

工作经历

2021.01 旷视研究院 AIC 组 | 算法研究员(实习) | Mentor: <u>周舒畅</u>至今 全职实习 26 个月,并参与多项研究工作,研究领域含: 视频插帧,神经 渲染,CLIP,3D 人体,Talking Head,和强化学习自动驾驶决策等。积 累了模型训练和论文写作的经验。

2021.03~ 北航×旷视 深度学习课程 | 开发助教 | 教师: 王静远

2021.06 编写基于 MegEngine 框架的 <u>代码和 tutorial</u>,包括模型蒸馏,对抗样本生成,白盒攻击,ISP,强化学习 (基于 DQN 的 2048AI,数独 AI,多智能体 AI 等)。 线上课程

项目经历

2022.11 - 参加正在进行中的自动驾驶决策比赛,负责使用强化学习 PPO 等算法生成决策信息,其中包括换道决策,最远距离决策等,实现车辆前进,避障,驶向终点等功能。

2022.06 旷视 2022 MegTech Open Day 展览:展出手绘图片转动画的 demo,该 demo 接收多张单人物的手绘图像,和一个动作序列,生成该人物的动画视频。

负责 demo 的制作和现场展示。GitHub Repo

2021.12 使用 MegEngine 复现插帧 RIFE 算法,并优化 StyleCLIP / HyperStyle 等文字驱动的人脸编辑算法。制作文字驱动的人脸编辑 pipeline,并上线微信小程序 MegLab。其中,RIFE 算法被用于将编辑前后的人脸进行融合,以增加编辑前后人脸的相似程度。

复现 RIFE 的 GitHub Repo 人脸编辑的 GitHub Repo: Oh-My-Face

研究经历

Detachable Intermediate Flow Estimation for Space-Time Video Super-Resolution (Submitted to CVPR2023)

Zhewei Huang, Ailin Huang, Xiaotao Hu, Shuchang Zhou

使用光流对视频的时间和空间进行扩展,即视频插帧和超分辨率。

我们提出了一个简单而有效的结构,并消融单个组件以评估其有效性。我们专注于构建基于光流的高级特征传播的关键组件,并在几个公共指标中实现了 SOTA。

2022 A Dynamic Multi-scale Voxel Flow Network for Video Prediction

(Submitted to CVPR2023)

Xiaotao Hu, Zhewei Huang, Ailin Huang, Shuchang Zhou

使用光流进行视频预测。我们提出了一个多尺度 Voxel-Flow 块,使我们的模型能够处理不同尺度的运动,以提高模型泛化性。 实验证明,在推理时,DMVFN 可以根据输入自适应地选择合适的子网络,加快推理时间并保持模型性能。

Perceptual Conversational Head Generation with Regularized Driver and Enhanced Renderer

* Paper | GitHub

* Equal Contribution

* Equal Contribution

ACM Multimedia* 2022 Conversational Head Generation Challenge 冠军,**论文已接收**。 * CCF-A 类会议

该比赛方案训练一个通用的音频-行为驱动,模拟一个高质量的渲染器,将行为表示转化为视频。

我们对 baseline 进行了精细的调整,并使用一个前后景融合模块将渲染结果进行后处理,探索多种集成学习方法。负责绝大部分实验工作和部分论文撰写。.

荣誉奖励

2022 ACM MM2022 Conversational Head Generation Challenge | <u>Track1</u> 银牌 <u>Track2</u> 金牌

2020 - 2021 ICPC2020 南京赛区银牌, ICPC2021 银川赛区银牌, CCPC2020 威海赛区银牌

2020 旷视人工智能开源大赛 初赛<u>第一名</u>(未入职时参加)

2018 高中信息学奥林匹克竞赛(NOIP)省级一等奖

其他

Skill C++/C, Python3, JavaScript | PyTorch, MegEngine, Qt

Part-time job ACM-ICPC 2021 昆明站出题人,算法竞赛教学 100+ 课时