

基于 YOLO V5 的安全帽快速检测

覃剑永*, 朱明增, 覃秋勤, 莫梓樱, 覃景涛
(广西电网有限责任公司贺州供电局, 广西 贺州 542800)

摘要: 安全帽的佩戴检测可有效防范因未佩戴安全帽而造成的施工事故, 在安全生产过程中发挥重要作用。该文提出一种基于 YOLO V5 的安全帽自动、高效检测方法。首先, 基于 Density Peak 聚类算法对训练数据进行聚类分析, 根据 Density Peak 决策图优化 anchor 尺度。然后, 综合 anchor 的位置和面积关系引入 DIOU 损失函数来提高模型的训练精度。最后, 建立安全帽佩戴样本数据库, 采用 YOLO V5 模型进行训练和测试。为验证算法的有效性, 该方法在自建的安全帽佩戴数据集上进行了测试。实验结果表明, 该方法的 mAP_{0.5} 可达 77.2%, 图像推理速率可达 140 帧/秒, 具有精度较高、检测速度快等优点。此外, 该方法也可以有效检测多目标图像, 能够满足复杂环境下安全帽的自动检测需求, 且容易部署和推广到嵌入式应用设备中。

关键词: YOLO V5; 安全帽; 深度学习; DIOU; Density Peak 聚类

中图分类号: TP39

文献标识码: A

文章编号: 1673-629X(2021)0126-05

An Efficient Safety Helmet Detection Based on YOLO V5

QIN Jian-yong*, ZHU Ming-zeng, QIN Qiu-qin, MO Zi-ying, QIN Jing-tao
(Hezhou Power Supply Bureau, Guangxi Power Grid Co., Ltd., Hezhou 542800, China)

Abstract: The detection of safety helmet plays an important role in the prevention of safety accidents, and we propose an automatic and efficient detection method of safety helmet. First, a Density Peak algorithm is used to analyze the size feature of safety helmet, which helps to increase the accuracy of detection. Then, according to area and distance of prediction anchor and truth anchor, a DIOU (distance-IoU) function is introduced to build the loss function of YOLO v5. Finally, the deep learning framework YOLO v5 is used to train and test the safety helmet. Experimental results on the created dataset show the efficiency and effectiveness of the proposed method on safety helmet detection, which obtains mAP_{0.5}, image inference ratio with 77.2% and 140 frame/s. In addition, the proposed method is also able to detect the images with multi-object, and can easily be applied for the embedded device.

Key words: YOLO V5; safety helmet; deep learning; distance-IoU; Density Peak cluster

0 引言

近年来, 随着中国大力加快基础设施建设, 推进经济社会高质量发展, 安全生产的制度化、常态化、持续化的切实举措受到人们的广泛关注。进入变电站、矿山、建筑工地等领域, 佩戴安全帽作业是基本制度。然而, 由于部分工人安全意识不强, 以及监管的疏忽, 未佩戴安全帽作业导致安全事故的情况频频发生^[1]。因此安全帽的佩戴检测在确保安全生产过程中具有重要意义。

早期的安全帽佩戴检测依赖于人工巡检, 或是根据现场视频进行实时监控, 这些方法不仅费时费力, 而且易出现漏判。传统的计算机目标检测方法需要基于安全帽的先验知识, 如颜色、形状手动提取安全帽特征, 存在模型泛化性不佳、精度不高问题^[2]。近年来,

由于具有自适应特征提取能力, 深度学习 (deep learning) 方法迅速引起了学者的高度重视, 各种基于深度学习的目标检测方法层出不穷^[3], 受到学术界和工业界的高度重视。2006年, Hinton 提出利用卷积神经网络 (CNN) 来学习样本特征。Girshick 等综合 region proposal networks (RPN) 和 CNN^[4] 提出一种 RCNN 的目标检测和语义分割方法^[5]。首先在每张图片上抽取若干候选区域, 将每个候选区域输入到神经网络提取图像特征, 然后, 基于 SVM 获取最终的分类结果。该方法在 Pascal VOC、MSCOCO 和 ImageNet 数据集上均取得良好的效果。为提高算法效率, Girshick 通过池化将 ROI 区域调整到特定大小, 融合分类和回归任务, 提出了 RCNN 的改进版本 - Fast RCNN^[5]。进一步, 文献 [6] 将特征抽取 (feature

收稿日期: 2021-01-05

作者简介: 覃剑永 (1974-), 男, 研究方向为数字图像处理、电力系统及其自动化。

extraction)、Region Proposal 提取, Bounding Box Regression 和分类整合到一个网络中,使得网络性能和效率均有较大提高。针对 RCNN 的 Proposal 的重复计算,Redmon 等^[7]将输入图片划分成若干个网格,并在每个网格中做预测,提出了一个基于 YOLO 的目标检测方法。在此基础上,通过融合借鉴 YOLO 网络设计和 Faster RCNN 的 anchor 思想,SSD 深度网络基于多尺度特征层进行检测,一次完成目标定位与分类^[8]。

在安全帽检测领域,徐守坤等^[9]运用多尺度训练和增加锚点数量来增强 Faster RCNN 网络的检测能力,并引入挖掘策略解决安全帽正负样本不均衡的问题。吴冬梅等^[10]在 Faster RCNN 基础上将特征层进行融合和多尺度检测,提出了一种基于改进 Faster RCNN 的安全帽检测方法。肖体刚等^[11]针对安全帽佩戴检测算法存在检测速度慢等问题,采用轻量型卷积神经网络 MobileNetV3-small 替换 SSD 检测算法的卷积神经网络 VGG-16,以减少模型参数,提升检测速率。为增强模型的泛化能力,刘欣等^[12]引入注意力机制来强化丰富特征之间的信息传播和增强模型的泛化能力,提出了一种矿井安全帽的检测方法。针对 Faster RCNN 网络对小尺寸目标和部分遮挡目标检测效果不佳的问题,文献^[13]等运用多层卷积特征融合技术优化区域建议网络产生候选区域特征图。综合利用安全帽的颜色、轮廓,以及基于多层神经网络分类建立的统计模型,刘喜文等^[14]设计了一种基于智能视频图像分析的安全帽识别系统,该系统能够获得较好的检测效果。针对由于施工现场光照多变、背景复杂,传统安全帽检测方法鲁棒性差的问题,方明等^[15]引入密集连接结构实现多层特征的融合,提高了网络对于小目标检测的敏感性。林俊等^[16]通过优化分类器张量维度和损失函数,研究了一种基于 YOLO V3 的安全帽检测模型。实验结果表明,该模型具有良好的精度和效率。刘霞等^[17]提出将分辨率增强模块融合 VGG16 网络中不同层次的特征图,利用融合后的特征图生成特征金字塔。杨莉琼等^[18]通过 YOLO V3 网络检测人脸区域,综合人脸区域 HOG 特征向量和 SVM 分类算法判断是否佩戴安全帽;刘晓慧等^[19]首先利用肤色检测方法定位到人脸区域,然后,引入 Hu 矩构建图像的特征向量,最后运用 SVM 来检测安全帽;刘喜文等^[14]综合利用安全帽外在特征,如颜色、轮廓,基于多层神经网络分类建立安全帽佩戴的统计模型;张春堂等^[20]将 SSD 的特征提取网络 VGG16 替换成 MobileNet 网络,构建了 SSD-MobileNet 的安全帽检测模型。针对低分辨率目标图像,郭敬东等^[21]利用稀疏表示模型进行超分辨率增强,提出了一种基于面向电力智能安监的低分辨率安全帽检测算法。

基于上述研究成果,为进一步提高算法效率,该文提出一种基于 YOLO V5 网络模型的自动检测方法。首先,基于 Density Peak 聚类算法对训练数据进行聚类分析,优化 anchor 尺度。然后,根据 anchor 面积和距离分布,引入 DIOU 损失函数来提高模型的训练精度。最后,建立安全帽佩戴样本数据库,基于 YOLO V5 模型进行训练和测试。实验结果表明,该方法具有检测速度快、精度高,能够满足复杂环境下安全帽的自动检测需求。

1 方法

由于具有灵活部署和高效率运算等优点,YOLO V5^[22]的 Pytorch 源代码一经公布,便受到的众多学者的广泛关注和高度重视。YOLO V5 模型共有 4 个版本,分别为 YOLO V5s, YOLO V5l, YOLO V5m 和 YOLO V5x。其中,YOLO V5s 是网络深度最小、特征图宽度最小的网络,其权重文件仅 27.7Mbyte,相应地,YOLO V5s 也是运行速度最快的网络。

1.1 YOLO V5s 简介

YOLO V5s 由 Input、CSP backbone、PA-NET neck 和 Prediction 四部分构成,如图 1 所示。为有效解决模型训练中“小对象”问题,Glen Jochery^[22]引入马赛克(mosaic augmentation)增强技术,对训练数据采用缩放和色彩空间调整技术来提高目标检测的鲁棒性。考虑到网络框架 Backbone 中的梯度信息重复,YOLO V5 采用 CSPDarknet 作为 Backbone,从输入图像中提取信息特征,并将梯度信息融合到特征图中,有效减少模型的参数和运算时间。为增强网络对不同尺寸的目标检测,YOLO V5 采用 Neck-路径聚合网络生产特征金字塔,强化低层特征的传播。在图 1 中,Prediction 结构用于在特征图上定位锚框(anchor box),并预测目标所属类别及对应得分。

1.2 基于 Density Peak 聚类的 anchor 尺度优化

为使锚框的尺度适应自建安全帽数据库,该文采用 Density Peak 聚类算法对锚框尺度进行分析。针对经典聚类算法无法有效处理非球面的类簇数据、对离群点噪声敏感等问题,Rodriguez 等^[23]提出了一种 Density Peak 聚类算法。该算法的主要思想是聚类簇中心被较低的局部密度邻域点包围,且与比聚类中心密度更高的样本点之间存在较大的距离。设 i 为样本点,局部密度值 ρ_i 定义如下:

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \quad (1)$$

式中, d_c 为截断距离,根据 d_c 定义以样本 i 为中心、 d_c 为半径的搜索区域; d_{ij} 为样本点 i 到样本点 j 的距离,当 $d_{ij} - d_c \geq 0$, $\chi(\cdot) = 0$; 否则, $\chi(\cdot) = 1$ 。此外, Density

Peak 还定义了样本点到其更高密度样本点的最小距离 δ_i :

$$\delta_i = \min_{j:\rho_j > \rho_i} (d_{ij}) \quad (2)$$

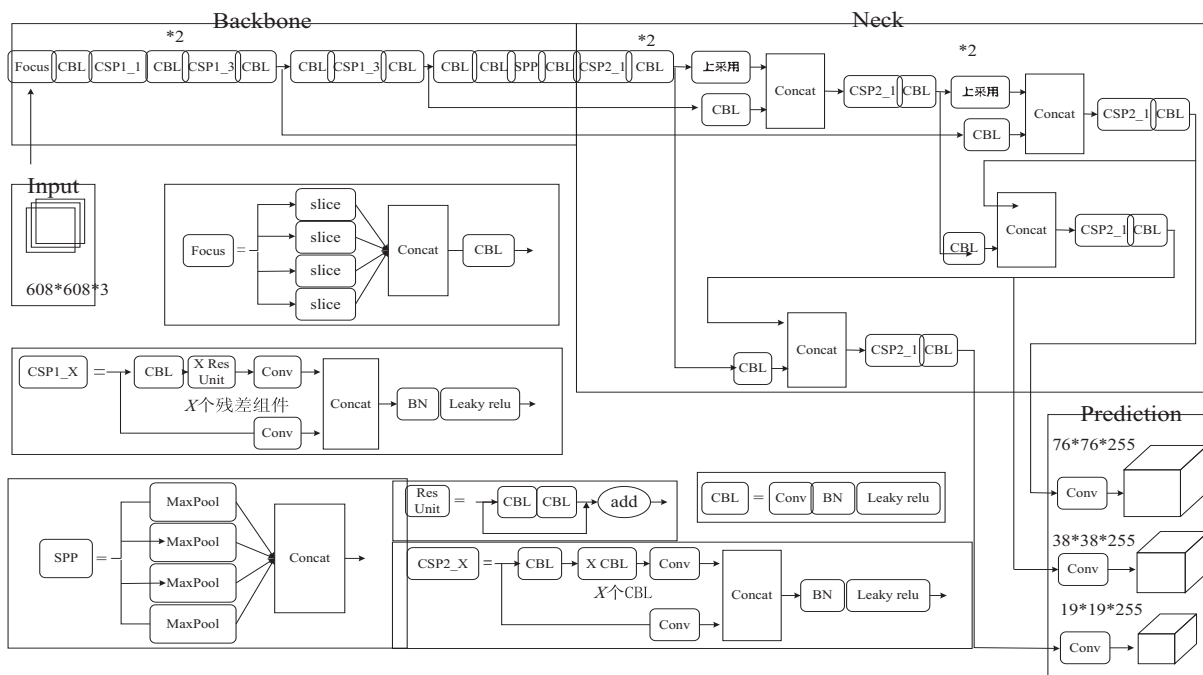


图 1 YOLO V5s 网络架构

对于局部密度值最大的样本点定义 $\delta_i = \max_j(d_{ij})$ 。当样本点的局部密度 ρ_i 为全局最大或者局部最大时,该样本点的 δ_i 值才会高于数据集中其他样本点的最小邻域距离。图 2 给出了 $d_c = 80$ 时安全帽数据集锚框的尺度分布和基于 Density Peak 的聚类

决策图,文中选取 δ_i 和 ρ_i 均较大的锚点尺度作为安全帽的预测尺度,如图 2 (b) 中虚线框所示,将差异较小的 anchor 尺度合并为一类,得到最终的多尺度 anchor,分别为: $[[12, 13, 16, 21, 17, 17], [19, 23, 21, 25, 37, 40], [75, 130, 99, 55, 209, 220]]$ 。

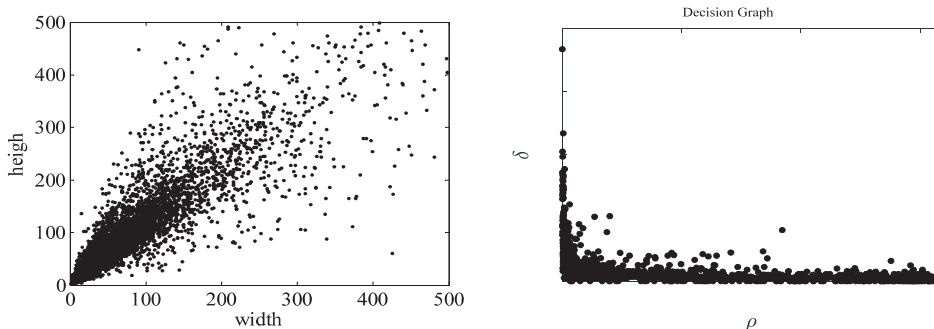


图 2 基于 Density Peak 的锚框尺度优化

1.3 优化 YOLO 的损失函数 DIoU

IoU (intersection over union) 是 YOLO 目标检测的重要评价,它根据预测和先验锚框 (anchor box) 重叠面积定义:

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{I}{U} \quad (3)$$

式中, A 和 B 分别表示预测锚框和先验锚框 (真实框), \cap 和 \cup 表示交集和并集运算。由于 IoU 无法优化无重叠区域的锚框,如果采用 IoU 作为损失函数 ($L_{IoU} = 1 - IoU$),无法辨识 A 和 B 之间的位置关系。例如,当 $A \cap B = \emptyset$,即使 A 和 B 的位置不一样, L_{IoU} 均

为 0。为解决 IoU 无法评价无重叠框的问题,DIoU 引入重叠面积、中心点距离定义损失函数:

$$L_{DIoU} = 1 - IoU + \frac{o^2(b_A, b_B)}{c^2} \quad (4)$$

式中, c 为预测框与先验框最小外接矩形的对角距离,如图 3 所示, b_A 和 b_B 分别为预测框和先验框和中心点, o 代表两点之间的欧氏距离。根据 DIoU 的定义,当 $A \cap B \neq \emptyset$ 时, $0 \leq L_{DIoU} \leq 1$,如图 3 (a) 所示;当 $A \cap B = \emptyset$ 时, $1 < L_{DIoU} < 2$,如图 3 (b) 所示。因此,与 IoU 相比,DIoU 具有更好的位置和面积辨识能力,文中采用 DIoU 作为 YOLOv5 的 anchor 损失函数。

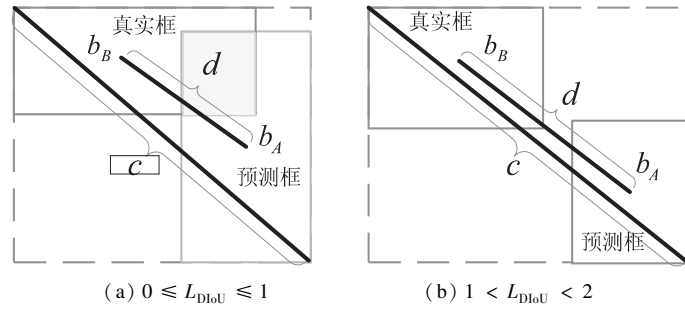


图 3 DIoU 示意图

2 实验结果

文中自建的安全帽佩戴检测数据集包含 6 400 个训练样本,数据主要来源来自互联网,部分为变电站的施工现场监控视频和图像。训练样本采用 LabelImg 开发软件进行标定。为提高算法精度,文中标定区域设置为头部区域,如果佩戴安全帽,则将该区域设置为正样本,否则,该区域定义为负样本。算法测试在

Ubuntu 18. 04 操作系统、PyTorch1. 6、NVIDIA GTX1070 8G GPU、i5 8400 CPU、16 GB 内存的开发环境下进行。

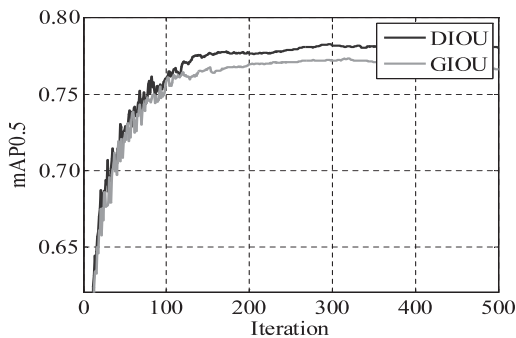
图 4 给出了文中方法的安全帽检测结果。可以看出,文中方法可以在施工现场、建筑工地、图像模糊、等复杂背景条件下准确检测安全帽的佩戴信息,对小目标(如第二行第三列中的小男孩所示)也有较好的检测效果。



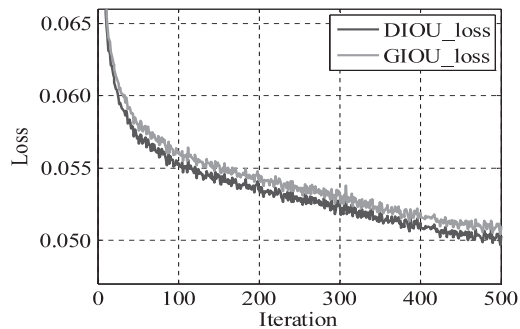
图 4 安全帽佩戴检测结果

图 5 给出了采用 DIoU 与 GIoU 作为损失函数的对比结果,如图所示,文中算法采用的 DIoU 在精度 mAP0.5 值达到 77. 2%,高于 GIoU。同时,由于锚框

中心距离参数的引入,DIoU 的锚框 loss 也略低于 GIoU 的锚框 loss,从而证明了该算法的有效性。



(a)基于 mAP0.5 的性能比较



(b)基于 DIoU_loss 和 GIoU_loss 的比较

图 5 DIoU 和 GIoU 的性能比较

为进一步量化文中算法的性能,表 1 给出了不同 YOLO 模型的权值文件大小、训练时间和图像推理速

率(为便于比较,以大小为 640 * 512 的图像作为帧速率参考标准)。可以看出,文中算法的训练时间、模型

大小和图像推理帧速率分别为 22.0 小时, 22.7 MB 和 140 帧/s, 明显优于其他模型。此外, 文中方法的 mAP-0.5 (77.2%) 也接近其他 YOLO V5m (79.0%) 和

YOLO V3-spp (78.5%) 的 mAP0.5, 高于 YOLO 51, 从而进一步证实了算法的优越性。

表 1 不同 YOLO 模型^[23]的参数和性能比较

模型	模型大小/MB	训练时间/小时	帧速率/ $f \cdot s^{-1}$	mAP-0.5
文中算法 YOLOv5s	27.7	22.0	140	0.772
YOLOv5m	174.4	36.8	84	0.790
YOLOv5l	399.5	84.9	40	0.658
YOLOv3-spp	250.9	102.2	43	0.785

3 结束语

提出了一种基于 YOLO V5 的快速安全帽佩戴检测方法。首先, 采用 Density Peak 聚类方法优化 anchor 的尺度, 综合 anchor 的面积和位置偏差引入 Distance-IoU 损失函数, 以提高算法精度。然后, 基于自建的数据集采用 YOLO V5s 模型进行训练和分类, 获取安全帽的检测结果。实验结果表明, 该方法可高效检测安全帽的佩戴情况, 其图像推理速率可以达到 140 帧/s, mAP0.5 为 77.2%, 能够满足复杂环境下安全帽的自动检测需求。此外, 与其他 YOLO 模型相比, YOLO V5s 模型的计算量小, 权值文件大小仅有 27.7 MB, 因此, 该方法也很容易部署和推广到嵌入式设备中。

参考文献:

[1] 王 兵, 李文璟, 唐 欢. 改进 YOLO v3 算法及其在安全帽检测中的应用[J]. 计算机工程与应用, 2020, 56(9): 33-40.

[2] 施 辉, 陈先桥, 杨 英. 改进 YOLO v3 的安全帽佩戴检测方法[J]. 计算机工程与应用, 2019, 55(11): 213-220.

[3] 余 凯, 贾 磊, 陈雨强, 等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展, 2013, 50(9): 1799-1804.

[4] GIRSHICK R. Fast R-CNN[C]//IEEE international conference on computer vision (ICCV). Santiago: IEEE, 2015: 1440-1448.

[5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE conference on computer vision and pattern recognition. Columbus: IEEE, 2014: 580-587.

[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: 1137-1149.

[7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 779-788.

[8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//European conference on computer vision. Amsterdam, The Netherlands: Springer, 2016: 21-37.

[9] 徐守坤, 王雅如, 顾玉宛, 等. 基于改进 Faster RCNN 的安全帽佩戴检测研究[J]. 计算机应用研究, 2020, 37(3): 901-905.

[10] 吴冬梅, 王 慧, 李 佳. 基于改进 Faster RCNN 的安全帽检测及身份识别[J]. 信息技术与信息化, 2020(1): 17-20.

[11] 肖体刚, 蔡乐才, 汤科元, 等. 改进 SSD 的安全帽佩戴检测方法[J]. 四川轻化工大学学报: 自然科学版, 2020, 33(4): 68-76.

[12] 刘 欣, 张灿明. 基于卷积神经网络的矿井安全帽佩戴检测[J]. 电子技术应用, 2020, 46(9): 38-42.

[13] 徐守坤, 王雅如, 顾玉宛. 基于改进区域卷积神经网络的安全帽佩戴检测[J]. 计算机工程与设计, 2020, 41(5): 1385-1389.

[14] 刘喜文, 何 岗, 杨贤文, 等. 基于智能视频图像分析的安全帽识别[J]. 计算机工程与设计, 2020, 41(5): 1464-1471.

[15] 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, 27(5): 1196-1205.

[16] 林 俊, 党伟超, 潘理虎, 等. 基于 YOLO 的安全帽检测方法[J]. 计算机系统应用, 2019, 28(9): 174-179.

[17] 刘 霞. 基于 SE Detection Net 的安全帽检测方法[J]. 计算机技术与发展, 2020, 30(6): 156-159.

[18] 杨莉琼, 蔡利强, 古 松. 基于机器学习方法的安全帽佩戴行为检测[J]. 中国安全生产科学技术, 2019, 15: 152-157.

[19] 刘晓慧, 叶西宁. 肤色检测和 Hu 矩在安全帽识别中的应用[J]. 华东理工大学学报: 自然科学版, 2014, 40: 365-370.

[20] 张春堂, 管利聪. 基于 SSD-MobileNet 的矿工安保穿戴设备检测系统[J]. 工矿自动化, 2019, 45: 96-100.

[21] 郭敬东, 李晓林. 面向电力智能安监的极低分辨率目标检测算法[J]. 计算机工程与设计, 2020, 41: 3188-3192.

[22] Jocher, yolov5 [DB/OL]. 2020. <https://github.com/ultralytics/yolov5>.

[23] RODRIGUEZ A, LAIO A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344: 1492-1496.