# 一种基于YOLOv5的口罩佩戴检测方法

## 【摘要】

1、本发明公开了一种将注意力机制与YOLOv5网络模型相结合的口罩佩戴检测方法，所述方法首先对训练集图片使用图像增强算法进行预处理，然后将图片送入到引入了注意力机制的YOLOv5网络中进行迭代训练，从而有效的增强人脸和口罩等关键点信息的提取。本发明在注意力机制的加持下，模型的鲁棒性显著提高，在光线昏暗条件下对口罩佩戴的检测准确率能达到92%，同时该方法具有较强可扩展性，对推进口罩佩戴检测的自动化、智能化，实现疫情防控和公共卫生安全具有重要的现实意义。

2、本文主要针对在可见度不高、光线昏暗的场景下，利用图像增强算法对图片进行预处理，将通道注意力和空间注意力结合，充分挖掘人脸口罩等关键特征点，同时对YOLOv5网络的损失函数进行相应的改进，提高模型在昏暗条件下的鲁棒性。

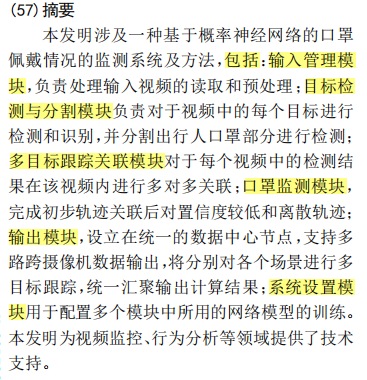
3、在可见度不高、光照强度不大的昏暗条件下，本文方法使用图像增强能改善图片的质量，然后通过注意力机制能更加准确的提取人脸口罩关键点特征，从而使得检测的准确率更高，具有较强的鲁棒性和扩展性，基本能够达到视频图像实时性的要求。

4、本文提出的基于注意力机制的光线昏暗条件下的口罩检测方法可应用在口罩佩戴识别系统中，具有较强的鲁棒性和可扩展性，对推进口罩佩戴检测的自动化、智能化，实现疫情防控和公共卫生安全具有重要的现实意义。

【简述模型效果】，如：本发明在××算法模型上结合口罩检测模型，以对人脸口罩进行检测，且提高了口罩的检测速度和精度。

说明算法、模型、方法的步骤

在图像增强和注意力机制的加持下，经过改进的YOLOv5模型实现了口罩佩戴的高效检测，对是否正确佩戴口罩作出了正确的判断



【本发明公开了一种+专利名+的方法，然后简述方法对数据的处理过程或者算法过程，简述算法特征及创新点，最后是：本发明提供的××方法具有的优点】

本发明的方法提高了小人脸的检测准确率，运行内存小、检测速度快、模型小，可以实时检测监控图像中众多小人脸是否佩戴口罩。

本文方法不仅成功地检测到了人脸信息，而且正确地检测出口罩佩戴的状态，并给出了对应的置信度。

【结论】实验结果表明，在可见度不高、光照强度不大的昏暗条件下，与参考文献[5]和AIZOO的方法相比，本文方法使用图像增强能改善图片的质量，然后通过注意力机制能更加准确的提取人脸口罩关键点特征，从而使得检测的准确率更高，具有较强的鲁棒性和扩展性，基本能够达到视频图像实时性的要求。

## **【权利要求书】**

1、一种基于YOLOv5的口罩佩戴检测方法，其特征包括：

在原始YOLOv5的网络基础上，引入卷积注意力模块CBAM(Convolutional Block Attention Module Network)。在卷积神经网络中，注意力机制作用在特征图之上，通过获取特征图中的可用注意力信息，能够达到更好的任务效果。

通过网络爬取与自行拍摄相结合制作数据集，数据集中80%来源源于网络，20%来源于实际拍摄。实际拍摄主要获取的是在昏暗条件下的口罩佩戴图片，实验过程中从楼道、室内等光线昏暗的场所进行了图片采集，同时也在傍晚和清晨等光线较弱的环境下进行了拍摄。

2、模型损失函数(Loss)由分类损失(Classification Loss)、定位损失(Localization Loss)和目标置信度损失(Confidence Loss)组成，如公式(2)所示。本文采用二元交叉熵损失函数来计算分类损失和目标置信度损失，分别如公式(3)和公式(4)所示。

3、根据权利要求1所述的基于YOLOv5的口罩佩戴检测方法，其特征在于，卷积注意力模块CBAM包含两个子模块，分别是通道注意力模块CAM(Channel Attention Module)和空间注意力模块SAM(Spatial Attention Module)。CAM汇总通道注意力信息，SAM汇总空间注意力信息。与通道注意力不同，SAM主要关注于目标在图像上的位置信息，它将CAM的输出特征图作为本模块的输出特征图。由于CBAM模型在通道注意力模块中加入了全局最大池化操作，它能在一定程度上弥补全局平均池化所丢失的信息。其次，生成的二维空间注意力图使用卷积核大小为7的卷积层进行编码，较大的卷积核对于保留重要大的卷积核对于保留重要的空间区域有良好的帮助。添加了CBAM的YOLOv5网络不仅能更为准确的对目标进行分类识别，而且能更为精准的定位目标所在的位置。

4、根据权利要求1所述的基于YOLOv5的口罩佩戴检测方法，其特征在于，引入注意力机制的YOLOv5模型总计367层，7150056个parameters。

5、根据权利要求1所述的基于YOLOv5的口罩佩戴检测方法，其特征在于，在网络模型训练阶段，迭代批量大小设置为32，总迭代次数为600次。初始学习率设置为0.001，采用小批量梯度下降法，并使用Adam优化器计算每个参数的自适应学习率。

6、根据权利要求2所述的模型损失函数，其特征在于，采用CIOU Loss作为目标框回归的损失函数，CIOU Loss如公式(5)所示。

【权利要求1中将方法的主要特征进行分类阐述，简明即可，后面会详细解释】，如：分为三部分：输入、特征融合、输出层卷积及检测。在后面的对特征提取、特征融合、检测单元的作用作简要阐述，作为从权。

【权力要求1分别简介数据集、网络结构、算法主要的创新点】【权利要求2及以后的内容对权利要求1进行深刻解释、细化】

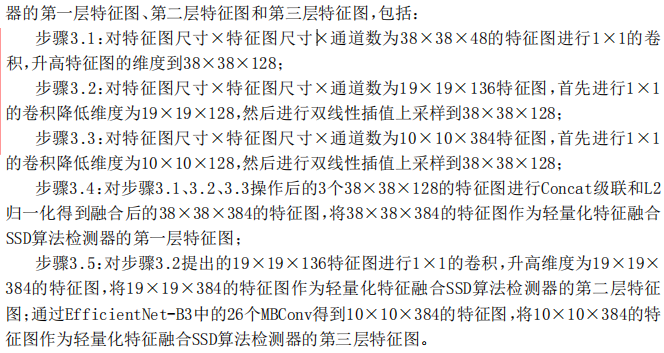
可以添加公式

【分层描述】，如：口罩识别模型包括三个卷积层、三个池化层和两个全连接层，三个卷积层和三个池化层交替连接。

【各单元使用的模型或者结构】，如：其特征在于，所述人脸识别单元采用Retinaface算法模型；Retinaface算法模型中特征提取网络采用MobileNet V1(0.25)结构，特征融合层采用FPN结构。

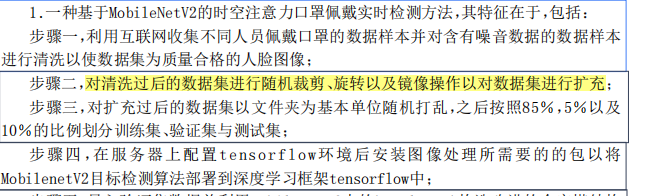
【算法包括那些步骤要按顺序说清楚】

【从权对特征图尺寸升维、降维进行步骤说明】，如：



训练步骤，重复佩戴口罩检测网络训练过程，不断修正佩戴口罩检测网络的参数，直至口罩检测网络学会找出图像中的人脸位置并能够正确的判断检测出来的人脸是否佩戴口罩，保存训练得到的参数。

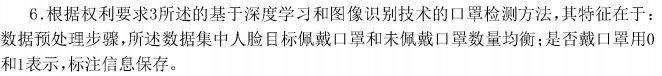
【添加数据采集、数据预处理内容，并将其作为从权，即从属于`独立权利`的权利】



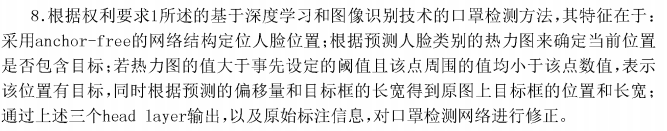
【迭代次数也可以作为从权】【添加数据标注描述，并将其作为从权，即从属于`独立权利`的权利】

网络模型训练阶段，迭代批量大小设置为32，总迭代次数为600次。初始学习率设置为0.001，采用小批量梯度下降法，并使用Adam优化器计算每个参数的自适应学习率。大约在350次迭代后，模型开始逐渐收敛。

当迭代次数接近400次左右时，平均精度均值的数值接近于0.996；从图 9b中可以看出，当迭代次数接近450次左右时，召回率的数值接近于1；从图 9c中可以看出，当迭代次数接近500次时，准确率的数值接近于0.995。

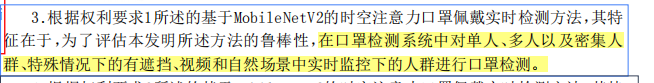


对目标定位、目标定位使用的技术细节进行描述，作为从权

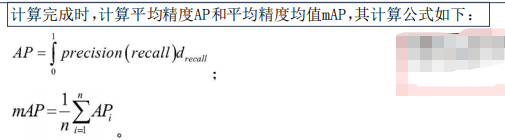


在从权中对“昏暗条件”进行详细表述

模型训练完成后，将得到的模型与参考文献[5]中的方法和AIZOO方法的口罩检测模型进行对比实验，分别在光照强度为30-75Lux(昏暗)，75-250Lux(较昏暗)和250-1000Lux(正常光照)的条件下进行对别实验，实验结果如图 10所示。其中，光照强度是指单位面积上所接受可见光的能量，常用于指示光照的强弱和物体表面积被照明程度的量，单位是Lux，光照强度越大，表明光照越强，物体表面被照的越亮。

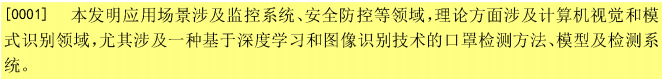


使用【4.2 模型评估指标】



## **【说明书】**

【技术领域】本发明涉及口罩佩戴检测技术领域，具体涉及一种基于YOLOv5的口罩佩戴检测方法。

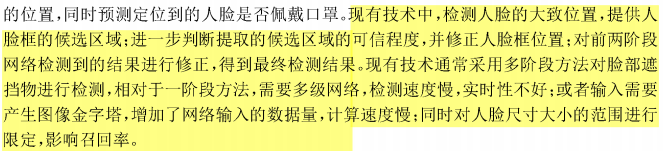




本发明应用场景涉及监控系统、安全防控等领域，理论方面涉及计算机视觉和模式识别领域，尤其涉及一种基于深度学习和图像识别技术的口罩检测方法、模型及检测系统。

【背景技术】新冠肺炎疫情暴发后，人类健康受到巨大威胁，为避免新冠肺炎疫情继续传播，人们在外出时规范佩戴口罩便成为了一项有效的防控手段。仅靠人工方式对流动人员进行口罩佩戴检测不仅效率较低，且会耗费大量人力物力，同时由于新冠肺炎传染性极强，近距离接触待检测人员更是存在较大的安全风险。因此，构建口罩自动检测系统检测活动人员在各类车站、大型商场等公共场合的口罩佩戴情况，对疫情防控具有重要的现实意义。

针对口罩佩戴检测这一特定问题，目前国内外有多位学者进行了研究。邓黄潇等人[4]提出了运用迁移学习和ReinaNet网络对口罩佩戴进行检测，验证集下AP达到86.5%。肖俊杰等人[5]将YOLOv3和YCrCb方法结合，正确佩戴口罩的识别率的达到82.5%。牛作东等人[6]对RetinaFace算法进行了相应的优化，同时引入了自注意力机制，实验结果表明检测效果较好。AIZOO也提出了一种基于Fast-RCNN和YOLOv3的目标检测方法，在口罩佩戴检测任务上取得了不错的成绩，但对可见度不高、光照强度不强的昏暗条件下，其检测精度仍有待提高

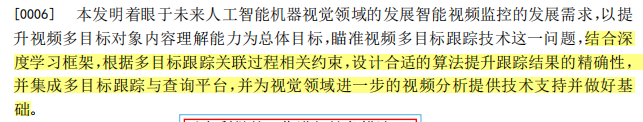


介绍YOLOv5：YOLO是一个高性能的通用目标检测模型，可完成目标位置定位和目标物体分类两个任务，后续版本在继续保持处理速度的基础上，从预测更准确、速度更快、识别对象更多这三个方面进行了改进，同时通过引入多尺度预测、改进基础网络和损失函数等进一步加速了目标检测在工业界的落地。YOLOv5相较于之前的版本，其结构更加小巧灵活，图像推理速度更快，能够满足视频图像实时性检测的需求。

YOLOv5随着其版本的迭代更新，网络内部的各个模块针对物体检测中的常见问题都做了优化处理。

在昏暗条件下，光照强度不大，可见度低，难以对人脸进行精确定位，口罩佩戴检测任务的难度也更为困难，因此需要对YOLOv5的网络做进一步的结构优化和调整。





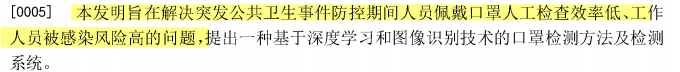
【发明内容】

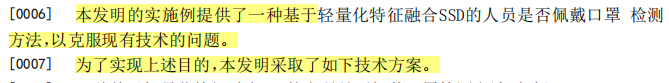
对权利要求书中的每项进行解释，说白了就是换一种说法再说一遍，使得句子通顺。

描述算法步骤，分点按顺序说明。

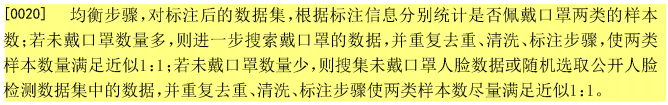
在昏暗条件下，光照强度不大，可见度低，难以对人脸进行精确定位，口罩佩戴检测任务的难度也更为困难，因此需要对YOLOv5的网络做进一步的结构优化和调整。

开头：





注意力机制最早在2014年率先被Google Deep Mind团队引入到RNN模型上来实现图像的分类[13]，实现了图像中多个物体对象的高效准确的识别。注意力机制作用在特征图之上，通过获取特征图中的可用注意力信息，能够达到更好的任务效果。



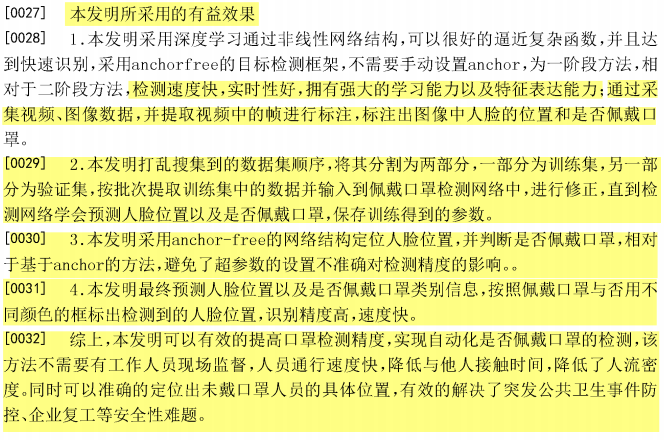
对数据采集过程、预处理、标注过程、获取坐标、数据均衡等均可加入说明书

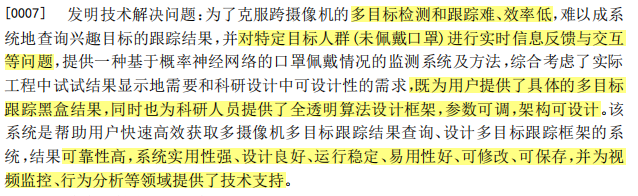
本文使用图像增强技术，对标签为good的样本图片进行图像平移、翻转、旋转、缩放，分离三个颜色通道并添加随机噪声来有效缓解类间不平衡的问题，可以看出增强后的数据集物体分布相对均匀，且中小物体占比更大，存在物体间遮挡的情况，符合日常实际应用场景。

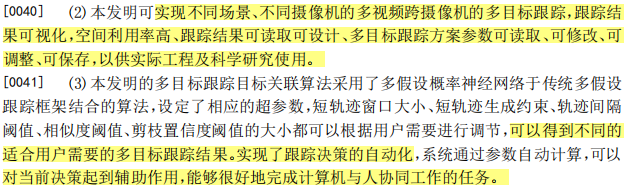
实验数据集共包含9000张图片，对其进行手工标注。该数据集分为两种类别，分别是bad和good，bad表示人员未佩戴或未按规范佩戴口罩，good表示正确佩戴口罩。数据集采用YOLO格式，图片标注使用LabelImg，标注后的文件以.txt作为后缀，文件名称和图片名称一致。训练集和测试集所占的比例为8：1

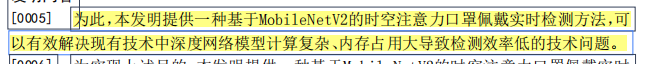
模型损失函数(Loss)由分类损失(Classification Loss)、定位损失(Localization Loss)和目标置信度损失(Confidence Loss)组成，如公式(2)所示。本文采用二元交叉熵损失函数来计算分类损失和目标置信度损失，分别如公式(3)和公式(4)所示。其中，K表示网络最后输出的的特征图划分为K×K个格子，M表示每个格子对应的锚框的个数，表示有目标的锚框，表示没有目标的锚框，表示对没有目标锚框的损失系数。

由于CIOU Loss从重叠面积、中心点距离和长宽比三个角度进行衡量，故预测框回归的效果更佳。

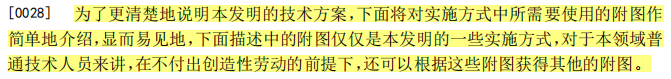






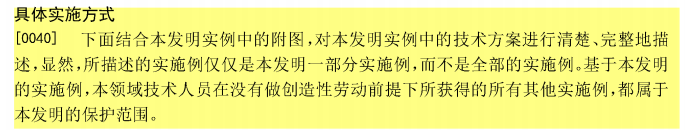


【附图说明】



【具体实施方式】

【开头可用】：



【说明书整个文档最后可以用】：

