

徘徊检测赛题冠军方案 | PRCV2022计算机视觉算法应用技术挑战赛

原创

CV开发者都爱看的

极市平台

2023-03-02 22:00:58

发表于广东

手机阅读

𐄞

↑ 点击蓝字 关注极市平台



作者 | 凝升

编辑 | 极市平台

极市导读

本文为PRCV2022比赛中获得徘徊检测赛题的冠军方案分享，从赛题分析道模型选取以及比赛中问题的解决都有详细的介绍。 >>加入极市CV技术交流群，走在计算机视觉的最前沿

一、介绍

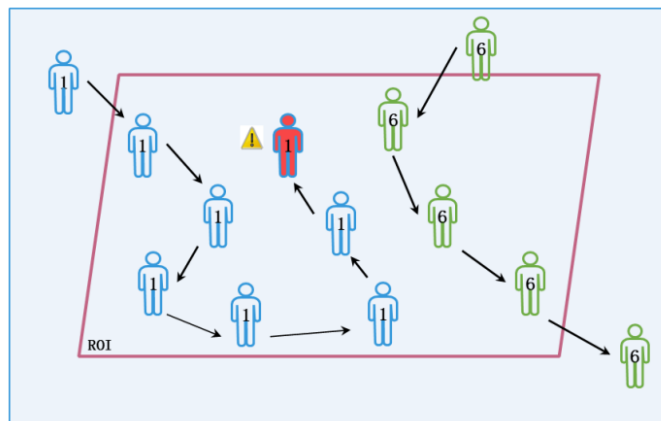
作为一位计算机视觉方向的在读专业型硕士，一直在寻找各种机会累积技术，提升自己。徘徊检测该赛题作为技术应用类赛题之一被选入PRCV2022计算机视觉算法应用技术挑战赛，本人独自组队参与了该比赛，在本导师的指导下，非常幸运能够获得本次比赛的冠军，且将解决方案进行了封装与推送以及OpenVINO计算架构的移植。

二、赛题分析

任务描述

本赛题的任务是对视频流中的人员进行跟踪，记录ROI区域人员的位置信息，识别出其超过时间阈值的徘徊行为，给出报警信号。

如图片中ID为1的人员，其在ROI中滞留时间超过了给定的时间阈值，需要对其进行报警处理。



图一. 徘徊超时行为预警

赛题难点

- 本赛题的不仅评估模型的精确程度，更注重其实际场景下的报警准确度，因此既要优化模型对目标的追踪效果，又要优化报警逻辑以包容模型应用缺陷；
- 标注数据类型单一，虽然是人员追踪的任务，但只给了人员目标检测的标注数据，不同图片中不同人员的匹配信息并没有给出标注，因此模型方案的选择和后续优化是存在瓶颈的；
- 性能分占比不小，对模型推理速度要求较高，赛题的评分计算规则如下：

$$Score = 0.8 \times M + 0.2 \times \text{性能分}$$

M ：衡量算法报警精度

$$M = 1 - \frac{(gt - pred)^2}{gt^2}$$

gt ：真实徘徊人数

$pred$ ：预测徘徊人数

性能分：衡量算法速度

性能分 = FPS/100，若FPS>100，则当满分计算

数据情况

训练数据：1万张左右，标注格式为目标检测，标注类型为VOC，标签类型为bounding box；



图二. 数据示例

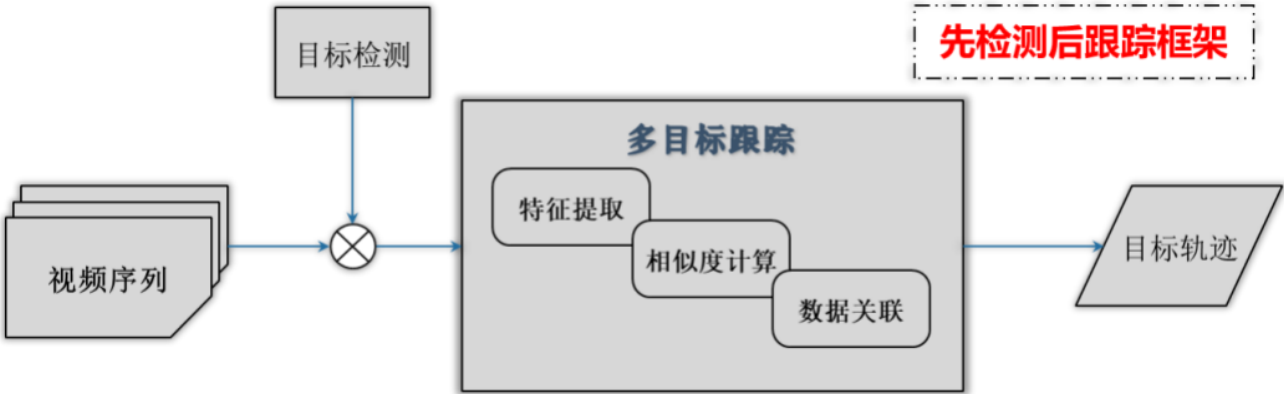
平台给定的训练用硬件配置为CPU4核，Mem12G，Tesla T4*1。

三、赛题思路

1.模型选择、数据处理方式、训练技巧等模型训练测试过程中的整体流程方法。

任务分析

本赛题可以归结为多目标追踪(MOT, Multiple Object Tracking)问题，由于只提供了目标检测的训练数据，因此只能采用MOT方法中的先检测后跟踪的处理框架，因此该问题应该分为两个部分，分解为：人员目标检测 + 多目标追踪



图三. 先检测后跟踪框架

该检测框架，需要先定位需要跟踪的目标，然后匹配不同帧之间相同的目标对象。

模型选择

拿到赛题后，首先进行baseline的构建，最先选定的方案是Yolov5s和DeepSort的组合版本。

Yolov5在工程化方面对应用部署的支持很灵活，非常适合对性能和精度都有要求的场景。其提供多个不同体积的预训练模型，可根据实际需要进行选择。

追踪算法DeepSort基于Sort改进而来，主要模块有：

- 目标检测模块：通过目标检测网络，获取输入每一帧图片中的目标框
- 轨迹跟踪模块：通过卡尔曼滤波进行轨迹预测和更新，获取新的轨迹集合
- 数据匹配模块：通过级联匹配和IOU匹配将轨迹和目标框关联

DeepSort主要流程：检测器获取视频当前帧中目标框 → 卡尔曼滤波根据当前帧的轨迹集合预测下一帧轨迹集合 → 预测轨迹与下一帧检测目标框进行 → 卡尔曼滤波更新匹配成功的轨迹

最初的成绩情况是：

- 数据处理：所有数据既作为训练集也作为验证集
- 目标检测：Yolov5，采用网络体积最小的yolov5s权重
- MOT算法：DeepSort，使用ImageNet 预训练的ReID权重
- 线上提交成绩：
 - MOTA: 0.7588
 - 性能分: 11.54
 - 成绩分: 0.68
- 总结：此时未采用任何针对化改进措施，榜单排名几近垫底。

在看到成绩后，认真梳理了几个改进的方向，首先是从数据集上入手进行分析，在统计训练数据情况时，有发现数据难点情况可归纳如下：

- 场景画面和人员目标角度多样化；
- 画面存在高曝光和低亮度的情况；
- 存在严重的遮挡，且有多目标聚集遮挡；
- 目标尺寸跨度较大，单目标单图片面积占比有5%~50%；
- 存在多个分辨率尺寸的数据：

分辨率	图片数量	比例
1920*1080	4720	53%
406*720	1337	15%
1280*720	1014	12%
2704*2028	892	10%
2560*1440	880	10%

在数据处理阶段，有经历过以下节点情况：

- 尝试划分一个更加合理验证集，使得验证集分布接近测试集，引导模型往推理场景靠近。
- 低分辨率和小目标常常阻碍目标检测精度的提升，针对以上两个问题结合数据集的曝光过暗等特点进行针对性的解决。
- 训练轮次增大导致成绩降低说明模型过拟合，同时，验证集与测试集分布存在差异。
- 训练轮次增大，成绩较为稳定，说明当前权重可能无法学习到更多的特征以提高精度。

最终的数据处理情况，可以总结如下：



在对数据处理手段尝试过后，根据任务本身难点对模型方案进行了针对性的升级与优化，期间换过许多不同版本的MOT模型，最后选定的是上限空间巨大的Yolov5+ByteTrack方案。

ByteTrack算法是一种利用低分检测框和跟踪轨迹之间相似性的追踪算法，从低分检测结果中去背景，挖掘出正确的物体。选定该方案的理由为：

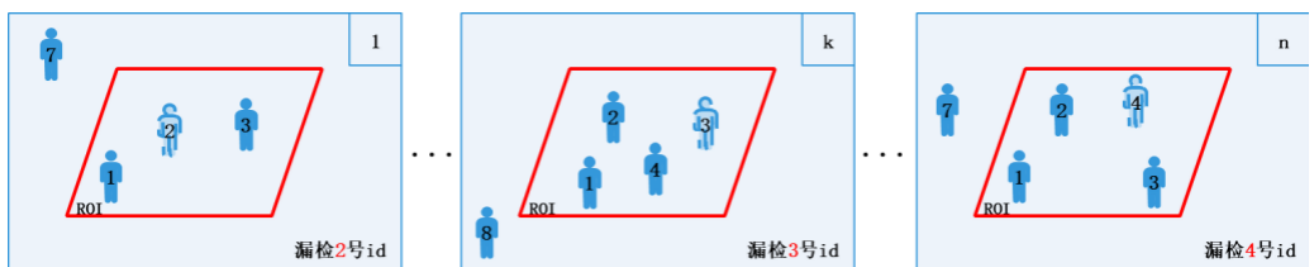
- 该算法对低分检测框充分利用，可以弥补检测模型的不足。
- 分析样例数据的特点发现，存在人员重叠与遮挡非常严重的情况。
- 该算法是纯运动模型，并没有使用ReID特征来计算外观相似度，适合对速度有要求的场景。

之后便是模型精度与速度的trade-off和细节优化以及调参工作，包括优化器的选择以及数据增强中参数的选定等等。为了尽可能地利用低分检测框，对于NMS的筛选逻辑进行优化，使得有更多的低分检测框能够被MOT模型利用。

算法逻辑优化

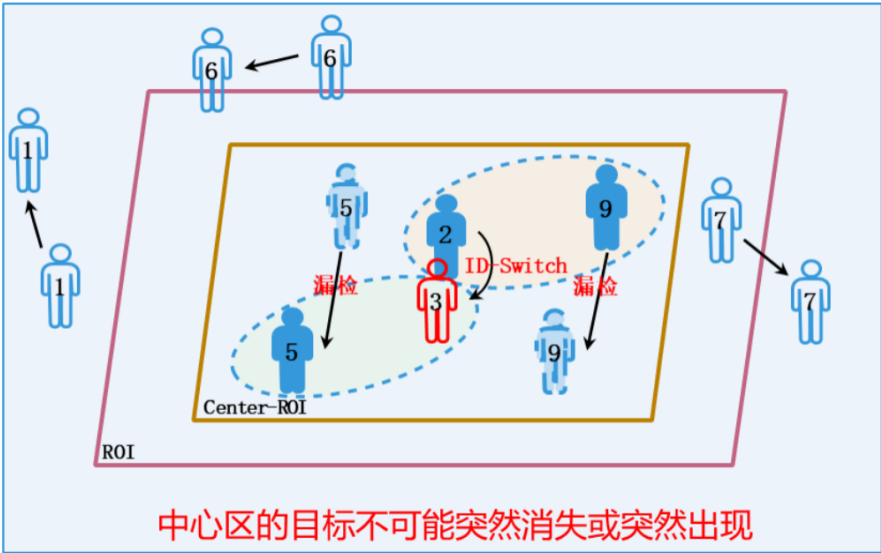
针对追踪任务本身，归纳出两个模型应用缺陷：

- 在需要报警时间段的开头、中间、结尾都可能发生目标漏检
- 随时都可能发生ID-Switch的问题

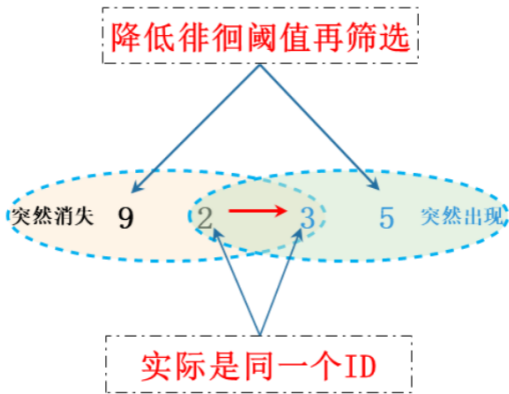


对于这两个问题，进行针对性的优化，

- 目标漏检：起止帧计算时长，出现徘徊后降低徘徊阈值再筛选。
- ID-Switch：以IOU匹配ROI中心区突然消失和突然出现的目标。同时，圈定了一个中心ROI区域，对中心ROI区域的目标信息进行了重点关注，以解决漏检和ID切换的问题。

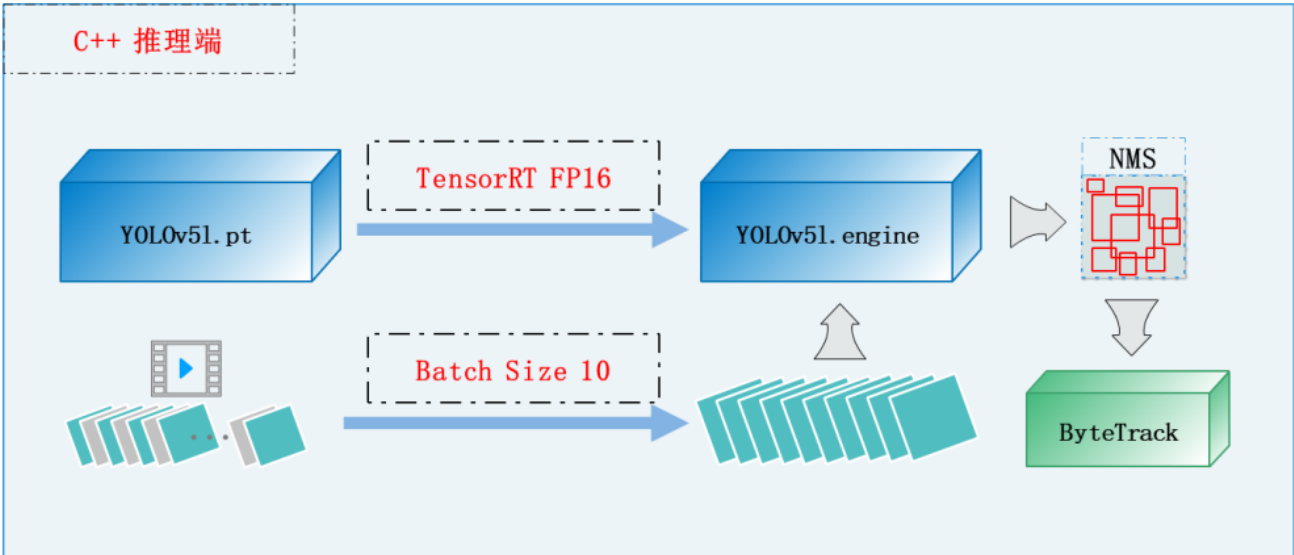


中心区，突然出现的目标ID队列和突然消失的ID队列实际上会出现重合的情况，突然消失的除了漏检还可能是ID-Switch中消失的目标，而突然出现的除了之前漏检的还可能是ID-Switch中突然出现的目标。



速度优化

从模型加速和隔帧检测两方面缩短推理时间



最终成绩为榜单榜首，获得该比赛的冠军，并将技术方案进行了封装推送。

四、总结

本次竞赛经验可以总结为以下几点：

- 从数据集特征分析入手，合理选择数据增强和设计Baseline
- 分模块应用工程技巧对模型方案进行迭代提升和优化
- 合理权衡模型精度和推理速度之间的关系
- 设计具有模型算法缺陷包容度的应用逻辑，并合理设定超参数

五、参考文献

1. <https://github.com/ultralytics/yolov5>
2. ByteTrack: Multi-Object Tracking by Associating Every Detection Box
3. <https://github.com/NVIDIA/trt-samples-for-hackathon-cn/blob/master/cookbook/01-SimpleDemo/TensorRT8>
4. https://github.com/ExtremeMart/ev_sdk_demo4.0_pedestrian_intrusion
5. Learning Efficient Convolutional Networks through Network Slimming

公众号后台回复“**极市直播**”获取**100+**期极市技术直播回放+PPT**极市平台**

为计算机视觉开发者提供全流程算法开发训练平台，以及大咖技术分享、社区交流、竞...
848篇原创内容

公众号

极市干货

技术干货：损失函数技术总结及Pytorch使用示例 | 深度学习有哪些trick？ | 目标检测正负样本区分策略和平衡策略总结

实操教程：GPU多卡并行训练总结（以pytorch为例） | CUDA WarpReduce 学习笔记 | 卷积神经网络压缩方法总结

极市CVPR2023交流群已成立

△长按扫码进群

添加极市小助手微信 (ID : cvmart4)

备注: 姓名-学校/公司-研究方向-城市 (如: 小极-北大-目标检测-深圳)

即可申请加入极市目标检测/图像分割/工业检测/人脸/医学影像/3D/SLAM/自动驾驶/超分辨率/姿态估计/ReID/GAN/图像增强/OCR/视频理解等技术交流群

每月大咖直播分享、真实项目需求对接、求职内推、算法竞赛、干货资讯汇总、与 10000+来自港科大、北大、清华、中科院、CMU、腾讯、百度等名校名企视觉开发者互动交流~



极市平台

为计算机视觉开发者提供全流程算法开发训练平台，以及大咖技术分享、社区交...

848篇原创内容

公众号

△点击卡片关注极市平台，获取最新CV干货

点击阅读原文进入CV社区

收获更多技术干货

阅读原文

喜欢此内容的人还喜欢

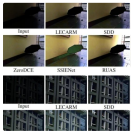
YOLOv5帮助母猪产仔？南京农业大学研发母猪产仔检测模型并部署到Jetson Nano开发板

极市平台



ICCV23 | 将隐式神经表征用于低光增强，北大张健团队提出NeRC

极市平台



ICCV 2023 | 南开程明明团队提出适用于SR任务的新颖注意力机制（已开源）

极市平台

