

doi:10.3969/j.issn.1001-4616.2023.02.014

# 面向复杂场景的基于改进 YOLOX\_s 的安全帽检测算法

江新玲<sup>1</sup>, 杨 乐<sup>1</sup>, 朱家辉<sup>1</sup>, 陶 磊<sup>1</sup>, 刘 峰<sup>1,2</sup>, 段倩倩<sup>1</sup>

(1. 太原理工大学信息与计算机学院, 山西 晋中 030600)

(2. 中国煤炭工业协会, 北京 100010)

**[摘要]** 在工业生产过程中, 安全帽是生产工人重要的安全保护工具. 针对现有安全帽检测算法在复杂应用场景下对小目标、密集目标以及遮挡目标存在漏检、检测精度较低等问题, 提出了一种基于 YOLOX\_s 的改进算法. 首先, 通过改进 YOLOX\_s 算法的模型结构, 在原有网络结构的基础上新设立了一个预测特征层, 其尺寸为 160×160, 该预测特征层通过将高层语义信息和低层传递的位置信息进行有效融合来预测小目标; 其次, 针对复杂的安全帽检测环境, 将 obj\_loss 的 BCE\_Loss 改为 Focal\_Loss, 即用 Focal\_Loss 来训练 obj 分支来降低漏检; 最后, 将 CSP1\_X 中的残差块改为 shuffleNet 基本单元以缩减参数量. 改进后的算法 mAP 和 recall 分别提高了 1.25% 和 2.32%, 参数量缩减为 3.61MB. 改进后的算法有效降低了复杂环境下安全帽的漏检率 and 提高了检测精度, 对实际生产过程中保障企业和工人的生命财产安全起到了一定的促进作用.

**[关键词]** 安全帽检测, YOLOX\_s, 遮挡目标, 小目标, 密集目标

**[中图分类号]** TP391 **[文献标志码]** A **[文章编号]** 1001-4616(2023)02-0107-08

## Safety Helmet Detection Algorithm Based on Improved YOLOX\_s for Complex Scenes

Jiang Xinling<sup>1</sup>, Yang Le<sup>1</sup>, Zhu Jiahui<sup>1</sup>, Tao Lei<sup>1</sup>, Liu Feng<sup>1,2</sup>, Duan Qianqian<sup>1</sup>

(1. College of Information and Computer, Taiyuan University of Technology, Jinzhong 030600, China)

(2. China National Coal Association, Beijing 100010, China)

**Abstract:** In the process of industrial production, safety helmet is an important safety protection tool for production workers. Aiming at the problems of missing detection and low detection accuracy of the existing safety helmet detection algorithm for small targets, dense targets and occlusion targets in complex application scenarios, an improved algorithm based on YOLOX\_s is proposed. Firstly, by improving the model structure of YOLOX\_s algorithm, a new prediction feature layer with a size of 160×160 is established on the basis of the original network structure, the prediction feature layer can predict small targets by effectively fusing high-level semantic information and low-level location information. Secondly, for the complex helmet detection environment, BCE\_Loss of obj\_loss is changed to Focal\_Loss, that is, Focal\_Loss is used to train obj branch to reduce missed detection. Finally, residual block is changed in CSP1\_X to the shuffleNet base unit to reduce parameters number, mAP and recall of the improved algorithm are increased by 1.25% and 2.32% respectively, and parameters number is reduced to 3.61MB. The improved algorithm effectively reduces the missed detection rate and improves the detection accuracy of safety helmet in complex environment, and plays a certain role in the actual production process to ensure the property and life safety of enterprises and workers.

**Key words:** safety helmet detection, improved YOLOX\_s, occlusion targets, small targets, dense targets

随着科技的快速进步以及经济的飞速发展, 安全帽佩戴检测在智能视频监控领域应用广泛. 在特定作业区域里, 高空坠物有时会伤及工作人员, 因此任何进入这些特定工作区域的人员都应该正确佩戴安全

收稿日期: 2022-10-16.

基金项目: 国家重点研发计划项目(2020YFB1314001)、山西省基础研究计划项目(20210302124029)、山西省重点研发计划项目(202102030201012)、山西省重点研发计划项目(202102100401015).

通讯作者: 杨乐, 博士, 副教授, 研究方向: 计算机视觉、图像处理和三维显示. E-mail: yangle@bupt.cn

帽以防止发生意外,正确佩戴安全帽是保障工人和企业生命、财产安全的重要安全措施。然而,在作业现场的安全管理中存在着许多问题。首先,对于施工人员,安全教育的全面覆盖是困难的,某些施工人员风险意识低,不按要求佩戴安全帽;其次,由于作业环境复杂、作业现场分布广、作业地点多,实现对生产现场的全方位、实时的监督管理困难;最后,大多数企业和监管部门仍然依靠专人监管施工人员是否戴安全帽,人工监管效率较低,实时性差。在安全帽佩戴检测问题上,计算机视觉技术可以在有效监管的同时大大降低人工成本。

安全帽检测主要分为传统的方法和基于深度学习的目标检测算法。基于前者,Wen等<sup>[1]</sup>提出了一种基于改进 Hough 变换的 ATM 监控系统安全帽检测算法,由于安全帽的轮廓是一个圆弧,当检测到 ATM 监控视频画面中存在圆弧,且圆弧半径大于某个阈值时则认为检测到安全帽。Du等<sup>[2]</sup>提出了一种结合机器学习和图像处理的视频序列安全帽检测方法,该方法主要包含三部分:第一部分是基于类 Haar 特征的人脸检测,在此阶段,检测所有可能的人脸区域。第二部分是运动和颜色过滤。在人脸检测之前,首先检测视频序列中的运动。如果视频中没有运动或几乎没有运动,就不进行任何操作。如果存在运动,则获取运动区域,并仅对这些区域进行人脸检测,以避免扫描整个图像。这可以节省计算时间并消除后台的假警报。人脸检测后,可能仍然会有一些人脸的假警报,使用面部肤色过滤掉非面部块来消除假警报。第三部分是安全帽颜色检测,若人脸正上方区域的颜色与安全帽的颜色相似,就认为佩戴了安全帽。Waranusast等<sup>[3]</sup>介绍了一种摩托车安全帽检测的机器视觉技术,先用背景减法提取出运动对象,然后用 KNN (K-Nearest Neighbor)将提取的运动对象分为摩托车或其他对象。为了减少计算量,作者仅处理摄像头范围内出现摩托车的一个实例,而非处理摄像头拍摄的所有图像帧,并从该实例的上部统计骑手头的个数,然后将头部进行分类,分为佩戴头盔和未佩戴头盔的,该系统的检测率为 74%,有效减轻了警务执法的工作量。然而,以上传统的目标检测算法大多计算量大,难以部署在移动端,实现实时检测。

基于深度学习的目标检测算法主要分为两阶段(two-stage)算法和一阶段(one-stage)算法<sup>[4]</sup>。前者首先利用 RPN<sup>[5]</sup>(Region Proposal Networks)网络生成感兴趣区域 RoI(Region of Interest),然后再把生成的区域进行分类,最后完成物体的检测,这种方法显著提升了检测精度,但也限制了检测速度,其代表算法有 RCNN<sup>[6]</sup>(Region with Convolutional Neural Network features),Fast RCNN<sup>[7]</sup>(Fast Region with CNN features),Faster RCNN<sup>[5]</sup>(Faster Region with CNN features)等。后者代表算法有 SSD<sup>[8-9]</sup>(Single Shot Multibox Detector)和 YOLO<sup>[10]</sup>(You Only Look Once)系列。SSD 算法利用多层特征信息并使用固定框进行区域生成。YOLO 算法使用一阶网络直接完成了物体检测,速度很快。YOLO 系列包含 YOLOv1<sup>[10-14]</sup>,YOLOv2<sup>[15]</sup>,YOLOv3<sup>[16]</sup>,YOLOv4<sup>[17]</sup>,YOLOv5<sup>[18]</sup>以及 YOLOX<sup>[19]</sup>等 6 个版本。

Geng等<sup>[20]</sup>提出一种基于 YOLOv3 改进版的算法用来解决数据集不平衡的问题,使用高斯模糊方法来处理不平衡数据集:图像上的每个点的像素取其周围点像素的平均值,使像素值更加平滑,图像更加模糊。该算法比直接使用 YOLOv3 算法提升了 1%~2%,同时提高了小头盔的召回率和大头盔的定位准确率。Deng等<sup>[21]</sup>提出了一种基于改进 YOLOv4 的安全帽检测方法。通过采集自制的施工现场视频数据集,利用 K-means 算法对数据集进行聚类,得到合适的先验框尺寸,获得更有针对性的边缘信息。随后,在网络训练过程中采用多尺度训练策略,以提高模型在不同检测尺度下的适应性。实验结果表明,在安全帽佩戴检测任务中,模型 mAP 达到 92.89%,检测速度达到 15f/s,与原 YOLOv4 算法相比,其检测精度和检测速度都有所提高,能够满足安全帽检测任务的实时性要求。Tan等<sup>[22]</sup>在 YOLOv5 的基础上进行改进,利用新添加的功能检测模块,解决了小目标难以检测的问题,同时使用 Diou NMS 代替 NMS,使边界框预测得更加准确,这些改进措施使得改进后的算法比原 YOLOv5 算法具有更高的检测精度,检测速度可达到 98 f/s,有效改善了安全帽检测过程中时效性差、漏检问题的现状。YOLOX 作为目前性能最佳的目标检测算法,尚未发现其被应用于安全帽检测任务当中。

针对复杂环境中小目标、密集目标以及遮挡目标的安全帽检测应用,本文提出一种基于改进 YOLOX<sub>s</sub>的算法,改进点主要体现在三方面:首先,由于现有算法在小目标检测上仍存在漏检等问题,因此在原算法的基础上新增加一个预测特征层,新增的预测特征层更加适合预测小目标;其次,用 Focal\_Loss 来训练 obj 分支,以优化对复杂环境中密集目标、遮挡目标的识别准确度和降低漏检率;最后,将 CSP1\_X 中的残差块改为 shuffleNet 基本单元以缩减参数量,在保证检测精度的前提下提高检测效率。

## 1 算法框架

### 1.1 改进后的YOLOX\_s算法框架

作为目前性能最好的算法,YOLOX算法已经应用于各种目标检测任务.为了使YOLOX算法更好地应用于环境复杂的安全帽检测任务,本文对原有算法进行了如下改进:(1)在原有算法的基础上新添加一个预测特征层,用来预测由于遮挡或距离远而难以检测的小目标;(2)使用Focal\_Loss替代obj\_loss的BCE\_Loss;(3)将CSP1\_X中的残差块改为shuffleNet基本单元以缩减参数量.改进后的算法结构如图1所示.

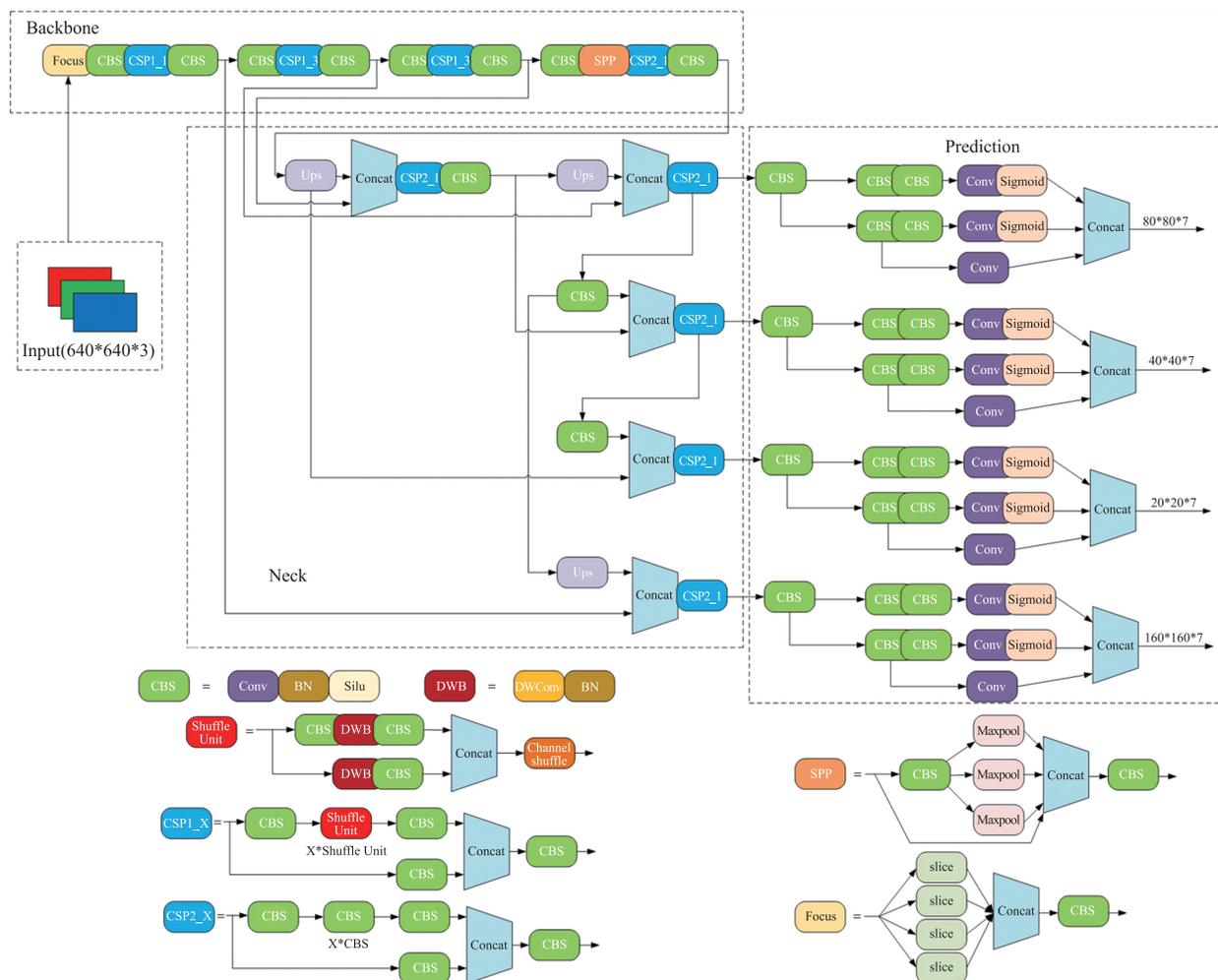


图1 改进后的YOLOX\_s算法框架

Fig.1 Algorithm framework of improved YOLOX\_s

图1所示算法结构主要分为4个部分,分别是Input、Backbone、Neck以及Prediction部分.输入尺寸为 $640 \times 640 \times 3$ ,图片大小默认为 $640 \times 640$ ,3为RGB颜色三通道. Backbone部分由基本的CBS(Conv2D-BatchNormal-Silu)单元、Focus结构、CSP(Cross Stage Partial network)结构以及SPP(Spatial Pyramid Pooling network)结构组成. Focus是一种切片操作,具体做法是对输入的一张图片每隔一个像素取一个值,一张图片经Focus操作后可以得到4张图片.然后骨干网络对得到的新图像进行卷积运算,以降低原有图像主要特征的丧失程度. CSP结构主要分为两种:CSP1\_X和CSP2\_X,前者带有残差块,而后者没有. CSP结构不仅可以有效防止梯度消失,还能提取到更多的细粒度特征. SPP结构显著增加了网络的感受野,可以关注到不同尺度的特征. Neck部分由FPN(Feature Pyramid Networks)和PAN(Path Aggregation Network)两种结构组成,FPN采取自顶向下,上采样的方式向下传递强大的语义信息,而PAN则相反,采取自底向上,下采样的方式向上传递强大的位置信息,FPN和PAN一起将网络获取的语义信息和位置信息高效融合在一

起,有效提高了网络的分类和定位能力. Prediction 部分也叫解耦头部分,使用解耦头可以有效提高网络的收敛速度和检测精度. 本文使用的最终输出预测特征层尺寸分别为:20×20×7、40×40×7、80×80×7、160×160×7,如图中 Prediction 部分所示,这里的 20、40、80、160 是指预测特征层的高和宽,7 是通道数,分别指 2 个类别信息,1 个置信度信息,4 个位置参数信息.

CSP1\_X 中的残差块结构如图 2 所示,输入的特征经 CBS 基本单元进行卷积操作,再与输入进行逐点相加操作. 本文将残差块修改为 shuffleNet 基本单元结构,如图 3 所示,输入特征一方面经深度可分离卷积和 CBS 单元计算,另一方面通过 CBS 单元提取特征,最后两分支通过通道拼接操作并经通道混洗进行通道间信息沟通. 修改后的结构避免了大量的逐点相加操作,降低了模型参数量,同时提升了模型训练速度.

### 1.2 Focal\_Loss

由于安全帽检测环境较为复杂,存在大量难样本容易造成漏检,本文应用 Focal\_Loss 来调节难易样本,公式如下:

$$FL(p_i) = -(1-p_i)^\gamma \log(p_i), \tag{1}$$

$$p_i = \begin{cases} p & \text{if}(y=1) \\ 1-p & \text{otherwise} \end{cases}, \tag{2}$$

式中, $y$  代表样本标签, $y \in \{\pm 1\}$ , $p$  代表样本标签  $y=1$ ,即预测为正样本时的预测概率, $\gamma$  是一个聚焦因子, $\gamma \geq 0$ .

对于该损失函数,有如下特性:

(1)引入 $(1-p_i)^\gamma$ 来调节简单样本与难样本的权重. 当边框被错误分类时, $p_i$ 较小,则调制因子 $(1-p_i)^\gamma$ 接近于 1,其损失几乎不受影响;当 $p_i$ 越接近于 1,表明其分类预测越好,样本越简单,调制因子 $(1-p_i)^\gamma$ 越接近于 0,损失越小.

(2)引入聚焦因子 $\gamma$ , $\gamma$ 越大,当 $p_i$ 也越大,即样本越简单时,则 $(1-p_i)^\gamma$ 越接近于 0,简单样本的损失贡献占比会越低,因此模型更容易集中在难样本上.



图 2 残差单元结构

Fig. 2 Residual unit structure

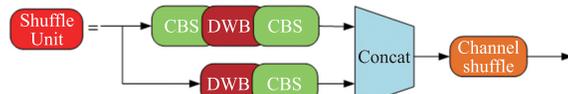


图 3 ShuffleNet 基本单元结构

Fig. 3 ShuffleNet basic unit structure

### 1.3 数据集与实验环境配置

本文使用的数据集通过网络爬虫获取,一共 7 829 张含有未佩戴安全帽的人和佩戴安全帽的人的图片,如图 4 所示. 数据集被划分为 3 部分,划分比例约为 8:1:1,分别是训练集、验证集和测试集. 训练过程中使用了随机翻转、亮度改变等数据增强手段. 数据集包括两类:“person”类和“helmet”类.“person”类代表未佩戴安全帽的人,“helmet”类代表佩戴了安全帽的人. 数据集通过 LabelImg 工具标记,被标记的文件后缀为 xml,格式为 PASCAL VOC,文件与被标记图片具有相同名称.

本实验使用 PyTorch 框架搭建网络模型,使用的实验环境配置如表 1.

表 1 实验环境配置

Table 1 Experimental environment configuration

| Parameter                | Configuration                                |
|--------------------------|--|
| CPU                      | Intel(R) Core(TM) i9-12900K<br>CPU@ 3.19 GHz |
| GPU                      | NVIDIA GeForce RTX 3 090                     |
| System Environment       | Ubuntu20.04                                  |
| Language                 | Python 3.8                                   |
| Acceleration Environment | CUDA 11.1,Cudnn8.0.5                         |

## 2 结果和分析

### 2.1 改进后的 YOLOX\_s 算法的表现

改进后的 YOLOX\_s 算法的训练过程如图 5 所示,损失曲线在 50 epochs 左右趋于收敛,虽然前期验证集的损失波动较剧烈,但是在训练后期,训练集的损失和验证集的损失非常接近,验证集的损失只比训练集的损失高一点,表明修改后的模型结构是合理的,不存在过拟合或者欠拟合现象. 训练的相关细节信息有:一共训练了 100 epochs,训练之前在 COCO 数据集上预训练过,正式训练的初始学习率为  $3 \times 10^{-3}$ ,学习率下降策略

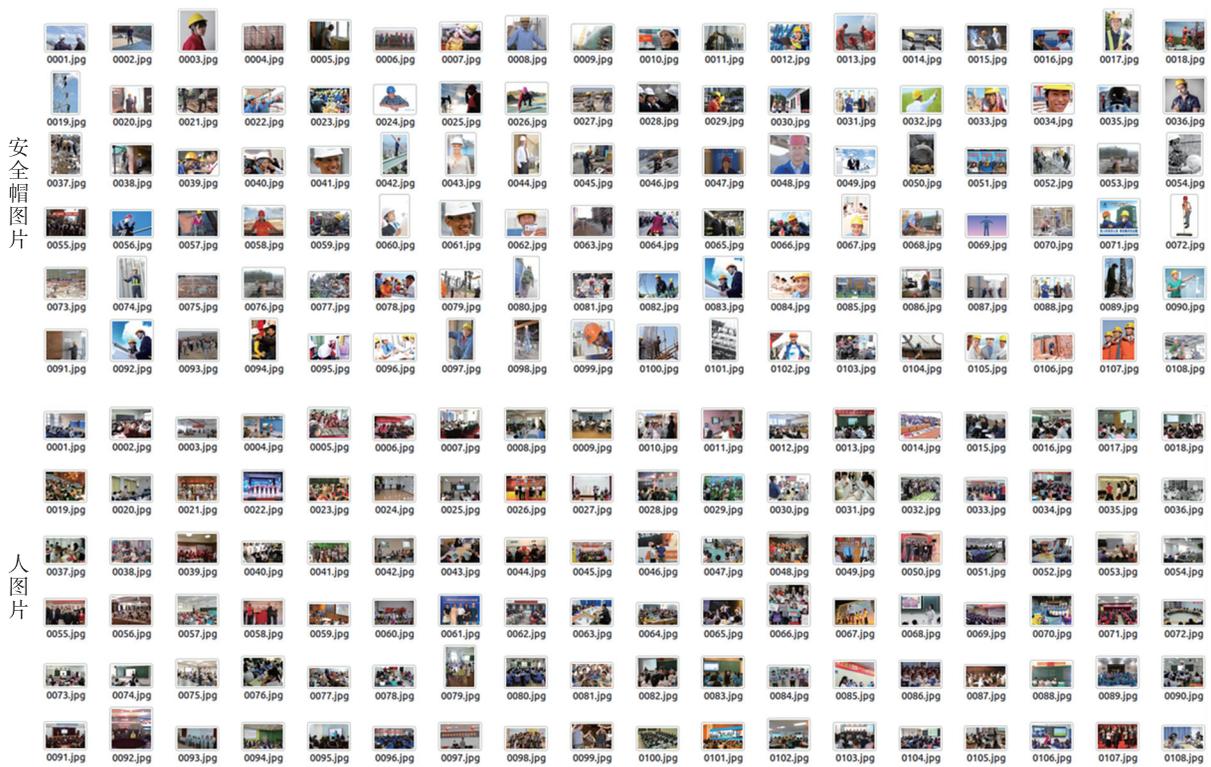


图4 训练中使用的图片样本

Fig. 4 Examples of images used in training

为 ReduceLRonPlateau, batchsize 为 16, 优化器为 Adam.

改进前和改进后的检测结果如图 6 所示, 检测结果中不易查看的目标均用红色矩形框标出进行提示, 小目标的对比检测结果被放大, 置于对应图片的左下方. 从第 I 组可以看出, 检测环境比较复杂, 图中的背景和前景光照差别非常大, 且局部存在曝光严重以及被检测对象之间出现严重相互遮挡的问题. 改进前, 左侧佩戴黄色安全帽的人并未被检测出来, 且左侧未佩戴安全帽的人也存在漏检, 改进算法后, 这些目标均能够正确被检测到. 在第 II 组中, 对于上方的小目标, 前三种算法检测结果或是出现漏检或是出现误检, 而改进后的

YOLOX\_s 算法并未出现这些问题, 且检测精度均高于前三者. 在第 III 组中, 目标较为密集且整体目标尺寸非常小, 对于不同的小目标, 前三种检测算法均出现不同程度的漏检, 且对于矩形框中提示的极小目标, 前三种算法均未检测到, 而改进后的算法没有出现漏检现象. 在第 IV 组中, 背景干扰较为严重, 前三种算法均未检测出后方的小目标, 而改进后的算法可以将其准确定位出来, 且前方的大目标的检测精度远远超过前三种算法. 在第 V 组中, 前三种算法均出现不同程度的误检, 将未被佩戴的安全帽识别为正确佩戴安全帽的人, 这显然是误检, 且左侧的女性目标甚至对应多个检测结果, 而改进后的算法不存在这些问题. 在第 VI 组中, 后方的小目标的检测结果出现和第 V 组相似的问题, 但是均能被改进后的算法很好地解决, 且检测精度更高. 上述检测结果验证了改进后的算法能够有效解决复杂背景下的小目标检测问题.

## 2.2 与其他算法比较

为了验证改进后的 YOLOX\_s 算法的有效性, 本节评估了安全帽数据集上的 YOLOX\_l、YOLOX\_m、YOLOX\_s 以及改进的 YOLOX\_s 算法. 以 mAP (mean Average Precision)、Recall、参数量和 FPS (Frames Per Second) 作为评价指标. 各评价指标的公式如下:

$$P = \frac{TP}{TP+FP}, \quad (3)$$

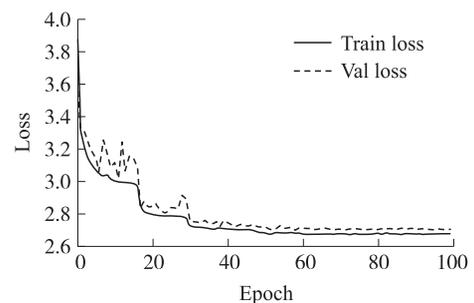


图5 改进后的YOLOX\_s训练过程

Fig. 5 Training process of improved YOLOX\_s



图 6 基于改进 YOLOX\_s 的安全帽识别结果

Fig. 6 Helmet recognition results based on improved YOLOX\_s

$$R = \frac{TP}{TP + FN}, \tag{4}$$

$$AP = \int_0^1 PdR, \tag{5}$$

式中, TP(True Positive) 代表正确检测框: 预测框正确地与标签框匹配了, 两者间的 IoU (Intersection of Union) 大于设定的阈值 0.3. FP (False Positive) 代表误检框: 将背景检测成了物体. FN (False Negative) 代表漏检框: 应检测出的物体却未被检测出来. P (Precision) 指准确率, 即正确预测框数量与当前遍历过的预测框数量的比值. R (Recall) 指召回率, 即当前一共检测出的标签框数量与所有标签框数量的比值. 当遍历到一个预测框时, 都可以生成一个对应的 P 与 R, 这两个值可以组成一个点 (R, P), 利用这些点绘制成曲线, 即形成了 P-R 曲线, AP (Average Precision) 即表示 P-R 曲线与两坐标轴所围成的面积. 每个类别的 AP 是相互独立的, 将所有类别的 AP 求平均值, 即可得到 mAP.

从表2中可以看出,改进后的YOLOX\_s算法的mAP为88.96%,相比于改进前,有1.25%的提升,Recall为83.04%,比改进前提升了2.32%。改进后的算法的mAP和Recall两项评估指标均是4种算法中最高的,可见改进后的算法具有更好的有效性。改进前YOLOX\_s算法的参数量为8.94MB,改进后参数量缩减为3.61MB,有效改善了由于增加小目标预测特征层而带来的参数增加的问题。缩减参数量后,原YOLOX\_s算法的FPS为147.94f/s,改进后为140.37f/s,二者相近,可以满足实时检测的要求,可见本文对原YOLOX\_s算法的改进是有效的。

### 2.3 消融实验

本文在整体实验的基础上继续做了独立的消融实验,以验证新添加的160×160预测特征层以及新的误差函数Focal\_Loss的有效性,结果如表3所示。

表3中打了“√”的表示添加了左侧对应的模块,从表3中可以看出,仅仅添加160×160预测特征层,改进后的算法的mAP和Recall分别为87.93%和81.43%,较改进前的87.71%和80.72%分别提升了0.22%和0.71%;

仅仅更换新的误差函数Focal\_Loss,改进后的算法的mAP和Recall分别为88.25%和81.56%,较改进前的87.71%和80.72%分别提升了0.54%和0.84%;而同时添加160×160预测特征层和更换新的误差函数Focal\_Loss,改进后的算法的mAP和Recall分别为88.96%和83.04%,较改进前的87.71%和80.72%分别提升了1.25%和2.32%。由此可见,改进后的算法中新添加的160×160预测特征层以及更换新的误差函数Focal\_Loss的做法是有效的。

## 3 结论

由于目标聚集,光照、距离等因素造成安全帽检测环境较为复杂,在这种复杂的环境下容易造成漏检严重,针对在此复杂环境下现有算法无法解决漏检以及检测精度较低的问题,本文提出改进版本的YOLOX\_s算法,改进后的算法在原有算法的基础上新增了一个更大尺度的预测特征层用来检测小目标,并用新的误差函数Focal\_Loss来训练obj分支以降低漏检,同时将CSP1\_X中的残差块改为shuffleNet基本单元以缩减参数量。改进后的算法无论是从视觉效果上还是从测试指标上均较改进前有所提升。且通过与YOLOX\_l、YOLOX\_m、YOLOX\_s算法比较,改进后的YOLOX\_s算法的mAP和Recall分别为88.96%和83.04%,较改进前的算法分别提升了1.25%和2.32%,改进后的参数量从8.94MB降到3.61MB,而检测速率并未受到较大影响,仍可以满足实时检测。实验表明,改进后的YOLOX\_s算法能有效解决复杂环境下的安全帽漏检严重的问题。

### [参考文献]

- [1] WEN C Y. The safety helmet detection technology and its application to the surveillance system[J]. Journal of forensic sciences, 2004, 49(4): 770-780.
- [2] DU S, SHEHATA M, BADAWY W. Hard hat detection in video sequences based on face features, motion and color information[C]//Proceedings of International Conference on Computer Research & Development, Shanghai, 2011: 25-29.
- [3] WARANUSAST R, BUNDON N, TIMTONG V, et al. Machine vision techniques for motorcycle safety helmet detection[C]//Proceedings of 2013 28th International Conference on Image and Vision Computing New Zealand, Wellington, New Zealand, 2013: 35-40.
- [4] JIAO L, ZHANG F, LIU F, et al. A survey of deep learning-based object detection[J]. IEEE Access, 2019, 7: 128837-128868.
- [5] REN S, HE K, GIRSHICK R B, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6): 1137-1149.

- [6] GIRSHICK R B, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014:580-587.
- [7] GIRSHICK R B. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015:1440-1448.
- [8] HU X, LI H, LI X, et al. MobileNet-SSD MicroScope using adaptive error correction algorithm: real-time detection of license plates on mobile devices[J]. Iet intelligent transport systems, 2020, 14(2):110-118.
- [9] CHEN Z, WU K, LI Y, et al. SSD-MSN: an improved multi-scale object detection network based on SSD[J]. IEEE access, 2019, 7:80622-80632.
- [10] REDMON J, DIVVALA S K, GIRSHICK R B, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016:779-788.
- [11] YIN Y, LI H, FU W. Faster-YOLO: an accurate and faster object detection method[J]. Digit signal process, 2020, 102:102756.
- [12] SADYKOVA D, PERNEBAYEVA D, BAGHERI M, et al. IN-YOLO: real-time detection of outdoor high voltage insulators using UAV imaging[J]. IEEE transactions on power delivery, 2020, 35:1599-1601.
- [13] NGUYEN D T, NGUYEN T N, KIM H, et al. A high-throughput and power-efficient FPGA implementation of YOLO CNN for object detection[J]. IEEE transactions on very large scale integration (VLSI) systems, 2019, 27(8):1861-1873.
- [14] HUANG Z, WANG J. DC-SPP-YOLO: dense connection and spatial pyramid pooling based YOLO for object detection[J]. Information sciences, 2020, 522:241-258.
- [15] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017:6517-6525.
- [16] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08). <https://doi.org/10.48550/arXiv.1804.02767>.
- [17] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23). <https://doi.org/10.48550/arXiv.2004.10934>.
- [18] FELZENSZWALB P F, MCALLESTER D A, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 2008:1-8.
- [19] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. (2021-08-06). <https://doi.org/10.48550/arXiv.2107.08430>.
- [20] GENG R, MA Y, HUANG W. An improved helmet detection method for YOLOv3 on an unbalanced dataset[C]//Proceedings of 2021 3rd International Conference on Advances in Computer Technology, Information Science and Communication (CTISC), Shanghai, China, 2021:328-332.
- [21] DENG B, LEI X, YE M. Safety helmet detection method based on YOLOv4[C]//Proceedings of 2020 16th International Conference on Computational Intelligence and Security (CIS), Nanning, China, 2020:155-158.
- [22] TAN S, LU G, JIANG Z, et al. Improved YOLOv5 network model and application in safety helmet detection[C]//Proceedings of 2021 IEEE International Conference on Intelligence and Safety for Robotics (ISR), Tokoname, Japan, 2021:330-333.

[责任编辑:陆炳新]