

DOI:10.19651/j.cnki.emt.2108620

基于迁移学习的轻量化YOLOv2口罩佩戴检测方法^{*}

张烈平 李智浩 唐玉良

(桂林理工大学机械与控制工程学院 桂林 541006)

摘要: 针对当前佩戴口罩数据集样本数量较少、硬件条件受限的情况下,本文提出了一种基于迁移学习的轻量化YOLOv2口罩佩戴检测方法。该方法以YOLOv2目标检测方法为基础,利用参数迁移学习的MobileNetV2作为特征提取网络,简化了网络模型并提高了训练速度。预训练的MobileNetV2特征提取网络与YOLOv2目标检测网络结合构成口罩佩戴检测网络模型。收集并建立了1000张人脸佩戴口罩图片数据集对网络模型进行训练和测试。实验结果表明,与YOLOv2、SSD300模型相比,MobileNetV2-YOLOv2模型口罩佩戴检测平均准确率提高3.8%、2.7%,检测速度提升2.5和2.4倍。并且在光线不足和密集检测条件下,MobileNetV2-YOLOv2依然可以有效进行口罩佩戴检测,相较于R-CNN和Faster-RCNN具有更好的检测效果,体现了更强的鲁棒性。

关键词: 神经网络;口罩佩戴检测;迁移学习;YOLOv2;MobileNetV2

中图分类号: TP391.9 **文献标识码:** A **国家标准学科分类代码:** 520.2

Light-YOLOv2 mask wearing detection method based on transfer learning

Zhang Lieping Li Zhihao Tang Yuliang

(College of Mechanical and Control Engineering, Guilin University of Technology, Guilin 541006, China)

Abstract: In view of the small number of samples in the current wearing mask data set and the limited hardware conditions, this paper proposes a lightweight YOLOv2 mask wearing detection method based on transfer learning. Based on the YOLOv2 target detection method, this method uses the MobileNetV2 of parameter transfer learning as the feature extraction network, which simplifies the network model and improves the training speed. The pre trained MobileNetV2 feature extraction network and YOLOv2 target detection network are combined to form a mask wearing detection network model. Collects and establishes a data set of 1000 pictures of face wearing masks to train and test the network model. The experimental results show that compared with YOLOv2 and SSD300 models, the average accuracy of mask wearing detection of MobileNetV2-YOLOv2 model is improved by 3.8%, 2.7%, and the detection speed increased by 2.5 and 2.4 times. Moreover, under the condition of insufficient light and dense detection, MobileNetV2-YOLOv2 can still effectively detect mask wearing, which has better detection effect and stronger robustness than R-CNN and Faster-RCNN.

Keywords: computer neural network; mask wearing detection; transfer learning; YOLOv2; MobileNetV2

0 引言

随着社会的不断发展,不仅是在新冠疫情期间的公共场合,在医疗领域、生产领域、甚至是在要求苛刻的无菌实验室都有佩戴口罩的规定。佩戴口罩在一定程度上也是对自身的一种保护,特别是在新冠疫情期间^[1],对行人进行是否佩戴口罩的高效率检测关系着整个疫情防控的进展。因此,如何检查人们在特殊时期或特定场所是否按要求佩戴口罩具有十分重要的意义^[2]。

目前,对口罩的佩戴情况的监督还主要是依靠佩戴口罩本人的自觉性和人工的检查。这种人工的方式不仅费时费力,容易漏检,而且很难满足全天候高效率的检测需求。鉴于检测口罩佩戴情况的机械性和重复性,再进一步考虑到现在的监控摄像头的广泛应用,利用目标检测网络对摄像头采集到的视频数据进行实时的佩戴口罩检测将是一个方便、快捷、可迅速大范围推广的方法。近些年来,随着深度学习算法的不断优化,以及庞大标注数据的开源和硬件算力的提升,使得采用计算机检测口罩佩戴情况变得更加

收稿日期:2021-12-17

^{*} 基金项目:国家自然科学基金(61741303)项目资助

准确和快速^[5]。基于深度学习的目标检测算法主要有基于候选区域的二阶段目标检测方法,如R-CNN^[4]、Fast-RCNN^[5]、Mask R-CNN^[6]等,以及基于回归的一阶段目标检测方法,如YOLO^[7]、SSD^[8]等。在PASCAL VOC2007数据集上相比与两阶段的目标检测算法,单阶段的目标检测算法在YOLOv1在检测准确率小幅下降的情况下明显提升了检测速度,YOLOv2同时兼顾了检测速度与准确率^[9]。因此,本文选择YOLOv2目标检测算法为基础进行口罩佩戴检测。目前,很多研究者已经将YOLO目标检测算法用于佩戴口罩的检测应用中。曹城硕等^[10]将注意特征网络SENet引入YOLO算法中,提出一种YOLO-Mask算法,提升了不同场景下口罩佩戴目标的检测能力。王艺皓等^[11]在YOLO算法中引入改进的空间金字塔池化结构和CIou损失函数提升佩戴口罩目标的检测精度。虽然以上改进算法明显提升了口罩佩戴检测准确性,但是对硬件依赖较高且需要大量数据集进行模型训练。口罩佩戴检测网络轻量化也是提升口罩佩戴检测算法性能的方向之一^[12]。牛作东等^[13]利用多任务卷积神经网络MTCNN结合轻量级卷积神经网络MobileNet在小样本数据集口罩佩戴检测准确率可达83%以上。薛均晓等^[14]提出一种轻量级卷积神经网络口罩佩戴检测方法,能够实现复杂条件下口罩佩戴的高精度检测。由于佩戴口罩人脸的数据集样本数量受限,运用迁移学习进行小样本训练能够取得较好的目标检测效果^[15]。Francesco等^[16]使用迁移学习的MobileNetV2作为口罩佩戴检测的网络模型,在4095张

口罩佩戴数据集中取得了更高的检测准确率。

本文提出了一种基于YOLOv2算法的口罩佩戴检测方法。考虑到口罩佩戴数据集数量和硬件条件有限,本文选取轻量化的MobileNetV2代替YOLOv2的Darknet-19,并与YOLOv2目标检测网络结合构成MobileNetV2-YOLOv2目标检测网络模型,同时使用参数迁移学习对特征提取网络预训练。将训练好的模型分别在光线不足和密集检测的条件下进行实验,对本模型进行口罩佩戴检测准确率和检测速度的验证。

1 基于YOLOv2的目标检测模型

1.1 YOLOv2目标检测网络模型

YOLOv2目标检测网络模型如图1所示。其中YOLOv2参考SSD和YOLOv1网络结构设计了Darknet-19作为网络的BackBone(包含了19个卷积层和5个最大池化层)。使用了常用 3×3 的卷积并借鉴了network in network的思想,将 1×1 卷积置于两个 3×3 的卷积间进行特征压缩。YOLOv2加入了Batch Normalization(BN层)加快训练收敛^[17]。YOLOv2参考了Faster-RCNN^[18]的方法引入了Anchor机制来预测候选框,并通过K-means算法对训练集标注框进行聚类处理产生符合要求的先验框。通过修正预测候选框的直接预测目标位置,提升YOLOv2网络检测速度。YOLOv2通过passthrough的拆分合并操作使得输出整合了底层细粒度特征信息以及高层的语义特征信息,提高了YOLOv2对于不同分辨率图像的检测能力。

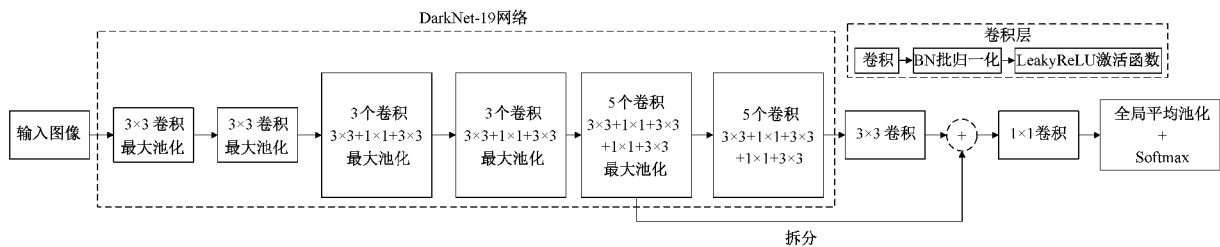


图1 YOLOv2目标检测网络模型

1.2 基于MobileNetV2参数迁移的YOLOv2目标检测网络模型

YOLOv2的基础网络为Darknet-19,然而针对口罩佩戴检测这种单一类别的目标检测任务中,Darknet-19较为复杂,具有较高的冗余性。为简化模型,提高训练速度并降低模型对于硬件性能的依赖。选取轻量级的MobileNetV2代替YOLOv2的Darknet-19作为骨干网络。MobileNetV2通过深度可分离卷积(Depthwise Separable Conv)中的DW卷积和两个 1×1 卷积分别进行升维和降维操作。相比与Darknet-19中常规的 3×3 卷积,深度可分离卷积可以大量减少模型的参数量和计算量,从而可以提升网络的运算速度。通过ImageNet数据集对MobileNetV2模型进行预训练,并使用模型参数迁移学习,即只替换训练分类层而保留源模型的全部特征提取

能力。将预训练的MobileNetV2特征提取网络与YOLOv2的目标检测网络构成MobileNetV2-YOLOv2目标检测网络模型,网络结构如图2所示。其中Bottleneck倒残差网络层由一个 1×1 卷积的上采样(RELU6激活函数)、一个 3×3 的DW卷积(RELU6激活函数)和一个 1×1 卷积的下采样(线性激活函数)组成。

2 基于YOLOv2迁移学习的口罩佩戴检测方法

2.1 口罩佩戴检测实现流程

本文基于YOLOv2迁移学习的口罩佩戴检测实现流程如图3所示。首先获取口罩佩戴图片数据集,进行是否佩戴口罩的数据集标签和口罩佩戴位置标注,将制作完成的数据集按比例划分为实验的训练集和测试集。实验数据集进行预处理操作后加载到基于MobileNetV2参数迁

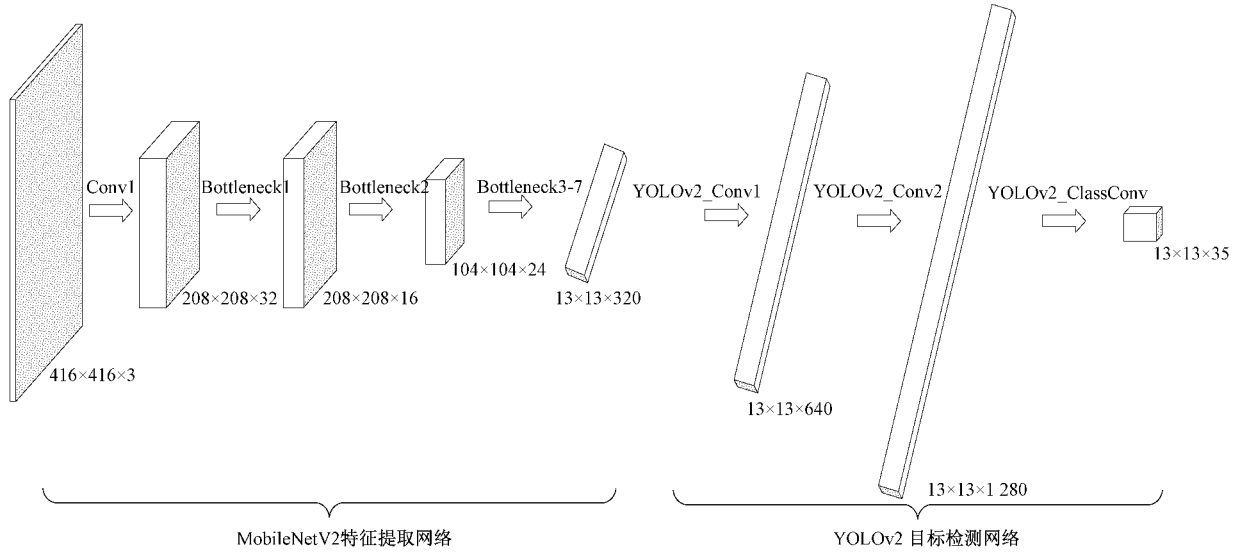


图 2 MobileNetV2-YOLOv2 目标检测网络模型

移的 YOLOv2 目标检测网络模型进行训练。得到性能最优的口罩佩戴检测模型后,对测试集数据进行检测评估,并在测试图片中标记口罩佩戴识别结果与准确率。

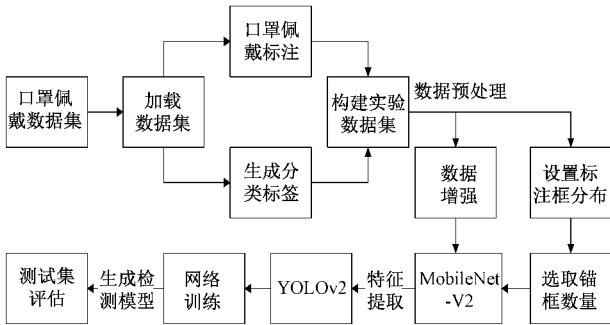


图 3 基于 YOLOv2 迁移学习的口罩佩戴检测实现流程图

2.2 实验数据集构建

由于目前人脸佩戴口罩图片的数据集比较少,因此本文根据互联网搜索图片自制小样本的口罩数据集。数据集中包含不同场景下人脸佩戴口罩图片并将口罩数据集分为两个标签类别分别为佩戴口罩 Mask 和未佩戴口罩 no_Mask,为保证模型的训练和检测效果,两种标签图片数量应大概保持一致。通过 Face Mask Detector 口罩标注算法进行口罩数据集标注并进行人工修正。最终获得数据集 1 000 张,其中训练集和测试集的比例为 7 : 3。部分标注图片如图 4 所示。

2.3 数据预处理

输入网络模型的数据集的预处理操作分为数据集标记、锚框设置和数据增强。其中数据集标记归一化操作计算规则如式(1)~(4)所示。式中 (x_{max}, y_{max}) 表示输入数据图像标注框左上角在图像中的具体坐标, (x_{min}, y_{min}) 表示输入数据图像标注框右下角在图像中的具体位置坐标。 (x, y) 表示处理后图像中心点的坐标, (w, h)



图 4 部分口罩佩戴数据集标注图片

表示处理后图像的长和宽。

$$x = (x_{max} + x_{min}) / 2 \times width \quad (1)$$

$$y = (y_{max} + y_{min}) / 2 \times height \quad (2)$$

$$w = (x_{max} - x_{min}) / width \quad (3)$$

$$h = (y_{max} - y_{min}) / height \quad (4)$$

YOLOv2 使用 K-means 算法聚类锚框,在聚类锚框时希望候选框与边界框之间有更高的交并比(intersection over union, IOU),距离度量公式为

$$IOU(box, centroid) = \frac{box \cap centroid}{box \cup centroid} \quad (5)$$

$$d(box, centroid) = 1 - IOU(box, centroid) \quad (6)$$

其中,centroid 表示中心点的边界框,box 表示样本框。IOU(box, centroid)表示中心店边界框与样本框的交并比,为得到最优的交并比,通过逐渐递增锚框个数计算中心点平均 IOU,当锚框数为 13 时平均 IOU 最大为 0.9。因此选取最佳锚框个数为 13。为了增强检测网络模型的泛化能力,需要对输入的数据进行数据增强处理,在颜色空间中随机调整图像的色调、饱和度、亮度和对比度。为了证明调整的随机性,在原始图像上重复操作四次,数据增强过程如图 5 所示。

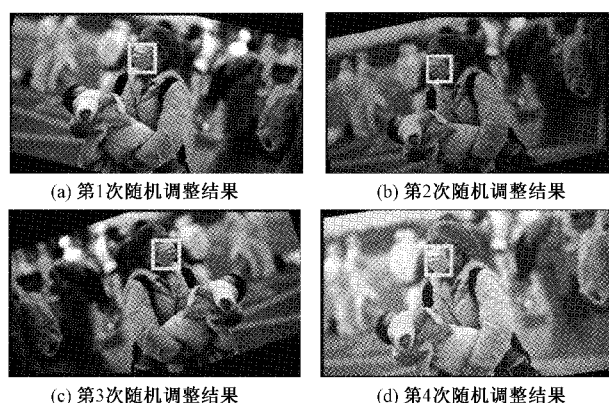


图5 图像数据增强过程

3 口罩佩戴检测实验

3.1 实验环境与测试环境

本文实验训练环境为 Win11 64 位操作系统, 16 G 内存, Intel Core i5-11600KF CPU @ 4.6 GHz, NVIDIA GEFORCE GTX3060TI, MATLAB2021a。选取 sgdmm 优化算法并设置 MobileNetV2-YOLOv2、YOLOv2、SSD300、R-CNN 和 Faster-RCNN 训练参数初始学习率为 0.001, 批处理尺度 MiniBatchSize 为 32, 最大迭代轮次为 50 轮。对于 MobileNetV2-YOLOv2, 选择加载预训练好的 MobileNetV2 特征提取网络, 输入预处理后的图片进行特征提取并进行 YOLOv2 目标的分类与检测。并分别将 YOLOv2、SSD300、R-CNN 和 Faster-RCNN 四种目标检测算法都按照预先设定的训练参数进行迭代训练, 直到最大迭代轮次停止模型训练输出其最佳性能的口罩佩戴检测模型, 运用训练好的模型对测试集数据进行测试并标注是否佩戴口罩的检测结果及置信度。模型测试环境选取移动端 Win10 64 位操作系统, 16 G 内存, AMD R7-4800U CPU@1.8 GHz, Matlab2021a。

3.2 模型评价指标

本文主要选用平均准确率 (average precision, AP)、P-R 曲线和检测速度 FPS (frames per second) 作为 MobileNetV2-YOLOv2 网络模型口罩佩戴检测性能评价指标。其中基于分类结果混淆矩阵进行模型准确率、查准率和召回率的计算。混淆矩阵中分为真正例 True Positive (TP)、假正例 False Positive (FP)、假反例 False Negative (FN) 和真反例 True Negative (TN)。3 种佩戴口罩检测性能指标的计算公式为

$$Accuracy = (TP + TN) / (P + N) \quad (7)$$

$$R_{recall} = TP / (TP + FN) \quad (8)$$

$$P_{precision} = TP / (TP + FP) \quad (9)$$

其中, TP 表示检测到目标标签与目标真实标签一致的样本数量; FP 表示检测到目标标签与目标真实标签不一致的样本数量; FN 表示真实目标存在并未检测出的样本数量。 P 和 N 为预测正确和预测错误的样本数量。根据计算得出 $P_{precision}$ 和 R_{recall} 绘制 P-R 曲线, 其中 P-R 曲线下方面积为目标检测的 AP 值。

3.3 实验结果与分析

首先选出一些样本分别输入到 YOLOv2、SSD300、R-CNN、Faster-RCNN 和本文提出的 MobileNetV2-YOLOv2 检测模型中进行检测, 图 6 和 7 分别为测试样本 1 和测试样本 2 在各个模型下的测试结果, 在样本 1 的单目标检测中各模型检测结果较为理想, 其中 R-CNN 和 Faster-RCNN 检测精度最高均为 100%, 并且 Faster-RCNN 的人脸佩戴口罩标记框更为准确, 在两阶段目标检测算法中 MobileNetV2-YOLOv2 准确率 92% 较 YOLOv2 准确率降低 4%, 但是高于 SSD300 的 80%。在样本 2 的多目标检测中, R-CNN 3 个人脸准确率均为 100%, 检测效果最好。Faster-RCNN 出现了口罩漏检现象, 且远处的小目标检测精度明显下降。在单阶段目标检测算法中 MobileNetV2-YOLOv2 准确率 3 个目标平均准确率达到

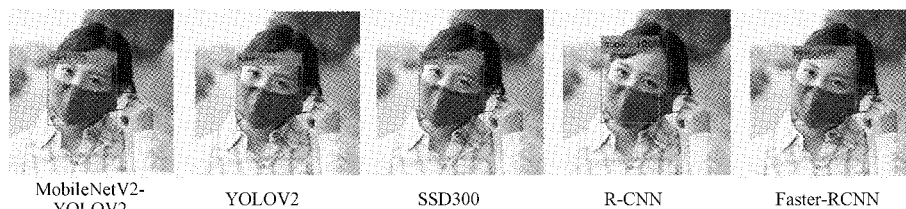


图6 样本1各模型检测结果

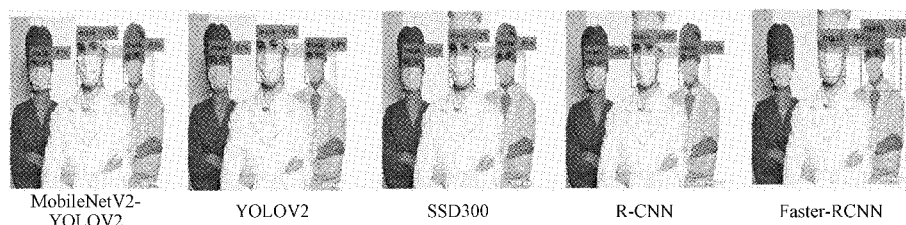


图7 样本2各模型检测结果

4 结 论

为提高口罩佩戴检测的准确性和效率,提出了一种基于迁移学习的轻量化YOLOv2口罩佩戴检测方法。本文从口罩佩戴检测模型轻量化角度出发,选用MobileNetV2代替YOLOv2的Darknet-19简化特征提取网络,减少网络训练参数,提升网络的运算速度。同时通过ImageNet训练MobileNetV2并进行参数迁移学习,降低网络模型的训练时间。使用小样本口罩佩戴数据对本文提出的MobileNetV2-YOLOv2目标检测网络模型进行训练,并分别在测试集和两种复杂环境下进行实验验证。实验结果表明,MobileNetV2-YOLOv2模型较原模型平均准确率提升3.8%,检测速度提高2.5倍,模型大小减少75.9%,同时相比其他算法能够更好的平衡检测精度和检测速度,具有良好的鲁棒性,更加适合应用到移动端和嵌入式设备实现口罩佩戴检测。未来将进一步在硬件条件有限的情况下提升本文口罩佩戴检测方法的检测精度和检测速度,实现在不同复杂环境下更加快速准确的实时口罩佩戴检测,提升检测模型的实用性。

参考文献

- [1] 冯基花,张剑峰.新型冠状病毒疫苗潜在不良反应的研究进展[J].广西科学,2021,28(2):103-112.
- [2] 张修宝,林子原,田万鑫,等.全天候自然场景下的人脸佩戴口罩识别技术[J].中国科学:信息科学,2020,50:1110-1120.
- [3] 刘启刚,孙向阳,徐伟.针对实时场景的口罩检测模型设计[J].实验技术与管理,2021,38(8):76-81.
- [4] GIRSHICK R, 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, OH, USA: IEEE, 2014: 580-587.
- [5] GIRSHICK R. Fast R-CNN[C]. Proceedings of 2015 IEEE International Conference on Computer Vision, Santiago, Chile: IEEE, 2015: 1440-1448.
- [6] HE K, GKIOVARI G, DOLLAR P, et al. Mask R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2980-2988.
- [7] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Press, 2017: 6517-6525.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multi-box detector[C]. Proceedings of 2016 European Conference on Computer Vision, Berlin, Germany: Springer, 2016: 21-37.
- [9] 严忱,严云洋,高尚兵,等.基于多级特征融合的视频火焰检测方法[J].南京师大学报(自然科学版),2021,44(3):131-136.
- [10] 曹城硕,袁杰.基于YOLO-Mask算法的口罩佩戴检测方法[J].激光与光电子学进展,2021,58(8):211-218.
- [11] 王艺皓,丁洪伟,李波,等.复杂场景下基于改进YOLOv3的口罩佩戴检测算法[J].计算机工程,2020,46(11):12-22.
- [12] 宋秦年,秦伟伟,梁卓,等.面向轻量化网络的改进双通道机制图像分类方法[J].航空兵器.2021,28(5):81-85.
- [13] 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.
- [14] 薛均晓,程君进,张其斌,等.改进轻量级卷积神经网络的复杂场景口罩佩戴检测方法[J].计算机辅助设计与图形学学报,33(7):1045-1053.
- [15] 赵凯琳,靳小龙,王元卓.小样本学习研究综述[J].软件学报,2021,32(2):349-369.
- [16] FRANCESCO M, ANTOELLA S. Transfer learning for mobile real-time face mask detection and localization[J]. Journal of the American Medical Informatics Association, 2021, 28(7): 1548-1554.
- [17] LOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. International Conference on Machine Learning, PM-LR, 2015: 448-456.
- [18] 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.

作者简介

张烈平,博士,教授,主要研究方向为检测与传感技术。

E-mail: 25761108@qq.com

李智浩,硕士研究生,主要研究方向为智能算法及其应用。

E-mail: lzh15639786020@qq.com

唐玉良,硕士研究生,主要研究方向为计算机应用技术。

E-mail: ty_l_shanghaiigong@163.com