



(12)发明专利申请

(10)申请公布号 CN 110490171 A

(43)申请公布日 2019. 11. 22

(21)申请号 201910795129.9

(22)申请日 2019.08.26

(71)申请人 睿云联(厦门)网络通讯技术有限公司

地址 361000 福建省厦门市火炬高新区创业园伟业楼北楼104K室

(72)发明人 黄懋哲 林昱

(74)专利代理机构 厦门创象知识产权代理有限公司 35232

代理人 吴慧敏

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/34(2006.01)

G06K 9/62(2006.01)

G06N 3/04(2006.01)

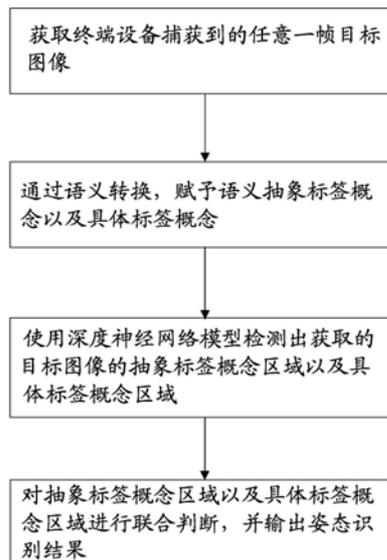
权利要求书3页 说明书10页 附图2页

(54)发明名称

一种危险姿态识别方法、装置、计算机设备及存储介质

(57)摘要

本发明提供一种危险姿态识别方法,包括获取终端设备捕获到的任意一帧目标图像;通过语义转换,赋予语义抽象标签概念以及具体标签概念;使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。本发明还提供一种危险姿态识别系统、计算机设备以及存储介质。本发明优点在于:可大大增加检测的准确度,减少错误判别的概率;同时还可以免除多模型预测带来的额外预测耗时。



1. 一种危险姿态识别方法,其特征在于:所述方法包括:
获取终端设备捕获到的任意一帧目标图像;
通过语义转换,赋予语义抽象标签概念以及具体标签概念;
使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;
对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。
2. 根据权利要求1所述的一种危险姿态识别方法,其特征在于:所述方法还包括:
获取输出的连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果。
3. 根据权利要求1所述的一种危险姿态识别方法,其特征在于:所述赋予语义抽象标签概念以及具体标签概念具体为:
对识别目标体赋予正常标签和异常标签两个抽象标签概念,对识别目标体赋予头部标签、上半身标签以及下半身标签三个具体标签概念;
所述使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域具体为:
使用MASK R-CNN深度神经网络模型对获取的目标图像进行图像语义分割识别,且将从目标图像中分割出的识别目标体识别为正常区域或者异常区域;同时,通过MASK R-CNN深度神经网络模型检测出识别目标体的头部、上半身以及下半身三个区域的最大外接矩形区域。
4. 根据权利要求1所述的一种危险姿态识别方法,其特征在于:所述对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果具体为:
如果将识别目标体识别为正常区域,则直接将该识别目标体忽略掉,并输出姿态正常的识别结果;如果将识别目标体识别为异常区域,则使用贪心算法将属于该识别目标体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域三个区域分离出来;
对识别为异常区域的目标识别体进行误判判断,且如果判断结果为误判,则输出姿态误判的识别结果;
如果判断结果为非误判,则取下半身的最大外接矩形区域的底边中点作为第一坐标点P0,取下半身的最大外接矩形区域的左上角作为第二坐标点P1,取下半身的最大外接矩形区域的右侧边中点作为第三坐标点P2;同理,取上半身的最大外接矩形区域的左上角作为第四坐标点P3,取上半身的最大外接矩形区域的右侧边中点作为第五坐标点P4;取头部的最大外接矩形区域的左上角作为第六坐标点P5,取头部的最大外接矩形区域的右侧边中点作为第七坐标点P6;并以第一坐标点P0为角的顶点,分别计算出P1P0P2、P3P0P4、P5P0P6三个夹角的余弦值;
根据计算出的三个夹角的余弦值,使用支持向量机SVM对目标识别体的姿态进行分类,并输出对应的姿态分类识别结果。
5. 根据权利要求4所述的一种危险姿态识别方法,其特征在于:所述对识别为异常区域的目标识别体进行误判判断具体为:
计算在目标图像中检测出的目标识别体的面积与目标图像的整体面积的比例,且如果

比例超过设定的第一阈值或者小于预设的第二阈值,则将该目标识别体判定为误判;

同时,根据固定的摄像头位置推断出目标图像的可视范围,依据正常的目标识别体的头部区域、上半身区域以及下半身区域之间的比例,对检测出的目标识别体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域进行综合判断,且如果头部的最大外接矩形区域超过所述可视范围的某一预设高度,则计算出头部的最大外接矩形区域与上半身的最大外接矩形区域的第一面积比值以及头部的最大外接矩形区域与下半身的最大外接矩形区域的第二面积比值,且如果第一面积比值和第二面积比值均符合正常的目标识别体的躯干比例,则将该目标识别体判定为误判;

如果未检测出目标识别体的下半身的最大外接矩形区域或者上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度达到预设的第三面积比值,则将该目标识别体判定为误判。

6. 一种危险姿态识别系统,其特征在于:所述系统包括图像获取模块、语义转换模块、区域检测模块以及联合判断模块;

所述图像获取模块,用于获取终端设备捕获到的任意一帧目标图像;

所述语义转换模块,用于通过语义转换,赋予语义抽象标签概念以及具体标签概念;

所述区域检测模块,用于使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;

所述联合判断模块,用于对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。

7. 根据权利要求6所述的一种危险姿态识别系统,其特征在于:所述系统还包括最终结果生成模块;

所述最终结果生成模块,用于获取输出的连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果。

8. 根据权利要求6所述的一种危险姿态识别系统,其特征在于:所述赋予语义抽象标签概念以及具体标签概念具体为:

对识别目标体赋予正常标签和异常标签两个抽象标签概念,对识别目标体赋予头部标签、上半身标签以及下半身标签三个具体标签概念;

所述使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域具体为:

使用MASK R-CNN深度神经网络模型对获取的目标图像进行图像语义分割识别,且将从目标图像中分割出的识别目标体识别为正常区域或者异常区域;同时,通过MASK R-CNN深度神经网络模型检测出识别目标体的头部、上半身以及下半身三个区域的最大外接矩形区域。

9. 根据权利要求6所述的一种危险姿态识别系统,其特征在于:所述对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果具体为:

如果将识别目标体识别为正常区域,则直接将该识别目标体忽略掉,并输出姿态正常的识别结果;如果将识别目标体识别为异常区域,则使用贪心算法将属于该识别目标体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域三个区域分离出来;

对识别为异常区域的目标识别体进行误判判断,且如果判断结果为误判,则输出姿态误判的识别结果;

如果判断结果为非误判,则取下半身的最大外接矩形区域的底边中点作为第一坐标点P0,取下半身的最大外接矩形区域的左上角作为第二坐标点P1,取下半身的最大外接矩形区域的右侧边中点作为第三坐标点P2;同理,取上半身的最大外接矩形区域的左上角作为第四坐标点P3,取上半身的最大外接矩形区域的右侧边中点作为第五坐标点P4;取头部的最大外接矩形区域的左上角作为第六坐标点P5,取头部的最大外接矩形区域的右侧边中点作为第七坐标点P6;并以第一坐标点P0为角的顶点,分别计算出P1P0P2、P3P0P4、P5P0P6三个夹角的余弦值;

根据计算出的三个夹角的余弦值,使用支持向量机SVM对目标识别体的姿态进行分类,并输出对应的姿态分类识别结果。

10. 根据权利要求9所述的一种危险姿态识别系统,其特征在于:所述对识别为异常区域的目标识别体进行误判判断具体为:

计算在目标图像中检测出的目标识别体的面积与目标图像的整体面积的比例,且如果比例超过设定的第一阈值或者小于预设的第二阈值,则将该目标识别体判定为误判;

同时,根据固定的摄像头位置推断出目标图像的可视范围,依据正常的目标识别体的头部区域、上半身区域以及下半身区域之间的比例,对检测出的目标识别体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域进行综合判断,且如果头部的最大外接矩形区域超过所述可视范围的某一预设高度,则计算出头部的最大外接矩形区域与上半身的最大外接矩形区域的第一面积比值以及头部的最大外接矩形区域与下半身的最大外接矩形区域的第二面积比值,且如果第一面积比值和第二面积比值均符合正常的目标识别体的躯干比例,则将该目标识别体判定为误判;

如果未检测出目标识别体的下半身的最大外接矩形区域或者上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度达到预设的第三面积比值,则将该目标识别体判定为误判。

11. 一种计算机设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现如权利要求1-5任一所述的危险姿态识别方法。

12. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于:该计算机程序被处理器执行时实现如权利要求1-5任一所述的危险姿态识别方法。

一种危险姿态识别方法、装置、计算机设备及存储介质

技术领域

[0001] 本发明涉及图像识别与处理技术领域,特别涉及一种危险姿态识别方法、装置、计算机设备及存储介质。

背景技术

[0002] 日常生活中,在某些场景下,一个人摔倒之后很难会被及时发现送医,这往往会造成一些不可逆转的伤害。这时,用户都会希望用一些技术来检测这些场景下的行为是否存在异常情况。

[0003] 目前,主流的危险姿态检测方案包括:从部署方式上看有传感器方案和摄像头监控方案两种;从数据处理方法上看有传统方法和基于人工智能与机器学习方法两种。

[0004] 其中,部署方式上,基于传感器的方案,由于成本高昂(通常会呈线性增加),适用场景狭隘(需要固定在某个固定场景下),灵活度差等原因,已逐渐被摄像头监控方案代替。而对数据的传统处理方法,由于识别率不高,逻辑复杂,误判率高等原因也已经逐渐被市场淘汰。

[0005] 基于人工智能与机器学习的方法又分为传统机器学习方法和神经网络学习方法两种。但是,传统机器学习方法由于无法直接处理高纬度的数据,需要对采集到的数据进行预处理,造成通用性不高,迁移性差,不够精准。故市场上采用神经网络深度学习的方式已成为主流。

[0006] 而根据深度学习不同模式的判定方式,目前大致又可分为如下两种:一是根据多帧之间的运动轨迹来进行判断;二是根据图像中人的骨骼姿势来进行判断。但是,根据多帧之间的运动轨迹来进行判断的实时性较差,消耗资源也较多。由于人体的灵活性和自由度,根据人体骨骼姿势来进行判断的准确度不高,如OPENPOSE骨架检测网络,由于其是针对关键点检测,因此丢失了大部分的像素细节信息,这导致对侧面坐姿、蹲下等姿势的分类效果不佳。

[0007] 基于以上分析可知,目前还未有一种使用场景宽泛且准确度高的危险姿态判别方案。

发明内容

[0008] 本发明要解决的技术问题,在于提供一种危险姿态识别方法、装置、计算机设备及存储介质,解决现有危险姿态检测方案存在的准确度不高的问题。

[0009] 本发明是这样实现的:一种危险姿态识别方法,所述方法包括:

[0010] 获取终端设备捕获到的任意一帧目标图像;

[0011] 通过语义转换,赋予语义抽象标签概念以及具体标签概念;

[0012] 使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;

[0013] 对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结

果。

[0014] 进一步地,所述方法还包括:

[0015] 获取输出的连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果。

[0016] 进一步地,所述赋予语义抽象标签概念以及具体标签概念具体为:

[0017] 对识别目标体赋予正常标签和异常标签两个抽象标签概念,对识别目标体赋予头部标签、上半身标签以及下半身标签三个具体标签概念;

[0018] 所述使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域具体为:

[0019] 使用MASK R-CNN深度神经网络模型对获取的目标图像进行图像语义分割识别,且将从目标图像中分割出的识别目标体识别为正常区域或者异常区域;同时,通过MASK R-CNN深度神经网络模型检测出识别目标体的头部、上半身以及下半身三个区域的最大外接矩形区域。

[0020] 进一步地,所述对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果具体为:

[0021] 如果将识别目标体识别为正常区域,则直接将该识别目标体忽略掉,并输出姿态正常的识别结果;如果将识别目标体识别为异常区域,则使用贪心算法将属于该识别目标体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域三个区域分离出来;

[0022] 对识别为异常区域的目标识别体进行误判判断,且如果判断结果为误判,则输出姿态误判的识别结果;

[0023] 如果判断结果为非误判,则取下半身的最大外接矩形区域的底边中点作为第一坐标点P0,取下半身的最大外接矩形区域的左上角作为第二坐标点P1,取下半身的最大外接矩形区域的右侧边中点作为第三坐标点P2;同理,取上半身的最大外接矩形区域的左上角作为第四坐标点P3,取上半身的最大外接矩形区域的右侧边中点作为第五坐标点P4;取头部的最大外接矩形区域的左上角作为第六坐标点P5,取头部的最大外接矩形区域的右侧边中点作为第七坐标点P6;并以第一坐标点P0为角的顶点,分别计算出P1P0P2、P3P0P4、P5P0P6三个夹角的余弦值;

[0024] 根据计算出的三个夹角的余弦值,使用支持向量机SVM对目标识别体的姿态进行分类,并输出对应的姿态分类识别结果。

[0025] 进一步地,所述对识别为异常区域的目标识别体进行误判判断具体为:

[0026] 计算在目标图像中检测出的目标识别体的面积与目标图像的整体面积的比例,且如果比例超过设定的第一阈值或者小于预设的第二阈值,则将该目标识别体判定为误判;

[0027] 同时,根据固定的摄像头位置推断出目标图像的可视范围,依据正常的目标识别体的头部区域、上半身区域以及下半身区域之间的比例,对检测出的目标识别体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域进行综合判断,且如果头部的最大外接矩形区域超过所述可视范围的某一预设高度,则计算出头部的最大外接矩形区域与上半身的最大外接矩形区域的第一面积比值以及头部的最大外接矩形区域与下半身的最大外接矩形区域的第二面积比值,且如果第一面积比值和第二面积

比值均符合正常的目标识别体的躯干比例,则将该目标识别体判定为误判;

[0028] 如果未检测出目标识别体的下半身的最大外接矩形区域或者上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度达到预设的第三面积比值,则将该目标识别体判定为误判。

[0029] 本发明是这样实现的:一种危险姿态识别系统,所述系统包括图像获取模块、语义转换模块、区域检测模块以及联合判断模块;

[0030] 所述图像获取模块,用于获取终端设备捕获到的任意一帧目标图像;

[0031] 所述语义转换模块,用于通过语义转换,赋予语义抽象标签概念以及具体标签概念;

[0032] 所述区域检测模块,用于使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;

[0033] 所述联合判断模块,用于对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。

[0034] 进一步地,所述系统还包括最终结果生成模块;

[0035] 所述最终结果生成模块,用于获取输出的连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果。

[0036] 进一步地,所述赋予语义抽象标签概念以及具体标签概念具体为:

[0037] 对识别目标体赋予正常标签和异常标签两个抽象标签概念,对识别目标体赋予头部标签、上半身标签以及下半身标签三个具体标签概念;

[0038] 所述使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域具体为:

[0039] 使用MASK R-CNN深度神经网络模型对获取的目标图像进行图像语义分割识别,且将从目标图像中分割出的识别目标体识别为正常区域或者异常区域;同时,通过MASK R-CNN深度神经网络模型检测出识别目标体的头部、上半身以及下半身三个区域的最大外接矩形区域。

[0040] 进一步地,所述对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果具体为:

[0041] 如果将识别目标体识别为正常区域,则直接将该识别目标体忽略掉,并输出姿态正常的识别结果;如果将识别目标体识别为异常区域,则使用贪心算法将属于该识别目标体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域三个区域分离出来;

[0042] 对识别为异常区域的目标识别体进行误判判断,且如果判断结果为误判,则输出姿态误判的识别结果;

[0043] 如果判断结果为非误判,则取下半身的最大外接矩形区域的底边中点作为第一坐标点P0,取下半身的最大外接矩形区域的左上角作为第二坐标点P1,取下半身的最大外接矩形区域的右侧边中点作为第三坐标点P2;同理,取上半身的最大外接矩形区域的左上角作为第四坐标点P3,取上半身的最大外接矩形区域的右侧边中点作为第五坐标点P4;取头部的最大外接矩形区域的左上角作为第六坐标点P5,取头部的最大外接矩形区域的右侧边中点作为第七坐标点P6;并以第一坐标点P0为角的顶点,分别计算出P1P0P2、P3P0P4、

P5P0P6三个夹角的余弦值；

[0044] 根据计算出的三个夹角的余弦值，使用支持向量机SVM对目标识别体的姿态进行分类，并输出对应的姿态分类识别结果。

[0045] 进一步地，所述对识别为异常区域的目标识别体进行误判判断具体为：

[0046] 计算在目标图像中检测出的目标识别体的面积与目标图像的整体面积的比例，且如果比例超过设定的第一阈值或者小于预设的第二阈值，则将该目标识别体判定为误判；

[0047] 同时，根据固定的摄像头位置推断出目标图像的可视范围，依据正常的目标识别体的头部区域、上半身区域以及下半身区域之间的比例，对检测出的目标识别体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域进行综合判断，且如果头部的最大外接矩形区域超过所述可视范围的某一预设高度，则计算出头部的最大外接矩形区域与上半身的最大外接矩形区域的第一面积比值以及头部的最大外接矩形区域与下半身的最大外接矩形区域的第二面积比值，且如果第一面积比值和第二面积比值均符合正常的目标识别体的躯干比例，则将该目标识别体判定为误判；

[0048] 如果未检测出目标识别体的下半身的最大外接矩形区域或者上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度达到预设的第三面积比值，则将该目标识别体判定为误判。

[0049] 本发明是这样实现的：一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现所述的危险姿态识别方法。

[0050] 本发明是这样实现的：一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现所述的危险姿态识别方法。

[0051] 本发明具有如下优点：1、采用语义分割模型来检测具体标签概念区域和抽象标签概念区域，并对抽象标签概念区域和具体标签概念区域进行联合判断，可大大增加检测的准确度，减少错误判别的概率；同时，由于具体标签概念和抽象标签概念的检测结果是并行得出的，因此，还可以免除多模型预测带来的额外预测耗时；

[0052] 2、通过对可能存在的误判情况进行判断，可避免误判情况带来的干扰，因此，可进一步提升对危险姿态识别的准确度；

[0053] 3、通过获取连续两帧目标图像的姿态识别结果进行比对分析，并根据比对分析结果生成最终的姿态识别结果，可进一步提升识别结果的置信度；

[0054] 4、使用人工智能深度学习的方式，部署成本低，使用场景宽泛，且不会随着使用年限的增加而老化，维护成本低。

附图说明

[0055] 下面参照附图结合实施例对本发明作进一步的说明。

[0056] 图1为本发明一种危险姿态识别方法的执行流程图。

[0057] 图2为本发明一种危险姿态识别系统的原理框图。

[0058] 图3为本发明在具体实现时各部件的组成示意图。

具体实施方式

[0059] 请参阅图1所示,本发明一种危险姿态识别方法的较佳实施例,下面,先对本发明中涉及到的一些名词做解释说明:

[0060] 贪心算法(又称贪婪算法):是指在对问题求解时,总是做出在当前看来是最好的选择。也就是说,不从整体最优上加以考虑,他所做出的仅是在某种意义上的局部最优解。贪心算法不是对所有问题都能得到整体最优解,但对范围相当广泛的许多问题他能产生整体最优解或者是整体最优解的近似解。

[0061] 语义分割:语义分割是一种典型的计算机视觉问题,其涉及将一些原始数据(例如,平面图像)作为输入并将它们转换为具有突出显示的感兴趣区域的掩模。许多人使用术语全像素语义分割(full-pixel semantic segmentation),其中图像中的每个像素根据其所属的感兴趣对象被分配类别ID。早期的计算机视觉问题只发现边缘(线条和曲线)或渐变等元素,但它们从未完全按照人类感知的方式提供像素级别的图像理解。语义分割将属于同一目标的图像部分聚集在一起来解决这个问题从而扩展了其应用领域。

[0062] MASK R-CNN深度神经网络模型:Mask R-CNN是实例分割的经典模型,Mask R-CNN是ICCV 2017的best paper,彰显了机器学习计算机视觉领域在2017年的最新成果。Mask R-CNN是一个非常灵活的框架,可以增加不同的分支完成不同的任务,例如,可以完成目标分类、目标检测、语义分割、实例分割、人体姿势识别等多种任务。Mask R-CNN算法步骤包括:首先,输入一幅你想处理的图片,然后进行对应的预处理操作,或者预处理后的图片;然后,将其输入到一个预训练好的神经网络中(ResNeXt等)获得对应的feature map;接着,对这个feature map中的每一点设定预定个的ROI,从而获得多个候选ROI;接着,将这些候选的ROI送入RPN网络进行二值分类(前景或背景)和BB回归,过滤掉一部分候选的ROI;接着,对这些剩下的ROI进行ROIAlign操作(即先将原图和feature map的pixel对应起来,然后将feature map和固定的feature对应起来);最后,对这些ROI进行分类(N类别分类)、BB回归和MASK生成(在每一个ROI里面进行FCN操作)。

[0063] 支持向量机SVM:SVM是常见的一种判别方法,在机器学习领域,是一个有监督的学习模型,通常用来进行模式识别、分类以及回归分析。SVM的主要思想可以概括为两点:1、它是针对线性可分情况进行分析,对于线性不可分的情况通过使用非线性映射算法将低维输入空间线性不可分的样本转化为高维特征空间使其线性可分从而使得高维特征空间采用线性算法对样本的非线性特征进行线性分析成为可能;2、它基于结构风险最小化理论之上在特征空间中构建最优超平面,使得学习器得到全局最优化,并且在整个样本空间的期望以某个概率满足一定上界。

[0064] 本发明的总体发明构思是:首先,通过语义转换,赋予语义各种抽象标签概念以及具体标签概念;然后,使用MASK R-CNN深度神经网络模型检测出各个抽象标签概念区域和具体标签概念区域;最后,通过对各个抽象标签概念区域和具体标签概念区域进行联合判断,从而准确的识别出人的危险姿态。在本发明的技术方案中,可以将姿态识别结果描述为:

[0065] $P_t = A_t \& T(A_t, H_t, U_t, L_t, \dots) t > 0$

[0066] 其中:

[0067] t 表示 t 时刻摄像头捕获到的一帧图像数据;

- [0068] A_t 表示t时刻检测到的抽象标签概念区域,比如摔倒动作;
- [0069] H_t, U_t, L_t 表示t时刻检测到的具体标签概念区域,比如头部、上半身等;
- [0070] $T(A_t, H_t, U_t, L_t)$ 表示是否保留 A_t ;
- [0071] P_t 表示t时刻是否存在危险行为。
- [0072] 所述方法包括如下步骤:
- [0073] 获取终端设备捕获到的任意一帧目标图像;其中,所述终端设备可以是各种摄像头,如监控摄像头等;
- [0074] 通过语义转换,赋予语义抽象标签概念以及具体标签概念;
- [0075] 使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;
- [0076] 对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。
- [0077] 本发明在具体实施时,为了提高姿态识别结果的置信度,所述方法还包括如下步骤:
- [0078] 获取输出的连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果。在具体实施时,需要获取连续两帧目标图像的姿态识别结果,并对两个姿态识别结果进行比对,且如果连续两帧目标图像的姿态识别结果都是有危险姿态,则判定最终的姿态识别结果为具有危险姿态;否则,如果连续两帧目标图像的姿态识别结果不全是有危险姿态,则判定为最终的姿态识别结果为误判或者不存在危险姿态,例如,如果连续两帧输出的姿态识别结果都是正常姿态,则判定最终的姿态识别结果为不存在危险姿态;如果连续两帧输出的姿态识别结果中存在有误判结果,则判定最终的姿态识别结果为误判;如果连续两帧输出的姿态识别结果中有一个是存在危险姿态,另一个是正常姿态,则判定最终的姿态识别结果为误判等等。
- [0079] 所述赋予语义抽象标签概念以及具体标签概念具体为:
- [0080] 对识别目标体(即人体)赋予正常标签和异常标签两个抽象标签概念,对识别目标体赋予头部标签、上半身标签以及下半身标签三个具体标签概念;
- [0081] 所述使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域具体为:
- [0082] 使用MASK R-CNN深度神经网络模型对获取的目标图像进行图像语义分割识别,且将从目标图像中分割出的识别目标体识别为正常区域或者异常区域,即将目标图像中的人体具体划分为正常人体或者异常人体;同时,通过MASK R-CNN深度神经网络模型检测出识别目标体的头部、上半身以及下半身三个区域的最大外接矩形区域,即检测出人体的头部、上半身和下半身三个区域。
- [0083] 所述对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果具体为:
- [0084] 如果将识别目标体识别为正常区域,则直接将该识别目标体忽略掉,并输出姿态正常的识别结果,实验表明,正样本几乎不会出现识别错误的情形,且由于实例分割网络也具备目标检测功能,因此细分检测“人体”这一任务的子任务,检测“正常人体”的置信度会更加提高。如果将识别目标体识别为异常区域,则使用贪心算法将属于该识别目标体的头

部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域三个区域分离出来(即将属于同一个人的头部、上半身以及下半身这三个最大外接矩形区域分离开来),以对异常区域、头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域这4个区域做进一步的分析;

[0085] 对识别为异常区域的目标识别体进行误判判断,且如果判断结果为误判,则输出姿态误判的识别结果;

[0086] 如果判断结果为非误判,则取下半身的最大外接矩形区域的底边中点作为第一坐标点 $P_0(x_0, y_0)$,取下半身的最大外接矩形区域的左上角作为第二坐标点 $P_1(x_1, y_1)$,取下半身的最大外接矩形区域的右侧边中点作为第三坐标点 $P_2(x_2, y_2)$;同理,取上半身的最大外接矩形区域的左上角作为第四坐标点 $P_3(x_3, y_3)$,取上半身的最大外接矩形区域的右侧边中点作为第五坐标点 $P_4(x_4, y_4)$;取头部的最大外接矩形区域的左上角作为第六坐标点 $P_5(x_5, y_5)$,取头部的最大外接矩形区域的右侧边中点作为第七坐标点 $P_6(x_6, y_6)$;并以第一坐标点 $P_0(x_0, y_0)$ 为角的顶点,分别计算出 $\angle P_1P_0P_2$ 、 $\angle P_3P_0P_4$ 、 $\angle P_5P_0P_6$ 三个夹角的余弦值,即分别计算出 $\cos\angle P_1P_0P_2$, $\cos\angle P_3P_0P_4$, $\cos\angle P_5P_0P_6$ 这三个余弦值;

[0087] 根据计算出的三个夹角的余弦值,使用支持向量机SVM对目标识别体的姿态进行分类,用以区分蹲下、坐姿等容易导致误判的姿态,并输出对应的姿态分类识别结果,例如,如果分类为蹲下姿态,则输出识别结果为蹲下姿态;如果分类为危险姿态,则输出识别结果为危险姿态等。

[0088] 所述对识别为异常区域的目标识别体进行误判判断具体为:

[0089] 计算在目标图像中检测出的目标识别体的面积与目标图像的整体面积的比例,且如果比例超过设定的第一阈值或者小于预设的第二阈值,则将该目标识别体判定为误判;例如,如果检测出的目标识别体的面积与目标图像的整体面积的比例超过1:3,则认为过于靠近摄像头导致的误识别,应当舍弃,此时就将该目标识别体判定为误判;又如,如果检测出的目标识别体的面积与目标图像的整体面积的比例小于1:20,则认为对于人体的检测产生错误,也应舍弃,此时也将该目标识别体判定为误判。

[0090] 同时,根据固定的摄像头位置推断出目标图像的可视范围,依据正常的目标识别体的头部区域、上半身区域以及下半身区域之间的比例,对检测出的目标识别体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域进行综合判断,且如果头部的最大外接矩形区域超过所述可视范围的某一预设高度(例如,超过可视范围高度的2/3处),则计算出头部的最大外接矩形区域与上半身的最大外接矩形区域的第一面积比值以及头部的最大外接矩形区域与下半身的最大外接矩形区域的第二面积比值,且如果第一面积比值和第二面积比值均符合正常的目标识别体的躯干比例,则认为误识别,此时,将该目标识别体判定为误判;

[0091] 如果未检测出目标识别体的下半身的最大外接矩形区域或者上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度达到预设的第三面积比值,则将该目标识别体判定为误判。例如,如果未检测出目标识别体的下半身的最大外接矩形区域,则认为误识别,此时,将该目标识别体判定为误判;又如,如果上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度IOU达到1/4,则认为误识别,此时,也将该目标识别体判定为误判。通过对可能存在的误判情况进行判断,可避免误判情况带来的干扰,因

此,可进一步提升对危险姿态识别的准确度。

[0092] 请参阅图3所示,本发明的技术方案在具体实现时,可由M1、M2和M3三个部件来构成;其中,M1为采集部件,主要用于采集图像帧,不需要做任何的额外处理,只需将采集的图像帧传输给M2即可;该M1需要包含至少一个摄像头,同时,还可以对M1进行进一步改善,如增加红外光用于活体识别,以进一步降低在M2进行区域识别时的错误率。M2为部署MASK R-CNN网络的服务器,用于将M1提供的图像进行归一化处理,检测出抽象标签概念区域以及具体标签概念区域并传输给M3。M3为危险姿态识别部件,用于对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。在具体实施时,M2可以单独工作,并实现单独的抽象标签概念区域以及具体标签概念区域的识别功能;M2也可以与M3组合成一个危险姿态检验系统进行使用。

[0093] 另外,需要说明的是:本发明提出的基于抽象标签概念和具体标签概念联合判断的检测,不仅仅适用于危险姿态的检测,对其它需要对人体进行抽象概念判别的领域,均可以使用本发明的技术方案;同时,本发明的技术方案还可以用于一些动物行为侦测等动作行为识别的领域。

[0094] 请参阅图2所示,与本发明方法对应的,本发明还提供一种危险姿态识别系统,所述系统包括图像获取模块、语义转换模块、区域检测模块以及联合判断模块;

[0095] 所述图像获取模块,用于获取终端设备捕获到的任意一帧目标图像;

[0096] 所述语义转换模块,用于通过语义转换,赋予语义抽象标签概念以及具体标签概念;

[0097] 所述区域检测模块,用于使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;

[0098] 所述联合判断模块,用于对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。

[0099] 为了提高姿态识别结果的置信度,所述系统还包括最终结果生成模块;

[0100] 所述最终结果生成模块,用于获取输出的连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果。

[0101] 在本发明中,所述赋予语义抽象标签概念以及具体标签概念具体为:

[0102] 对识别目标体赋予正常标签和异常标签两个抽象标签概念,对识别目标体赋予头部标签、上半身标签以及下半身标签三个具体标签概念;

[0103] 在本发明中,所述使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域具体为:

[0104] 使用MASK R-CNN深度神经网络模型对获取的目标图像进行图像语义分割识别,且将从目标图像中分割出的识别目标体识别为正常区域或者异常区域;同时,通过MASK R-CNN深度神经网络模型检测出识别目标体的头部、上半身以及下半身三个区域的最大外接矩形区域。

[0105] 在本发明中,所述对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果具体为:

[0106] 如果将识别目标体识别为正常区域,则直接将该识别目标体忽略掉,并输出姿态正常的识别结果;如果将识别目标体识别为异常区域,则使用贪心算法将属于该识别目标

体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域三个区域分离出来；

[0107] 对识别为异常区域的目标识别体进行误判判断，且如果判断结果为误判，则输出姿态误判的识别结果；

[0108] 如果判断结果为非误判，则取下半身的最大外接矩形区域的底边中点作为第一坐标点P0，取下半身的最大外接矩形区域的左上角作为第二坐标点P1，取下半身的最大外接矩形区域的右侧边中点作为第三坐标点P2；同理，取上半身的最大外接矩形区域的左上角作为第四坐标点P3，取上半身的最大外接矩形区域的右侧边中点作为第五坐标点P4；取头部的最大外接矩形区域的左上角作为第六坐标点P5，取头部的最大外接矩形区域的右侧边中点作为第七坐标点P6；并以第一坐标点P0为角的顶点，分别计算出P1P0P2、P3P0P4、P5P0P6三个夹角的余弦值；

[0109] 根据计算出的三个夹角的余弦值，使用支持向量机SVM对目标识别体的姿态进行分类，并输出对应的姿态分类识别结果。

[0110] 在本发明中，所述对识别为异常区域的目标识别体进行误判判断具体为：

[0111] 计算在目标图像中检测出的目标识别体的面积与目标图像的整体面积的比例，且如果比例超过设定的第一阈值或者小于预设的第二阈值，则将该目标识别体判定为误判；

[0112] 同时，根据固定的摄像头位置推断出目标图像的可视范围，依据正常的目标识别体的头部区域、上半身区域以及下半身区域之间的比例，对检测出的目标识别体的头部的最大外接矩形区域、上半身的最大外接矩形区域以及下半身的最大外接矩形区域进行综合判断，且如果头部的最大外接矩形区域超过所述可视范围的某一预设高度，则计算出头部的最大外接矩形区域与上半身的最大外接矩形区域的第一面积比值以及头部的最大外接矩形区域与下半身的最大外接矩形区域的第二面积比值，且如果第一面积比值和第二面积比值均符合正常的目标识别体的躯干比例，则将该目标识别体判定为误判；

[0113] 如果未检测出目标识别体的下半身的最大外接矩形区域或者上半身的最大外接矩形区域与下半身的最大外接矩形区域之间的重叠度达到预设的第三面积比值，则将该目标识别体判定为误判。

[0114] 这里需要说明的是：上述系统的实施例中，所包括的各个模块只是按照功能逻辑进行划分的，但并不局限于上述的划分，只要能够实现相应的功能即可；另外，各功能模块的具体名称也只是为了便于相互区分，并不用于限制本发明实施例的保护范围。

[0115] 与本发明方法对应的，本发明还提供一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现所述的危险姿态识别方法；所述方法包括：

[0116] 获取终端设备捕获到的任意一帧目标图像；

[0117] 通过语义转换，赋予语义抽象标签概念以及具体标签概念；

[0118] 使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域；

[0119] 对抽象标签概念区域以及具体标签概念区域进行联合判断，并输出姿态识别结果。具体实现的详细内容请参照所述的危险姿态识别方法。

[0120] 其中，存储器的数量可以是一个或者多个，所述处理器的数量也可以是一个或者

多个。所述存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。

[0121] 与本发明方法对应的,本发明还提供一种计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现所述的危险姿态识别方法;所述方法包括:

[0122] 获取终端设备捕获到的任意一帧目标图像;

[0123] 通过语义转换,赋予语义抽象标签概念以及具体标签概念;

[0124] 使用深度神经网络模型检测出获取的目标图像的抽象标签概念区域以及具体标签概念区域;

[0125] 对抽象标签概念区域以及具体标签概念区域进行联合判断,并输出姿态识别结果。具体实现的详细内容请参照所述的危险姿态识别方法。具体实现的详细内容请参照所述的危险姿态识别方法。

[0126] 其中,计算机可读存储介质可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本发明中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0127] 综上所述,本发明具有如下优点:1、采用语义分割模型来检测具体标签概念区域和抽象标签概念区域,并对抽象标签概念区域和具体标签概念区域进行联合判断,可大大增加检测的准确度,减少错误判别的概率;同时,由于具体标签概念和抽象标签概念的检测结果是并行得出的,因此,还可以免除多模型预测带来的额外预测耗时;

[0128] 2、通过对可能存在的误判情况进行判断,可避免误判情况带来的干扰,因此,可进一步提升对危险姿态识别的准确度;

[0129] 3、通过获取连续两帧目标图像的姿态识别结果进行比对分析,并根据比对分析结果生成最终的姿态识别结果,可进一步提升识别结果的置信度;

[0130] 4、使用人工智能深度学习的方式,部署成本低,使用场景宽泛,且不会随着使用年限的增加而老化,维护成本低。

[0131] 虽然以上描述了本发明的具体实施方式,但是熟悉本技术领域的技术人员应当理解,我们所描述的具体的实施例只是说明性的,而不是用于对本发明的范围的限定,熟悉本领域的技术人员在依照本发明的精神所作的等效的修饰以及变化,都应当涵盖在本发明的权利要求所保护的范围内。

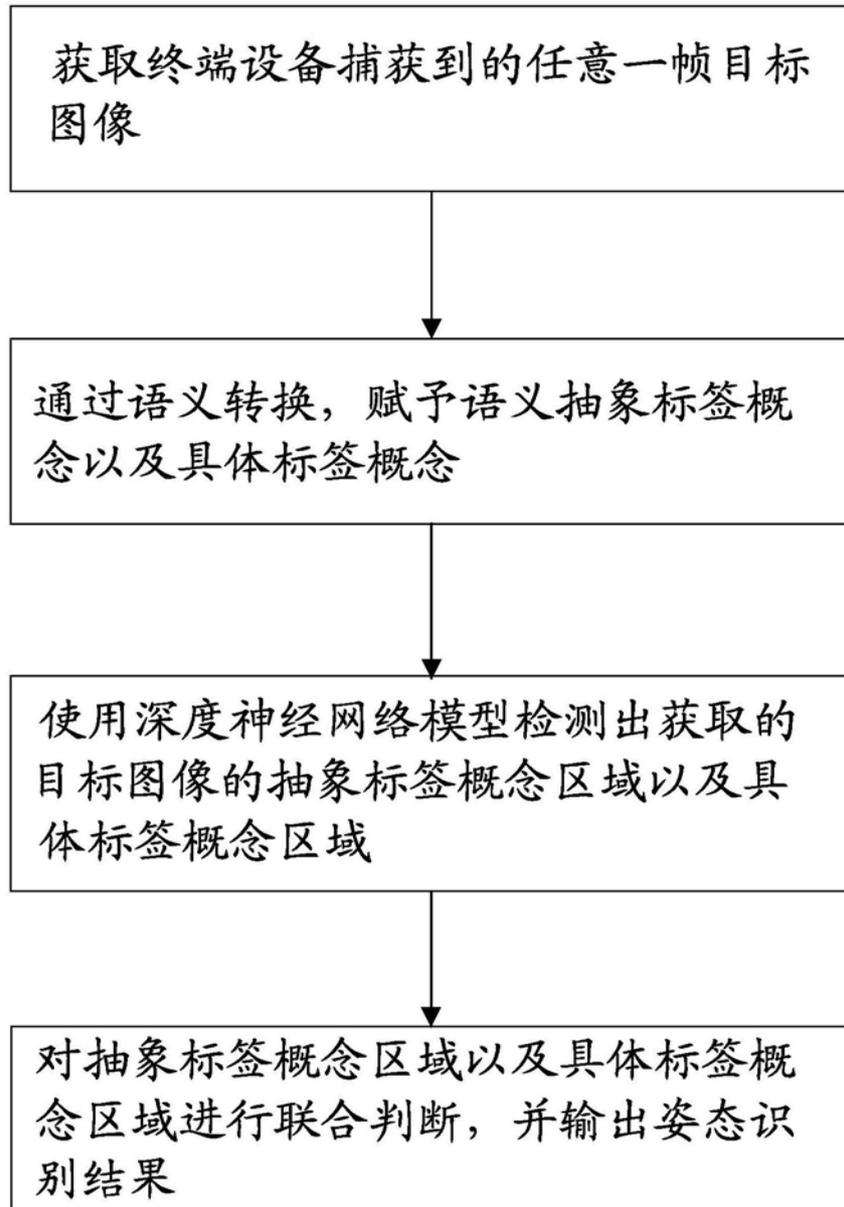


图1

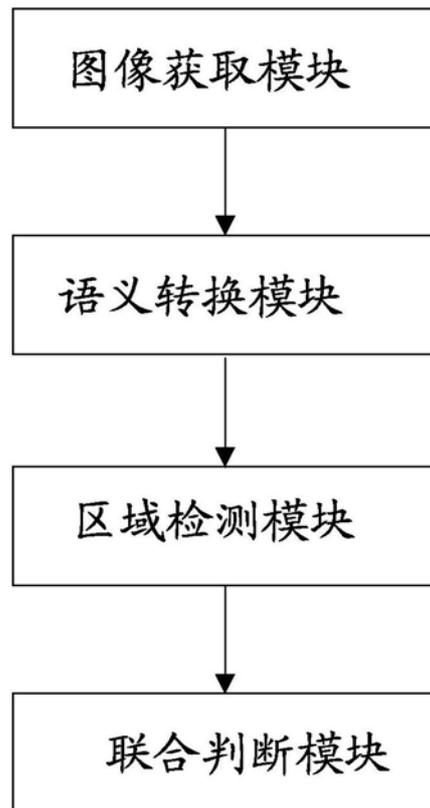


图2

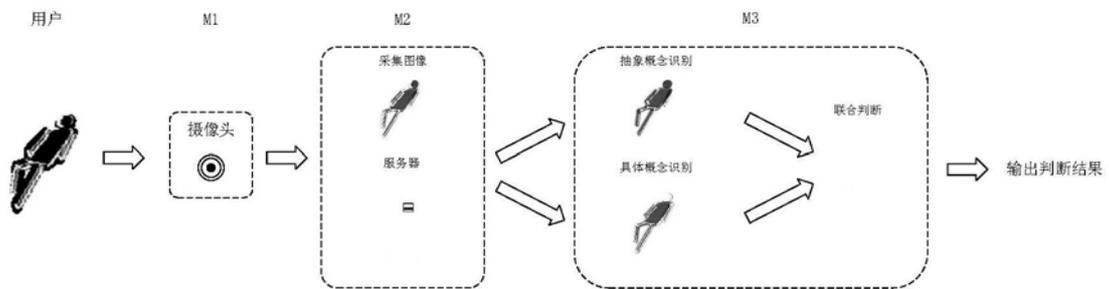


图3