

DOI:10.19651/j.cnki.emt.2108363

基于改进 YOLOX-S 的安全帽反光衣检测算法^{*}

程换新¹ 蒋泽芹¹ 程力² 成凯¹

(1. 青岛科技大学自动化与电子工程学院 青岛 266061; 2. 中国科学院新疆理化技术研究所 乌鲁木齐 830011)

摘要: 在工业生产和交通工程中,安全帽和反光衣都是员工重要的生命安全保障。针对传统安全帽反光衣识别方法只能检测单一颜色反光衣、检测效率低的问题,提出一种基于改进 YOLOX-S 网络模型的安全帽反光衣检测方法。使用简化 BiFPN 模块替换原加强特征提取网络,提高网络对不同尺度的特征提取能力;使用 Mosaic 方法进行训练,提高网络在复杂场景下的检测能力;使用 GIoU 损失函数,进一步提高模型的识别准确率。在扩充后的安全帽反光衣数据集上实验表明,所提算法在保持较高推理速度的情况下,mAP 达 83.74%,与原 YOLOX-S 相比,对戴安全帽、穿反光衣和行人的检测 AP 值有 1%~3% 不等的提高,对反光衣颜色无依赖性,有效实现了快速准确的安全帽反光衣检测。

关键词: 安全帽反光衣检测; YOLOX-S; BiFPN; Mosaic 方法

中图分类号: TP391.4 文献标识码: A 国家标准学科分类代码: 520.6040

Helmet and reflective clothing detection algorithm based on improved YOLOX-S

Cheng Huanxin¹ Jiang Zeqin¹ Cheng Li² Cheng Kai¹

(1. College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China;

2. Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Urumqi 830011, China)

Abstract: In industrial production and traffic engineering, safety helmets and reflective clothing are important safety protection. Aiming at the problem that the traditional helmet reflective clothing recognition method can only detect single color reflective clothing and low detection efficiency, we proposed a helmet reflective clothing detection method based on improved YOLOX-S network model. The simplified BiFPN module is used to replace the original enhanced feature extraction network to improve the feature extraction ability of the network for targets with different scales. The Mosaic method is used for model training to improve the detection ability of the network in complex scenes. The GIoU loss function is used to further improve the recognition accuracy of the model. Experiments on the expanded helmet reflector data set show that the proposed algorithm can achieve 83.74% mAP while maintaining a high detection speed. Compared with the original YOLOX-S, the detection AP of wearing helmets, reflective clothing and pedestrians is improved by 1%~3%, and there is no dependence on the color of reflective clothing, which effectively realizes the rapid and accurate detection of helmets and reflective clothing.

Keywords: helmet and reflective clothing detection; YOLOX-S; BiFPN; Mosaic method

0 引言

安全帽和反光衣在人类生产生活中,特别是工业生产和交通工程中起着特殊的安全保障作用,广泛用于造船行业、钢铁行业、机械行业、石油行业、救护、机场、交警及户外工作或施工等,是很好的安全警示用品。然而实际生产工作中,总存在员工安全意识薄弱、巡逻人员监察不到位的情

况,时有发生因员工不按规定穿戴安全帽或反光衣从而导致的安全事故,因此利用摄像头自动检测员工是否佩戴安全帽及穿着反光衣具有重要意义。

目前常用的着装检测方法有传统的图像处理法和基于深度学习的目标检测算法。传统方法一般通过检测对象的 HSV 颜色(色相 Huc、饱和度 Saturation、亮度 Value)或形状特征进行识别。文献[1]通过提取图像中人的头盔、上衣

收稿日期:2021-11-16

*基金项目:国家海洋局重大专项项目(国海科字[2016]491号 No. 30)资助

和裤子 3 部分的形状和颜色特征向量,再利用样本 Q 邻域敏感度的径向基神经网络进行着装识别,针对人员直立姿态时的着装识别准确率高,但在弯腰、下蹲姿态下很容易发生误检,且过于依赖服装的颜色特征,易受光线和背景等环境因素干扰。文献[2]将图像通过 HSV 颜色模型转化为二值图像后进行服装检测,另外利用卷积神经网络和迁移学习进行安全帽检测,但只能检测单一颜色的工作服,应用范围窄。针对这些问题,基于深度学习的目标检测算法为着装检测提供了更多的解决方案。

基于深度学习的目标检测算法主要分为一阶段算法(one-stage)和两阶段算法(two-stage)两类^[3]。较经典的 two-stage 目标检测算法有 R-CNN、Fast R-CNN 和 Faster R-CNN 等^[4-6],具有准确率高运算速度慢的特点。文献[7]通过引入 L2 正则项的回归损失函数优化 Faster R-CNN 模型,对污染场地作业人员进行安全帽和防毒面具检测,受光照和拍摄角度影响大。相较而言,one-stage 目标检测算法运行速度快,更符合实际应用需求,代表算法有 SSD 算法和 YOLO 系列算法^[8-12]。文献[13]使用 ResNet50 网络和可形变卷积模块对 SSD 算法进行改进,检测安全头盔和反光衣,但存在多个边界框重叠现象。文献[14]利用 Densenet 取代 YOLOv3 网络的主干进行特征提取,解决了边界框重叠问题,实现人员密集条件下的安全帽检测,但没有进行工作服装检测。针对以上问题,本文提出使用具有更高性能的 YOLOX 检测器进行安全帽和反光衣的检测。

为进一步提高模型在不同尺度下识别安全帽和反光衣的准确度,以及在复杂场景下的检测能力,本文提出基于改

进 YOLOX-S 算法的安全帽反光衣检测方法。在原 YOLOX-S 网络基础上,对 Neck 部分使用简化的 BiFPN 网络进行改进,采用 GIOU_Loss 作为边界框损失函数,经过训练得到适用于识别安全帽反光衣的目标检测模型。最后在自制数据集上对改进前后的网络以及 YOLOX 系列其他网络进行对比实验,测试本文模型的有效性。实验证明本文模型的具有较好的准确性和实时性,在复杂场景下表现出更好的检测能力,可以在施工现场等多种场景下进行应用。

1 相关算法及改进

1.1 YOLOX-S 算法

YOLOX 算法^[13]是旷视科技研究院 2021 年提出的高性能检测器,它将解耦头、数据增强、无锚点以及标签分类等目标检测领域的优秀进展与 YOLO 进行了巧妙的组合,不仅实现了超越 YOLOv3、YOLOv4 和 YOLOv5 的 AP,且取得了极具竞争力的推理速度。YOLOX 系列主要包括 YOLOX-S、YOLOX-M、YOLOX-L、YOLOX-X,其中参数量最少的 YOLOX-S 整体网络结构如图 1 所示,共分为 3 个部分,其中 Backbone 部分即主干特征提取网络使用的是 CSPDarknet 网络,Neck 部分即加强特征提取网络使用的是特征金字塔网络(feature pyramid networks, FPN),Prediction 部分即检测头使用了 3 个解耦头(Decoupled Head)。输入图片在 Backbone 部分进行浅层特征提取,输出 3 个特征层传入 Neck 部分进行深层特征提取,输出分别传入 3 个 Decoupled Head 进行目标框预测,得到检测结果。

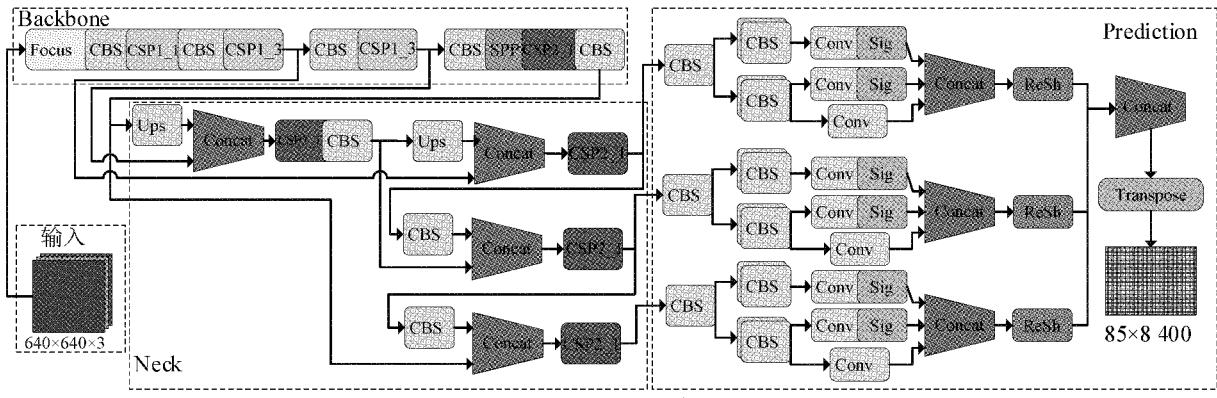


图 1 YOLOX-S 网络结构

1.2 BiFPN

BiFPN 是谷歌团队 2019 年发表的 EfficientDet 算法^[14]中所提出的一种加权双向特征金字塔网络。其网络结构如图 2 所示,其中 $P_3 \sim P_7$ 是 5 个输入特征,经过 BiFPN 中的上采样、下采样和叠加等运算,输出 5 个提取后的特征。在实际网络构建中,可根据特征提取需求,将 BiFPN 当作一个特征网络层,进行多次堆叠。

在特征提取时,不同的输入特征其分辨率也不同,而它们对输出特征的贡献明显是不相等的,因此 BiFPN 网络

为每个输入特征额外添加了可学习的权重,使网络不断调整权重确定每个输入特征对输出特征的重要性。BiFPN 网络采用精度与 softmax-based fusion 相似但速度快 30% 的快速归一化方法,如式(1)所示。

$$Out = \sum_i \frac{\omega_i}{\epsilon + \sum_j \omega_j} \cdot In_i \quad (1)$$

式中: In 和 Out 分别表示输入和输出特征,权重 ω_i 都使用 ReLU 激活函数以保证 $\omega_i \geq 0$, $\epsilon = 0.0001$ 是一个保证数值稳定的小量。

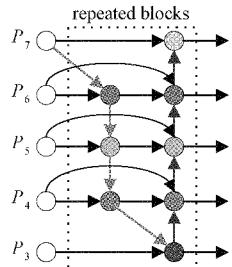


图2 BiFPN网络结构

1.3 改进的 YOLOX-S 网络

对于YOLOX-S主干网络部分,本文保留原来的网络结构,也同样对中间层、中下层和底层这3个特征层进行特征提取,然后传入BiFPN加强特征提取网络。但由于BiFPN网络有5个输入特征层,所以将BiFPN网络简化为3个输入特征层(记为BiFPN-mini),以减少计算量和适应YOLOX网络,BiFPN-mini网络结构如图3所示。

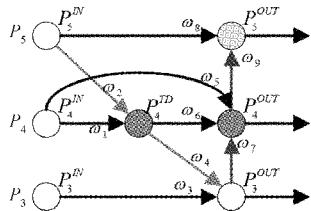


图3 BiFPN-mini网络结构

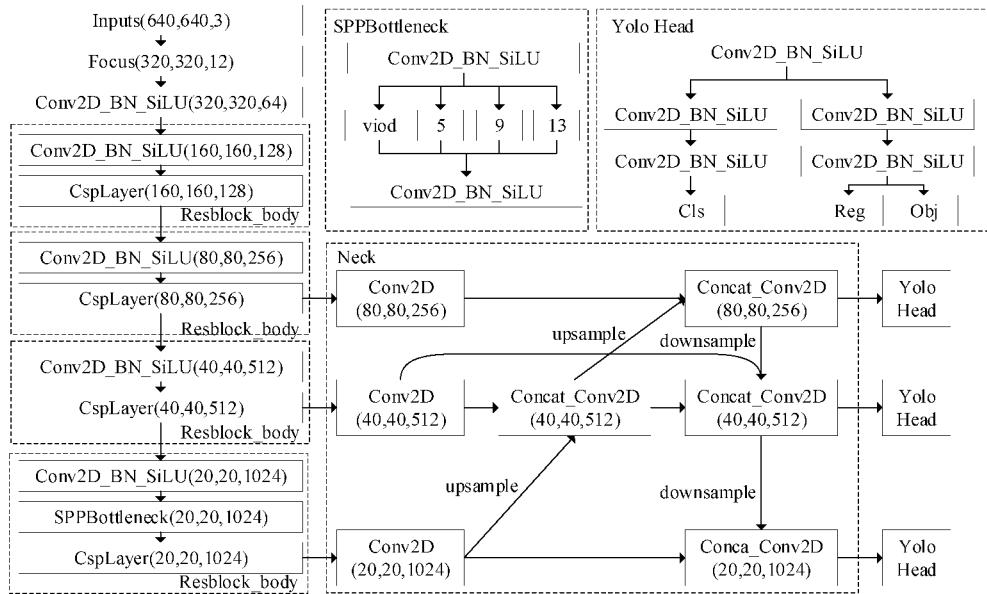


图4 改进后的 YOLOX-S 网络结构

图4中,Reg预测结果的卷积通道数为4,代表预测框的中心点相较于该特征点的偏移情况和预测框的宽高相较于对数指数的参数;Obj预测结果的卷积通道数为1,代表每一个特征点预测框内部包含物体的概率。Cls预测结果的卷积通道数为num_classes,代表每一个特征点对应

当输入大小为(640,640,3)时,BiFPN-mini网络的3个输入特征层分别为 $P_3^{IN} = (80,80,256)$ 、 $P_4^{IN} = (40,40,512)$ 、 $P_5^{IN} = (20,20,1024)$,BiFPN-mini网络的每个特征融合节点都会给每个输入特征分别加权 ω_i ,同时使用快速归一化公式训练这些权重。各融合节点输出的计算公式如式(2)~(5)所示。

$$P_4^{TD} = \text{Conv}\left(\frac{\omega_1 \cdot P_4^{IN} + \omega_2 \cdot \text{Resize}(P_5^{IN})}{\omega_1 + \omega_2 + \epsilon}\right) \quad (2)$$

$$P_3^{OUT} = \text{Conv}\left(\frac{\omega_3 \cdot P_3^{IN} + \omega_4 \cdot \text{Resize}(P_4^{TD})}{\omega_3 + \omega_4 + \epsilon}\right) \quad (3)$$

$$P_1^{OUT} = \text{Conv}\left(\frac{\omega_5 \cdot P_1^{IN} + \omega_6 \cdot P_4^{TD} + \omega_7 \cdot \text{Resize}(P_3^{OUT})}{\omega_5 + \omega_6 + \omega_7 + \epsilon}\right) \quad (4)$$

$$P_5^{OUT} = \text{Conv}\left(\frac{\omega_8 \cdot P_5^{IN} + \omega_9 \cdot \text{Resize}(P_1^{OUT})}{\omega_8 + \omega_9 + \epsilon}\right) \quad (5)$$

式中:Conv表示卷积操作,Resize表示对输入进行上采样(upsample)或下采样(downsample)操作, $\omega_i \geq 0$ 是可学习的权重, $\epsilon = 0.0001$ 是一个保证数值稳定的小量。

改进后的YOLOX整体结构如图4所示。BiFPN-mini网络输出3个加强特征提取后的特征层,其shape大小分别为 $P_3^{OUT} = (80,80,256)$ 、 $P_4^{OUT} = (40,40,512)$ 和 $P_5^{OUT} = (20,20,1024)$ 。经过Decoupled Head预测后,每个特征层获得3个预测结果,分别是目标框的坐标(Reg)、目标框前景背景判断(Obj)、目标框的类别(Cls)。

某类物体的概率,最后一维度num_classes中的预测值代表属于每一个类的概率。

1.4 损失函数

对于预测Obj和Cls用到的损失函数,都采用原YOLOX网络中的二分类交叉损失函数(BCE_loss),但对

于预测 Reg 的边界框损失函数,本文采用了广义交并比损失函数(GIoU_Loss)^[17]。与交并比(intersection over union,IoU)相比,广义交并比(generalized intersection over union,GIoU)具有 4 个特点^[18-19]:

- 1)与 IoU 一样,具有非负性、尺度不变性等特性。
- 2)GIoU 对尺度不敏感。
- 3)GIoU 是 IoU 的下界,取值为 [-1, 1],由于惩罚项的引入,在不重叠的情况下,预测框会向目标框移动。
- 4)GIoU 除了关注重叠区域不同,还关注了非重叠区域,能够更好地反映重合度。

由此可知,只有当预测框和真实框重合时,GIoU_Loss 才会为 0,相比 IoU_Loss,GIoU_Loss 在任意情况下都可以进行训练。GIoU_Loss 计算公式如式(6)、(7)所示。

$$GIoU_Loss = IoU - \frac{A^c - U}{A^c} \quad (6)$$

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{I}{U} \quad (7)$$

式中: A 是预测矩形框, B 是真实矩形框, I 是 A 与 B 交集面积, U 是 A 与 B 并集面积, A^c 表示 A 与 B 的最小外接矩形面积。

2 实验过程

2.1 实验环境

本实验所用计算机处理器型号为 i5-9300H,显卡型号为 NVIDIA GeForce GTX 1600 Ti,内存为 8 GB。操作系统为 Windows10 64 位,神经网络部分在 Pytorch 框架下搭建,软件编程环境为 python3.7。

2.2 数据集与预处理

对于本文的安全帽和反光衣检测方法,为使所构建的模型取得更好的结果,采用百度 AI 平台开源的安全帽反光衣 VOC2007 数据集,共 1 083 张图片,标签包括 helmet、without_helmet、Reflective_clothes、Other_clothes、people,分别代表戴安全帽、没戴安全帽、穿反光衣、穿其他衣服、行人。为满足数据在多样性方面的要求,对现有数据集进行数据预处理,扩充数据集的广度和深度,从而提高模型的鲁棒性。

本文选取 3 种图像处理措施,随机对原图的纹理结构和几何学特征进行改变^[20]:

1)对图片进行水平镜像翻转,增加整个网络在方向上的不变性。

2)对图片添加一定范围的随机椒盐噪声或高斯噪声,提高网络对摄像头失真的不变性。

3)对图片随机改变亮度,模拟同一地点不同亮度下的安全帽和反光衣情况。

经过不同措施预处理的图片如图 5 所示。预处理扩充后的数据集共有 7 581 张图片,符合网络训练需求。将数据集图片按照 9 : 1 的比例划分为训练集和测试集,再将训练集中 20% 图片作为验证集。

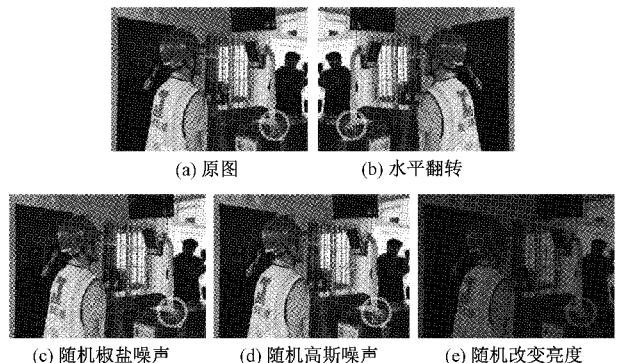


图 5 预处理结果

2.3 训练过程

采取冻结训练方法,以提高训练效率,加速收敛。阈值设为 0.5,迭代次数(epoch)共 500 代,在前 90% 的 epoch 中采取 Mosaic 数据增强方法^[11]训练,其他相关参数设置如表 1 所示,表 1 中 Freeze 和 Unfreeze 分别表示冻结阶段和解冻阶段,Batch_size 表示批次大小。

表 1 训练参数设置

训练阶段	epoch	Batch_size	学习率	优化器
Freeze	100	32	0.000 1	Adam
Unfreeze	400	8	0.000 01	Adam

训练过程的 loss 曲线如图 6 所示,横纵坐标分别为 epoch 和 loss 值。解冻前训练集和验证集 loss 值分别收敛于 0.37、0.40,解冻后训练集和验证集 loss 值分别收敛于 0.27、0.30。

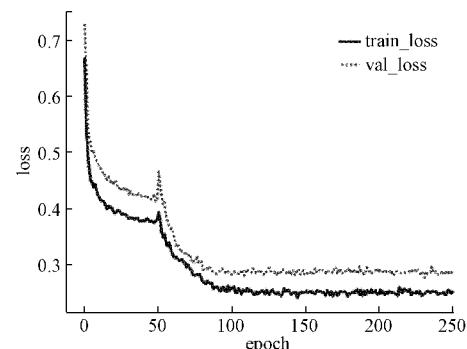


图 6 训练过程 loss 曲线

3 实验结果分析

3.1 评价指标

为了判断模型对穿戴安全帽和反光衣的检测效果,采用平均精度(average precision,AP)、平均精度均值(mean average precision,mAP)、召回率均值(mean recall,mR)、模型大小、每帧推理时间作为评价指标。

其中 AP 是在单个类别下的平均精准度, mAP 是 AP 值在所有类别下的均值, 召回率(Recall)是被正确地识别出来的一类目标个数与测试集中所有此类目标个数的比值, mR 是 Recall 在所有类别下的均值, 式(8)为 mR 计算公式。

$$mR = \frac{N^{tp}}{(N^{tp} + N^{fn}) \cdot num} \quad (8)$$

式中: N^{tp} 为正样本识别正确总数; N^{fn} 为正样本识别错误总数; num 为总类别数。

3.2 实验结果

为验证模型有效性, 在训练参数相同情况下对改进后的 YOLOX-S 网络进行横向纵向对比实验, 比较不同模型

的 AP、mAP、mR、模型大小(Size)和每帧推理时间(t), 在测试集进行实验, 结果如表 2 所示。由于本文改进 YOLOX-S 模型融合了 BiFPN_mini 模块、采用了 90% 代 Mosaic 训练以及 GIoU_Loss 网络损失函数, 因此对这三方面进行横向对比实验。表 2 中, YOLOX-S 表示改进前的原网络(使用 FPN), 改进 YOLOX-S(NM) 和改进 YOLOX-S(IoU) 表示在本文改进 YOLOX-S 基础上分别进行如下操作: 在训练中不采用 Mosaic 方法、iou_loss 部分选择 IoU_loss 损失函数。 AP^h 、 AP^w 、 AP^r 、 AP^o 、 AP^p 分别表示在阈值为 50 时类别 helmet、without_helmet、Reflective_clothes、Other_clothes、people 的 AP 值。

表 2 对比实验结果

模型	Map/%	mR/%	Size/MB	t /ms	AP^h /%	AP^w /%	AP^r /%	AP^o /%	AP^p /%
YOLOX-L	80.74	72.58	207	69	92.14	93.27	79.95	50.94	87.41
YOLOX-M	82.51	76.43	96	53	95.17	93.00	80.63	56.00	87.73
YOLOX-S	83.20	75.99	34	37	94.11	94.75	78.42	61.92	86.78
改进 YOLOX-S(NM)	66.81	39.30	35	39	84.87	83.06	52.80	41.34	71.99
改进 YOLOX-S(IoU)	80.18	66.37	35	39	92.20	93.27	75.00	52.31	88.13
改进 YOLOX-S(本文)	83.74	77.57	35	39	95.58	93.24	79.90	60.31	89.65

从表 2 中可以看出, 本文改进的 YOLOX-S 网络比原网络的 Size 略大, 约有 1 MB, 每帧推理速度增加 2 ms, 但是 mAP 更高, 尤其是针对戴安全帽、穿反光衣和行人这 3 个类别的识别准确率有 1%~3% 不等的提高。与 YOLOX-M 模型相比, mAP 差别不大, 但 YOLOX-M 模型在实时性方面不如本文模型, 因此通过横纵向对比实验, 可以说明本文改进的 YOLOX-S 网络在安全帽反光衣检测方面更具有效性。

mAP 较高的 3 个模型在测试集上的部分检测效果对

比图 7 所示, 图中的检测结果用矩形框表示, 框上是识别的类别标签和相应的概率。通过对比, YOLOX-M 模型对复杂场景下安全帽和反光衣的漏检率较高, 对暗背景下的目标检测能力较差; YOLOX-S 模型对行人检测有较大优势, 但对特征不明显的反光衣存有误检情况; 本文的改进模型对以上问题都有明显改善, 尤其对不同颜色的反光衣和安全帽检测效果著有成效, 进一步证明了本文模型的有效性和实用性。



图 7 测试集检测效果

4 结 论

本文针对安全帽反光衣检测任务提出一种改进 YOLOX-S 算法。为提高网络对不同尺度特征提取能力, 提出使用简化的 BiFPN 网络作为加强特征提取网络, 在保证较高推理速度前提下, 提高了目标识别准确率。采用能更好地反映预测框与先验框重合度的 GIOU 损失函数, 提高了检测精度。增加行人目标检测和使用 Mosaic 方法训练模型, 进一步提高了模型在复杂背景下的检测性能, 使模型更具实用性。实验结果表明, 本文改进的 YOLOX-S 模型能够检测不同颜色的反光衣安全帽, 且在复杂背景和不同亮度条件下表现良好, 与 YOLOX 系列其他模型相比, 在推理速度和检测准确度的平衡上表现更佳, 能更好地应用在安全帽反光衣检测领域。下一步将考虑如何在保证较高准确率前提下, 进一步提高推理速度。

参考文献

- [1] 李聪, 胡金磊, 周俊煌, 等. 基于 S-RBFNN 的变电作业人员着装图像识别算法[J]. 广东电力, 2019, 32(9): 124-130.
- [2] 袁一丹. 基于图像识别的工作人员穿戴规范性检测技术研究[D]. 成都:电子科技大学, 2019.
- [3] JIAO L, ZHANG F, LIU F, et al. A survey of deep learning-based object detection [J]. IEEE Access, 2019, 7:128837-128868.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. IEEE Computer Society. IEEE Computer Society, 2013:580-587.
- [5] GIRSHICK R. Fast R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [6] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39 (6): 1137-1149.
- [7] 刘欣宜, 张宝峰, 符烨, 等. 基于深度学习的污染场地作业人员着装规范性检测[J]. 中国安全生产科学技术, 2020, 16(7): 169-175.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [9] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.
- [10] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv Preprint, 2018, ArXiv: 1804.02767.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv: 2004.10934.
- [12] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single shot multiBox detector [C]. European Conference on Computer Vision, Springer, Cham, 2016:21-37.
- [13] 韩泽佳, 肖秦琨, 张立旗. 基于改进 SSD 安全头盔反光衣检测算法[J]. 自动化与仪表, 2021, 36(9): 63-68.
- [14] WU F, JIN G, GAO M, et al. Helmet detection based on improved YOLOv3 deep model[C]. 2019 IEEE 16th International Conference on Networking, Sensing and Control(ICNSC), IEEE, 2019:363-368.
- [15] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021 [J]. ArXiv Preprint, 2021, ArXiv: 2107.08430.
- [16] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection [J]. ArXiv Preprint, 2019, ArXiv: 1905.11946.
- [17] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019:658-666.
- [18] 刘丹, 汪慧兰, 曾浩文, 等. VoVNet-FCOS 道路行人目标检测算法研究[J]. 国外电子测量技术, 2021, 40(11): 64-71.
- [19] 王宸, 唐禹, 张秀峰, 等. 基于改进 EfficientNet 的锻件磁粉探伤智能检测方法研究[J]. 仪器仪表学报, 2021, 42(9): 89-96.
- [20] 程换新, 蒋泽芹, 成凯. 基于改进 SegNet 模型的斑马线图像语义分割[J]. 电子测量技术, 2020, 43(23): 104-108.

作者简介

程换新,工学博士,教授,主要研究方向为控制科学与工程、图像处理与识别、人工智能等。

蒋泽芹,硕士研究生,主要研究方向为人工智能、深度学习、图像处理与识别等。

程力,工学博士,研究员,主要研究方向为大数据分析、人工智能、互联网网络安全等。

成凯,硕士研究生,主要研究方向为人工智能、深度学习、图像处理与识别等。

E-mail:17854273630@163.com