

【发明名称】一种基于YOLOv5的口罩佩戴检测方法

0 前言

全文是根据6篇专利文档的总结的一些笔记、批注：

- 1、一种基于Retinaface算法的人脸口罩检测系统和方法-重庆大学
- 2、基于深度学习和图像识别技术的口罩检测方法及检测系统-深延科技（北京）有限公司
- 3、一种基于YOLOV3的人脸口罩检测方法、系统、设备及介质-广州大学华软软件学院
- 4、一种基于概率神经网络的口罩佩戴情况的监测系统及方法-北京航空航天大学
- 5、基于轻量化特征融合SSD的人员是否佩戴口罩检测方法-北京交通大学
- 6、基于MobileNetV2的时空注意力口罩佩戴实时检测方法-上海工程技术大学

1 摘要

【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。

【最好能说明算法、模型、方法的步骤】，如：

(57)摘要

本发明公开了一种基于深度学习和图像识别技术的口罩检测方法及检测系统,包括,数据采集步骤、数据采集模块采集视频、图像数据,并提取视频中的帧;数据预处理步骤、对得到的数据进行标注,标注出图像中人脸的位置和是否佩戴口罩;训练集划分、打乱搜集到的数据集顺序,将其分割为训练集、验证集;训练步骤、按批次提取训练集中的数据并输入到佩戴口罩检测网络中进行修正网络参数;输出步骤,直到检测网络学会预测人脸位置以及是否佩戴口罩;验证集输入到训练得到的佩戴口罩检测网络中得到最终预测人脸位置以及是否佩戴口罩类别信息,按照佩戴口罩与否用不同颜色的框标出检测到的人脸位置。

【方法分为几个模块，最后简要提一下专利的有益效果】，如：

(57) 摘要

本发明涉及一种基于概率神经网络的口罩佩戴情况的监测系统及方法，包括：输入管理模块，负责处理输入视频的读取和预处理；目标检测与分割模块负责对于视频中的每个目标进行检测和识别，并分割出行人口罩部分进行检测；多目标跟踪关联模块对于每个视频中的检测结果在该视频内进行多对多关联；口罩监测模块，完成初步轨迹关联后对置信度较低和离散轨迹；输出模块，设立在统一的数据中心节点，支持多路跨摄像机数据输出，将分别对各个场景进行多目标跟踪，统一汇聚输出计算结果；系统设置模块用于配置多个模块中所用的网络模型的训练。本发明为视频监控、行为分析等领域提供了技术支持。

【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】，如：

本发明公开了一种基于YOLOV3的人脸口罩检测方法、系统、设备及介质，所述方法为将待测数据输入至目标YOLOV3算法模型，通过DarkNet53特征提取网络进行特征提取，得到不同格式的特征数据；所述待测数据为格式统一的人脸图像，所述目标YOLOV3算法模型为结合IOU Loss函数的YOLOV3算法模型；将所述特征数据输入至特征融合层，通过卷积和长采样进行特征融合；通过输出层对所述特征融合后的数据进行卷积操作，并对所述人脸图像的边界框进行检测，得到人脸口罩检测结果。本发明提供的基于YOLOV3的人脸口罩检测方法，具有检测速度快，结果准确度高的优点。

【摘要最后可用】：

果。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。

2 权利要求书

【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】

【权利要求书中可以添加公式】，如下是某篇专利的内容：

所述复合维度缩放方法的具体公式如下：

深度： $d = a^p$

通道数： $w = b^p$

分辨率： $r = c^p$

约束条件：

【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：其特征不在于，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，自己必须清楚】，如：

1. 一种基于深度学习和图像识别技术的口罩检测方法，包括获取人脸框的中心坐标步骤、训练集划分、训练集修正、训练步骤，其特征在于：

【从权对特征图尺寸升维、降维进行步骤说明】，如：

器的第一层特征图、第二层特征图和第三层特征图，包括：

步骤3.1：对特征图尺寸×特征图尺寸×通道数为 $38 \times 38 \times 48$ 的特征图进行 1×1 的卷积，升高特征图的维度到 $38 \times 38 \times 128$ ；

步骤3.2：对特征图尺寸×特征图尺寸×通道数为 $19 \times 19 \times 136$ 特征图，首先进行 1×1 的卷积降低维度为 $19 \times 19 \times 128$ ，然后进行双线性插值上采样到 $38 \times 38 \times 128$ ；

步骤3.3：对特征图尺寸×特征图尺寸×通道数为 $10 \times 10 \times 384$ 特征图，首先进行 1×1 的卷积降低维度为 $10 \times 10 \times 128$ ，然后进行双线性插值上采样到 $38 \times 38 \times 128$ ；

步骤3.4：对步骤3.1、3.2、3.3操作后的3个 $38 \times 38 \times 128$ 的特征图进行Concat级联和L2归一化得到融合后的 $38 \times 38 \times 384$ 的特征图，将 $38 \times 38 \times 384$ 的特征图作为轻量化特征融合SSD算法检测器的第一层特征图；

步骤3.5：对步骤3.2提出的 $19 \times 19 \times 136$ 特征图进行 1×1 的卷积，升高维度为 $19 \times 19 \times 384$ 的特征图，将 $19 \times 19 \times 384$ 的特征图作为轻量化特征融合SSD算法检测器的第二层特征图；通过EfficientNet-B3中的26个MBConv得到 $10 \times 10 \times 384$ 的特征图，将 $10 \times 10 \times 384$ 的特征图作为轻量化特征融合SSD算法检测器的第三层特征图。

【训练步骤，我们的专利可以改一改直接使用】，如下：

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。

1. 根据权利要求1所述的基于深度学习和图像识别技术的口罩检测方法，其特征在于：

【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】，如：

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；

【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】

【是否戴口罩的标签：bad、good也可放在权利要求书中】，如下是某篇专利的内容：

6.根据权利要求3所述的基于深度学习和图像识别技术的口罩检测方法，其特征在于：数据预处理步骤，所述数据集中人脸目标佩戴口罩和未佩戴口罩数量均衡；是否戴口罩用0和1表示，标注信息保存。

【对目标定位、目标定位使用的技术细节进行描述，作为从权】，如下是某篇专利的内容：

8.根据权利要求1所述的基于深度学习和图像识别技术的口罩检测方法，其特征在于：采用anchor-free的网络结构定位人脸位置；根据预测人脸类别的热力图来确定当前位置是否包含目标；若热力图的值大于事先设定的阈值且该点周围的值均小于该点数值，表示该位置有目标，同时根据预测的偏移量和目标框的长宽得到原图上目标框的位置和长宽；通过上述三个head layer输出，以及原始标注信息，对口罩检测网络进行修正。

【在从权中对“昏暗条件”进行详细表述】，如：

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。

【精确率与召回率的公式也可以写在权利要求书中】，如下是某篇专利的内容：

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：

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

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



【总之要把算法的创新点描述清楚】

3 说明书

3.1 技术领域

【可以借鉴】：

[0001] 本发明应用场景涉及监控系统、安全防护等领域，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。

【可以使用】：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。

【可以借鉴】：

[0001] 本发明涉及一种基于概率神经网络的口罩佩戴情况的监测系统及方法，属于计算机视觉视频监控侦察信息化建设方向。

3.2 背景技术

【研究的背景与意义，介绍YOLOv5，对其他检测方法进行简要描述，主要阐述其不足】

【现有算法的不足】，如：目前监控系统的人脸识别算法主要用来检测人脸并将其框出来，例如经典的Retinaface算法，但无法分辨该人脸是否戴了口罩，难以起到安全监控的作用。

【对当前技术的主要问题进行简要介绍】，如：

的位置，同时预测定位到的人脸是否佩戴口罩。现有技术中，检测人脸的大致位置，提供人脸框的候选区域；进一步判断提取的候选区域的可信程度，并修正人脸框位置；对前两阶段网络检测到的结果进行修正，得到最终检测结果。现有技术通常采用多阶段方法对脸部遮挡物进行检测，相对于一阶段方法，需要多级网络，检测速度慢，实时性不好；或者输入需要产生图像金字塔，增加了网络输入的数据量，计算速度慢；同时对人脸尺寸大小的范围进行限定，影响召回率。

【口罩检测的难点】：

[0005] 在科学研究方面由于目标的多样性和多变性,场景的多样性和多变性以及目标间交互的复杂性使得多目标跟踪的研究成为难点,大多数跟踪方法在跨摄像机的多目标检测

【对专利做的工作进行抽象描述】：

[0006] 本发明着眼于未来人工智能机器视觉领域的发展智能视频监控的发展需求,以提升视频多目标对象内容理解能力为总体目标,瞄准视频多目标跟踪技术这一问题,结合深度学习框架,根据多目标跟踪关联过程相关约束,设计合适的算法提升跟踪结果的精确性,并集成多目标跟踪与查询平台,并为视觉领域进一步的视频分析提供技术支持并做好基础。

3.3 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。

【在前面先用一句话介绍发明目的】,如:

[0005] 本发明旨在解决突发公共卫生事件防控期间人员佩戴口罩人工检查效率低、工作人员被感染风险高的问题,提出一种基于深度学习和图像识别技术的口罩检测方法及检测系统。

【开头可以借鉴】：

[0006] 本发明的实施例提供了一种基于轻量化特征融合SSD的人员是否佩戴口罩检测方法,以克服现有技术的问题。

[0007] 为了实现上述目的,本发明采取了如下技术方案。

【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】,我们也有数据均衡过程,如下是某篇专利的内容:

[0020] 均衡步骤,对标注后的数据集,根据标注信息分别统计是否佩戴口罩两类的样本数;若未戴口罩数量多,则进一步搜索戴口罩的数据,并重复去重、清洗、标注步骤,使两类样本数量满足近似1:1;若未戴口罩数量少,则搜集未戴口罩人脸数据或随机选取公开人脸检测数据集中的数据,并重复去重、清洗、标注步骤使两类样本数尽量满足近似1:1。

【有益效果】， 如下是某篇专利的内容：

[0027] 本发明所采用的有益效果

[0028] 1. 本发明采用深度学习通过非线性网络结构,可以很好的逼近复杂函数,并且达到快速识别,采用anchorfree的目标检测框架,不需要手动设置anchor,为一阶段方法,相对于二阶段方法,检测速度快,实时性好,拥有强大的学习能力以及特征表达能力;通过采集视频、图像数据,并提取视频中的帧进行标注,标注出图像中人脸的位置和是否佩戴口罩。

[0029] 2. 本发明打乱搜集到的数据集顺序,将其分割为两部分,一部分为训练集,另一部分为验证集,按批次提取训练集中的数据并输入到佩戴口罩检测网络中,进行修正,直到检测网络学会预测人脸位置以及是否佩戴口罩,保存训练得到的参数。

[0030] 3. 本发明采用anchor-free的网络结构定位人脸位置,并判断是否佩戴口罩,相对于基于anchor的方法,避免了超参数的设置不准确对检测精度的影响。。

[0031] 4. 本发明最终预测人脸位置以及是否佩戴口罩类别信息,按照佩戴口罩与否用不同颜色的框标出检测到的人脸位置,识别精度高,速度快。

[0032] 综上,本发明可以有效的提高口罩检测精度,实现自动化是否佩戴口罩的检测,该方法不需要有工作人员现场监督,人员通行速度快,降低与他人接触时间,降低了人流密度。同时可以准确的定位出未戴口罩人员的具体位置,有效的解决了突发公共卫生事件防控、企业复工等安全性难题。

【发明技术所解决的问题及有益效果】， 如：

[0007] 发明技术解决问题:为了克服跨摄像机的多目标检测和跟踪难、效率低,难以成系统地查询兴趣目标的跟踪结果,并对特定目标人群(未佩戴口罩)进行实时信息反馈与交互等问题,提供一种基于概率神经网络的口罩佩戴情况的监测系统及方法,综合考虑了实际工程中测试结果显示地需要和科研设计中可设计性的需求,既为用户提供了具体的多目标跟踪黑盒结果,同时也为科研人员提供了全透明算法设计框架,参数可调,架构可设计。该系统是帮助用户快速高效获取多摄像机多目标跟踪结果查询、设计多目标跟踪框架的系统,结果可靠性高,系统实用性强、设计良好、运行稳定、易用性好、可修改、可保存,并为视频监控、行为分析等领域提供了技术支持。

【描述了本发明与现有技术相比的优点】， 如：

[0040] (2) 本发明可实现不同场景、不同摄像机的多视频跨摄像机的多目标跟踪,跟踪结果可视化,空间利用率高、跟踪结果可读取可设计、多目标跟踪方案参数可读取、可修改、可调整、可保存,以供实际工程及科学研究使用。

[0041] (3) 本发明的多目标跟踪目标关联算法采用了多假设概率神经网络于传统多假设跟踪框架结合的算法,设定了相应的超参数,短轨迹窗口大小、短轨迹生成约束、轨迹间隔阈值、相似度阈值、剪枝置信度阈值的大小都可以根据用户需要进行调节,可以得到不同的适合用户需要的多目标跟踪结果。实现了跟踪决策的自动化,系统通过参数自动计算,可以对当前决策起到辅助作用,能够很好地完成计算机与人协同工作的任务。

【可以借鉴】：

[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] 本发明还提供一种基于MobileNetV2的时空注意力口罩佩戴实时检测系统,包括处理器、存储器、通信接口及通信总线,所述存储器用于存储计算机程序,所述通信接口用于接收外部数据并传输数据,所述处理器用于执行所述计算机程序,以实现如权利要求1-4所述的方法。

【最后说明算法性能的提高】

3.4 附图说明

【可以借鉴句式】：

【0028】 为了更清楚地说明本发明的技术方案，下面将对实施方式中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施方式，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

3.5 具体实施方式

【开头可用】：

具体实施方式

【0040】 下面结合本发明实例中的附图，对本发明实例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明的实施例，本领域技术人员在没有做创造性劳动前提下所获得的所有其他实施例，都属于本发明的保护范围。

【说明书整个文档最后可以用】：

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。

END