

ECV 2021 冠军方案解读：渣土车识别方案

原创

CV开发者都爱看的

极市平台

2021-11-11 22:00:00

手机阅读

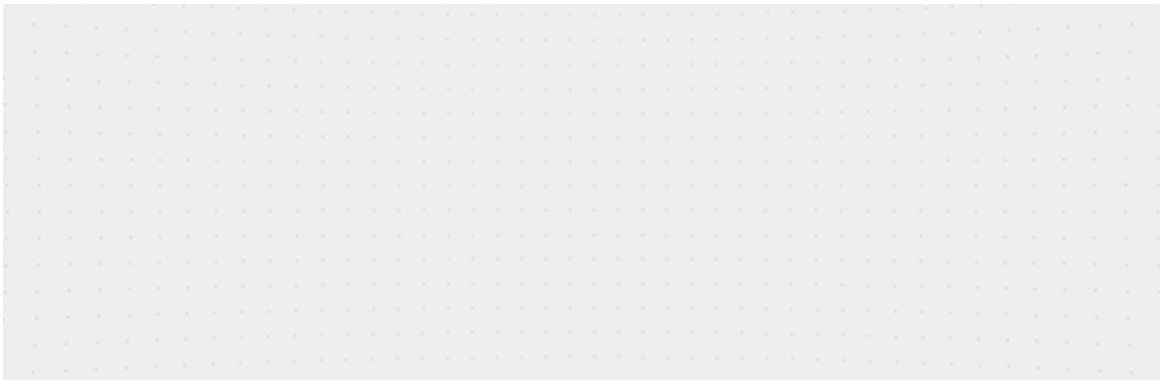
囧

收录于话题

#计算机视觉

5个

↑ 点击蓝字 关注极市平台



作者 | 胡可飞团队

编辑 | 极市平台

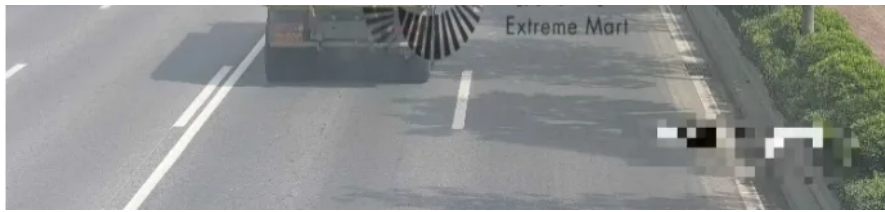
极市导读

本文为获得渣土车识别冠军的胡可飞团队方案解读，团队选用了yolov5s网络来作为此次比赛任务的检测模型，最后优化的模型推理速度从9FPS最高到52FPS，已具备CPU上实时推理的能力。>>加入极市CV技术交流群，走在计算机视觉的最前沿

一、赛题回顾

本赛题希望基于计算机视觉技术可以实时监测到“黑车”，能及时、安全地管理渣土车，实现智能化管理，保证无“黑车”上岗，为创造智慧工地提供大力保障。 赛题目标是使用计算机视觉算法对图片中的渣土车的车牌进行识别，并且输出车牌的文字内容。观察样例数据推测图片主要采集自交通监控摄像头。



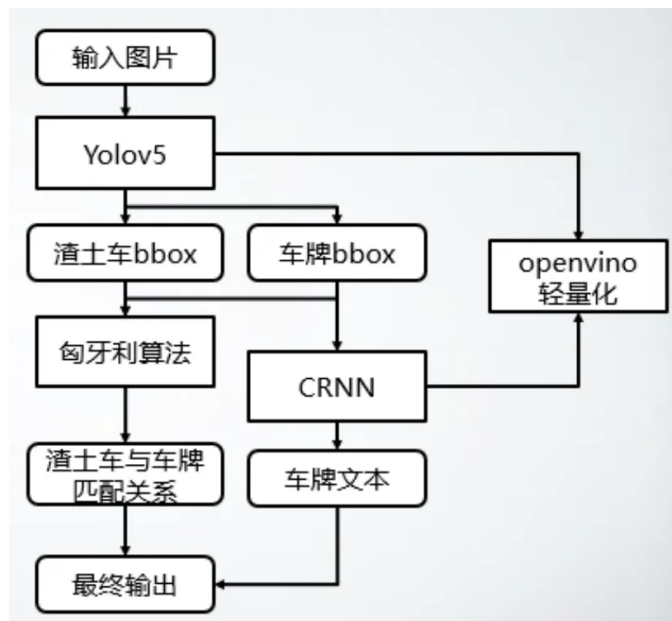


比赛的评价指标由F1 score和性能分FPS组成，这要求算法设计要兼顾精度和速度，算法的推理性能要具备实时性（>30FPS）。本题主要难点如下：

- 采集自交通摄像头的图片存在分辨率低，对比度低，环境光污染等问题。
- 部分渣土车严重脏污导致车牌难以分辨。
- 人工标注存在错标（车牌中不应出现的字符），漏标，标准不一等问题（如遮挡严重的车牌）。
- 实时性要求高。

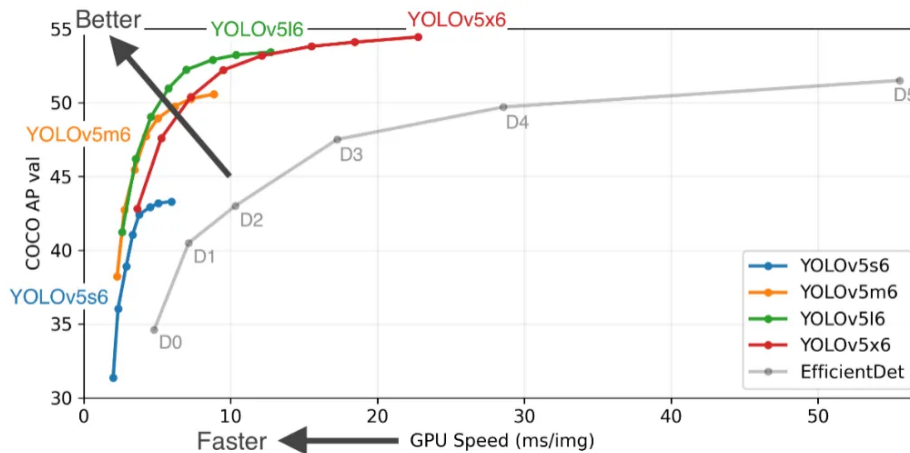
二、算法设计

主要算法设计如下图所示：

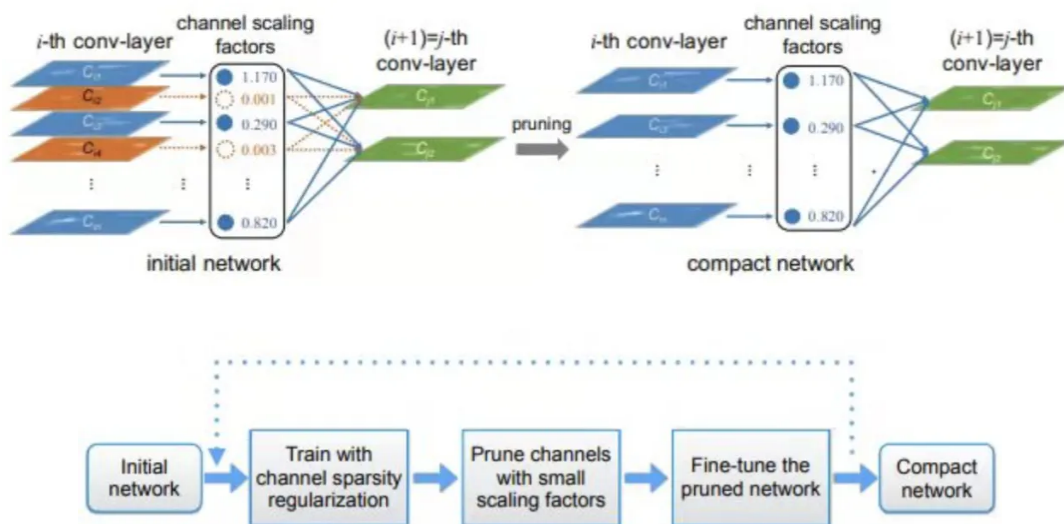


1. 使用比赛数据标注训练检测模型，同时检测渣土车目标和车牌目标
2. 将车牌部分抠图训练OCR识别模型
3. 使用匈牙利算法对渣土车和车牌目标进行匹配
4. 对算法进行轻量化等优化

其中检测模型我们选用了yolov5s网络，该网络具有较好的精度和速度的trade-off，常用于落地部署项目中。该检测网络直接使用数据标注的渣土车和车牌类进行训练，同时检出两类目标。最终输出中并不需要输出车牌框bbox，所以此处输出的车牌bbox仅用于提取车牌内容。



由于最终推理设备是cpu，所以对模型的轻量化要求较高，未经优化的yolov5s的推理速度越10FPS+，很难达到实时推理，所以我们对yolov5s网络进行剪枝，方法参考了slim-yolov3，通过bn层缩放系数，对模型各层的通道数进行结构化剪枝。剪枝的原理和操作流程如下：

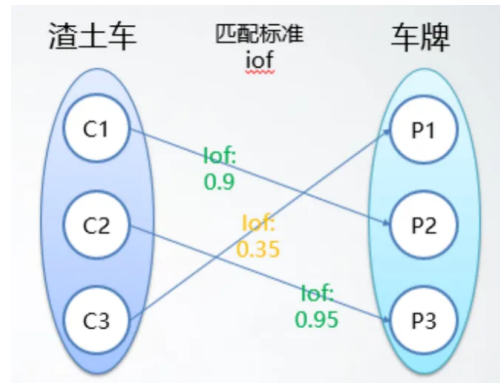


1. 稀疏化训练yolov5s
2. 由bn层scaling factor对通道数进行排序
3. 生成8倍数通道剪枝后模型结构config
4. 使用剪枝后模型重新训练

由于检测网络同时输出渣土车和车牌的bbox，丢失了两者的关联关系，所以我们使用匈牙利算法对两者进行匹配，获得每个渣土车目标对应的车牌bbox。匹配的cost使用的是IoF，即Intersecti

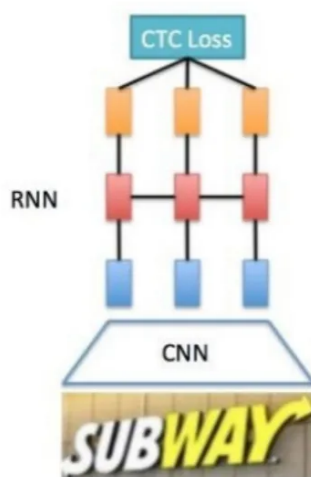
on over foreground，车牌即为foreground。理论上，车牌对于其所在的渣土车的IoF应该为1，实际匹配中我们设置了一个更宽松的阈值0.8。主要处理流程如下：

- Conf过滤+NMS去除误检框
- 计算渣土车目标和车牌目标的iof代价矩阵
- 使用匈牙利算法得到与渣土车匹配的车牌
- 删除低质量匹配 ($\text{iof} < 0.8$)



OCR部分我们使用了经典的CRNN网络，backbone等组成部分都选用了轻量化网络结构：

- MobilenetV3
- RNN
- CTC head



因为检测网络已经输出了车牌bbox，且车牌bbox基本等于车牌的文字区域，考虑到算法整体的实时性要求，我们未使用专门的文字区域检测网络，而是将检测网络输出的车牌bbox直接抠图送入OCR网络进行推理。OCR的训练数据来自于原数据中使用车牌bbox抠图得到的车牌图片patch，数据的准备流程如下：

1. Bbox jitter：对原来的bbox标注添加随机抖动，模拟检测网络的回归误差，提高泛化性，防止OCR网络过拟合。
2. Data Cleaning：根据我国车牌的文字排列规则，对标注错误的数据进行剔除等。
3. Data augment：通用的OCR数据增强，如旋转，HSV，模糊等。
4. Weighted batch sample：由于车牌文字分布不均，在数据集采样时对出现频率低的字符提高权重。

整体算法流程即上文所述方法，为了提高推理速度，我们的轻量化策略主要如下：

- Openvino转换
- 减小输入图片尺寸
- 检测模型yolov5剪枝
- 减小CRNN中间层通道数
- 前后处理性能优化



优化后模型的推理速度从9FPS最高到52FPS，已具备CPU上实时推理的能力。

消融实验的结果如下表：

	F1 score	FPS	score
Baseline(pytorch)	0.542	6.1607	0.4459
+ openvino	0.6169	18.703	0.5309
+ rule-based postprocessing			
+ small input size	0.636	25.4975	0.5598
+ optimized postprocessing parameters			
+ network pruning	0.639	46.2634	0.6037
+ CRNN augmented finetuning	0.6565	52.7139	0.6306

三、总结

极市平台的比赛提供了免费算力，并对数据进行了加密，使比赛更加公平，选手可以将更多注意力放在算法设计和优化上。为了提高开发效率，减少bug的发生。在比赛初期，大家可以将样例数据下载下来，在本地的设备上完成主要流程的开发后，再复制到平台的开发环境中。

由于线上运行时触发代码bug，需要选手重新debug和排队，会严重影响开发效率，所以代码的debug十分重要，可以利用好logger功能，将主要信息写入日志。

由于比赛需要兼顾性能分，算法的轻量化工作十分重要，可以通过模型选型，剪枝，量化，模型复用等方法提高推理性能。

我们设计的算法兼顾了精度和速度，可以部署落地在边缘设备中，但精度仍存在提升空间。

如果觉得有用，就请分享到朋友圈吧！



极市平台

专注计算机视觉前沿资讯和技术干货，官网：www.cvmart.net

624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

公众号后台回复“88”获取严彬：STARK-
基于时空Transformer的视觉目标跟踪PPT下载

CV技术社群邀请函



△长按添加极市小助手

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~

觉得有用麻烦给个在看啦~

阅读原文

喜欢此内容的人还喜欢

15个目标检测开源数据集汇总

极市平台