

基于 YOLO-Mask 算法的口罩佩戴检测方法

曹城硕, 袁杰*

新疆大学电气工程学院, 新疆 乌鲁木齐 830047

摘要 常态化疫情防控形势下,通过口罩佩戴检测可以及时提醒人们正确佩戴口罩,从而降低公共场合人员交叉感染的风险。针对口罩佩戴检测任务中被遮挡目标和小目标检测困难的问题,提出一种 YOLO-Mask 算法。该算法以 YOLOv3 为基础,在特征提取网络中引入注意力机制,以提升模型对显著特征的表达能力;然后使用特征金字塔和路径聚合策略进行特征融合,使细节特征信息得到增强,实现不同层次特征信息的充分利用;最后优化了损失函数。实验表明:对不同场景下的口罩佩戴目标进行检测,YOLO-Mask 算法的平均精度均值达到 93.33%,相比于原始 YOLOv3 算法提高 7.62%;与其他主流算法相比,该算法具有更好的检测效果和鲁棒性。

关键词 图像处理; 口罩佩戴检测; YOLO-Mask; 注意力机制; 特征融合; 疫情防控

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0810019

Mask-Wearing Detection Method Based on YOLO-Mask

Cao Chengshuo, Yuan Jie*

School of Electrical Engineering, Xinjiang University, Urumchi, Xinjiang 830047, China

Abstract Under normalized epidemic prevention and control, mask-wearing detection can promptly remind people to wear masks correctly, thus reducing the risk of cross-infection of people in public places. Aiming at the difficulty of detecting obscured and small targets in the mask-wearing detection task, a YOLO-Mask algorithm is proposed. The proposed algorithm is based on YOLOv3. It introduces an attention mechanism into the feature extraction network to enhance the model's ability to express salient features. Moreover, it uses feature pyramid and path aggregation strategies for feature fusion to enhance detailed feature information and utilize different levels of feature information. The loss function is optimized. The experimental results show that the average accuracy of the YOLO-Mask algorithm is 93.33% for mask-wearing detection targets in different scenarios, which is 7.62% higher than that of the existing YOLOv3 algorithm. The proposed algorithm has better detection results and robustness compared with other mainstream algorithms.

Key words image processing; mask-wearing detection; YOLO-Mask; attention mechanism; feature fusion; epidemic prevention and control

OCIS codes 100.4996; 150.1135; 150.1835

1 引言

自 2019 新型冠状病毒 (COVID-19) 肺炎疫情暴发以来,社会经济的发展和人们的生产生活都受到不同程度的影响^[1]。面对外部环境不断变化的形势,我国进入常态化疫情防控新阶段,疫情防控工作

面临新的巨大挑战^[2]。疫情期间,正确佩戴口罩是阻止病毒在人与人之间通过飞沫传播的重要举措,但公众主动佩戴口罩的意识还有待增强^[3]。因此,在公共场合进行有效的口罩佩戴检测是十分必要的,此举能够及时提醒人们正确佩戴口罩,降低人员交叉感染的风险,保障人员生命安全。相比于行人

收稿日期: 2020-08-10; 修回日期: 2020-09-08; 录用日期: 2020-09-20

基金项目: 国家自然科学基金(61863033)、新疆维吾尔自治区“天山青年计划”—优秀青年科技人才培养项目(2019Q018)

* E-mail: yuanjie222@126.com

检测、人脸检测等检测任务,口罩佩戴检测不仅需要识别出人脸目标,还需要对目标正确佩戴口罩和错误佩戴口罩情况进行准确区分,因此对检测算法学习和利用细节特征的能力有更高的要求。口罩佩戴检测的难点在于小目标和被遮挡目标的检测,除此之外,有关口罩佩戴的公开数据集较少且数据不完善,需要重新建立数据集并进行标注。

目前,专门用于口罩佩戴检测任务的算法极少,文献[4]通过改进 RetinaFace 人脸识别算法,实现了口罩佩戴检测。但事实上,通用检测算法具有更好的泛化能力,在通用算法的基础上,面向检测对象进行算法改进,有望取得更好的检测效果。从目标检测领域来看,随着深度学习在领域内应用与发展,先后出现了基于区域建议的 R-CNN 系列^[5-7]和基于回归的 SSD^[8]、YOLO 系列^[9-11]等主流通用目标检测方法。比较而言,基于区域建议的目标检测算法在检测精度上表现更好,基于回归的目标检测算法,检测速度更快,但检测精度略有不足。上述算法中,YOLOv3 凭借特征金字塔结构进行不同层次的特征信息融合,在小目标检测上较其他几种主流算法有一定的优势。

YOLOv3 是一种以 Darknet-53 为基础网络,通过 FPN (Feature Pyramid Networks)^[12] 进行多尺度特征融合的端对端目标检测算法。YOLOv3 具有检测速度快、综合性能强的特点,但直接将其应用于某些特定的检测对象时,由于受到场景复杂性和特征多样化等因素的影响,检测效果不能满足要求,需对其进行针对性改进。文献[13]利用空洞卷积对 YOLOv3 的基础网络进行改进,提高了机场中飞机

目标的检测精度。文献[14]通过设计新的损失函数替代平方和损失函数,减少模型误差,提高了收敛速度。文献[15]采用反残差网络提取特征,使计算复杂度降低,实现了实时车辆的检测。文献[16]通过评估卷积核的权重并进行剪枝处理,实现了模型的压缩和加速,满足了工业上对织物缺陷检测的要求。

在实际的口罩佩戴目标检测过程中,由于场景较为复杂,正确佩戴口罩和错误佩戴口罩的差异较小,原始 YOLOv3 算法表现并不理想,尤其在被遮挡目标和小目标检测上,检测效果较差。本文尝试在 YOLOv3 算法的基础上进行改进,提出一种针对口罩佩戴检测任务的检测算法。SENet (Squeeze-and-Excitation Networks)^[17] 是一种注重特征通道关系的网络结构,通过建模卷积特征各个通道之间的作用关系来改善网络模型的表达能力。将其应用在口罩佩戴的检测上,可以增强重要特征的通道权重,从而提高检测的效果。本文将 SENet 引入 YOLOv3 基础网络,提出一种以 SEDarknet-53 为基础网络,以采用路径聚合策略的 FPN 作为特征融合网络的 YOLO-Mask 算法。实验结果表明,YOLO-Mask 算法提高了对小目标和被遮挡目标的检测率,与 SSD、Faster R-CNN、原始 YOLOv3 等主流算法进行对比,检测效果更好,鲁棒性更强。

2 YOLO-Mask 算法相关模型

2.1 YOLOv3 算法原理

原始的 YOLOv3 以去除全连接层的 Darknet-53 作为其前置网络,用 FPN 特征金字塔替代全连接层进行特征融合。整体网络结构如图 1 所示。

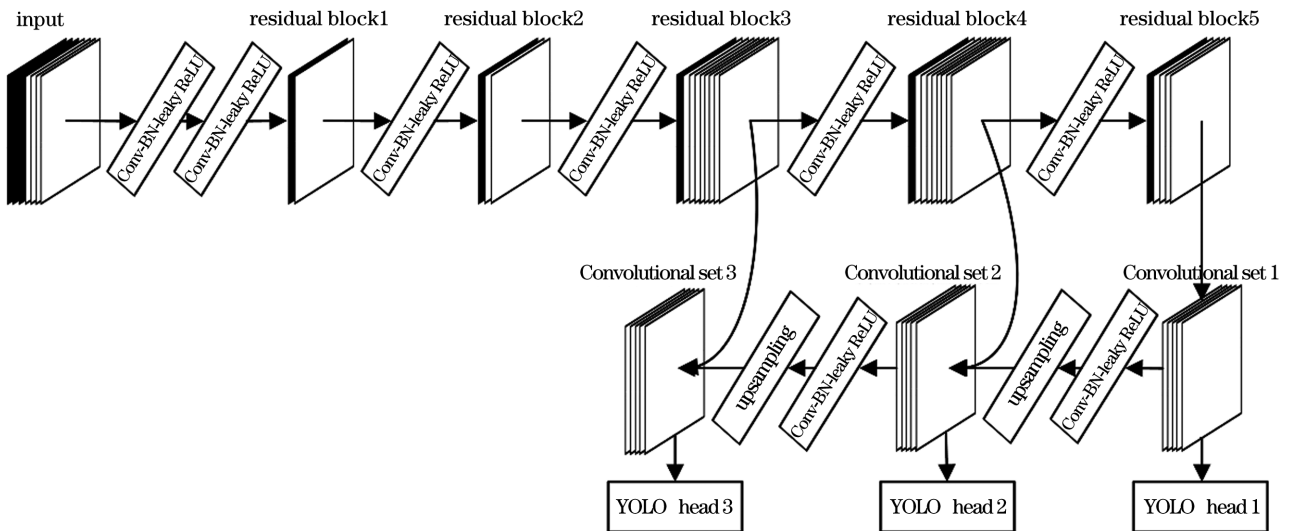


图 1 YOLOv3 网络结构图

Fig. 1 YOLOv3 network structure diagram

该模型中输入图片尺寸选取 416×416 , 前置网络中 Residual block3、Residual block4、Residual block5 这三个残差块输出的特征层作为 FPN 的输入。在 FPN 中, 特征图经过通道和尺寸调整后, 进行特征融合, 其中 Convolutional set 包含 5 层 1×1 、 3×3 、 1×1 、 3×3 、 1×1 的卷积层。三个 Convolutional set 分别输出尺度为 52×52 、 26×26 、 13×13 的特征图, 然后通过 YOLO head 得出预测结果。YOLO head 生成的预测结果包括类别、置信度信息和预测框的调整参数。利用置信度阈值和非极大值抑制对预测框进行筛选, 得到最终的检测结果。

2.2 SEDarkNet-53 主干网络

SEDarkNet-53 在原有 Darknet-53 的基础上, 引入了 SENet 通道注意机制, 其目的是通过预测输出通道的权重, 对各个通道进行加权, 从而关注信息量大的通道特征, 抑制对当前任务不太重要的通道特征, 进而改善网络性能。SENet 主要由压缩 (Squeeze) 和激励 (Excitation) 两部分组成。SENet 首先对输入的特征图进行压缩操作, 即用全局平均池化层将特征图的维度压缩成一维, 得到通道级的全局特征, 以获得更广的感受区域。然后对全局特征进行激励操作, 即通过全连接层学习各个通道间的关系, 得到不同通道的权重。接下来通过 Sigmoid 函数将获得的权重归一化到 $(0, 1)$ 之间; 最后通过 Scale 的操作将归一化后的权重加权到每个通道的特征上。

将 SENet 引入到 Residual 残差单元的残差分支中, 可得到 SE-residual 注意力残差单元, 二者结构如图 2 所示。图中 Residual unit 为传统的残差结构, SE-residual unit 为引入通道注意力机制的注意力残差结构。方框右侧的符号代表该层的输出维度信息, W 、 H 分别为特征图的长和宽, C 为通道数, r 为缩减系数。实验表明: 当 r 选取为 16 时, 可以得到较好的性能并且计算量相对较小。与传统 Residual unit 相比, SE-residual unit 在不改变输出维度的情况下, 增强了网络非线性特性, 提升了模型的泛化能力。SEDarkNet-53 采用 SE-residual unit 作为其基本结构单元, 实现通道间的信息交互, 抑制无用特征对模型的影响, 提升了检测准确率。

2.3 改进的特征融合网络

特征提取过程中, 浅层网络和深层网络提取到的特征有所不同。浅层网络旨在学习更多的细节信息, 深层网络旨在学习更多的语义信息。特征融合

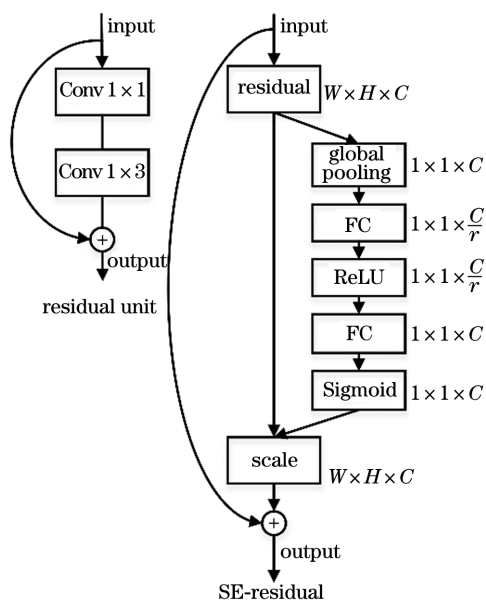


图 2 残差单元与注意力残差单元

Fig. 2 Residual unit and SE-residual unit

网络将两种特征信息进行融合, 实现特征增强, 同时避免了使用单一特征信息而造成大量信息丢失的现象。原始 YOLOv3 算法的特征融合网络为 FPN, 其通过特征金字塔结构融合不同尺度特征信息, 从而实现多尺度检测。在口罩佩戴检测的任务中, 为了提高对小目标和遮挡目标的检测能力, 网络模型需要利用更多的细节信息, 因此需要对 FPN 进行改进。PANet (Path Aggregation Network)^[18] 是一种最早应用于图像实例分割任务的路径聚合网络, 其通过自下而上的路径聚合策略, 利用低层次特征对其他层次特征进行加强, 提升了网络对细节的感知能力。将路径聚合策略与 FPN 结合, 可以更有效地利用细节特征信息, 改善模型对小目标和被遮挡目标的检测效果。网络结构如图 3 所示。

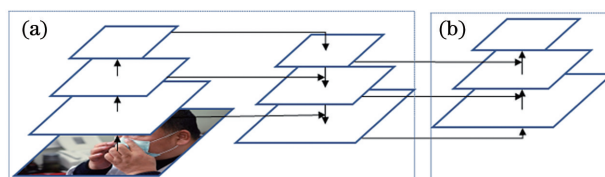


图 3 改进的特征融合网络结构。(a) 特征金字塔网络; (b) 自下而上路径聚合

Fig. 3 Improved feature fusion network structure. (a) FPN; (b) bottom-up path aggregation

3 YOLO-Mask 口罩佩戴检测算法

3.1 YOLO-Mask 网络结构

本研究用 SEDarknet-53 作为 YOLOv3 的基础

网络。首层是 3×3 的卷积层,然后经过 5 个分别包含有 1、2、8、8、4 个注意力残差单元的 SE-residual block 进行残差学习。每个 SE-residual block 之前是 Conv(3×3) + BN + Leaky ReLU 的组合,卷积核步长为 2,作用是下采样。SEDarknet-53 输出 3 种不同尺度的有效特征层,并传给采用路径聚合策略的 FPN 网络进行特征融合。SE-residual block5 输出的特征层经过 Convolutional set1 实现通道调整,并在卷积和上采样后,与 SE-residual block4 输出的张量拼接融合,同样地,融合后的信息经过 Convolutional set2、卷

积和上采样,然后与 SE-residual block3 的输出进行融合,此次融合后的信息一方面通过 YOLO head3 进行尺度为 52×52 的预测,一方面进行下采样,再次融合 Convolutional set2 的输出信息,5 次卷积处理后,作为 YOLO head2 的输入,进行尺度为 26×26 的预测。同理,可得到 YOLO head1 的输入,进行尺度为 13×13 的预测。YOLO head 包含一层 Conv(3×3) + BN + Leaky ReLU 卷积和一层普通 Conv(3×3) 卷积,利用其输出结果对先验框进行调整,可得到实际预测框。网络结构如图 4 所示。

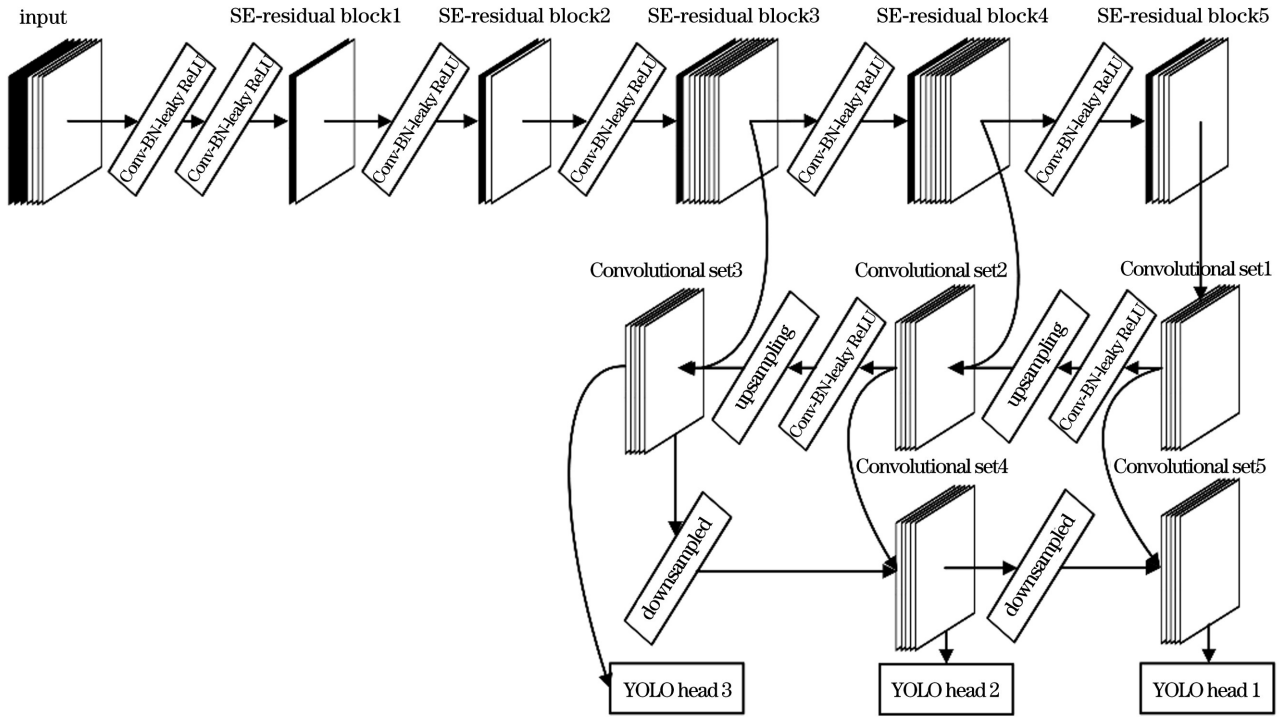


图 4 YOLO-Mask 算法网络结构

Fig. 4 YOLO-Mask algorithm network structure

与原始的 YOLOv3 算法相比, YOLO-Mask 网络利用对卷积特征各个通道之间的作用关系来建模,使网络具有更强的表达能力,同时能够更充分地利用不同尺度特征层的信息,提升算法的检测精度。

3.2 YOLO-Mask 的损失函数

YOLO-Mask 的损失函数 L 包含三部分:目标置信度损失 L_{conf} 、目标类别损失 L_{cla} 、目标定位损失 L_{loc} 。其中,目标定位损失 L_{loc} 采用 DIoU Loss^[19] 代替原有的均方误差损失,使损失函数 L 得到优化。DIoU Loss 综合考虑预测框与真实框的重叠程度、中心点距离两个主要几何要素,改善了回归的精度。损失函数的表达式为

$$L = L_{\text{conf}} + L_{\text{cla}} + L_{\text{loc}}, \quad (1)$$

$$L_{\text{conf}} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[\hat{C}_i^j \log C_i^j + (1 - \hat{C}_i^j) \times \log(1 - C_i^j) \right] - \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} \left[\hat{C}_i^j \log C_i^j + (1 - \hat{C}_i^j) \log(1 - C_i^j) \right], \quad (2)$$

式中: S^2 有 3 种取值,分别为 13×13 , 26×26 , 52×52 , 代表着 YOLOv3 输出的 3 种不同尺度特征图上的网格数; B 表示先验框个数; I_{ij}^{obj} 表示第 i 个网格的第 j 个先验框是否有负责预测的目标,如果有则 $I_{ij}^{\text{obj}} = 1$, 否则 $I_{ij}^{\text{obj}} = 0$; I_{ij}^{noobj} 表示第 i 个网格的第 j 个先验框是否没有负责预测的目标,如果没有则 $I_{ij}^{\text{noobj}} = 1$, 否则 $I_{ij}^{\text{noobj}} = 0$; \hat{C}_i^j 和 C_i^j 分别表示置信度的真实值与预测值。 λ_{noobj} 为常数系数,一般取 0.5,目

的是平衡正负样本。

$$L_{\text{cla}} = - \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in N} \left\{ \hat{p}_i^j(c) \log [p_i^j(c)] + [1 - \hat{p}_i^j(c)] \log [1 - p_i^j(c)] \right\}, \quad (3)$$

式中: S^2 、 B 和 I_{ij}^{obj} 与(2)式中一致; c 是目标类别; N 是目标总类别数; $\hat{p}_i^j(c)$ 与 $p_i^j(c)$ 分别第 i 个网格中第 j 个先验框中目标属于某一类别真实与预测的概率值。

$$L_{\text{loc}} = 1 - I_{\text{IoU}} + \frac{\rho^2(b, b^{\text{gt}})}{c^2}, \quad (4)$$

式中: $\rho^2(b, b^{\text{gt}})$ 代表预测框和真实框中心点的欧氏距离; I_{IoU} 表示预测框和真实框面积的交并比; c 代表能够同时包含预测框和真实框的最小闭包区域的对角线距离。DIoU Loss 的采用, 改善了回归精度, 降低了预测误差。



图 5 数据集图片示例

Fig. 5 Example of dataset images

针对数据集中错误佩戴口罩图片相对较少而产生的样本失衡的问题, 利用随机裁剪、旋转、色彩变换等方法对部分数据进行样本增强, 将数据集扩展为 4500 张, 使得样本多样性增加, 样本质量提高, 同时有效缓解了训练过程中的过拟合现象。根据实验要求将数据集转换成 VOC2007 数据集格式, 并用 LabelImg 标注软件进行数据标注, 标注类别分别为 right_mask、wrong_mask、no_mask 三类。

在数据集中随机选取 4000 张图片作为训练样本, 500 张图片作为测试样本。使用 Adam 优化器对网络进行优化, 初始学习率(learn rate)设置为 0.001, 并采用自适应调整策略对学习率进行动态调整, 其中调整倍率(factor)为 0.5, 忍受值(patience)为 2, 当连续两个 Epoch 指标不发生变化时, 学习率以调整倍率调整当前学习率。为了防止过拟合, 当连续 6 个 Epoch 评价指标不变化时, 训练过程终止。训练时一个批次包含 4 张图片, 共进行 120 个 Epoch 的训练, 最终 loss 值收敛到 3 左右。参数设置如表 1 所示。

4 实验与结果分析

4.1 实验平台

实验环境配置如下: Intel i7-9750H 处理器, Nvidia GeForce RTX 2060 显卡, 16 GB 内存, 操作系统为 Ubuntu 16.04。实验采用 Python 编程语言, 深度学习框架为 Keras。

4.2 数据集构建和模型训练

由于目前人脸口罩佩戴图片的数据集还比较少, 照片质量不高, 因此本文使用 WIDER Face、MAFA 两个开源的人脸图片数据集部分图片, 并通过个人拍摄和互联网搜集, 自制了人脸口罩佩戴数据集。数据集包含各种场景下人们佩戴口罩的图片, 共 3000 张。数据集图片包含不同场景下人们口罩佩戴正确、错误和无佩戴口罩三种情况, 其中还包括部分遮挡、阴影等情形, 如图 5 所示。

表 1 关键参数选取

Table 1 Selection of key parameters

Batch size	Factor	Patience	Learn rate
4	0.5	2	0.001

4.3 评价指标

口罩佩戴检测中各个类别的检测精度都很重要, 尤其在疫情防控的形势下, 误检和漏检都可能会增大疫情传播的风险。因此本文选择平均精度(AP, X_{AP})和平均精度均值(mAP, X_{mAP})作为目标检测算法的评价指标。这两种评价指标兼顾了精确率(P)和召回率(R), 表达式为

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \times 100\%, \quad (5)$$

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \times 100\%, \quad (6)$$

式中: 以本文所检测目标中的 right_mask 类别为例, N_{TP} 表示检测模型将正确佩戴口罩目标检测为 right_mask 的数量, N_{FP} 表示误把错误佩戴和无佩戴口罩目标检测为 right_mask 的数量, N_{FN} 表示误

把正确佩戴口罩目标检测为 wrong_mask 或 no_mask 的数量。分别取召回率和精确率的值作为横坐标和纵坐标,绘制了一条 $P-R$ 曲线,曲线下的面积即为 AP。对所有类别求 AP 并取均值即为平均精度均值。mAP 是评估模型性能的重要指标,可以反映网络模型的整体性能,避免评价过程中出现某些类别性能极端化而弱化其他类别性能的问题。

4.4 结果分析

以数据集中随机选取的 500 张图片作为测试集进行测试。将实验结果与原始 YOLOv3、Faster R-CNN、SSD 等主流目标检测算法的检测结果进行对比分析,检测结果如图 6 所示。

从各检测算法的检测效果来看,SSD 算法的检测效果不太好,尤其是对小目标的检测效果较差,如图 6(a)所示;图 6(d)中,Faster R-CNN 在小目标检测上优于 SSD 算法,但对被遮挡目标检测效果并不好,如图 6(f)所示;原始 YOLOv3 整体检测效果明显优于前两种算法,尤其是对小目标检测,检测率相对较高,但其存在置信度低和误检、漏检的情况,图 6(g)左右方站立的男士被误检,同时在左侧窗帘后的被遮挡目标也未检出,图片 6(i)被遮挡目标的置信度低;图 6(j)、(k)、(l)为本文算法的检测结果,相较于其他 3 种算法,本文算法在小目标和遮挡目标上有更好的检测效果且置信度较高,成功检测出了图 6(j)中左侧窗帘后被严重遮挡的目标,置信度为 0.78。相比于其他算法,本文算法在保证检测精度的同时,改善了小目标和被遮挡目标的检测率,鲁棒性较强。

改进 YOLOv3 算法检测结果的 $P-R$ 曲线如图 7 所示。图中,横坐标为召回率,纵坐标为精确率。通过计算 $P-R$ 曲线阴影部分面积可得到 AP 值。right_mask、wrong_mask、no_mask 3 种类别的 AP 值分别达到了 94.30%、91.13%、94.56%。将本文



图 6 不同算法检测结果对比。(a)(b)(c) SSD 检测结果;(d)(e)(f) Faster R-CNN 检测结果;(g)(h)(i) YOLOv3 检测结果;(j)(k)(l) YOLO-Mask 算法检测结果

Fig. 6 Comparison of detection results of different algorithms. (a) (b) (c) SSD detection results; (d) (e) (f) Faster R-CNN detection results; (g) (h) (i) YOLOv3 detection results; (j) (k) (l) YOLO-Mask detection results

算法与其他主流目标检测算法性能指标进行对比,对比结果如表 2 所示。

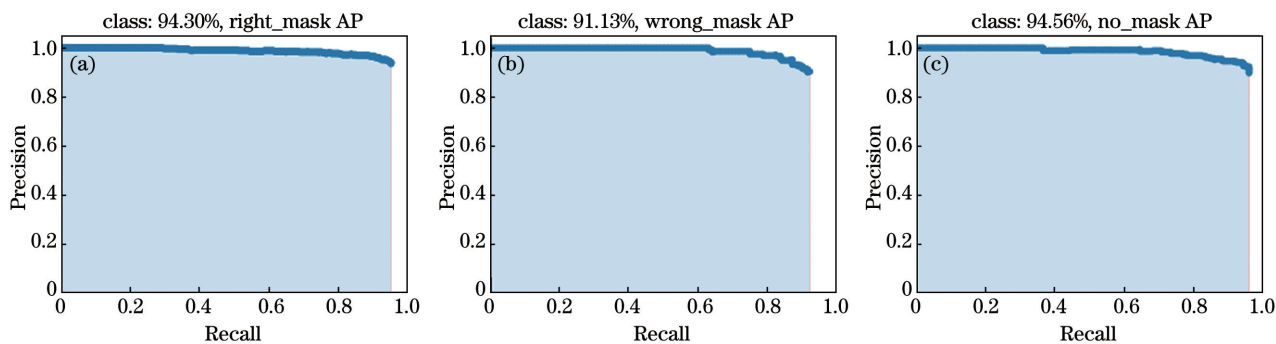


图 7 YOLO-Mask 算法各类目标 $P-R$ 曲线。(a) Right_mask; (b) wrong_mask; (c) no_mask

Fig. 7 Single target $P-R$ curves of YOLO-Mask algorithm. (a) Right_mask; (b) wrong_mask; (c) no_mask

表 2 不同检测算法性能指标对比

Table 2 Comparison of performance indicators of different detection algorithms

unit: %

Method	AP			mAP
	Right_mask	Wrong_mask	No_mask	
SSD	67.33	73.34	65.17	68.61
Faster R-CNN	75.78	71.27	72.94	73.33
YOLOv3	89.81	79.97	87.34	85.71
YOLO-Mask	94.30	91.13	94.56	93.33

从检测算法的性能指标来看:与 Faster R-CNN、SSD 等算法相比, YOLO-Mask 算法具有更高的检测精度。与原始 YOLOv3 相比, 本文算法在 3 种类别检测目标的平均检测精度上分别提高 4.49 个百分点、11.16 个百分点和 7.22 个百分点, 在 mAP 上提高了 7.62 个百分点。综上所述, 相比于其他 3 种算法, YOLO-Mask 算法在性能上有一定优势。模型能够更充分地使用低层次特征信息, 提高了对小目标的检测率; 注意力机制减少了无用特征对模型的干扰, 然后通过路径聚合策略的 FPN 网络实现了特征增强, 利用增强后的特征信息进行目标检测, 改善了遮挡目标的检测效果, 提升了模型性能。实验结果证明了算法的有效性。

5 结 论

本文提出一种基于 YOLO-Mask 算法的口罩佩戴检测方法, 用于解决小目标和被遮挡目标检测困难的问题。首先, 采用通道注意力机制对 YOLOv3 的特征提取网络进行改进, 从而得到 SEDarkNet-53 主干网络, 提高有用特征的通道权重, 抑制无用特征对模型的干扰。然后, 利用 FPN 和路径聚合策略进行特征融合, 加强低层特征的使用, 改善了网络对小目标和被遮挡目标的检测效果。在损失函数中引入 DIoU 损失, 减少预测框回归的误差, 提高了模型性能。最后, 本文建立了自然场景下口罩佩戴数据集, 并在数据集上进行算法验证和对比实验。实验结果表明, 本文提出的 YOLO-Mask 算法平均精度均值达到 93.33%, 在小目标和被遮挡目标检测上取得不错的检测效果, 相比于其他主流算法, 本文算法有较好的准确性和鲁棒性, 在疫情防控工作上具有良好的应用前景。

参 考 文 献

[1] Xia S, Li J, Ni Z L. The impact of the new crown pneumonia epidemic on my country's social and

economic development [J]. Financial Supervision, 2020(10): 5-9.

夏上, 李俊, 倪志良. 新冠肺炎疫情对我国社会经济发展的影响 [J]. 财政监督, 2020(10): 5-9.

[2] Yu Q M, Zheng D H. Study on the normalization of epidemic prevention and control of the corona virus disease 2019 [J]. China Public Security (Academy Edition), 2020(1): 65-68.

虞乔木, 郑东桦. 新冠肺炎疫情防控常态化研究 [J]. 中国公共安全(学术版), 2020(1): 65-68.

[3] Zhou Y P, Jiang Y, Rao H, et al. Investigation on status quo using masks among the public during the outbreak of COVID-19 [J]. Chinese Nursing Research, 2020, 34(11): 2041-2044.

周艳萍, 姜怡, 饶翻, 等. 新型冠状病毒肺炎疫情期间公众使用口罩现状调查 [J]. 护理研究, 2020, 34(11): 2041-2044.

[4] Niu Z D, Qin T, Li H D, et al. Improved algorithm of RetinaFace for natural scene mask wear detection [J]. Computer Engineering and Applications, 2020, 56(12): 1-7.

牛作东, 覃涛, 李捍东, 等. 改进 RetinaFace 的自然场景口罩佩戴检测算法 [J]. 计算机工程与应用, 2020, 56(12): 1-7.

[5] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587

[6] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.

[7] Ren S Q, He K M, Girshick R, 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.

[8] Liu W, Anguelov D, Erhan D, et al. SSD: single

- shot MultiBox detector[M]//Computer Vision-ECCV 2016. Cham: Springer, 2016: 21-37.
- [9] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [11] Redmon J, Farhadi A. YOLOV3: an incremental improvement [C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT. New York: IEEE Press, 2018: 1-6.
- [12] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [13] Guo J X, Liu L B, Xu F, et al. Airportscene aircraft detection method based on YOLOv3 [J]. Laser & Optoelectronics Progress, 2019, 56(19): 191003.
郭进祥, 刘立波, 徐峰, 等. 基于 YOLOv3 的机场场面飞机检测方法 [J]. 激光与光电子学进展, 2019, 56(19): 191003.
- [14] Lyu S, Cai X, Feng R. YOLOv3 network based on improved loss function [J]. Computer Systems & Applications, 2019, 28(2): 1-7.
- 吕铄, 蔡烜, 冯瑞. 基于改进损失函数的 YOLOv3 网络 [J]. 计算机系统应用, 2019, 28(2): 1-7.
- [15] Li H B, Xu C Y, Hu C C. Improved real-time vehicle detection method based on YOLOv3 [J]. Laser & Optoelectronics Progress, 2020, 57(10): 101507
李汉冰, 徐春阳, 胡超超. 基于 YOLOv3 改进的实时车辆检测方法 [J]. 激光与光电子学进展, 2020, 57(10): 101507.
- [16] Zhou J, Jing J F, Zhang H H, et al. Real-time fabric defect detection algorithm based on S-YOLOv3 model [J]. Laser & Optoelectronics Progress, 2020, 57(16): 161001.
周君, 景军锋, 张缓缓, 等. 基于 S-YOLOV3 模型的织物实时缺陷检测算法 [J]. 激光与光电子学进展, 2020, 57(16): 161001.
- [17] Hu J, Sun G, Wu E H, et al. Squeeze and excitation networks [C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT. New York: IEEE Press, 2018: 7132-7141.
- [18] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation [C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT. New York: IEEE Press, 2018: 8759-8768.
- [19] Zheng Z H, Wang P, Liu W, et al. Distance-IoU Loss: faster and better learning for bounding box regression [C] // Proceedings of the 2020 AAAI Conference on Artificial Intelligence, February 7-12, 2020, Hilton Midtown, New York. Menlo Park, CA: AAAI Press, 2020.