

李慧

+86 18844121033 | crystallee0418@gmail.com



教育经历

复旦大学	2024年06月 - 2028年06月
计算机科学与技术 博士 计算机科学技术学院	
吉林大学	2020年09月 - 2023年06月
数学 硕士 人工智能学院	
吉林大学	2016年09月 - 2020年06月
信息与计算科学 本科 数学学院	

研究经历

以人物为中心的视频生成 2024年03月 - 2025年05月

- OpenHumanVid**: 发布了首个大规模 (13.2M clips, 16.7K hours) 以人物为中心视频数据集, 提出了高质量的人体运动数据筛选管线, 并在Video diffusion model 上验证数据集有效性, 目前数据集已收到600+国内外大厂、创业公司、高校的申请, 该工作发表在CVPR 2025 (一作), 并选为highlight (top 5%)。
- Hallo**: 提出了基于扩散生成模型的音频驱动人脸视频生成网络架构, 使得能够控制音频与唇部、表情和姿态之间的映射。并完全开源了项目训练、推理和推理权重, 目前github repo已获**8.4K stars** (共同一作, 提交了两万行代码), github人脸视频生成领域开源项目排名前3。
- Hallo2**: 提出了基于patch drop来解决长时间视频误差累积的策略, 可以生成时长为1小时的音频驱动人脸视频。目前github repo获**4.3K stars**, 该工作发表在ICLR 2025 (共同一作)。
- Hallo3**: 提出了基于pretrained video diffusion model来解决talking视频的泛化性问题(生动的背景, 自然的肢体动作等), 目前github repo获**1.3K stars**, 该工作发表在CVPR 2025 (二作, 主要负责数据的筛选)。

图像生成基础技术 2025年05月 - 2025年11月

- Pyramidal patchification**: 提出了在 diffusion / flow-base model的noise schedule中, 使用不同大小的patchification的策略, 例如, 在高noise-level时, 使用4x4的patchification layer, 在低noise-level时, 使用2x2的patchification layer。可以实现推理提速1.6-2.0倍的情况下, 生成质量与baseline持平, 或更好。该工作发表在ICLR 2026 (一作)。
- MixFlow**: 发现了slowflow现象, 并通过mixflow来缓解diffusion / flow-base model中的train-test discrepancy, 在目前主流的SiT, REPA, RAE模型架构上均可以提高生成质量。在SDE few-step采样设置上, FID可以提升200 - 300%。该工作发表在CVPR2026 (一作)。

科研论文 (一作及共同一作)

Hui Li*, Mingwang Xu*, Yun Zhan, Shan Mu, Jiaye Li, Kaihui Cheng, Yuxuan Chen, Tan Chen, Mao Ye, Jingdong Wang, Siyu Zhu. OpenHumanVid: A Large-Scale High-Quality Dataset for Enhancing Human-Centric Video Generation. CVPR 2025 highlight.

Hui Li, Baoyou Chen, Liwei Zhang, Jiaye Li, Jingdong Wang, and Siyu Zhu, Pyramidal Patchification Flow for Visual Generation. ICLR 2026.

Hui Li, Jiayue Lyu, Fuyun Wang, Kaihui Cheng, Siyu Zhu and Jingdong Wang. MixFlow Training: Alleviating Exposure Bias with Slowed Interpolation Mixture. CVPR 2026.

Jiahao Cui*, **Hui Li***, Yao Yao, Hao Zhu, Hanlin Shang, Kaihui Cheng, Hang Zhou, Siyu Zhu, Jingdong Wang. Hallo2: Long-duration and high-resolution audio-driven portrait image animation. ICLR 2025.

Mingwang Xu*, **Hui Li***, Qingkun Su*, Hanlin Shang, Liwei Zhang, Ce Liu, Jingdong Wang, Yao Yao and Siyu Zhu. hallo: hierarchical audio-driven visual synthesis for portrait image animation. In arXiv: 2406.08801. 2024.

Hui Li, Zihao Li, Rui Ma, Tieru Wu. FD-CAM: Improving Faithfulness and Discriminability of Visual Explanation for CNNs. In International Conference on Pattern Recognition (ICPR). 2022. Oral.

工作经历

之江实验室 2023年07月 - 2023年12月

算法工程师

ControlGIF: 图生视频, 视频编辑。基于Animatediff和Controlnet, 在WebVid-10M数据集子集上训练Controlnet, 以驱动一张给定的图像或编辑视频, 模型在Github及HuggingFace平台上赢得了开源社群的广泛赞誉, 成功融入Charlex Inc.等动画制作企业的生产流程中。