

ECV2023 | 安全带规范佩戴检测冠军方案分享

CV开发者都爱看的 极市平台 2023-07-29 22:23:54 发表于广东 手机阅读 𐄞

收录于合集
#冠军方案

9个

↑ 点击蓝字 关注极市平台



作者 | 山东大学万有引力

编辑 | 极市平台

极市导读

本文为ECV2023安全带规范佩戴检测赛题的冠军方案分享，详细的阐述了该赛题的分析和算法实现流程。>>加入极市CV技术交流群，走在计算机视觉的最前沿

1.赛题分析：题目解读

项目背景：安全带规范佩戴检测算法主要用在建筑工地等施工现场，检测施工人员是否佩戴了安全带，从而提升施工区域内的管控工作效率，保证施工人员的人身安全。

训练集：训练数据包含9000张图片；测试集包含1000张图片；

评价指标：模型榜和算法榜都采用准确率、算法性能绝对值综合得分的形式，其中IoU为0.5. $Score = 算法精度 * 0.95 + 算法性能值 * 0.05$

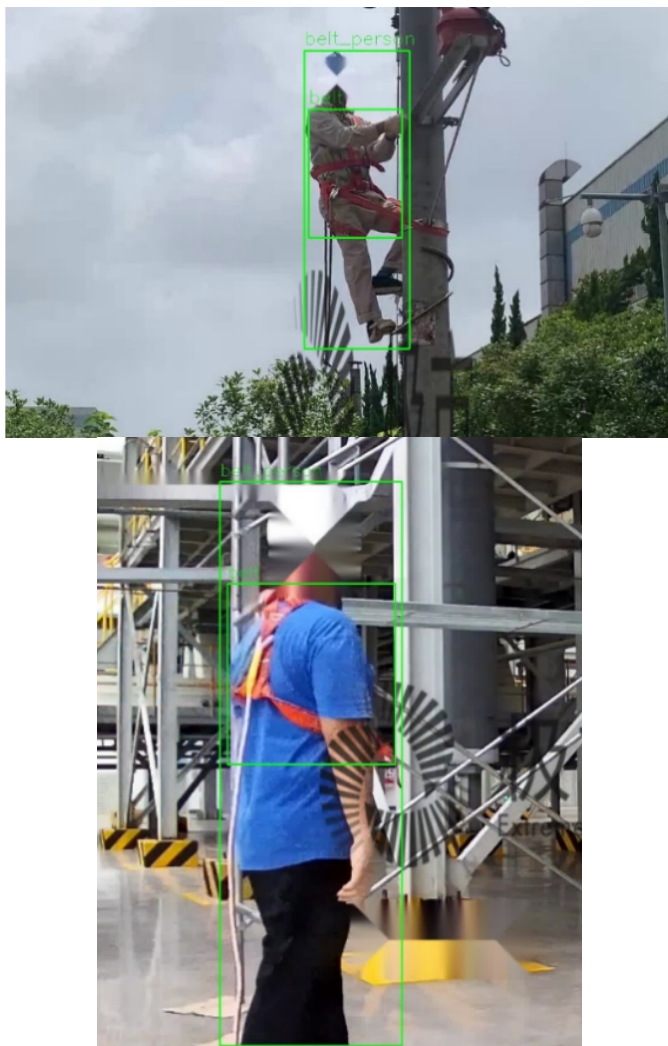
从赛题中可以得到如下的重要信息：

1. 目标检测类问题

2. 类别种类共三类：belt, belt_person, no_belt_person

通过上面的分析，我们有了一个简单的猜测：

安全带类别相比于其他两个类别，属于是“小目标”，如下面两个图所示，我们认为可能是影响最终排名的决定性因素。



2. 赛题分析：模型选择

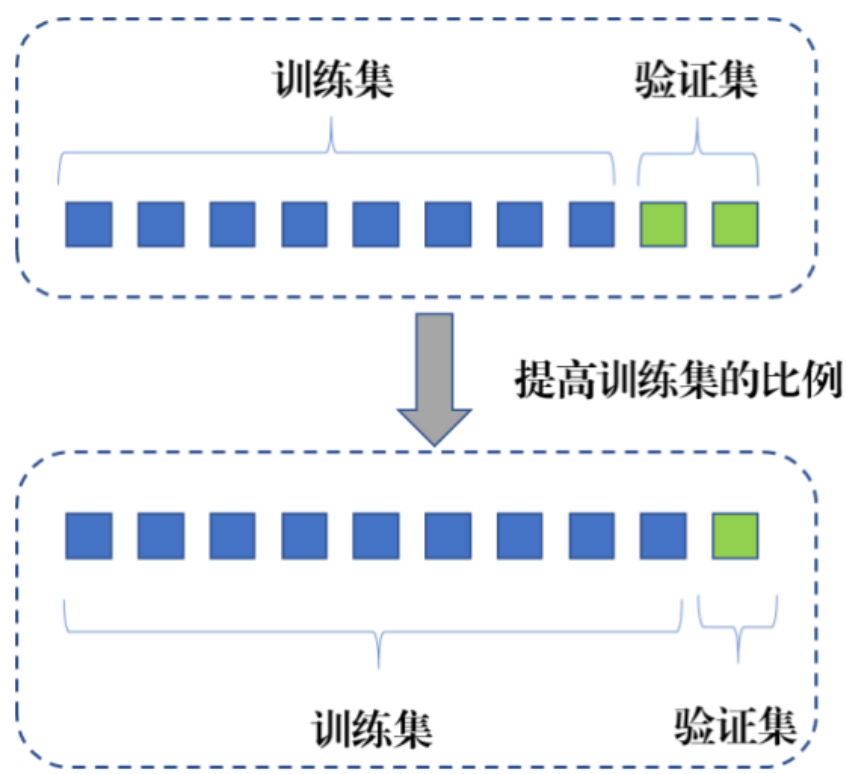
一个合适的模型将直接决定最终的结果，我们以赛题提供的代码作为baseline（极市目标检测开发套件）先熟悉极视平台训练流程（我们是第一次参加极市的比赛）；然后跑通baseline；

极市开发套件包含了三个模型的运行脚本：yolov5, yolox, yolov7。这里我们选择了yolov5：

1. YOLOv5的部署比较完善，推理速度快。
2. 参数量也比较合适

3. 算法流程及实现

3.1 划分训练集和测试集



我们首先调整的参数是，提高训练集和验证集的比例：8:2 → 9:1

进行这个参数的特征，主要有如下两点考虑：

- 1. 训练数据一共有9000张图片，总体比较少；
- 2. 希望模型能接触到更多训练数据；

使用更多的训练数据进行模型的训练，也有可能带来一定的危害，如测试集合过少，验证出的最优模型可能在测试集上并不是最优的，也就是验证集过少，可能不能验证模型的真实性能，但是值得一试。

具体的实现方式：采用随机划分

实验表明，我们的修改是正确的，提高了模型的f-score

3.2 标签平滑

$$y_i = \begin{cases} 1 - \varepsilon & \text{if } i = \text{true}, \\ \frac{\varepsilon}{K - 1} & \text{otherwise} \end{cases}$$



为了更好的性能，我们使用了标签平滑的技术，标签平滑是一种正则化的策略，通过soft label来加入噪声，减少真实样本的类别在计算损失时的权重，起到抑制过拟合的效果。除此之外，该方法还可以提高模型的泛化能力。

我们的平滑参数设置： $\varepsilon=0.1$ ，获得了模型榜第一名

3.3数据增强

数据增强有多种方式，这里我们几乎能用的数据增强方式都用到了：

- 随机旋转、随机裁剪、随机观察角度
- 随机模糊、随机灰度化、降低图像质量
- 在HSV颜色空间上进行数据增强，像素级别数据增强
- Mosaic：一张主图和随机三张图片平成一张图片
- MixUp：将原图和随机选取的图片进行混合

数据增强对于提高模型的泛化能力和提高模型识别的准确率非常重要。

4 测试结果分析

118542	标准测试	测试完成	发起:04-23 20:59:09 开始:04-23 21:00:33 结束:04-23 21:05:03	5	74.2196	0.9565	详情 中止 撤销
--------	------	------	---	---	---------	--------	--

采用上面的方法改进后，我们对模型进行了训练和测试，结果如上图所示，可以看到，模型的f-score是0.9565，已经是比较高的了。这说明训练和测试的数据并不复杂，简单的yolov5s模型就可以学习到比较好的特征。

然后我们还分析了每个类别的精确率和召回率，我们发现，belt和belt_person两个类别无论是精确率还是召回率都非常高，no_belt_person反而是最低的。

标签名	标签数量	精确率	召回率	f-score
belt	956	0.9285	0.9644	0.9461
belt_person	953	0.9833	0.9895	0.9864
no_belt_person	895	0.9138	0.9598	0.9362

这和我们最初的猜测是相违背的，我们猜测，可能是因为：

1. belt和belt_person两个类别是有关联的，模型可学习到了这种关联，因此belt的识别率并不低，而no_belt_person需要分别出没有带belt，反而性能降低了；
2. belt和belt_person是“相容”的，而no_belt_person与之不“相容”；

5 其他模型尝试

除了yolov5s之外，我们也尝试了yolov7和yolox。

在精度方面：yolov7精度表现和yolov5s基本相同，而yolox精度则要低一些。

在性能方面：yolov5s的性能几乎是他们的两倍。

综上，我们考虑使用yolov5s作为最终的检测框架。

6 模型测试阶段调试

标签名	标签数量	精确率	召回率	f-score
belt	956	0.9285	0.9644	0.9461
belt_person	953	0.9833	0.9895	0.9864
no_belt_person	895	0.9138	0.9598	0.9362

在查看测试结果的时候，我们发现每个类型的精确率都是低于召回率的。考量模型性能的一个指标是f-score，而f-score是精确率和召回率的调和平均值，如下面公式所示：

$$F - score = \frac{2 * Precision * Recall}{Precision + Recall}$$

可以通过简单的数据分析可以知道，f-score的值只有在Precision和Recall都很高的情况下才会变得高。目前我们是召回率很高，通过提高置信度的阈值，由原来的0.2提高到0.45，损失一点召回率，提高了模型的精确率，最终提高f-score也是非常值得的。实验的结果也证明了我们的想法。

7 模型部署

起初，我们使用了极市平台提供的C++部署套件，并按照赛题的要求，修改其中的业务逻辑，并使用tensorRT进行加速。但是最终的测试性能，却远没有我们预想的那么快，可能是出现了bug。后期，因为平台也支持python，所以我们最终采用python代码来进行测试，先将pytorch模型转换成onnx模型，然后再使用tensorRT进行优化。



配置好一切的环境后，在部署榜上的性能结果基本上接近模型榜上的性能。

8. 总结及可能的改进

总结：

1. 数据并不复杂，不需要特别大的模型

2. 使用多个数据增强，提高模型的泛化能力
3. 可以增加正则损失，避免模型过拟合
4. 要根据训练结果调整训练参数，置信度阈值等等参数
5. 推理性能，使用TensorRT进行推理优化。

可能的改进：

1. 尝试使用半精度推理，进一步提高模型的性能
2. 尝试模型融合，多个模型的结果进行合并，进一步提高f-score
3. 采用交叉验证方式，选择更好的模型参数



极市 | 2023年7月6-26日

EVDeploy 部署套件

让算法落地更简单

解锁 AI 推理部署新体验，EVDeploy 公测中

长按识别二维码 即刻参加

限时有奖公测

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

极视角动态：2023GCVC全球人工智能视觉产业与技术生态伙伴大会在青岛圆满落幕！ | 极视角助力构建城市大脑中枢，芜湖市湾沚区智慧城市运行管理中心上线！

数据集：面部表情识别相关开源数据集资源汇总 | 打架识别相关开源数据集资源汇总（附下载链接） | 口罩识别检测开源数据集汇总

经典解读：多模态大模型超详细解读专栏

点击阅读原文进入CV社区
收获更多技术干货

收录于合集 #冠军方案 9

上一篇
ECV2023 | 睡岗识别冠军方案解读

下一篇
ECV2023 | 皮带跑偏识别冠军方案解读

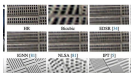
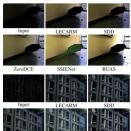
阅读原文

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NeRCo
极市平台



极市平台

