



(12) 发明专利申请

(10) 申请公布号 CN 113139444 A

(43) 申请公布日 2021. 07. 20

(21) 申请号 202110376357.X

(22) 申请日 2021.04.06

(71) 申请人 上海工程技术大学

地址 201620 上海市松江区龙腾路333号

(72) 发明人 赵晓丽 尹明臣 陈正 张佳颖

(74) 专利代理机构 北京翔石知识产权代理事务
所(普通合伙) 11816

代理人 李勇

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/34 (2006.01)

G06K 9/62 (2006.01)

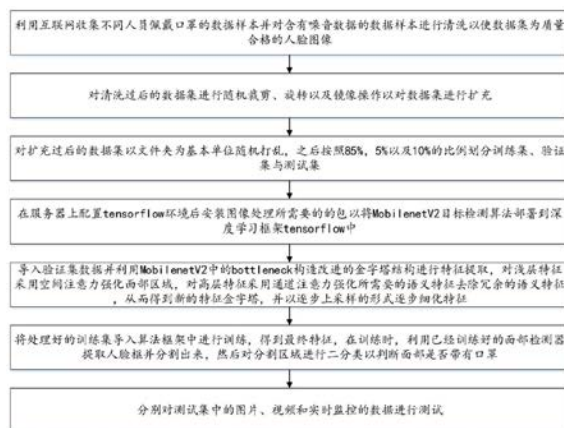
权利要求书2页 说明书6页 附图4页

(54) 发明名称

基于MobileNetV2的时空注意力口罩佩戴实时检测方法

(57) 摘要

本发明涉及一种基于MobileNetV2的时空注意力口罩佩戴实时检测方法,包括:步骤一,利用互联网收集不同人员佩戴口罩的数据样本并对含有噪音数据的数据样本进行清洗以使数据集为质量合格的人脸图像;步骤二,数据集扩充;步骤三,划分数数据集;步骤四,在服务器上配置tensorflow环境后将MobilenetV2目标检测算法部署到深度学习框架tensorflow中;步骤五,导入验证集数据并利用MobilenetV2中的bottleneck构造改进的金字塔结构进行特征提取,对浅层特征采用空间注意力强化面部区域,对高层特征采用通道注意力强化所需要的语义特征去除冗余的语义特征,从而得到新的特征金字塔,并以逐步上采样的形式逐步细化特征等,从而能够实现高效的人脸口罩检测任务。



1. 一种基于MobileNetV2的时空注意力口罩佩戴实时检测方法,其特征在于,包括:

步骤一,利用互联网收集不同人员佩戴口罩的数据样本并对含有噪音数据的数据样本进行清洗以使数据集为质量合格的人脸图像;

步骤二,对清洗过后的数据集进行随机裁剪、旋转以及镜像操作以对数据集进行扩充;

步骤三,对扩充过后的数据集以文件夹为基本单位随机打乱,之后按照85%,5%以及10%的比例划分训练集、验证集与测试集;

步骤四,在服务器上配置tensorflow环境后安装图像处理所需要的的包以将MobilenetV2目标检测算法部署到深度学习框架tensorflow中;

步骤五,导入验证集数据并利用MobilenetV2中的bottleneck构造改进的金字塔结构进行特征提取,对浅层特征采用空间注意力强化面部区域,对高层特征采用通道注意力强化所需要的语义特征去除冗余的语义特征,从而得到新的特征金字塔,并以逐步上采样的形式逐步细化特征;

步骤六,将处理好的训练集导入算法框架中进行训练,得到最终特征,在训练时,利用已经训练好的面部检测器提取人脸框并分割出来,然后对分割区域进行二分类以判断面部是否带有口罩;

步骤七,分别对测试集中的图片、视频和实时监控的数据进行测试。

2. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法,其特征在于,所述测试为测试精确率precision和召回率recall,其计算公式如下:

$$\text{precision} = \text{TP} / (\text{TP} + \text{FP});$$

$$\text{recall} = \text{TP} / (\text{TP} + \text{FN});$$

式中,TP表示正确预测正样本,FP表示错误预测正样本,TN表示正确预测负样本,FN表示错误预测负样本;

计算完成时,计算平均精度AP和平均精度均值mAP,其计算公式如下:

$$AP = \int_0^1 \text{precision}(\text{recall}) d_{\text{recall}};$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i。$$

3. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法,其特征在于,为了评估本发明所述方法的鲁棒性,在口罩检测系统中对单人、多人以及密集人群、特殊情况下的有遮挡、视频和自然场景中实时监控下的人群进行口罩检测。

4. 根据权利要求3所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法,其特征在于,所述单人检测是为了便于观察及分析改进模型对于人脸口罩检测的精度及可操作性;所述多人测试是为了验证模型的泛化能力;所述特殊情况下的有遮挡包括捂口和遮嘴;所述在视频中进行测试是为了验证本发明所述算法在自然场景下口罩佩戴检测的效果。

5. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法,其特征在于,所述步骤一中利用互联网收集指的是在网上进行爬取和下载已经公布出来的数据。

6. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法, 其特征在于, 所述步骤一中对含有噪音数据的数据样本进行清洗采用双边滤波函数。

7. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法, 其特征在于, 所述步骤三中划分比例的原则为采用留出法进行划分。

8. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法, 其特征在于, 所述步骤五中进行特征提取是依赖卷积操作, 使用局部感受区域的思想融合空间信息和通道信息来提取包含信息的特征。

9. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法, 其特征在于, 所述步骤六中利用已经训练好的面部检测器提取人脸框并分割出来采用的是基于锚点的分割; 所述二分类包括戴口罩和不戴口罩两类。

10. 根据权利要求1所述的基于MobileNetV2的时空注意力口罩佩戴实时检测方法, 其特征在于, 所述步骤七中进行测试采用的是把测试样本输入到训练好的网络中进行测试。

基于MobileNetV2的时空注意力口罩佩戴实时检测方法

技术领域

[0001] 本发明涉及口罩佩戴实时检测领域,尤其涉及基于MobileNetV2的时空注意力口罩佩戴实时检测方法。

背景技术

[0002] 传染病毒目前还是主要通过飞沫传播和接触传播,在某些特殊的条件下才可能发生气溶胶传播,一般的工作生活条件下,正确佩戴口罩,足以满足日常防护需求。疫情期间,普通公众出门正确戴口罩显得尤为重要。为全力做好传染病毒的防控工作,有效切断病毒传播途径,坚决遏制疫情蔓延势头,确保人民群众生命安全和身体健康,需要在各小区、学校、单位、食堂、车站等公共场合增设大量的防疫人员和检测点,逐一检测进出人员是否佩戴口罩,但是该检查方法需要耗费大量人工资源,并且在人流量大时可能出现漏检等情况。而且由于疫情的传染性强及公共场所的复杂性,仅仅通过人力检查不可避免地存在感染风险高、工作强度大、覆盖面窄、实时性差等弊端,迫切需要机器自动检测口罩佩戴情况。基于计算机视觉实现佩戴口罩的自动检测,可有效缓解人工检测所面临的一系列问题,具有重要研究意义。

[0003] 近年来,专门应用于人脸口罩佩戴检测的相关算法研究较少,Deng等人提出的‘Retinaface’、牛等人提出的‘改进RetinaFace的自然场景口罩佩戴检测算法’和王等人提出的‘复杂场景下基于改进YoloV3的口罩佩戴检测算法’,受自然场景中存在的遮挡、密集、小尺度等复杂因素影响,检测效果不理想。

[0004] 目前,已经有针对以往改进的深度网络模型计算复杂、内存占用大的缺点的一些口罩佩戴实时检测方法,但是普遍效率低,且使用的网络模型不能在保证良好效果的前提下有效应用于移动端。

发明内容

[0005] 为此,本发明提供一种基于MobileNetV2的时空注意力口罩佩戴实时检测方法,可以有效解决现有技术中深度网络模型计算复杂、内存占用大导致检测效率低的技术问题。

[0006] 为实现上述目的,本发明提供一种基于MobileNetV2的时空注意力口罩佩戴实时检测方法,包括:

[0007] 步骤一,利用互联网收集不同人员佩戴口罩的数据样本并对含有噪音数据的数据样本进行清洗以使数据集为质量合格的人脸图像;

[0008] 步骤二,对清洗过后的数据集进行随机裁剪、旋转以及镜像操作以对数据集进行扩充;

[0009] 步骤三,对扩充过后的数据集以文件夹为基本单位随机打乱,之后按照85%,5%以及10%的比例划分训练集、验证集与测试集;

[0010] 步骤四,在服务器上配置tensorflow环境后安装图像处理所需要的的包以将MobilenetV2目标检测算法部署到深度学习框架tensorflow中;

[0011] 步骤五,导入验证集数据并利用MobilenetV2中的bottleneck构造改进的金字塔结构进行特征提取,对浅层特征采用空间注意力强化面部区域,对高层特征采用通道注意力强化所需要的语义特征去除冗余的语义特征,从而得到新的特征金字塔,并以逐步上采样的形式逐步细化特征;

[0012] 步骤六,将处理好的训练集导入算法框架中进行训练,得到最终特征,在训练时,利用已经训练好的面部检测器提取人脸框并分割出来,然后对分割区域进行二分类以判断面部是否带有口罩;

[0013] 步骤七,分别对测试集中的图片、视频和实时监控的数据进行测试。

[0014] 进一步地,所述测试为测试精确率precision和召回率recall,其计算公式如下:

[0015] $precision = TP / (TP + FP)$;

[0016] $recall = TP / (TP + FN)$;

[0017] 式中,TP表示正确预测正样本,FP表示错误预测正样本,TN表示正确预测负样本,FN表示错误预测负样本;

[0018] 计算完成时,计算平均精度AP和平均精度均值mAP,其计算公式如下:

[0019] $AP = \int_0^1 precision(recall) d_{recall}$;

[0020] $mAP = \frac{1}{n} \sum_{i=1}^n AP_i$ 。

[0021] 进一步地,为了评估本发明所述方法的鲁棒性,在口罩检测系统中对单人、多人以及密集人群、特殊情况下的有遮挡、视频和自然场景中实时监控下的人群进行口罩检测。

[0022] 进一步地,所述单人检测是为了便于观察及分析改进模型对于人脸口罩检测的精度及可操作性;所述多人测试是为了验证模型的泛化能力;所述特殊情况下的有遮挡包括捂口和遮嘴;所述在视频中进行测试是为了验证本发明所述算法在自然场景下口罩佩戴检测的效果。

[0023] 进一步地,所述步骤一中利用互联网收集指的是在网上进行爬取和下载已经公布出来的数据。

[0024] 进一步地,所述步骤一中对含有噪音数据的数据样本进行清洗采用双边滤波函数。

[0025] 进一步地,所述步骤三中划分比例的原则为采用留出法进行划分。

[0026] 进一步地,所述步骤五中进行特征提取是依赖卷积操作,使用局部感受区域的思想融合空间信息和通道信息来提取包含信息的特征。

[0027] 进一步地,所述步骤六中利用已经训练好的面部检测器提取人脸框并分割出来采用的是基于锚点的分割;所述二分类包括戴口罩和不戴口罩两类。

[0028] 进一步地,所述步骤七中进行测试采用的是把测试样本输入到训练好的网络中进行测试。

[0029] 与现有技术相比,本发明的有益效果在于,针对以往改进的深度网络模型计算复杂、内存占用大的缺点,对近几年发表的轻量级网络进行研究,受Google所提出的

MobileNet的启发,将MobileNetV2作为主干网络,并在其基础上进行了部分改进以实现高效的人脸口罩检测任务。改进后的网络模型可有效应用于移动端,同时又保证了良好的效果。

[0030] 进一步地,针对现有监控技术灵活性差的缺点,本发明使用可移动的摄像头来对目标进行检测和分类。可移动的摄像头具有高度灵活性,可以在大范围的场景之下对目标进行实时监控,并做出反应,有效解决了固定摄像头灵活性差、监控区域小等问题。

附图说明

[0031] 图1为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的流程示意图;

[0032] 图2为为本发明基于MobileNet V2的时空注意力口罩佩戴实时检测网络结构图;

[0033] 图3为本发明为MobileNet V2的Inverted Residual Block卷积过程图;

[0034] 图4为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的清洗过后的高清图像;

[0035] 图5为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的单人口罩检测结果图;

[0036] 图6为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的多人口罩检测结果图;

[0037] 图7为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的特殊情况下的口罩检测结果图;

[0038] 图8为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的自然场景下的视频口罩检测结果图;

[0039] 图9为本发明基于MobileNetV2的时空注意力口罩佩戴实时检测方法的自然场景下实时监控中口罩检测结果图。

具体实施方式

[0040] 为了使本发明的目的和优点更加清楚明白,下面结合实施例对本发明作进一步描述;应当理解,此处所描述的具体实施例仅仅用于解释本发明,并不用于限定本发明。

[0041] 下面参照附图来描述本发明的优选实施方式。本领域技术人员应当理解的是,这些实施方式仅仅用于解释本发明的技术原理,并非在限制本发明的保护范围。

[0042] 需要说明的是,在本发明的描述中,术语“上”、“下”、“左”、“右”、“内”、“外”等指示的方向或位置关系的术语是基于附图所示的方向或位置关系,这仅仅是为了便于描述,而不是指示或暗示所述装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。

[0043] 此外,还需要说明的是,在本发明的描述中,除非另有明确的规定和限定,术语“安装”、“相连”、“连接”应做广义理解,例如,可以是固定连接,也可以是可拆卸连接,或一体地连接;可以是机械连接,也可以是电连接;可以是直接相连,也可以通过中间媒介间接相连,可以是两个元件内部的连通。对于本领域技术人员而言,可根据具体情况理解上述术语在本发明中的具体含义。

[0044] 请参阅图1-9所示,本发明提供一种基于MobileNetV2的时空注意力口罩佩戴实时检测方法,包括:

[0045] 步骤一,利用互联网收集不同人员佩戴口罩的数据样本并对含有噪音数据的数据样本进行清洗以使数据集为质量合格的人脸图像;

[0046] 步骤二,对清洗过后的数据集进行随机裁剪、旋转以及镜像操作以对数据集进行扩充;

[0047] 步骤三,对扩充过后的数据集以文件夹为基本单位随机打乱,之后按照85%,5%以及10%的比例划分训练集、验证集与测试集;

[0048] 步骤四,在服务器上配置tensorflow环境后安装图像处理所需要的的包以将MobilenetV2目标检测算法部署到深度学习框架tensorflow中;

[0049] 步骤五,导入验证集数据并利用MobilenetV2中的bottleneck构造改进的金字塔结构进行特征提取,对浅层特征采用空间注意力强化面部区域,对高层特征采用通道注意力强化所需要的语义特征去除冗余的语义特征,从而得到新的特征金字塔,并以逐步上采样的形式逐步细化特征;

[0050] 步骤六,将处理好的训练集导入算法框架中进行训练,得到最终特征,在训练时,利用已经训练好的面部检测器提取人脸框并分割出来,然后对分割区域进行二分类以判断面部是否带有口罩;

[0051] 步骤七,分别对测试集中的图片、视频和实时监控的数据进行测试。

[0052] 本发明所述实施例中对数据集进行扩充后可以增强样本的多样性,使模型能够保持较高的泛化能力以防止训练中出现过拟合;所述质量合格的人脸指的是人脸在图中所占比例合适。所述对清洗过后的数据集进行随机裁剪、旋转以及镜像操作在代码中进行操作,为本领域常规技术手段;所述对数据集进行扩充指的是对数据进行增强;所述随机打乱的目的在于使数据呈现一定的随机性;所述训练集用于网络的学习,验证集是用来找出学习最好的模型参数,测试集则是用来检验算法性能的;所述对高层特征采用通道注意力强化采用的是注意力机制;所述去除冗余的语义特征指的是在注意力机制中使有用的特征权重增大,无用的特征权重减小;所述判断面部是否带有口罩基于本发明所述算法可直接得出。

[0053] 具体而言,所述测试为测试精确率precision和召回率recall,其计算公式如下:

[0054] $precision = TP / (TP + FP)$;

[0055] $recall = TP / (TP + FN)$;

[0056] 式中,TP表示正确预测正样本,FP表示错误预测正样本,TN表示正确预测负样本,FN表示错误预测负样本;

[0057] 计算完成时,计算平均精度AP和平均精度均值mAP,其计算公式如下:

[0058] $AP = \int_0^1 precision(recall) d_{recall}$;

[0059] $mAP = \frac{1}{n} \sum_{i=1}^n AP_i$ 。

[0060] 本发明所述实施例中的精确率表示预测为正样本的样本中真正为正样本所占的比例,召回率表示正样本中被成功预测为正样本的比例,平均精度AP(Average

Precision), 平均精度均值mAP (Mean Average Precision) 表示所有类别平均精度的均值, 反映了总体上的目标检测效果。

[0061] 具体而言, 为了评估本发明所述方法的鲁棒性, 在口罩检测系统中对单人、多人以及密集人群、特殊情况下的有遮挡、视频和自然场景中实时监控下的人群进行口罩检测。

[0062] 具体而言, 所述单人检测是为了便于观察及分析改进模型对于人脸口罩检测的精度及可操作性; 所述多人测试是为了验证模型的泛化能力; 所述特殊情况下的有遮挡包括捂口和遮嘴; 所述在视频中进行测试是为了验证本发明所述算法在自然场景下口罩佩戴检测的效果。

[0063] 具体而言, 所述步骤一中利用互联网收集指的是在网上进行爬取和下载已经公布出来的数据。

[0064] 具体而言, 所述步骤一中对含有噪音数据的数据样本进行清洗采用双边滤波函数。

[0065] 具体而言, 所述步骤三中划分比例的原则为采用留出法进行划分。

[0066] 具体而言, 所述步骤五中进行特征提取是依赖卷积操作, 使用局部感受区域的思想融合空间信息和通道信息来提取包含信息的特征。

[0067] 具体而言, 所述步骤六中利用已经训练好的面部检测器提取人脸框并分割出来采用的是基于锚点的分割; 所述二分类包括戴口罩和不戴口罩两类。

[0068] 具体而言, 所述步骤七中进行测试采用的是把测试样本输入到训练好的网络中进行测试。

[0069] 实施例1

[0070] 为了验证实时检测, 本发明采用一个外接摄像头进行实时测试, 得到了非常好的检测效果(如图9所示)。

[0071] 实施例2

[0072] 为了验证模型的泛化性, 本发明与现有的人脸口罩检测算法进行对比, 由表1可知, 对于人脸佩戴口罩目标检测来讲, 本发明所述算法取得了显著的检测效果, 与RetinaFace算法和Attention-RetinaFace相比, 其AP值分别提升了21.3%和13.1%, 而mAP值分别提升了12.3%和16.5%;

[0073] 与YoloV3和Improve YoloV3相比, 其AP值分别提升了17.3%和2.4%, 而mAP值分别提升了15.1%和4%。但对于人脸目标检测来讲, 本文算法也取得了较好的结果。

[0074] 表1实验对比数据

Method	Face/%	Mask/%	mAP/%
Deng J 等人的方法	87.3	76.5	81.9
牛等人的方法	90.6	84.7	87.7
YoloV3	77.6	80.5	79.1
王等人的方法	84.9	95.4	90.2
本发明方法	90.7	97.8	94.2

[0076] 至此, 已经结合附图所示的优选实施方式描述了本发明的技术方案, 但是, 本领域

技术人员容易理解的是,本发明的保护范围显然不局限于这些具体实施方式。在不偏离本发明的原理的前提下,本领域技术人员可以对相关技术特征做出等同的更改或替换,这些更改或替换之后的技术方案都将落入本发明的保护范围之内。

[0077] 以上所述仅为本发明的优选实施例,并不用于限制本发明;对于本领域的技术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

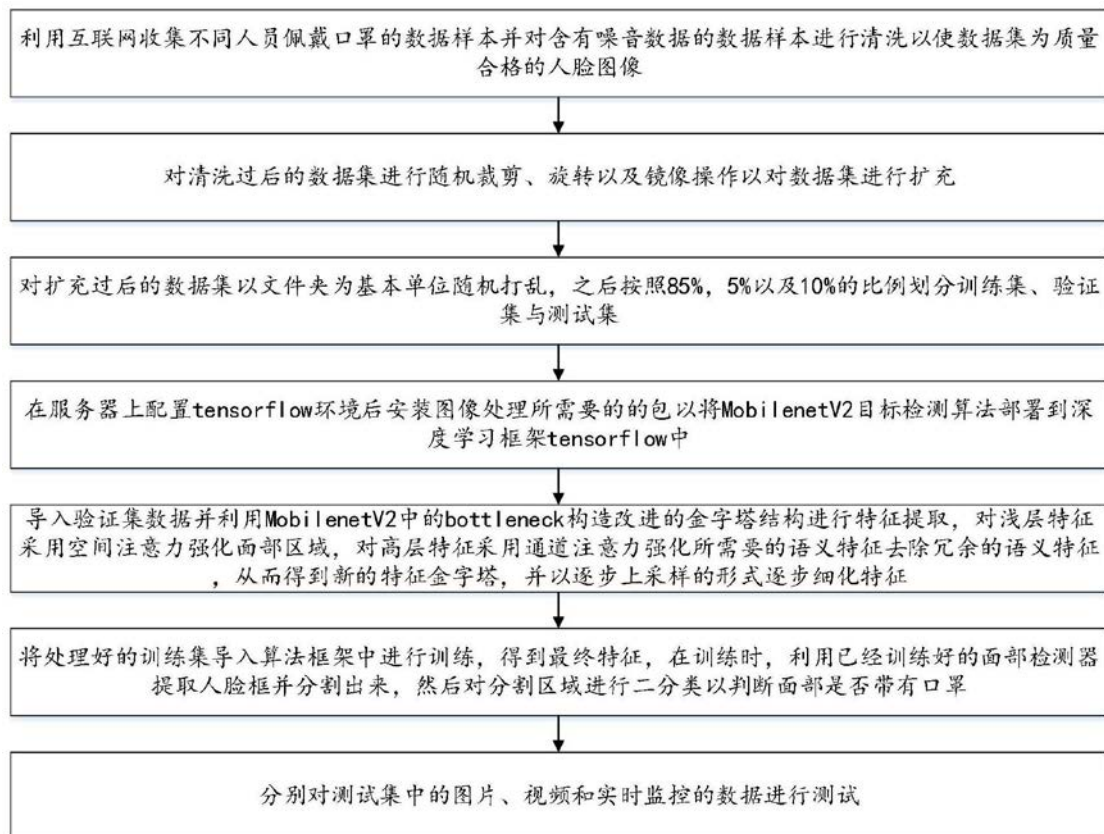


图1

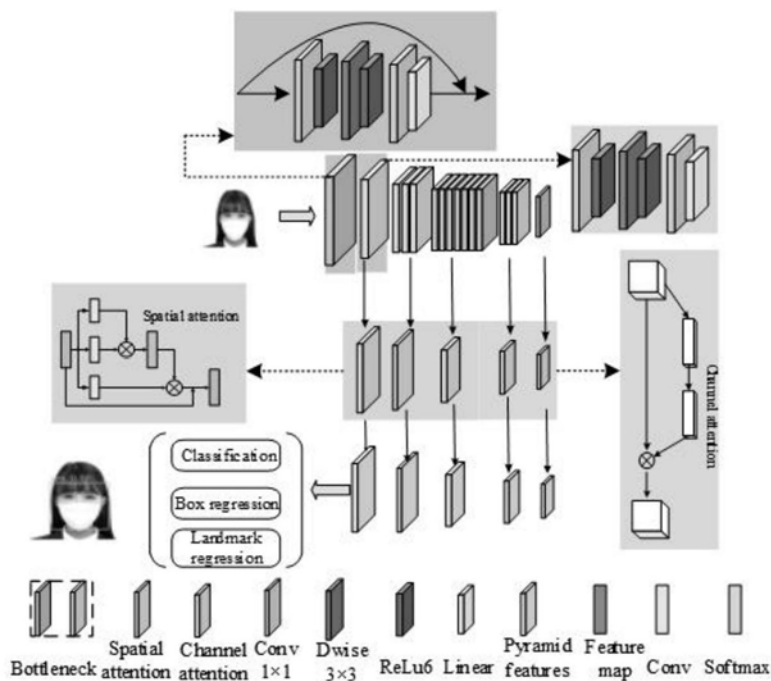


图2

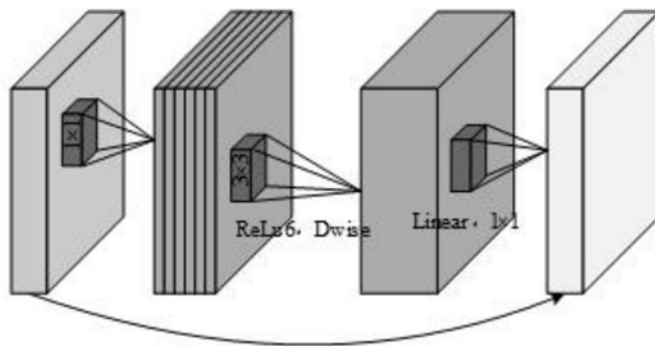


图3



图4



图5



图6



图7



图8



图9