

阳隆荣

✉ longrongyang@zju.edu.cn · 📞 (+86) 151-0835-1076 · 个人主页

🎓 教育背景

浙江大学	2022.09 – 至今
在读博士研究生，人工智能，预计 2026 年 3 月毕业	
电子科技大学	2019.09 – 2022.06
硕士，信息与通信工程	
电子科技大学	2015.09 – 2019.06
学士，电子信息工程	

⚙️ 研究方向

- 多模态 agent: 多模态统一 agent, 环境反馈的 agent RL
- 多模态大模型: prompt MoE (ECCV 2024), MoE (ICLR 2025), token 压缩 (CVPR 2025)
- 目标检测模型: 噪声标签学习 (ECCV 2020), 蒸馏算法 (ICCV 2023), 小样本学习 (T-IP 2024)

⚙️ 部分论文

多模态 agent

- Wenkang Han, Zhixiong Zeng, Jing Huang, Shu Jiang, Liming Zheng, **Longrong Yang**, Haibo Qiu, Chang Yao, Jingyuan Chen, Lin Ma.
“GUIRoboTron-Speech: Towards Automated GUI Agents Based on Speech Instructions.”
一个高性能的语音 GUI agent。

多模态大模型

- Longrong Yang**, Dong Shen, Chaoxiang Cai, Kaibing Chen, Fan Yang, Tingting Gao, Di Zhang, Xi Li.
“Libra-Merging: Importance-redundancy and Pruning-merging Trade-off for Acceleration Plug-in.”
IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR) 2025
一种 token 压缩技术。用 PTI 避免选择重复的重要 token, 再根据 token 特点动态使用裁剪/合并。
- Longrong Yang**, Dong Shen, Chaoxiang Cai, Fan Yang, Tingting Gao, Di Zhang, Xi Li.
“Solving Token Gradient Conflict in Mixture-of-Experts for Large Vision-Language Model.”
International Conference on Learning Representations (ICLR) 2025
MoE。用 token 级梯度建模 token 间关系, 根据 token 间关系优化 MoE 路由。
- Longrong Yang**, Hanbin Zhao, Yunlong Yu, Xiaodong Zeng, Xi Li.
“RCS-Prompt: Learning Prompt to Rearrange Class Space for Prompt-based Continual Learning.”
European Conference on Computer Vision (ECCV) 2024
Prompt MoE。建模任务间的样本关系, 增大新旧任务空间的 margin, 减少不同任务的冲突。

目标检测模型

- Longrong Yang**, Hanbin Zhao, Hongliang Li, Liang Qiao, Ziwei Yang, Xi Li.
“GCSTG: Generating Class-confusion-aware Samples with a Tree-structure Graph for FSOD.”
IEEE Transactions on Image Processing (T-IP) 2024
用对抗样本生成技术生成困难的新类样本, 让模型从生成的困难样本中学习怎么识别新类。
- Longrong Yang**, Xianpan Zhou, Xuwei Li, Liang Qiao, Zheyang Li, Ziwei Yang, Gaoang Wang, Xi Li.
“Bridging Cross-task Protocol Inconsistency for Distillation in Dense Object Detection.”
International Conference on Computer Vision (ICCV) 2023
对齐分类和蒸馏损失中使用的分类得分, 将教师网络的知识充分传递给学生网络。

- **Longrong Yang**, Hongliang Li, Qingbo Wu, Fanman Meng, Heqian Qiu, Linfeng Xu.
“Bias-correction Feature Learner for Semi-supervised Instance Segmentation.”
IEEE Transactions on Multimedia (T-MM) 2022
- **Longrong Yang**, Hongliang Li, Fanman Meng, Qingbo Wu, King Ngi Ngan.
“Task-Specific Loss for Robust Instance Segmentation with Noisy Class Labels.”
IEEE Transactions on Circuits Systems for Video Technology (T-CSVT) 2021
- **Longrong Yang**, Fanman Meng, Hongliang Li, Qingbo Wu, Qishang Cheng.
“Learning with Noisy Class Labels for Instance Segmentation.”
European Conference on Computer Vision (ECCV) 2020

🧑 实习经历

美团，北京 2025.05.20 – 2025.09.01

计算和智能平台部 / M17 美团大模型北斗实习生，mentor: 马林

- 多模态统一 agent: 将多模理解、GUI agent 和 robotic agent 统一到一个模型中，扩大模型的 action space，用 LoRA-MoE 进一步研究不同 agent 任务之间的关系。
- 环境反馈的 agent RL: 让 GUI agent 在环境中执行 action，得到环境反馈。在环境反馈的引导下，设计 RL 算法激发 GUI agent 搜索 action space，实现特定功能。

快手，杭州 2024.03.01 – 2024.12.01

大模型与多媒体技术部 / 多模态理解与应用组

- Token 压缩: 多模态大模型中的视觉 token 冗余，可以裁剪/合并视觉 token，减少 token 数量。我们提出一种新的视觉 token 合并方案，兼顾 token 的重要性/独特性，以及裁剪/合并技术路线的优势，性能几乎无损的情况下降低约 70% Flops。成果已发表于 CVPR 2025。
- MoE: 多模态大模型需要处理多种任务，任务之间会相互促进或冲突，当相互促进的任务过同一专家时，对学习有利，反之有害。我们提出用 token 级梯度判断 token 之间是否冲突，然后用新损失优化 token 路由以消除冲突，消除冲突后模型有置信的性能提升。成果已发表于 ICLR 2025。

♡ 获奖情况

中国电子学会优秀硕士学位论文:《有限标签场景下的对象解析研究》	2022 年
四川省优秀毕业生	2022 年
国家奖学金	2021 年
遥感图像稀疏表征与智能分析竞赛全国一等奖	2019 年

i 学术活动

- CVPR 审稿人, 2022, 2023, 2024, 2025
- ICCV 审稿人, 2023, 2025
- ECCV 审稿人, 2022, 2024
- ICLR 审稿人, 2024, 2025
- NeurIPS 审稿人, 2023, 2024, 2025
- ICML 审稿人, 2023, 2024, 2025

i 一些想法

我目前对“环境反馈如何影响模型的 action”比较感兴趣。在未来，agents 将突破人类生成数据的局限，越来越多地从自身与世界的交互中学习。Agents 将通过统一的 action space，自主与环境互动，它们会在贯穿一生的经验流中持续适应，制定以自身行动对环境造成的后果为依据的计划。最终，这种范式转变将帮助模型在众多领域解锁超越人类能力的新技能。