB **实现大规模线性规划求解器**，进行稀疏矩阵分解和牛顿迭代，并应用于点云、人脸和合成数据集。

**面向理论计算机问题求解的 Agent**

- ✧ 面向开放式数学问题的长期探索，设计由协调器、路线规划、结果构造及独立检查组成的多智能体架构，将研究方向判断与任务调度解耦，支持研究任务的分解、扩展与回溯。
- ✧ 设计“**探索树 + 逻辑依赖**”的研究状态管理机制，分别记录任务演化与证明引用，支持跨分支复用已检查结论，并追溯研究路线的决策依据。已应用于平面 Delaunay stretch factor 问题，取得更优近似上界。