

使用场景增强的安全帽佩戴检测方法研究

徐传运, 袁含香, 李 刚, 郑 宇, 刘 欢
重庆理工大学 计算机科学与工程学院, 重庆 400054

摘 要:为了解决现有安全帽佩戴数据集样本数量有限导致模型检测精度较低的问题,提出一种基于场景增强的样本扩充算法。该算法将训练集中随机抽取的图像中的检测目标随机缩放后,粘贴到另一随机场景图像上的任意位置,基于现有场景构建出拥有新的检测目标的增强场景,通过场景增强扩充安全帽佩戴训练数据集,增加训练数据集的多样性。为了验证该算法在安全帽佩戴检测中的有效性,使用场景增强算法扩充HelmetWear数据集,并用其训练基于YOLO v4的安全帽佩戴检测模型,通过检测精度评估场景增强算法。在HelmetWear数据集上检测精度达到93.81%,检测精度提升了6.39个百分点。实验结果表明该算法能有效提升安全帽佩戴检测的精度,特别是在样本最为欠缺的小目标上表现更为显著;场景增强算法对解决其他领域目标检测训练数据不足的问题有重要的借鉴意义。

关键词:安全帽检测;场景增强;HelmetWear数据集;YOLO v4

文献标志码:A **中图分类号:**TP183 **doi:**10.3778/j.issn.1002-8331.2103-0030

Research on Safety Helmet Wearing Detection Method Based on Scene Augment

XU Chuanyun, YUAN Hanxiang, LI Gang, ZHENG Yu, LIU Huan

School of Computer Science & Engineering, Chongqing University of Technology, Chongqing 400054, China

Abstract: In order to solve the problem of low model detection accuracy caused by the limited number of samples in the existing helmet dataset, a sample expansion algorithm based on scene augment is proposed. This algorithm randomly scales the small and medium targets in the foreground image randomly extracted from the training set, and pastes it to any position on another random scene image to construct a new real training image with border labels to improve the combination of background and foreground in the dataset diversity. In order to evaluate the effectiveness of the algorithm in the helmet wearing detection, the HelmetWear dataset is expanded using the scene augment algorithm and the scene augment algorithm is evaluated by detection accuracy using its training hard hat wearing detection model based on YOLO v4. Detection accuracy reaches 93.81% on the HelmetWear dataset and improves detection accuracy by 6.39 percentage points. The experimental results show that the algorithm can effectively improve the accuracy of hard hat wearing detection, especially on the small targets where the sample is most lacking. The scene augment algorithm is of great significance to solve the problem of insufficient target detection training data in other fields.

Key words: safety helmet detection; scene-augment; HelmetWear dataset; YOLO v4

在2019年,我国建筑工地中物体打击伤亡事故占所有伤亡事故的比例高达15.91%^[1]。部分原因是作业人员安全意识薄弱,未按规定佩戴安全帽。企业除了加强宣传、强制要求外,采用智能化技术实现作业人员安全帽佩戴的精确自动化监管,也是降低此类事故发生率的有效手段之一。

随着深度学习方法在安全帽佩戴检测中的深入应用,安全帽佩戴检测精度有较大提升。但是,为了提高检测精度,网络模型在逐渐加深过程中,参数量变大,需要的训练样本数量也随之增多。高昂的训练样本采集和标注成本,阻碍了安全帽佩戴智能检测方法的进一步发展。因此样本扩充和数据增强成为在不增加采集和

基金项目:重庆市技术创新与应用发展专项面上项目(cstc2020jscx-msxmX0086);重庆市教委科学技术研究项目(青年)(KJQN202001137);国家重点研发计划项目(2018YFB1308602);重庆市技术创新与应用发展(重大主题专项)项目(cstc2019jscx-zdztzx0031)。

作者简介:徐传运(1979—),男,博士,副教授,CCF会员,研究方向为计算机视觉、机器学习、机器视觉;袁含香(1995—),女,硕士研究生,研究方向为计算机视觉、机器学习,E-mail:1603267963@qq.com;李刚(1961—),男,博士,教授,研究方向为计算机视觉、机器视觉;郑宇(1995—),男,硕士,研究方向为机器学习;刘欢(1997—),女,硕士,研究方向为机器学习。

收稿日期:2021-03-02 **修回日期:**2021-05-07 **文章编号:**1002-8331(2022)19-0326-07

标注成本的情况下,缓解训练样本不足问题的一种有效解决方案。

数据增强通常可分为有监督增强和无监督增强两类^[2]。有监督增强又可分为单样本增强和多样本增强。单样本增强即增强操作围绕着一个样本进行,常见的有图像翻转、剪切、变形、缩放、旋转、添加噪声、图像模糊等。多样本增强则同时使用多个样本产生新的样本,如2017年,Dwibedi等人^[3]对室内环境检测,提出Cut-Paste算法实现不同实例目标和场景的组合,从而得到新的合成数据。同年,Zhang等人^[4]提出的Mixup算法在成对样本及其标签的凸组合上训练神经网络,增强了训练样本之间的线性表达,将CIFAR-10数据集上的错误率降低了1.5%;2018年,Inoue等人^[5]提出sample-pairing算法,从训练集中随机抽取两张图像,分别经过单样本增强处理后经像素取平均值而叠加合成一个新的样本,将CIFAR-10数据集上的错误率从8.22%降低到6.93%;2019年,Hussein等人^[6]提出的SMOTE算法为解决样本类别不平衡问题,对少数类样本进行分析并使用聚类算法根据少数类样本人工合成新样本添加到数据集中,一定程度上增加了少数类的样本数量;张家伟等人^[7]同样使用此种方法提高了少数类样本分类的准确性;2020年,Bochkovskiy等人^[8]提出的YOLO v4中则使用Mosaic方法将随机生成的四张图片进行拼接,以增加训练数据的多样性。无监督数据增强算法可分为:(1)利用模型学习数据的分布,随机生成与训练数据集分布一致的图像数据,如2020年,杜井然等人^[9]采用邻域差分滤波器和生成对抗网络(GAN)相融合的方法,对工业产品数据集进行样本重建;同年,孙晓等人^[10]利用生成器由某一表情下的面部图像生成同种静态图像数据的增强方法。(2)通过模型学习适合当前任务的数据增强方法,如2019年,Google提出的自动选择最优数据增强方案即AutoAugment^[11]算法,在CIFAR-10数据集上误差降低了0.65%。

在前期实验中,本文尝试采用以上算法对安全帽数据进行扩充。其中,基于Mosaic的增强方法未能提高安全帽佩戴的检测精度;基于GAN的增强方法需要大量数据来拟合样本的分布,且不能任意改变图像中目标的数量、大小和位置,因此不适合用于安全帽佩戴检测任务;基于Cut-Paste^[3]的增强方法需要把训练集中的目标进行精细分割,标注成本太高;基于sample-pairing的增强方法能提高安全帽佩戴的检测精度,但效果并不理想,精度还需进一步提升,相关结果也在实验部分进行了阐述。

因此本文针对安全帽数据集样本偏少而导致网络检测精度较低的问题,采用有监督的多样本增强方法,提出一种基于场景增强(scene-augment,SA)的样本扩充算法。该算法将从训练集中随机提取前景图像,对其

中小目标进行剪切、随机缩放后,粘贴到随机选择的另一源场景图像的任意位置,以较小的成本扩展出大量的可训练数据。另外,本文还将SA算法添加到YOLO v4网络模型(该组合称为YOLO v4(SA))中,对HelmetWear^[12]数据集(HW)进行安全帽佩戴检测。实验表明,该方法相比其他学者在HW数据集上的测试结果,检测精度得到了进一步提升。

1 场景增强

场景增强是指把训练数据中的前景和背景抽取出来,将二者随机组合以构建出数据集中原本不存在的场景,可以应用于训练样本不足或训练样本分布不均匀的情况。

本文使用HW数据集进行安全帽佩戴检测实验,训练集有2 272张图像,参照MS COCO^[13]数据集中按面积大小对不同尺度目标的定义见表1,统计出HW数据集内的小、中、大目标的数量分别为285例、2 736例、4 134例,其中目标分为不戴安全帽和戴安全帽两类,分别使用0、1标注。

表1 HW数据集中目标数量统计

Table 1 Statistics of number of targets in HW dataset

	小目标	中目标	大目标
样本面积	(0,32×32]	(32×32,96×96]	(96×96,+∞)
样本数量	285	2 736	4 134

该数据集目前存在以下两种问题:(1)0类主要由生活场景中的行人构成、1类主要由建筑工地场景中作业人员构成,后者作业人员90%都佩戴安全帽,且黄色安全帽佩戴者数量占一半以上,这种呈现重尾分布的数据会使得网络一方面倾向于将建筑工地场景中的工作人员预测为佩戴安全帽一类、且更注重对佩戴黄色安全帽工作人员特征的学习,另一方面倾向于将生活场景中的行人预测为不戴安全帽一类。(2)由于大目标的数量占比高达57.78%,训练时匹配的锚框数量也相对较多,因此网络极易倾向于预测大目标,而忽略中小目标。

本文针对上述两种问题提出了一种简单有效的解决方案,即将随机抽取的前景图像内的中小目标剪切、随机缩放后,粘贴到另一源场景图像的任意位置上,构建出新的增强场景,如图1所示。粗略估算,HW数据集增强后的场景图像可以多达500万张,能极大丰富训练样本的多样性。实验结果表明,该算法结合YOLO v4网络模型能够稳定提升安全帽佩戴检测模型的精度。

场景增强能构建出新的场景,但是会带来以下三个新的问题:(1)粘贴目标可能会遮挡源场景图像中的目标;(2)随机粘贴可能导致粘贴位置周围像素梯度差较大,目标不能无痕的融入到背景中;(3)过多的复制样本可能导致模型过拟合。

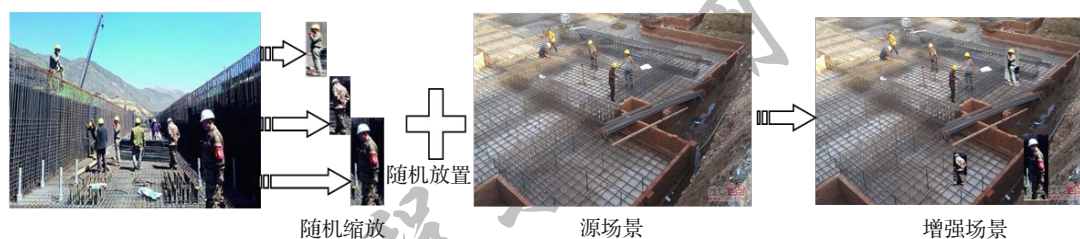


图1 场景增强

Fig.1 Scene augment

针对第一个问题,本文分析了四大类共14种不同的遮挡情况,并对每种遮挡情况进行处理。如图2所示,其中B为前景图像中的目标,A为源场景图像中的目标。第一类遮挡情况中,B在A的4个角的位置上遮挡,为了保持边界框的完整性,此标注框不予以处理;在第二及第三类遮挡情况中,B对A的4条边分别产生了部分及全部遮挡,为保留这两者中较大目标的个体完整性,因此删除较小目标被遮挡的部分;第4类遮挡情况中,B完全遮挡了A,A在增强的场景中没有任何显示,因此舍弃目标B;当B的面积较A目标较小时,能够模拟现实中目标正常遮挡情况,因此保留B目标。

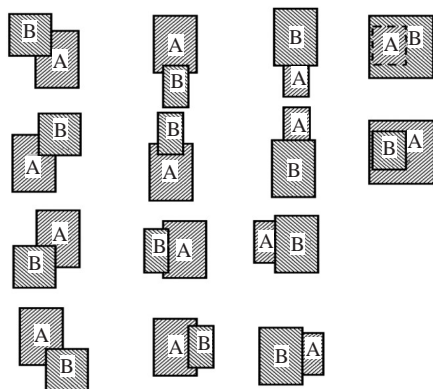


图2 4类共14种不同的遮挡情况

Fig.2 Four types of 14 different occlusion situations

针对第二个问题,文献[3]证明了目前的检测器相对于全局场景布局更加关注局部区域的特征,例如安全帽检测器主要关注工作人员的视觉外观及与背景的融合,而不是在场景中出现的位置(如天上或地面)。同样,文献[8]中提出的Mosaic方法,对于拼接处像素梯度突变的情况并未做任何处理,且在MS COCO数据集上检测精度提升了0.5个百分点。像素突变对检测模型的训练是个不利因素,但现有的研究成果和本文的实验结果表明,像素突变对模型检测精度的影响是有限的,场景增强对检测精度的提升足以抵消其不利影响。

针对第三个问题,在使用场景增强时需要控制增强的强度,以保证模型在不发生过拟合的情况下,提高网络的检测精度。不同增强强度的对比实验分析将在第3章中详细阐述。

在安全帽佩戴检测应用中,目标可以出现在施工场景中的任意位置,即不对目标的位置施加约束,不仅能

扩大安全帽佩戴检测技术的应用范围,也能提升检测方法的鲁棒性。在提出的场景增强方法中,我们假设检测目标可以以任意大小的形态出现在任意场景中的任意位置。以上假设使得增强场景和真实的安全帽佩戴检测场景更加一致,从而提高检测模型在真实应用场景中的检测精度,也能提高模型的鲁棒性。

2 场景增强算法

目标检测模型需要具有平移不变、旋转不变、缩放不变等基本特性,为了训练检测模型具有这些特性,需要使用不同位置、不同大小、不同倾斜角度的训练数据。场景增强SA(scene-augment)算法对安全帽佩戴数据集中抽取的检测目标和抽取的背景进行组合,以最小的代价生成大量拥有不同位置、不同大小、不同倾斜角度的检测目标的场景图像。

设训练集 X 中图像的数量为 N ,第 i 张图像为 x_i ,其中 $i=1,2,\dots,N$, S_i^{dj} 为 x_i 中 d 等尺度目标的第 j 个目标, $d \in \{s, m, l\}$, s, m, l 分别代表小、中、大尺度的目标, y_i^{dj} 是 x_i 图像内 d 等尺度目标的第 j 个目标的标签。 x_a 为待增强的源场景图像, x_b 为从训练数据集 X 中随机抽取含有检测目标的图像,其中 x_a 内待检测目标的数量有 t 个,分别为 $\{S_a^1, S_a^2, \dots, S_a^t\}$, x_b 内小目标数量有 k 个,分别为 $\{S_b^{s1}, S_b^{s2}, \dots, S_b^{sk}\}$,中目标数量有 q 个,分别为 $\{S_b^{m1}, S_b^{m2}, \dots, S_b^{mq}\}$,将 x_b 内的中小目标 $\{S_b^{m1}, S_b^{m2}, \dots, S_b^{mq}\}, \{S_b^{s1}, S_b^{s2}, \dots, S_b^{sk}\}$,复制粘贴到场景 x_a 中,并对粘贴的中目标进行 $\alpha(\alpha \in [0.5, 1])$ 倍的随机缩放,即缩放后的中目标为 $\{a_1 S_b^{m1}, a_2 S_b^{m2}, \dots, a_q S_b^{mq}\}$,增强后的场景的检测目标为 $\{S_a^1, S_a^2, \dots, S_a^t, S_b^{s1}, S_b^{s2}, \dots, S_b^{sk}, a_1 S_b^{m1}, a_2 S_b^{m2}, \dots, a_q S_b^{mq}\}$ 。粘贴的位置在源场景图像中随机选择,但需要确保粘贴后目标不能超出原始图像的范围。在源训练场景图像上添加新的目标后,需要根据前面讨论过的遮挡处理方法处理遮挡。

为保证训练的安全帽佩戴检测模型具有较高的泛化能力和鲁棒性,在训练模型时,场景增强的数据与未增强的数据需要混合使用。实验结果表明,过少的增强数据不能起到提升检测精度的目的,过多的增强数据反而会降低模型的精度。为了控制场景增强的强度,在模

型训练的每个 Mini-Batch 中以概率 p 选择部分图片作为源场景图像,经场景增强后替换原始图片,场景增强后图片数量的计算公式如下:

$$SUM=D\times BS\times p$$

(1)

其中, SUM 为使用 SA 增强后的图片数量, D 为训练迭代次数, BS 为每次迭代的 batch size 大小, p 为增强强度。本文中, D 为 8 000, BS 为 64, 如果 p 为 0.3, 那么扩充的图片数量为 153 600 张。

综合以上论述,设计出 SA 数据增强算法。具体实现如下:

- Begin
- 输入:输入 Mini-Batch 大小为 z 的图像。
- 输出:增强后 Mini-Batch 图像的图像 (\tilde{x}, \tilde{y}) 。
1. 从 Mini-Batch 中以 p 的概率抽取源场景图像进行有放回
- 的 SA 增强。
2. 对每张源场景图片 (x_a, y_a) 做以下处理。
3. 从训练集 X 中随机抽取一张 (x_b, y_b) 。
4. 将 x_b 中的中目标进行 $\alpha(\alpha \in [0.5, 1])$ 倍的随机缩放,生成 $\{a_1S_b^{m1}, a_2S_b^{m2}, \cdots, a_qS_b^{mq}\}$;
5. 将 x_b 中的 $\{S_b^{m1}, S_b^{m2}, \cdots, S_b^{mq}\}, \{S_b^{s1}, S_b^{s2}, \cdots, S_b^{sk}\}$ 粘贴到 x_a 的随机位置,构成增强场景中新的目标 $\{S_a^{l1}, S_a^{l2}, \cdots, S_a^{lq}, S_a^{s1}, S_a^{s2}, \cdots, S_a^{sk}, a_1S_b^{m1}, a_2S_b^{m2}, \cdots, a_qS_b^{mq}\}$, 需要确保粘贴后目标不能超出原始图像的范围。
6. 检查增强目标 $\{S_a^{l1}, S_a^{l2}, \cdots, S_a^{lq}, S_a^{s1}, S_a^{s2}, \cdots, S_a^{sk}, a_1S_b^{m1}, a_2S_b^{m2}, \cdots, a_qS_b^{mq}\}$ 的遮挡情况,根据图 2 的情况处理遮挡,构成场景增强图像 (\tilde{x}, \tilde{y}) 。
7. 用 (\tilde{x}, \tilde{y}) 替换 (x_a, y_a) 。
- End

如图 3 展示了把 SA 数据增强算法与目标检测网络模型结合的安全帽检测框架,训练集中样本数据经 SA 增强后得到新的训练图片及标签,对其进行基础变换后再将其送入 YOLO v4 网络模型中进行训练。

3 实验结果与分析

3.1 数据预处理

在实验中发现 HW 数据集中有部分目标没有标注,

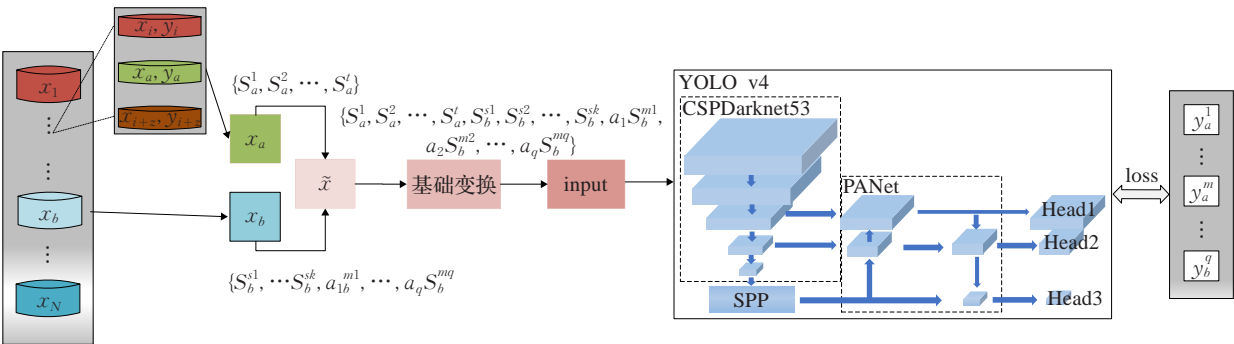


图3 基于场景增强的安全帽佩戴检测过程

Fig.3 Safety helmet wearing detection process based on scene augment

这无疑会影响安全帽佩戴检测模型的性能,为了提高 HW 数据集的质量,使用 LabelImg 工具对部分遗漏的目标进行了补充标注。经过标注后的 HW 数据集标识为 HW+, 其中, HW+ 训练集中图像数量总和不变, 检测目标增加了 570 例, 其中小目标 395 例、中目标 157 例、大目标 18 例, 测试集增加了检测目标 102 例, 其中小目标 66 例、中目标 29 例、大目标 7 例, 如表 2 所示。表中 Δ_t 、 Δ_c 分别代表训练集、测试集中新增的小、中、大目标的数量。图 4 中展示了补充标注前后的对比图, 其中图 (a)、(b)、(c) 各列分别为原图、原始标注情况、补充标注后的情况。

表2 HW 数据集补充标注后的目标数量

Table 2 Number of targets after supplementary annotation in HW

目标	训练集			测试集		
	HW	HW+	Δ_t	HW	HW+	Δ_c
小目标	285	680	395	20	86	66
中目标	2 736	2 893	157	347	376	29
大目标	4 134	4 152	18	735	742	7
0 类	3 646	3 914	268	532	587	55
1 类	3 511	3 813	302	570	617	47



(a)原图 (b)原始标注情况 (c)补充标注后的情况

图4 补充标注前后对比

Fig.4 Comparing before and after supplementary annotation

为了验证补充后是否提高了模型的检测精度,分别使用HW、HW+训练YOLO v4,并在两个数据集的测试集上交叉测试,结果见表3。使用HW训练出的模型的测试mAP分别为92.29%、89.21%,使用HW+训练出的模型的测试mAP分别为92.86%、91.29%。纵向对比两组结果可知,HW+数据集对每个测试集都有相应的提高,说明添加标签能够一定程度上提高模型的检测精度。

表3 补充标注前后的模型检测精度对比
Table 3 Comparison of model detection accuracy before and after supplementary annotation 单位:%

训练集	测试集	
	HW	HW+
HW	92.29	89.21
HW+	92.86	91.29

3.2 网络训练

实验基于PyTorch框架实现,硬件配置为:两张NVIDIA RTX 2080Ti,配备22 GB图形缓存,32 GB系统内存,Intel I9 9900K CPU;软件配置为ubuntu16.04,64位操作系统。实验测试评价指标为平均准确率(average precision,AP),描述了每个类别准确率的平均值,其中mAP、AP^s、AP^m、AP^l分别表示AP的均值和小、中、大3种不同尺度物体的AP值。

训练样本采用补充标注后的HW+数据集,并利用VOC2007数据集处理方法将其输入到YOLO v4网络模型中分两阶段训练^[8]:第一阶段冻结网络预训练权重只更新检测头的参数;第二阶段更新整个网络的参数。batch size分别为64、16,训练及测试图像大小均为416×416,每个实验训练迭代8 000次。表4的结果展示了分别使用余弦退火(cosine annealingLR)^[14]、自适应调整(reduce LR on Plateau)^[15]学习率下降策略及标签平滑(label smoothing)^[16]训练技巧对安全帽检测精度的影响。结果表明余弦退火、标签平滑对安全帽佩戴的检测精度起到了削弱作用,自适应调整能够有效提高模型检测精度,因此以下实验将采用自适应学习率下降策略,且取消标签平滑操作。

表4 模型调优
Table 4 Model tuning 单位:%

模型	mAP
余弦退火	92.86
自适应调整	93.02
标签平滑	92.34

为消除YOLO v4自带Mosaic数据增强算法的干扰,接下来的实验选择将其剔除,另外为和其他学者在HW数据集上的测试结果进行对比,以下将使用HW测试集验证测试。

本文还在同样的实验环境下替换多种数据增强算法对HW+训练集进行训练从而得到对应的检测模型,以便于在HW测试集上与SA的检测结果进行对比。

3.3 面向小目标的场景增强实验

小目标检测是计算机视觉中最具挑战性和重要的问题之一,为了验证场景增强在小目标检测上的有效性,设计了面向小目标的场景增强实验。实验只选择(x_b, y_b)中的小目标粘贴到源场景(x_a, y_a)中,测试了Mini-Batch中图像进行SA增强的概率 p 对检测精度的影响,表5展示了实验结果。当 $p=0.3$ 时,即10张图像中有3张进行SA增强,在测试集上检测的mAP由93.02%提高到93.41%,其中,小目标检测的mAP由33.65%提高到35.09%,由结果可知,当 $p=0.3$ 时,SA增强算法能够提高模型的检测精度,且在小目标检测上的表现尤为明显。

表5 小目标增强的对比实验结果
Table 5 Contrastive experimental results of small target enhancement 单位:%

p	mAP	AP ^s
0	93.02	33.65
0.1	92.39	35.05
0.3	93.41	35.09
0.5	92.88	39.12
0.7	91.91	33.36
1.0	91.99	29.09

为了验证算法的鲁棒性,训练了10组 $p=0.3$ 时的检测模型,并测试其检测精度,实验结果参见表6,可知,10组实验检测精度的平均值为93.42%,标准差为0.07%。从结果可以看出,场景增强算法能稳定提升安全帽佩戴检测模型的检测精度。

表6 $p=0.3$ 时多次重复实验结果
Table 6 Results of repeated experiments when $p=0.3$ 单位:%

次数	mAP	AP ^s
1	93.41	35.09
2	93.42	36.01
3	93.36	35.00
4	93.54	40.58
5	93.35	35.62
6	93.51	38.46
7	93.43	36.23
8	93.31	38.22
9	93.45	37.43
10	93.39	36.99
平均值	93.41	36.96
标准差	0.07	1.75

3.4 基于中小目标的场景增强实验

为增强中小目标的样本数量,尝试在小目标与场景融合的同时,对中目标采用剪切缩放并融合的方法。实验结果见表7,当 $p=0.1$ 及 $p=0.3$ 时模型检测精度均有提升,但随着 p 的增大,扩充样本增多,模型过拟合的概率越大,模型检测精度逐渐下降。当 $p\leq 0.3$ 时,能够在

保证模型不发生拟合的前提下,有效提高模型的检测精度,且这种设置比只融合小尺度目标效果更好。

表7 不同概率对中小目标增强的实验对比
Table 7 Experimental comparison of enhancement of small and medium targets with different probabilities

单位: %			
P	mAP	AP ^s	AP ^m
0	93.02	33.65	80.47
0.1	93.81	35.08	82.79
0.3	93.67	35.09	81.00
0.5	92.73	39.60	78.20
0.7	92.14	36.25	79.48
1.0	86.83	30.00	78.84

图5展示了不同安全帽佩戴检测方法的检测结果对比,其中图(a)、(b)、(c)各列分别为原始测试图像、YOLO v4原始模型检测结果、YOLO v4(SA)的检测结果。可以看出,相对于添加Mosaic的YOLO v4模型检测结果,YOLO v4(SA)模型能更加准确地检测出小目标的位置。



(a)原始测试图像 (b)YOLO v4原始模型检测结果 (c)YOLO v4(SA)的检测结果
图5 SA算法检测结果对比

Fig.5 Comparison of detection results of SA algorithm

3.5 数据增强方法对比实验

面向目标检测的数据增强方法有多种,为了验证这些方法对安全帽佩戴检测的有效性,本文将其分别添加到YOLO v4网络模型中对HW+数据集进行训练并对实验结果对比分析,结果如表8。从实验结果可见,

表8 不同数据增强算法实验对比
Table 8 Experimental comparison of different data enhancement algorithms

单位: %				
增强方式	mAP	AP ^s	AP ^m	AP ^l
基础变换	93.02	33.65	80.47	91.05
Mosaic ^[8]	93.01	33.9	81.9	92.03
randome rasing ^[17]	92.97	35.01	81.45	91.33
sample pairing ^[5]	93.40	38.66	81.97	90.37
SA(ours)	93.81	35.08	82.79	91.15

Mosaic增强^[8]、random erasing^[17]增强没有提高安全帽佩戴检测的精度,sample pairing^[5]增强能提高安全帽佩戴检测的精度,但是实验精度要低于SA增强,进一步说明SA算法相比目前主流数据增强算法在提升安全帽检测精度上具有一定的优势。

3.6 不同模型对比实验

图6分别显示了YOLO v2^[18],MD-YOLO v2^[12],YOLO v3^[19],YOLO v4^[8](SA)5个框架在HW测试集上的测试结果。其中MD-YOLO v2的实验结果由方明等人^[12]提供(是当前公开发布的最好检测结果),其他结果由测试得到。可知,YOLO v4(SA)相比YOLO v2、MD-YOLO v2、YOLO v3分别提高了7.29个百分点,6.39个百分点,5.57个百分点。YOLO v4(SA)方法相比YOLO v4,能够在参数量不变的情况下,对安全帽佩戴检测效果有较好的提升。

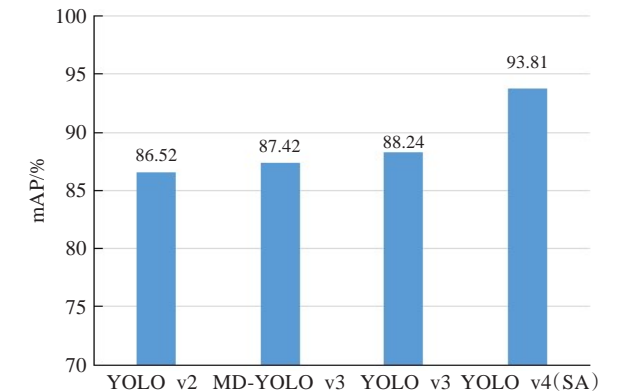


图6 不同模型的测试结果对比
Fig.6 Comparative test results of different models

4 结束语

为了解决现有安全帽佩戴数据集样本数量有限而导致模型检测精度较低的问题,本文提出了一种基于场景增强的样本扩充算法并将其应用到YOLO v4网络模型中对安全帽佩戴情况进行检测,结果表明,合适的场景增强强度结合YOLO v4网络模型能有效提高检测精度。

但由于在场景增强时,目标被粗暴的粘贴进源场景图像中,目标和背景的拼接处将会有较大的梯度突变,这可能会影响增强数据集的训练效果,在未来的研究中,计划将目标与背景融合成自然的增强场景并提高场景构建的随机性。

参考文献:

[1] 住房和城乡建设部办公厅关于2019年房屋市政工程生产安全事故情况的通报[EB/OL].(2020-06-19)[2021-02-20]. http://www.mohurd.gov.cn/wjfb/202006/t20200624_246031.html.
Circular of the General Office of the Ministry of Housing

- and Urban Rural Development on production safety accidents of municipal housing projects in 2019.[EB/OL]. (2020-06-19)[2021-02-20].http://www.mohurd.gov.cn/wjfb/202006/t20200624_246031.html.
- [2] 言有三.深度学习中的数据增强方法都有哪些?[EB/OL]. (2019-04-22)[2021-02-20].https://blog.csdn.net/weixin_43876801/article/details/102958247.
- YAN Y S.What are the data enhancement methods in deep learning?[EB/OL]. (2019-04-22)[2021-02-20].https://blog.csdn.net/weixin_43876801/article/details/102958247.
- [3] DWIBEDI D,MISRA I,HEBERT M.Cut,paste and learn: Surprisingly easy synthesis for instance detection[C]//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV), 2017.
- [4] ZHANG H,CISSE M,DAUPHIN Y N,et al.Mixup:Beyond empirical risk minimization[J].arXiv:1710.09412, 2017.
- [5] INOUE H.Data augmentation by pairing samples for images classification[J].arXiv:1801.02929, 2018.
- [6] HUSSEIN A S,DIALLO B,LIU J.EBSMOTE:Evaluation-based synthetic minority oversampling Technique for imbalanced dataset learning[C]//Proceedings of IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 2019:1226-1233.
- [7] 张家伟,郭林明,杨晓梅.针对不平衡数据的过采样和随机森林改进算法[J].计算机工程与应用, 2020, 56(11):39-45.
- ZHANG J W, GUO L M, YANG X M.Improved oversampling and random forest algorithm for imbalanced data[J].Computer Engineering and Applications, 2020, 56(11):39-45.
- [8] BOCHKOVSKIY A,WANG C Y,LIAO H Y M.YOLOv4: Optimal speed and accuracy of object detection[J].arXiv: 2004.10934, 2020.
- [9] 杜卉然,许亮,吕帅.基于邻域差分滤波生成式对抗网络的数据增强方法[J].计算机应用研究, 2020, 37(6):301-305.
- DU H R,XU L,LYU S.Data enhancement method based on neighborhood differential filter generative countermeasure network[J].Application Research of Computers, 2020, 37(6):301-305.
- [10] 孙晓,丁小龙.基于生成对抗网络的人脸表情数据增强方法[J].计算机工程与应用, 2020, 56(4):115-121.
- SUN X,DING X L.Data augmentation method based on generative adversarial networks for facial expression recognition sets[J].Computer Engineering and Applications, 2020, 56(4):115-121.
- [11] CUBUK E D,ZOPH B,MANÉ D,et al.AutoAugment: Learning augmentation strategies from data[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019:113-123.
- [12] 方明,孙腾腾,邵桢.基于改进YOLOv2的快速安全帽佩戴情况检测[J].光学精密工程, 2019(5):1196-1205.
- FANG M,SUN T T,SHAO Z.Fast helmet wearing condition detection based on improved YOLOv2[J].Optics and Precision Engineering, 2019(5):1196-1205.
- [13] LIN T,MAIRE M,BELONGIE S,et al.Microsoft COCO: Common objects in context[C]//Proceedings of European Conference on Computer Vision, 2014:740-755.
- [14] LOSHCHILOV I,HUTTER F.SGDR:Stochastic gradient descent with restarts, 2016[J].arXiv:1608.03983, 2016.
- [15] ZEILER M D.ADADELTA:An adaptive learning rate method[J].arXiv:1212.5701, 2012.
- [16] MO X,WEI T,ZHANG H,et al.Label-smooth learning for fine-grained visual categorization[C]//Proceedings of Asian Conference on Pattern Recognition, 2019:17-31.
- [17] ZHONG Z,ZHENG L,KANG G,et al.Random erasing data augmentation[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2017.
- [18] REDMON J,FARHADI A.YOLO9000: Better, faster, stronger[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017:6517-6525.
- [19] REDMON J,FARHADI A.YOLOv3: An incremental improvement[J].arXiv:1804.02767, 2018.