

基于YOLOv5s的智能口罩佩戴检测系统的设计与实现

许少军,唐浩博,曹领

(安阳工学院 计算机科学与信息工程学院,河南 安阳 编号 455000)

摘要:为实现对口罩佩戴检测的实时检测,提高检测效率,加强口罩佩戴检测在工业生产、医疗卫生等领域的应用,文章提出了基于YOLOv5s模型优化的口罩佩戴检测算法,通过对自适应缩放功能和网络层数的改进,在一定程度上提升了检测速度,实现了图片检测、视频检测、实时检测的功能的可视化,多场景多目标实时检测功能。

关键词:YOLOv5s; 口罩识别; 深度学习

中图分类号:TP18 文献标识码:A

文章编号:1009-3044(2024)03-0032-04



开放科学(资源服务)标识码(OSID)

0 引言

人工检查是否佩戴口罩存在一定的局限性，并且效率较低，因此对佩戴口罩进行自动实时识别是更好地解决问题的方法，尤其在工业生产、医疗健康等领域应用较为广泛。结合目前计算机视觉和人工智能的图像处理及目标检测技术已广泛应用于各种图像及视频场景。计算机视觉技术也愈发成熟，越来越多的领域都在引入智能化，智能口罩佩戴检测也应该在各领域中发挥它的作用。

本文展示的是基于 YOLOv5s 改进的戴口罩实时检测系统，能够实现图片检测，实时检测。以深度学习为基础的目标检测是目前比较主流的一种目标检测计划。现阶段主流优秀的目标检测算法主要分为两类，一类是基于 Region Proposal 的 Two stage 目标检测算法，双阶段目标检测算法过程主要将检测分为生成候选框和检测识别两个过程，并提取生成的候选框的信息。它利用卷积神经网络完成目标检测和目标识别，主流的网络模型有 SPPNET、Fast RCNN、Faster RCNN、R-FCN 等。第二类是以回归问题为基础的 One Stage 目标检测算法，将全过程统一，同时使用具有结构简单、运算速度快等优点的 CNN 进行特征提取和回归，是单一阶段目标检测算法过程。典型的网络模式有 YOLO、SSD、DSSD、CornerNet 等。正是由于 One stage 和 Two stage 这两种算法的差异，使得两者在性能上也存在差异，前者在检测精度和检测定位精度上占据优势，后者则在检测速度上占据优势。

1 项目的设计与实现

1.1 YOLOv5 网络框架的搭建

YOLOV5 根据其网络框架深度和宽度主要分为

YOLOv5s、YOLOV5M、YOLOV5L、YOLOV5X 四个版本,而 YOLOv5s 是这几个版本中宽度最小、深度最小的特征图网络,因此它有更快的检测速度,更适用于小型嵌入式系统。不同网络框架的对比如表1所示:

表1 特征图深、宽度对比

	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv5x
深度(depth)	0.33	0.67	1.0	1.33
宽度(width)	0.50	0.75	1.0	1.25

YOLOV5 的模型结构主要分为 4 个部分, 分别是 Input 输入端、Backbone Basic Network、Neck Network 和 Prediction 输出层。

1) Input 输入端可以对输入的图像进行自适应图片缩放、数据增强、锚框等处理后将结果输入 backbone 基础网络。

2) Backbone Basic Network 包含 CSP 架构与 FOCUS 架构。其中在 YOLOv5 中的 FOCUS 模块,是图片进入 Backbone 之前的切片操作。YOLOv5s 的 CSP 结构设计有 2 种。

3) Neck Network 由 csp2_x 网络来代替普通的卷积网络,使网络特征融合的能力得到了加强。

4) Prediction 输出层包括 Bounding Box Regression Loss 和 NMS。其中的 GIOU_LOOSS 被用在了 YOLOv5 中, 用于制作 Bounding Box 的损失函数。在目标检测的后处理过程中, 使用加权 NMS 的方式来筛选许多目标框。YOLOv5s 的网络架构如图 1 所示。

3) 在预测中, 使用非最大抑制(NMS)的加权方法 NMS, 利用 GeneralizedIoU(GIoU)Loss 作为 BBOX 的损失函数。如式(1)所示。

收稿日期:2023-11-15

基金项目:2022年大学生创新创业训练计划项目(编号:202211330015)

作者简介:许少军(1999—),男,河南洛阳人,在校本科生,研究方向为计算机视觉,机器学习;唐浩博(2002—),男,河南登封人,在校本科生,研究方向为深度学习、计算机视觉;曹领(1980—),男,河南安阳人。通信作者:讲师,研究方向为嵌入式系统、智能物联网。

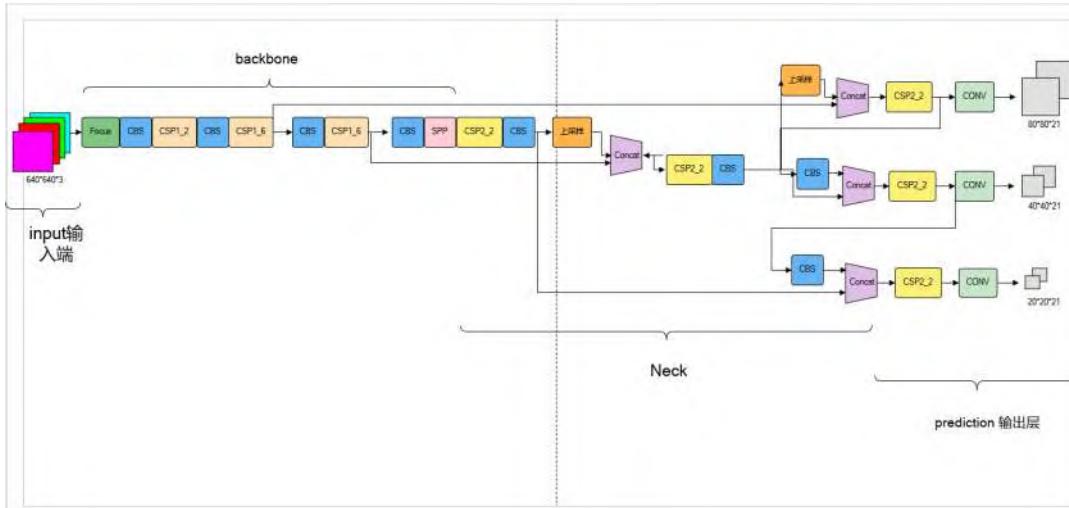


图1 YOLOv5s网络架构

$$\begin{aligned} L_{GloU} &= 1 - IoU + \frac{|C - (B - B^{gt})|}{|C|} \\ IoU &= \frac{|A \cap B|}{|A \cup B|} \end{aligned} \quad (1)$$

1.2 数据集的制作

数据集包含4 000张不同场景下佩戴面具和未佩戴面具的人面图片,通过旋转裁剪这些图片来实现数据集的增强后,使用Labelimg对图片进行标记,将人面的位置标注在图片中的长方形框标上,并将人面的类别标注为佩戴口罩和未佩戴口罩,之后标注为YOLO格式(YOLO格式),这里将数据集按照8:1:1来对训练集(train.txt)、测试集(test.txt)和验证集(val.txt)进行划分。训练集用来训练模型以及确定参数;而验证集是为了确定超参数的网络结构和调整模型而设计的。

1.3 模型的训练与环境的配置

本次实验在Windows操作系统下完成,采用PyTorch的框架,Python 3.8的版本环境, GPU型号为RTX3080。

2 模型检测的改进

2.1 目标检测算法的改进

早期的对象调查都是以滑动窗口的方式产生目标建议框,本质上与穷举法并无差异。而Fastr-CNN

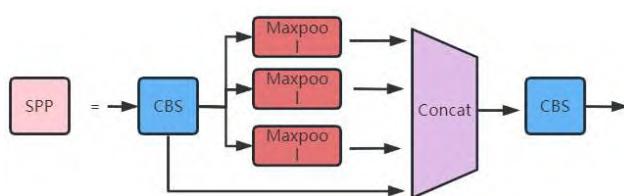


图2 SPP层架构

的出现,解决了冗余计算这一棘手问题。Fastr-CNN增加了简化的SPP层,使其培训和测试过程融为一体。SPP层架构如图2所示。

Fastr-CNN使用SelectiveSearch生成目标候选框架,但速度仍无法达到实时要求。FasterR-CNN则直接利用RPN(注册网络)网络生成

目标候选框架。RPN输入任意像素点的RAW影像,输出若干个对应目标坐标资讯与信心的矩形区域。从R-CNN到FasterR-CNN,是将传统检测的3步整合到同一深度网络模型中的一个合并过程。R-CNN原则框架如图3所示:

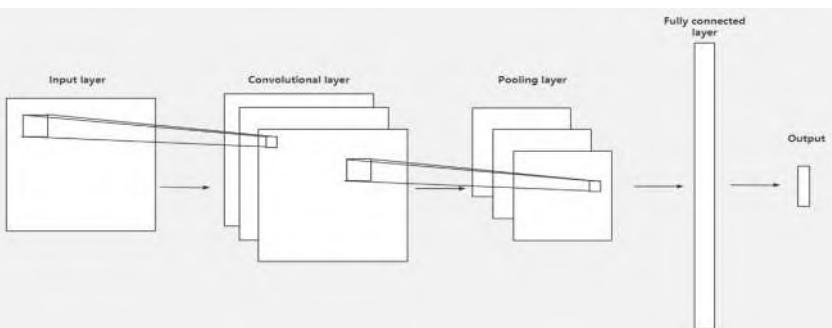


图3 R-CNN 原理框架

基于回归算法的YOLO和SSD检测模型,让检测领域再上一个新台阶。其中,以YOLO和SSD两种方式为代表的检测方式真正达到了立竿见影的效果。

2.2 对于YOLOv5本身检测的改进

1)自适应图片缩放改进。YOLOv5s代码中对自适应图片缩放进行改进,使得YOLOv5的检测速度能有所提升。

2)图片素材强化。抓取的图片进行素材强化,扩大图片的数量如旋转裁剪图片,增加无目标背景图片清晰的高速网络,通过背景和目标的对比,区分出图片中微小的物体,什么是无目标背景,什么是目标。

3)减少网络层数。避免多层卷积核混合后分不清多个小目标的情况发生。在图像缩放的过程中尽可能少地添加黑边,来提高模型的推理速度。步长适当减小,避免过多的提取小目标像素点。

4)将图片切分为多张图片,对每张图片进行单独检测然后汇总,可以增强对多目标的检测。

3 应用场景及意义

智能口罩检测系统在于节省人力、提高效率、智能提醒、信息化监控。在以下场景中有重要作用:

1) 工业领域。用于保护劳动者呼吸系统,防止其吸入有害粉尘、烟雾、气体等颗粒物,减少工作场所的安全隐患。举例如下:

粉尘口罩:用于防止吸入粉尘、花粉、灰尘、煤尘等。

化学防护口罩:用于防止工人吸入有害气体和蒸汽。

2) 医疗领域。医务人员在医疗过程中需要佩戴口罩,而智能检测口罩可以帮助医院检测医务人员佩戴口罩的情况,保证医务人员的防护措施达标,起到保护医务人员和患者身体健康的作用。举例如下:

预防医务人员感染:医务人员容易接触病人的体液、血液等分泌物,在诊治病人呼吸道感染性疾病及手术过程中,戴口罩可有效预防病人呼吸道感染。防止患者感染疾病的发生。在医院中,患者一般要长时间待在医疗环境里,戴上口罩可以有效地防止患者呼吸道感染传染给他人,包括医务人员和其他患者。

保护手术区域:在手术期间,医生需要保持一定的清洁度和卫生,佩戴口罩可以防止口腔和鼻腔的细菌进入手术部位,从而避免手术感染。

防止病毒传播:在面对一些高度传染性病毒,例如SARS、禽流感、新冠肺炎等时,佩戴口罩,可以在一定程度上防止医务人员和患者感染。

3) 交通运输领域。尽管疫情的情况已经好转,我们已经不用时时刻刻佩戴口罩,但在一些特殊场合仍需要佩戴口罩,特别是现在“二阳”与甲流的特殊时期。

综上,口罩检测依然有十分重要的作用。此系统能够实时监测画面中的人脸,并将其分为:戴口罩、不戴口罩两类,并且在具体场景中,可以有相应作用,其特点是灵活运用,适用面较广。

4 性能分析及测试结果

4.1 混淆矩阵

混淆矩阵是将分类问题的预测结果进行汇总,说明在预测时,哪个部分会被分类模型所混淆。它不仅可以让我们了解分类模型中所犯的错误和正在发生的错误是怎样的类型,同时也克服了分类精确度所造成的限制(regecence)。还可以得到辨识度的信息,该模型口罩的辨识度达到了0.86,人的辨识度在0.97。如表2、图4所示。

4.2 测试数据

在对数据进行训练后,准确度(Precision)、平均识别精度均值(MAP)、召回率(Recall)等参考量可以较直观地判断出所采用的模型对图片的辨识度。 TP :分类器预测结果为正样本,实际也为正样本,即正样本被

表2 混淆矩阵原理

混淆矩阵		实际(Actual)		
	200	类1	类2	类3
预测(Prediction)	类1	54	5	2
	类2	3	59	4
	类3	1	2	60

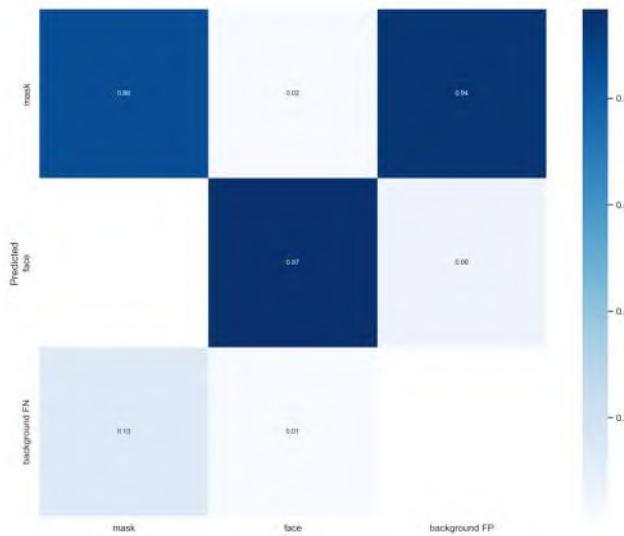


图4 混淆矩阵

正确识别的数量; FP :分类器预测结果为正样本,实际为负样本,即误报的负样本数量; FN :分类器预测结果为负样本,实际为正样本,即漏报的正样本数量。

精确度是指由正样本预测出正样本数与全部预测出正样本数之比的所有预测框的个数,即全部预测框的个数。它的计算公式是:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

召回率指的是所有正向样本数与所有正向样本数的预测比值,而所有的正向样本数都是所有标的物的框本数。其计算公式如下:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

平均识别精确度均值 MAP 顾名思义,就是所有平均识别精确度 AP 的平均值,以精确度和召回率为横纵坐标所围成的面积。它的计算公式是:

$$AP = \sum_{i=1}^{n-1} (r_i + 1 - r_i)p(r_i + 1) \quad (4)$$

p 表示精确度, r 表示召回率。

$$mAP = \frac{1}{n * \text{sum}(AP)} \quad (5)$$

n 表示类数。

图5为优化后的YOLOv5s精确度(精确度)、召回率(回复率)、平均识别精确度(MAP)。

4.3 检测结果

针对场景中的不同需求,本次设计涵盖了图片检测、视频检测和实时检测三大功能,且可以实现单一

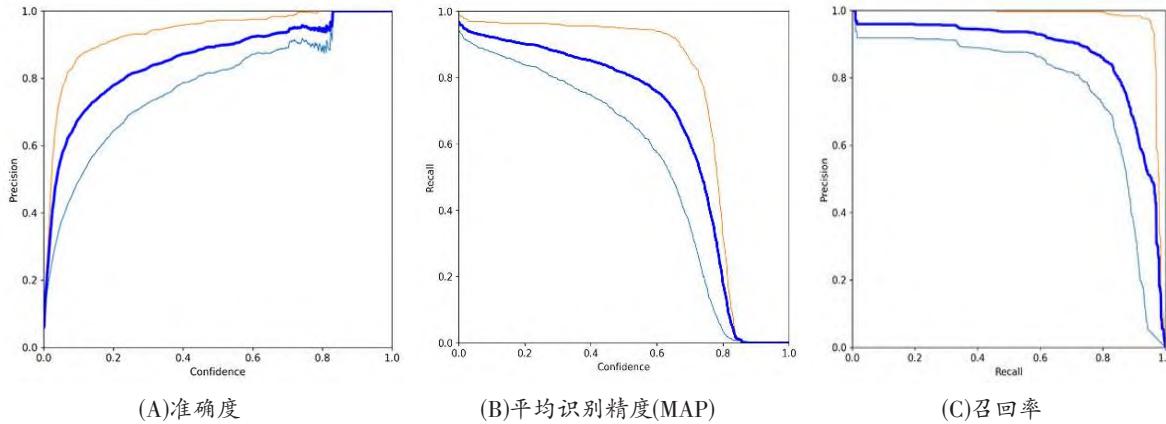


图5 测试曲线图



图6 不同场景目标检测效果

目标检测、多目标检测、较远距离中小目标的检测。检测实例如图6所示。

5 结束语

本设计针对口罩佩戴检测的问题以及提高口罩佩戴检测的效率,提出了一种基于YOLOv5s的口罩佩戴检测系统。本系统提出一种改进的YOLOv5s算法,使用数据集相关数据对模型进行训练。训练结果达到了预期的精度,模型的检测精度和速度同时均衡,达到了预期规定的实时性检测目标,同时,模型的检测精度和速度也达到了预期。此系统满足嵌入式终端设备的部署要求,希望在今后的学习中,在保证模型检测精度和速度均衡的前提下,进一步提高模型的精度,同时进一步优化网络结构、各种参数,使算法更出色。

参考文献:

- [1] 王烁宇,崔庆森.基于YOLOv5s网络模型的口罩佩戴检测系统[J].信息记录材料,2022,23(8):94-96.
- [2] 王赫,王刚,李莘诺,等.基于YOLOv5改进模型的口罩识别方法[J].黑龙江科学,2022,13(24):58-61.

- [3] 项融融,李博,赵桥.基于改进YOLOv5s的口罩佩戴检测算法[J].国外电子测量技术,2022,41(7):39-44.
- [4] 许德刚,王露,李凡.深度学习的典型目标检测算法研究综述[J].计算机工程与应用,2021,57(8):10-25.
- [5] 陈科圻,朱志亮,邓小明,等.多尺度目标检测的深度学习研究综述[J].软件学报,2021,32(4):1201-1227.
- [6] 谈世磊,别雄波,卢功林,等.基于YOLOv5网络模型的人员口罩佩戴实时检测[J].激光杂志,2021,42(2):147-150.
- [7] 宋谱怡,陈红,苟浩波.改进YOLOv5s的无人机目标检测算法[J/OL].计算机工程与应用,2022:1-10.(2022-06-22).https://kns.cnki.net/kcms/detail/11.2127.TP.20220621.1105.004.html.
- [8] 陈兆凡,赵春阳,李博.一种改进IoU损失的边框回归损失函数[J].计算机应用研究,2020,37(S2):293-296.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(1):142-158.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision and Pattern Recognition. IEEE, 2016:779-788.

【通联编辑:唐一东】