

- 1) 以 ICON Matting[1]抠图模型作为基准, 评估其在 ICM-57 数据集上的性能, 以该论文使用的几种指标作为性能评价标准 (如计算资源有限, 可适当降低图像分辨率评估模型);
- 2) 拓展 ICM-57 数据集, 考虑更多的类别, 更多的场景, 完成数据的标注。
- 3) 查找抠图相关文献, 改进 Icon Matting 模型, 改进样例上下文抠图的精度。(提示: i) 可将 alpha matte 分成确定区域 (0, 1) 和不确定区域 (0 和 1 之间的非整数), 先分类两类区域, 再预测不确定区域的 alpha 值; ii) 扩展训练数据的规模和多样性)
- 4) 寻找一个 Icon Matting 在现实中的应用场景, 制作视频演示 demo。

数据集描述:

Icon Matting 模型的训练采用了 RM1K 图像抠图数据集和部分 OpenImages 图像分割数据集。ICM-57 数据集是一个 in-context matting 的测试集, 包含 57 组同实例或同类前景的图像。

RM1K 数据集下载: 链接: <https://pan.baidu.com/s/19gACWY9JOSpY9CRkG42JLQ>

提取码: BA1c

ICM-57 数据集下载: 见[1]中的下载链接

参考文献:

[1] <https://tiny-smart.github.io/icm.github.io/>

[2] <https://github.com/wchstrife/Awesome-Image-Matting>

指导老师: 陆昊, e-mail: hlu@hust.edu.cn, QQ: 370875068

5.10 基于光流的密集人群运动分析

任务描述:

基于视频监控的人群运动分析是公共安全管控中的一个重要环节, 特别对于机场、火车站、地铁站等人流密集的场所尤为重要。解决该任务的一个核心在于光流场的精准提取和利用。光流 (Optical Flow) 是计算机视觉运动检测的重要基础, 其实现方法已由传统算法演变为由深度学习占主导的光流网络。给定输入序列, 经过模型处理后得到的像素运动场即描述了所谓的光流信息。本任务面向密集人群, 首先要求在公开数据集上评测当前多种光流模型, 并尝试分析密集人群场景下影响光流精度的难点所在; 基于评测结果, 基于一个合适的光流模型, 在具体地铁场景下完成密集人群的运动分析。



人群运动分析结果 (左) 与光流场可视化 (右)

题目要求:

- 1) 在 TUB CrowdFlow[1]数据集上对比例如 RAFT[2]、FlowNet2.0[3]等至少 (#小组人数) 的主流光流模型, 并在该数据集上报告 EPE、AE、IE 等光流评价指标的对比结果;
- 2) 基于一个合适的光流模型, 在 WuhanMetro 数据集上进行人群运动分析, 实现:
 - i) 人群运动方向的识别, 要求获取不同人群运动方向的光流向量, 输出不同人群运动方向的数量并按上图的形式区分显示, 以 MAE 和 MSE 为评测指标;
 - ii) 人群运动区域分割, 获取不同运动方向人群的分割图并评估分割区域的 Pixel Accuracy 和 mIOU 等指标;
- 3) 集成所选算法, 在 WuhanMetro 数据集选取典型场景完成 demo 的制作, 输出光流、人群运动场、人群分割图等处理结果。

数据描述:

TUB CrowdFlow 是专门面向密集人群运动分析而渲染的数据集, 包含 5 个场景 10 个序列, 每个场景渲染两次: 用静态视角和动态摄像机模拟基于无人机/无人机的监视。场景渲染使用虚幻引擎在高清分辨率(1280x720)以 25 fps 实现, 总帧数为 3200。下载链接: <https://hidrive.ionos.com/lnk/LUiCHfYG>

WuhanMetro 是从武汉地铁站采集的真实数据, 包含了不同监控视频视角下丰富的人群运动场景。本数据集具有良好的运动方向区分度, 适合使用光流算法进行处理, 该数据集上的定量评测需要实现部分数据标注工作。本题目的最终演示测试将基于该数据, 请

联系指导老师获取。

参考文献:

[1] <https://github.com/tsenst/CrowdFlow>

[2] Teed Z, Deng J. Raft: Recurrent all-pairs field transforms for optical flow, ECCV: 402-419,2020.

[3] Ilg, Eddy, et al. "FlowNet 2.0: Evolution of optical flow estimation with deep networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[4] Liu, Weizhe, Mathieu Salzmann, and Pascal Fua. "Counting people by estimating people flows." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.11 (2021): 8151-8166.

[5] Li, Xuelong, Mulin Chen, and Qi Wang. "Quantifying and detecting collective motion in crowd scenes." *IEEE Transactions on Image Processing* 29 (2020): 5571-5583.

指导老师: 陆昊, e-mail: hlu@hust.edu.cn, QQ: 370875068

5.11 时空一致的视频行人计数

任务描述:

本任务的主要目的是针对地铁视频监控场景中的行人数量进行统计。2024 年 1 月 1 日凌晨, 武汉江汉路地铁站发生了严重拥堵, 造成了重大的公共安全隐患。在公共交通日益繁忙的今天, 交通部门对于行人实时流量的监控变得尤为重要。在视觉中, 研究人员已提出了大量基于图像的人群计数与定位算法[1][2]。然而, 在视频流中逐帧应用基于图像的行人计数易产生帧间预测结果的剧烈抖动, 造成时空上的不一致。造成这类问题的一个重要原因在于未有效建模人群的运动信息。此外, 由于公共数据集中人群背面的样本较少, 易导致直接应用现有预训练计数模型时易出现欠估计现象。

