

基于改进 YOLOv5 算法的现场不安全行为识别方法研究

刘艺超¹,唐哲¹,方汀¹,赵小军¹,李存阳²,郑运昌²

(1.张家口卷烟厂有限责任公司,河北 张家口;2.河北建筑工程学院,河北 张家口)

摘要:在一些工厂的生产车间现场中,工人需要驾驶升降车作业,在驾驶升降车时存在着不安全行为,比如没有佩戴安全帽等。因此需要对升降车人员不安全行为进行识别。一般的算法在识别工人驾驶升降车作业时佩戴安全帽这种复杂背景的场景时容易漏检,误报。基于此,提出一种基于 YOLOv5 改进的目标检测算法。此算法用于对工厂的车间现场不安全行为进行检测。在 YOLOv5 改进的算法中增加了 CBAM 注意力机制,通过结合通道和空间的注意力机制模块,以增强算法对小目标的检测精度和对特征的提取能力。然后将原有的特征金字塔网络结构替换为加权特征金字塔网络结构(BIFPAN),实现高效的双向跨尺度连接和加权特征融合。通过自制的安全帽数据集实验结果显示,改进后的算法均值平均精度(mAP)达到了 92.4%,检测精度和召回率分别提高了 2.7%和 1.9%。

关键词:YOLOv5;工厂;检测;加权特征融合

中图分类号:TH211.+6;TP391.41

文献标识码:A

文章编号:2096-4390(2023)08-0005-04

引言

在工厂车间现场中工作人员需要驾驶升降车进行作业,但在工厂生产车间中存在很多的安全隐患和不安全行为,比如工人在驾驶时未佩戴安全帽。这就需要对其不安全行为进行监测^[1]。然而在工厂车间中一般采用人工监督的方式来监测,这种采用人工的方式耗时人工成本高,且工人在驾驶升降车时,人工在监测这种运动目标时很容易出现漏检的情况。所以现在一般采用深度学习辅助人工监测是否佩戴安全帽,基于传统人工监测耗时、耗力、成本高、容易漏检的缺点,采用深度学习辅助人工监测的方式更符合工厂车间的监测要求^[2]。在传统的目标检测算法中,基本上是基于特征和基于分割两个方向。其中,基于特征的目标检测算法主要是通过找到对象的某种属性(可以是手工设计的特征也可以是算法提取的抽象特征)来实现检测与识别。但传统目标检测算法特征的设计与选择极大程度上依赖于人工,其准确度、客观性与泛化性都受到了一定的制约^[3]。所以近些年来,目标检测算法主流的一般是关于卷积神经网络这种不依赖于人工选择特征的方法,并且得到很快的发展和

更迭^[4]。在目前的目标检测算法中,主要分为双阶段目标检测算法和单阶段目标检测算法。其中,双阶段目标检测算法为代表的有 Fast RCNN、Faster RCNN、Cascade RCNN 等。而单阶段算法中有 SSD、RetinaNet、YOLO 系列等^[5]。单阶段目标检测方法是指只需一次提取特征即可实现目标检测,其速度相比双阶段算法快,一般精度稍微低一些,但在实时性上,单阶段算法要优于双阶段算法,所以现在一般多采用检测速度更快的单阶段算法。其中 YOLO 是单阶段算法的代表,推理预测速度快,满足实时性的要求。基于此,本研究以 YOLOv5 模型为基础,进行了一些改进。

1 YOLOv5 算法简介

YOLOv5 是 YOLO 中的一种网络结构,是在之前 YOLOv3 和 YOLOv4 的基础上进行集成化的结构。它的网络结构由输入端、骨干网络(Backbone)、颈部(Neck)和预测端(Prediction)4 个部分组成^[6]。

(1) 输入端:由 Mosaic 数据增强、自适应锚框计算、自适应图片缩放组成。这一部分通过对图片进行预处理操作,丰富了数据集和减少了 GPU 计算^[7]。

作者简介:刘艺超(1991-),男,本科,助理工程师,研究方向:安全生产管理,安全生产标准化。

(2) 骨干网络(Backbone):主要是进行特征提取的网络结构,主要由 Focus 结构和 CSP 结构组成。

(3) 颈部(Neck):此结构中采用了特征金字塔网络 (FPN)+ 路径聚合网络 (PAN)的结构,从而对骨干网络中提取的特征信息进行进一步的提取和融合^[8]。

(4) 预测端:主要是目标框损失函数,选择了 GIOU U(Generalized Intersection over Union,广义交并比)作为边界框回归损失函数^[9]。

2 YOLOv5 算法的改进

2.1 骨干网络中加入 CBAM

由于在工厂车间中工人驾驶升降车时环境比较复杂,且安全帽的尺寸较小,图片中的信息不易被关注,所以针对此问题,在 YOLOv5 中的骨干网络中加入了注意力机制,以此来提高网络对信息特征的关注度,从而提升识别检测的速度跟精度^[10]。在人类大脑皮层中,人在观察图像时通过关注重点信息而忽略次要信息来进行筛选,从而提高识别效率的原理。通过使用注意力机制能够更快捷和高效地分析复杂场景信息^[11]。注意力机制便受此启发而来,从而被引入到计算机视觉中来提升性能。CBAM (Convolutional Block Attention Module)卷积注意力模块主要是从通道和空间两个维度的方向来关注特征信息。并且能够序列化地在通道和空间两个维度上产生注意力特征图信息,然后两种特征图信息在之前原输入特征图相乘进行自适应特征修正,产生最后的特征图^[12]。CBAM 分为通道注意力模块 CAM (Channel Attention Module)和空间注意力模块 SAM(Spatial Attention Module)两个模块。CBAM 模块能够有效减少计算量,比较轻量,可移植性高,作用效果很好。CBAM 卷积注意力模块见图 1。

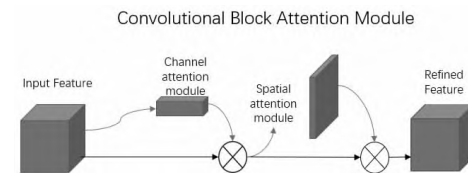


图1 CBAM 卷积注意力模块

2.1.1 CAM 通道注意力机制模块

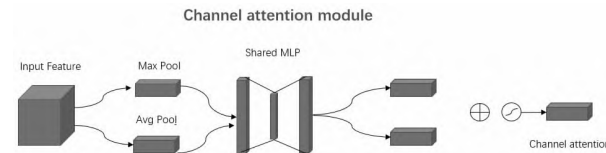


图2 CAM 通道注意力机制模块

CAM 通道注意力机制模块流程图见图 2。先将输入特征图分别进行全局最大池化和全局平均池化,对特征映射基于两个维度压缩,获得两张不同维度的特征描述。池化后的特征图共用一个多层感知器网络,先通过 11 卷积降维再 11 卷积升维^[13]。将两张特征图叠加,经过 sigmoid 激活函数归一化特征图的每个通道的权重。将归一化后的权重和输入特征图相乘。

2.1.2 SAM 空间注意力模块

SAM 空间注意力模块流程图见图 3,对通道注意力机制的输出特征图进行空间域的处理。首先,特征图分别经过基于通道维度的最大池化和平均池化,将输出的两张特征图在通道维度堆叠^[13]。然后使用 1*1 卷积调整通道数,最后经过 sigmoid 函数归一化权重。将归一化权重和输入特征度相乘。

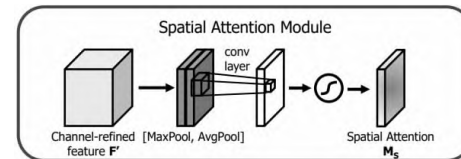


图3 SAM 空间注意力模块

2.2 特征金字塔网络结构替换为加权特征金字塔网络结构(BIFPN)

在 YOLOv5 算法最初版本中,颈部网络结构只包含 FPN 结构。然后在后续的版本中改进为 FPN+PAN 结构。通过对骨干网络提取到的特征信息进行进一步融合,以此来提升识别的精度。但在多尺度融合中可能会存在不同尺度特征信息不一致的问题,从而这种融合特征并不能达到很好的效果。同时只用 PAN+FPN 又会增大计算量等这些问题。针对这些问题,通过借鉴 EfficientDet 网络中的 Neck 网络结构:BiFPN (Bidirectional Feature Pyramid Network)加权双向特征金字塔网络,将 PAN 替换为 BiFPN。与 PAN 只能单一的自底向上传达特征不同,BiFPN 不仅可以自底向上传递特征,又可以自顶向

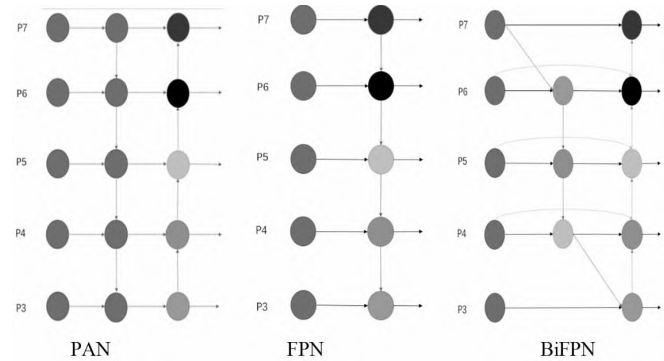


图4 PAN、FPN、BiFPN 结构



图5 标注图片

表1 自制数据集实验数据对比

Model	P / %	R / %	mAP@0.5	mAP@0.5:0.95
原 YOLOv5	88.5%	86.4%	90.7%	59.6%
改进的 YOLOv5	91.2%	88.3%	92.4%	60.4%

下传递特征^[4]。BiFPN 自顶向下的通路,传递的是高层特征的语义信息;自底向上的通路,传递的是低层特征的位置信息。通过采用这种双向加权特征网络能够更好的提取特征信息,更好的融合特征。图4为PAN、FPN、BiFPN 结构图。

3 实验结果

3.1 实验数据处理

本次实验采用的安全帽数据集是来自工厂自制图片和视频的数据集。包括 10 000 个佩戴有安全帽的人物头像与 10 000 个未佩戴安全帽的人物头像。数据集

的标注我们通过 Labeling 进行标注,见图5。我们将标签标记为两类:hat 和 person,将佩戴安全帽的人标记为 hat,而将未佩戴安全帽的人头标记为 person。按照 6:2 的比例划分训练集和验证集。

3.2 实验数据对比

将改进后的 YOLOv5 算法和原 YOLOv5 算法分别对自制数据集进行实验对比。最终得到的数据见表1。

通过自制安全帽数据集的进行实验,将改进的 YOLOv5 算法与原 YOLOv5 算法进行对比。精确度提升了 2.7%,召回率提升了 1.9%,mAP@0.5 提升了 1.7%。由最终对比的实验数据得出通过增加 CBAM 和 BiFPN,可以很好的提升安全帽检测的精度。

4 结论

本研究是在原 YOLOv5 算法的基础上进行改进,通过在骨干网络中增加 CBAM (Convolutional Block Attention Module)卷积注意力机制,以及在特征金字塔网络替换为加权特征金字塔网络 BIFPAN,从而提升了在检测一些工厂的工人驾驶升降车佩戴安全帽的精度,对小目标的检测效果有

了很好的提升,且对工人驾驶时密集的目标有了很好的效果。改进后的算法在安全帽的检测中有了进一步的提升,可以更好的应用于类似的工厂安全检测环境。

参考文献

- [1]周兆银,田书画,谢艳,等.计算机视觉技术在施工人员不安全行为管理中的应用[J].重庆大学学报,2022,45(S1):74-78.
- [2]李政谦,刘晖.基于深度学习的安全帽佩戴检测算法综述[J].计算机应用与软件,2022,39(6):194-202.
- [3]李继东,赵镕,黄佳,等.基于 Faster R-CNN 的环网柜电

- 缆相序检测 [J]. 杭州电子科技大学学报 (自然科学版), 2021,41(4):61-66.
- [4]李莎莎.基于深度回归神经网络的小目标和遮挡目标检测算法研究[D].郑州:河南大学,2022.
- [5]常志磊.基于 RGB-D 的双手手势识别方法研究及系统设计[D].呼和浩特:内蒙古大学,2022.
- [6]贾豆豆.基于 YOLOv5+DeepSort 的小目标跟踪方法研究[D].太原:中北大学,2022.
- [7]张欣怡,张飞,郝斌,等.基于改进 YOLOv5 的口罩佩戴检测算法[J].计算机工程,2022-12-08.
- [8]刘鹏,毕誉轩,张天翼,等.注意力机制优化的全尺寸目标检测方法[J].电子测量与仪器学报,2023-3-3.
- [9]何凌波,陈西曲.基于改进 YOLOv5 的安全帽佩戴检测算法[J].长江信息通信,2022,35(11):14-19.
- [10]高宝东.基于 YOLOv5 的小目标检测算法与应用研究[D].银川:宁夏大学,2022.
- [11]张大伟.基于深度学习的视觉目标跟踪算法研究[D].金华:浙江师范大学,2022.
- [12]Chao Chen, Bin Wu, Hongying Zhang. An Image Recognition Technology Based on Deformable and CBAM Convolution Resnet50 [J]. IAENG International Journal of Computer Science,2023,50(1).
- [13]张振国,邢振宇,赵敏义,等.改进 YOLOv3 的复杂环境下红花丝检测方法[J].农业工程学报,2023-3-13.
- [14]邹文武,许贵阳,白堂博.基于 EfficientDet 的轨道扣件识别与检测[J].武汉大学学报(工学版),2022-11-14.

Research on Unsafe Behavior Recognition Method based on Improved YOLOv5 Algorithm

Liu Yichao¹, Tang Zhe¹, Fang Ting¹, Zhao Xiaojun¹, Li Cunyang², Zheng Yunchang²

(1.Zhangjiakou Cigarette Factory Co., Ltd., Zhangjiakou, China;

2.Hebei University of Architecture, Zhangjiakou, China)

Abstract: In the production halls of some factories, workers need to drive lift trucks to work, and there are unsafe behaviors when driving lift trucks, such as not wearing safety helmets. Therefore, it is necessary to identify the unsafe behavior of lift personnel. The general algorithm is prone to missed detection and false alarm when identifying the complex background scene of workers wearing safety helmets when driving the lift truck. Based on this, proposes an improved object detection algorithm based on YOLOv5. This algorithm is used to detect unsafe behavior of people driving lift vehicles. The CBAM attention mechanism is added to the improved algorithm of YOLOv5, and the attention mechanism module of channel and space is combined to enhance the detection accuracy of small targets and the extraction ability of features. Then, the original feature pyramid network structure is replaced with the weighted feature pyramid network structure (BIFPAN) to achieve efficient two-way cross-scale connection and weighted feature fusion. The experimental results of self-made safety helmet dataset show that the mean mean accuracy (mAP) of the improved algorithm reaches 92.4%, and the detection accuracy and recall rate are improved by 2.7% and 1.9%, respectively.

Key words: YOLOv5; factory; detection; weighted feature fusion