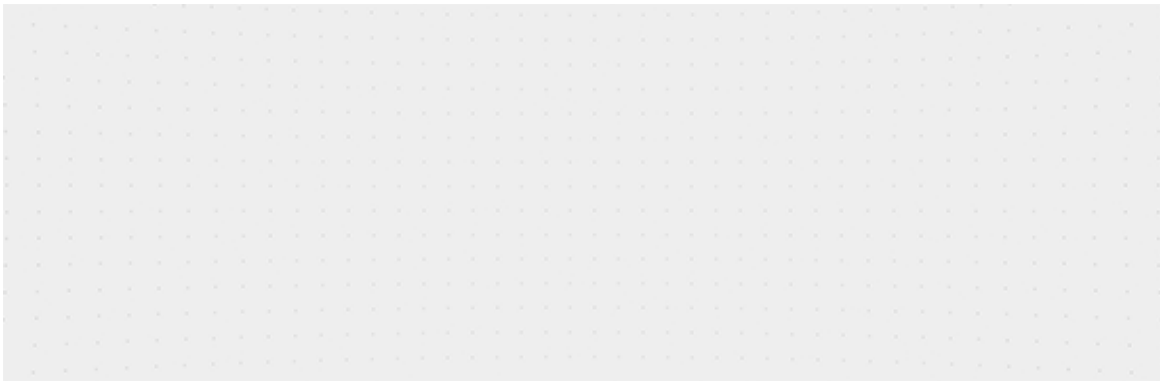


ECV 2021 冠军方案解读：驾驶员不良驾驶识别方案

原创 CV开发者都爱看的 极市平台 2021-11-10 22:00:00 手机阅读

↑ 点击蓝字 关注极市平台



作者 | 王世磊团队

编辑 | 极市平台

极市导读

本文为获得驾驶员不良驾驶识别方案冠军的王世磊团队的解读，团队选用了MobileNet系列的MobileNet v3来作为此次比赛任务的Baseline模型，最终成绩为0.9956。 >>加入极市CV技术交流群，走在计算机视觉的最前沿

一、赛题描述



ECV-2021聚焦于计算机视觉领域的前沿科技与应用创新，全面升级赛制，设立超百万现金奖励，旨在汇聚全球AI人才解决AI产业实际问题，促进人才技术交流，提升开发者人才的算法开发到落地应用的工程化能力，推动计算机视觉算法人才的专业工程化能力认证。

交通事故一直是社会痛点，其中由于驾驶员的危险驾驶行为导致的事故更是频频发生，为有效规范出行司机的不良驾驶行为，降低事故发生率。本赛题希望通过开发驾驶员不良驾驶行为识别算

法，实时甄别驾驶员的不良行为，及时、有效报警与提示，做到消除道路交通安全隐患。我们需要使用计算机机器学习方法，对不良驾驶的视频片段进行分类，输出其类别。

二、数据说明

数据集来自于由摄像头采集的视频片段，每段视频经过分帧后被转成JPG格式的图片并存入到一个文件夹中，每张图片采用frame_id.jpg的命名格式，其中frame_id表示以1为起始的帧序号。每个图片文件夹都会有一个对应的标注文件，文件名称与文件夹一致，格式为XML，包含的标注类别如下所示：

1. 抽烟 (smoke)
2. 打哈欠 (yawn)
3. 打手机 (phone)
4. 驾驶员异常 (a_driver)
5. 左顾右盼 (look_around)

样例数据集：每种类别分别会有几个视频帧集合，供参赛者了解赛题典型场景数据，可用于编码调试；

训练数据集：抽烟400，打哈欠 110，打手机 400，驾驶员异常 400，左顾右盼700，文件夹名称即为类别名称，参赛者需要在编码调试完成后，发起训练任务方可自动读取；

测试数据集：抽烟100，打哈欠 40，打手机 100，驾驶员异常 100，左顾右盼300，参赛者成功发起测试任务，即可自动读取；

三、自动测试

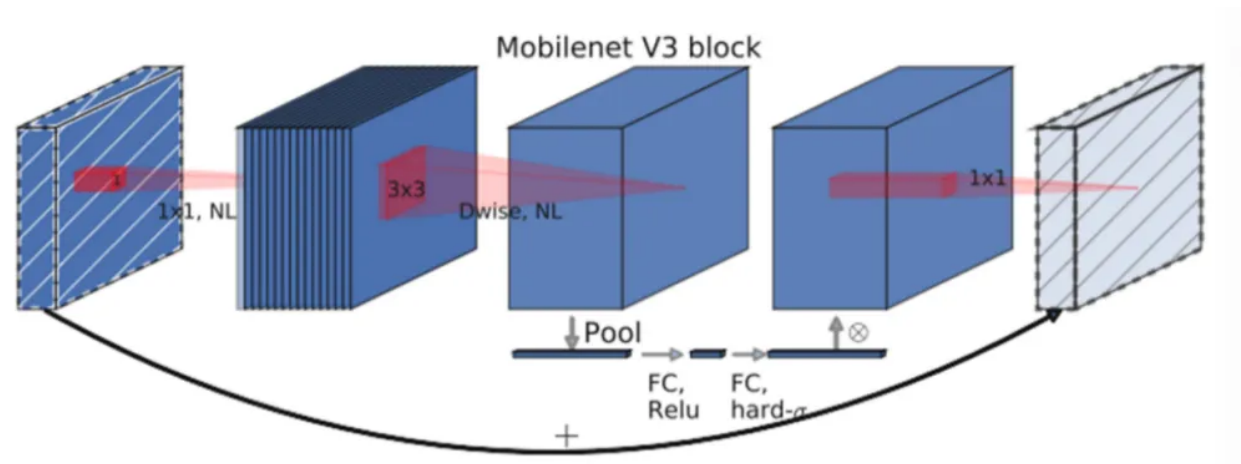
平台提供的自动测试服务，意在模拟算法真实场景下的落地过程，并且为大家提供方便的模型测试工具。因此，需要按照平台的要求封装SDK，也就是按照平台封装模型推理时的输入输出。使用python进行封装。

四、主要工作

1. Baseline的选择

此次比赛的任务为分类任务，测试方式为转换为open vino在cpu上进行测试。因此我们考虑选用参数利用率较高的模型，选用了MobileNet系列的MobileNet v3来作为此次比赛任务的Baseline

模型。Baseline 成绩为：0.7585



2.采视频代表帧进行分类

我们使用Mobile Net V3模型进行测试，能够得到92左右的准确率，但是FPS仅仅只有10左右。模型的速度方面有较大的提升空间，因为这次赛题是对视频进行分类，所以我们就自然的考虑到了从视频中抽取出代表帧来代表整段视频进行分类。使用该方法，在保证精度的同时FPS提升到了30，成绩提升到0.8118。

3.轻量化模型

我们使用相比于Mobile Net V3参数量更小的模型来进行实验，发现得到的精度与Mobile Net V3相同。因此我们根据一步步的测试将模型参数两降低到能保证精度的最小。成绩提升到0.8408。

4.减小乘加量

经过分析，我们发现提出模型的主要运算量来自第一层因此我们考虑对输入的数据进行降维来减小运算量进一步提升检测速度。通过数据降维模型减少了约40%的乘加量，进一步提升了检测速度，FPS达到了100以上。降维后成绩为：0.9398。

Layer	Weight	MADD	Layer	Weight	MADD
Conv1	2.4K	60M	Conv1	2.4K	20M
Conv2	18K	37M	Conv2	18K	37M
Conv3	1K	2M	Conv3	1K	2M
...	1M	2M	...	1M	2M
Total	1M	101M	Total	1M	61M

5.视频帧与标签融合

在FPS达到了100之后，我们考虑进一步提升算法的精度。首先我们尝试了数据增强技术（Auto Aug）。之后为网络增添协调注意力这些均对模型有小幅度的提升。

我们考虑到在真实应用场景我们在对驾驶员行为进行分类时，需要考虑他是否有吸烟习惯以及这个时间段是否容易疲劳驾驶等，因此可以利用驾驶员身份和时间来制作标签信息结合视频图像进行融合再进行分类。在这个赛题中我们为视频制作了统一的标签再与视频帧相融合。这个技巧使得我们的模型在这次比赛中的精度大幅度提升，取得最终成绩0.9956。

五、团队介绍

团队来自西北农林科技大学信息工程学院宁纪锋教授实验室。团队成员为：王世磊。

参考文献：

- 1.Hou, Q., Zhou, D. and Feng, J., 2021. Coordinate attention for efficient mobile network design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 13713-13722).
- 2.Cubuk, E.D., Zoph, B., Mane, D., Vasudevan, V. and Le, Q.V., 2018. Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv:1805.09501.
- 3.Kavyashree, P.S. and El-Sharkawy, M., 2021, January. Compressed MobileNet v3: a light weight variant for resource-constrained platforms. In 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 0104-0107). IEEE.

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“CVPR21检测”获取CVPR2021目标检测论文下载~

极市干货

- 神经网络：视觉神经网络模型优秀开源工作：timm库使用方法和最新代码解读
- 技术综述：综述：神经网络中 Normalization 的发展历程 | CNN轻量化模型及其设计原则综述
- 算法技巧（trick）：8点PyTorch提速技巧汇总 | 图像分类算法优化技巧

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

