



(12) 发明专利

(10) 授权公告号 CN 114677754 B

(45) 授权公告日 2025. 04. 15

(21) 申请号 202210239612.0

G06N 3/08 (2023.01)

(22) 申请日 2022.03.11

G06N 3/0464 (2023.01)

(65) 同一申请的已公布的文献号

G06V 10/762 (2022.01)

申请公布号 CN 114677754 A

G06V 10/82 (2022.01)

G06V 10/764 (2022.01)

(43) 申请公布日 2022.06.28

(56) 对比文件

(73) 专利权人 上海商汤智能科技有限公司

CN 113095257 A, 2021.07.09

地址 200233 上海市徐汇区桂平路391号3

CN 113205512 A, 2021.08.03

号楼1605A室

审查员 唐嘉

(72) 发明人 苏海昇

(74) 专利代理机构 北京派特恩知识产权代理有限公司 11270

专利代理师 王花丽 张颖玲

(51) Int. Cl.

G06V 40/20 (2022.01)

G06V 20/40 (2022.01)

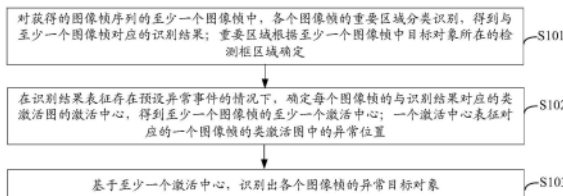
权利要求书3页 说明书14页 附图12页

(54) 发明名称

行为识别方法、装置、电子设备及计算机可读存储介质

(57) 摘要

本公开实施例公开了一种行为识别方法、装置、设备及计算机可读存储介质。该方法包括：对获得的图像帧序列的至少一个图像帧中，各个图像帧的重要区域分类识别，得到与至少一个图像帧对应的识别结果；重要区域根据至少一个图像帧中目标对象所在的检测框区域确定；在识别结果表征存在预设异常事件的情况下，确定每个图像帧的与识别结果对应的类激活图的激活中心，得到至少一个图像帧的至少一个激活中心；一个激活中心表征对应的一个图像帧的类激活图中的异常位置；基于至少一个激活中心，识别出各个图像帧的异常目标对象。通过本公开，能够提高识别准确度。



1. 一种行为识别方法,其特征在于,包括:

对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与所述至少一个图像帧对应的识别结果;所述重要区域根据所述至少一个图像帧中目标对象所在的检测框区域确定;

在所述识别结果表征存在预设异常事件的情况下,确定每个图像帧的与所述识别结果对应的一组激活值;所述一组激活值用于生成所述每个图像帧的类激活图,且所述类激活图中的每个像素位置与所述一组激活值中的一个激活值一一对应;确定所述每个图像帧的一组激活值中最大的激活值,并将所述最大的激活值所对应的像素位置确定为异常位置;基于所述每个图像帧的所述异常位置,得到所述每个图像帧的激活中心,对于所述至少一个图像帧,对应得到至少一个激活中心;一个激活中心表征对应的一个图像帧的所述类激活图中的异常位置;

基于所述至少一个激活中心,识别出所述各个图像帧的异常目标对象;

所述基于所述至少一个激活中心,识别出所述各个图像帧中的异常目标对象,包括:

确定每个激活中心在对应的一个图像帧中的位置;在所述位置位于所述图像帧中的任意一检测框区域的情况下,将所述任一检测框区域对应的目标对象,确定为所述图像帧中的异常目标对象,对于所述至少一个激活中心,对应得到所述各个图像帧的异常目标对象。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述至少一个激活中心,识别出所述各个图像帧中的异常目标对象,还包括:

根据所述至少一个激活中心,确定出目标激活中心;

根据所述目标激活中心,识别出所述各个图像帧中的异常目标对象。

3. 根据权利要求1所述的方法,其特征在于,所述基于所述至少一个激活中心,识别出所述各个图像帧中的异常目标对象,还包括:

通过无监督聚类方式对所述至少一个激活中心进行聚类,得到目标激活中心;

根据所述目标激活中心,识别出所述各个图像帧中的异常目标对象。

4. 根据权利要求3所述的方法,其特征在于,所述通过无监督聚类方式对所述至少一个激活中心进行聚类,得到目标激活中心,包括:

通过无监督聚类方式对所述至少一个激活中心进行聚类,得到包含的激活中心数量最多的中心簇;

确定所述中心簇中包含的激活中心的平均值,并将所述平均值确定为所述目标激活中心。

5. 根据权利要求1所述的方法,其特征在于,所述确定每个图像帧的与所述识别结果对应的一组激活值,包括:

生成所述每个图像帧的与所述识别结果对应的一组初始类激活映射值;

对所述初始类激活映射值进行上采样处理,得到一组中间激活值;所述一组中间激活值用于生成与所述每个图像帧的重要区域对应的类激活图;

将所述一组中间激活值与所述每个图像帧的原始像素值进行融合处理,得到所述一组激活值。

6. 根据权利要求1-5任一项所述的方法,其特征在于,在所述对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与所述至少一个图像帧对应的识

别结果之前,所述方法还包括:

对所述图像帧序列中的当前图像帧进行目标检测,确定存在的至少一个目标对象对应的至少一个检测框区域;

确定所述至少一个检测框区域中的中心检测框区域;

在已检测所述至少一个图像帧的情况下,基于所述至少一个图像帧对应的至少一个中心检测框区域,确定所述至少一个图像帧中各个图像帧的所述重要区域。

7. 根据权利要求6所述的方法,其特征在于,所述确定所述至少一个检测框区域中的中心检测框区域,包括:

将所述至少一个检测框区域扩大第一预设比例,得到与所述至少一个检测框区域对应的至少一个第一区域;

基于所述至少一个第一区域中任意两个第一区域之间的面积交并比,确定所述任意两个第一区域之间的匹配结果;所述匹配结果表征所述任意两个第一区域是否匹配;

基于所述匹配结果,确定每个第一区域的匹配次数,并确定匹配次数最大的第一区域;

根据所述匹配次数最大的第一区域,确定密集中心,并将所述密集中心对应的检测框区域,确定为所述中心检测框区域。

8. 根据权利要求6所述的方法,其特征在于,所述基于所述至少一个图像帧对应的至少一个中心检测框区域,确定所述至少一个图像帧中各个图像帧的所述重要区域,包括:

将所述至少一个图像帧对应的至少一个中心检测框区域扩大第二预设比例,得到与所述至少一个检测框区域对应的至少一个第二区域;

根据所述至少一个第二区域中各个第二区域在所属的图像帧中的位置,确定出概括区域位置;

对于每个图像帧,将所述概括区域位置在所述每个图像帧中对应的区域,确定为所述每个图像帧的所述重要区域;所述每个图像帧的重要区域包含所述每个图像帧对应的第二区域。

9. 根据权利要求7所述的方法,其特征在于,所述根据所述匹配次数最大的第一区域,确定密集中心,包括:

在存在一个所述匹配次数最大的第一区域的情况下,将所述匹配次数最大的第一区域确定为所述密集中心;

在存在两个或两个以上的所述匹配次数最大的第一区域的情况下,将所述两个或两个以上的所述匹配次数最大的第一区域中面积最大的第一区域,确定为所述密集中心。

10. 一种行为识别装置,其特征在于,包括:

识别单元,用于对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与所述至少一个图像帧对应的识别结果;所述重要区域根据所述至少一个图像帧中目标对象所在的检测框区域确定;

确定单元,用于在所述识别结果表征存在预设异常事件的情况下,确定每个图像帧的与所述识别结果对应的一组激活值;所述一组激活值用于生成所述每个图像帧的类激活图,且所述类激活图中的每个像素位置与所述一组激活值中的一个激活值一一对应;确定所述每个图像帧的一组激活值中最大的激活值,并将所述最大的激活值所对应的像素位置确定为异常位置;基于所述每个图像帧的所述异常位置,得到所述每个图像帧的激活中心,

对于所述至少一个图像帧,对应得到至少一个激活中心;一个激活中心表征对应的一个图像帧的所述类激活图中的异常位置;

所述识别单元,还用于基于所述至少一个激活中心,识别出所述各个图像帧的异常目标对象;

所述识别单元,还用于确定每个激活中心在对应的一个图像帧中的位置;在所述位置位于所述图像帧中的任一检测框区域的情况下,将所述任一检测框区域对应的目标对象,确定为所述图像帧中的异常目标对象,对于所述至少一个激活中心,对应得到所述各个图像帧的异常目标对象。

11.一种电子设备,其特征在于,包括:

存储器,用于存储可执行计算机程序;

处理器,用于执行所述存储器中存储的可执行计算机程序时,实现权利要求1至9任一项所述的方法。

12.一种计算机可读存储介质,其特征在于,存储有计算机程序,用于引起处理器执行时,实现权利要求1至9任一项所述的方法。

行为识别方法、装置、电子设备及计算机可读存储介质

技术领域

[0001] 本公开涉及终端技术,尤其涉及一种行为识别方法、装置、电子设备及计算机可读存储介质。

背景技术

[0002] 视频中的异常检测是计算机视觉领域的一个重要问题,在智能识别领域有着广泛的应用,例如检测非法行为、交通事故和一些不常见的事件等等。成千上万的摄像头在全世界范围内进行部署。然而,大多数的摄像头仅仅只是记录每刻的动态,而没有起到自动识别的能力(往往需要特殊人员来负责人工察看)。由于巨大的视频数量,仅靠人力去过滤视频中的内容显然是不太现实的。我们需要利用计算机视觉和深度学习的技术来自动检测发生在视频中的异常事件。

[0003] 识别视频中的异常事件是极其困难的。可能的挑战包括由于小概率事件导致标注数据的稀缺,类间/类内方差大,异常事件的主观定义差别,视频图像的低分辨率,等等。而作为人类,我们通过常识来进行异常的识别,例如,如果人群聚集在一条通常没有交通的街道,可能会是一个异常发生;例如,发生打架这种暴力事件。对于机器来说,它们没有常识,只有视觉的特征。通常来说,视觉特征越强,则期望的异常检测性能就越好。

[0004] 相关技术中已存在一些关于异常事件的识别方法,但采用相关技术中的识别方法进行异常识别时,识别准确度较低。

发明内容

[0005] 本公开实施例提供一种行为识别方法、装置、电子设备及计算机可读存储介质,能够提高识别准确度。

[0006] 本公开实施例的技术方案是这样实现的:

[0007] 本公开实施例提供一种行为识别方法,包括:

[0008] 对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与所述至少一个图像帧对应的识别结果;所述重要区域根据所述至少一个图像帧中目标对象所在的检测框区域确定;

[0009] 在所述识别结果表征存在预设异常事件的情况下,确定每个图像帧的与所述识别结果对应的类激活图的激活中心,得到所述至少一个图像帧的至少一个激活中心;一个激活中心表征对应的一个图像帧的所述类激活图中的异常位置;

[0010] 基于所述至少一个激活中心,识别出所述各个图像帧的异常目标对象。

[0011] 本公开实施例提供一种行为识别装置,包括:识别单元,用于对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与所述至少一个图像帧对应的识别结果;所述重要区域根据所述至少一个图像帧中目标对象所在的检测框区域确定;

[0012] 确定单元,用于在所述识别结果表征存在预设异常事件的情况下,确定每个图像帧的与所述识别结果对应的类激活图的激活中心,得到所述至少一个图像帧的至少一个激

活中心;一个激活中心表征对应的一个图像帧的所述类激活图中的异常位置;

[0013] 所述识别单元,还用于基于所述至少一个激活中心,识别出所述各个图像帧的异常目标对象。

[0014] 本公开实施例提供一种电子设备,包括:存储器,用于存储可执行计算机程序;处理器,用于执行所述存储器中存储的可执行计算机程序时,实现上述的行为识别方法。

[0015] 本公开实施例提供一种计算机可读存储介质,存储有计算机程序,用于引起处理器执行时,实现上述的行为识别方法。

[0016] 本公开实施例提供的行为识别方法、装置、设备及计算机可读存储介质,对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与该至少一个图像帧对应的识别结果;重要区域根据至少一个图像帧中目标对象所在的检测框区域确定;在识别结果表征存在预设异常事件的情况下,确定每个图像帧的与识别结果对应的类激活图的激活中心,得到至少一个图像帧的至少一个激活中心,一个激活中心表征对应的一个图像帧的类激活图中的异常位置;基于至少一个激活中心,识别出各个图像帧的异常目标对象。采用本技术方案,一方面,由于是对确定出的图像帧的重要区域进行分类识别,所以,相比于根据每个图像帧的全图的数据进行识别而言,有效缩小了识别时的搜索范围,同时降低了对识别的干扰,从而提高了识别准确度;另一方面,由于是通过确定每个图像帧的与识别结果对应的类激活图中的异常位置,来定位每个图像帧的异常中心(激活中心),并根据得到的异常中心识别出对应的图像帧中的异常目标对象,并且,与识别结果对应的类激活图能够表征出图像帧中与识别结果相关的目标对象,所以,使得识别出的每个图像帧的异常目标对象更准确,从而提高了对异常目标对象的识别准确度;因此,本公开实施例提供的行为识别方法能够提高识别准确度。

[0017] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,而非限制本公开。

附图说明

[0018] 此处的附图被并入说明书中并构成本说明书的一部分,这些附图示出了符合本公开的实施例,并与说明书一起用于说明本公开的技术方案。

[0019] 图1为本公开实施例提供的行为识别方法的一个可选的流程示意图;

[0020] 图2为本公开实施例提供的示例性的类激活图和激活中心的一个展示效果示意图;

[0021] 图3为本公开实施例提供的示例性的一个图像帧中的异常目标对象的展示效果示意图;

[0022] 图4为本公开实施例提供的行为识别方法的一个可选的流程示意图;

[0023] 图5为本公开实施例提供的行为识别方法的一个可选的流程示意图;

[0024] 图6为本公开实施例提供的示例性的类激活图的另一个展示效果示意图;

[0025] 图7A为本公开实施例提供的行为识别方法的一个可选的流程示意图;

[0026] 图7B为本公开实施例提供的行为识别方法的一个可选的流程示意图;

[0027] 图8为本公开实施例提供的示例性的图像帧对应的类激活图和目标激活中心的一个展示效果示意图;

- [0028] 图9为本公开实施例提供的行为识别方法的一个可选的流程示意图；
- [0029] 图10为本公开实施例提供的行为识别方法的一个可选的流程示意图；
- [0030] 图11为本公开实施例提供的行为识别方法的一个可选的流程示意图；
- [0031] 图12为本公开实施例提供的示例性的两个不同的第二区域的概括区域位置在图像帧中对应的区域的示意图；
- [0032] 图13为本公开实施例提供的行为识别装置的结构示意图；
- [0033] 图14为本公开实施例提供的电子设备的结构示意图。

具体实施方式

[0034] 为了使本公开的目的、技术方案和优点更加清楚,下面将结合附图对本公开作进一步地详细描述,所描述的实施例不应视为对本公开的限制,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例,都属于本公开保护的范围。

[0035] 本公开的发明人经研究发现,传统的行为识别方法通常对输入视频序列进行全图的数据增强或其他预处理后送入到行为识别模型中进行预测,然而这种方式只适用于以人为中心的视频行为识别,该类数据常见于公开的视频学术数据集中。对于监控摄像头拍摄的视频来说,往往包含更多的信息,覆盖的视野也更大;同时,目标事件的发生位置和人体尺度也具有随机性;因此,简单地以全图作为模型输入显然是不合理的,画面中的无关信息较多,会对行为识别造成干扰,从而影响行为识别的准确度,以及影响对行为执行人的识别准确度。

[0036] 对此,本公开的发明人经研究认为,可以以局部区域代替整张图像输入到行为识别模型中进行行为分类,如此,可减少画面中大多数无关信息的干扰,提高行为识别的准确度;但在此基础上还需考虑到在缩小了行为识别模型的搜索范围之后,后续如何准确定位到具体的行为执行者的位置这一层面。

[0037] 为此,本公开实施例提供一种行为识别方法,能够提高异常行为识别和与异常行为识别关联的异常目标对象(行为执行人)的识别准确度。对本公开实施例进行进一步详细说明之前,对本公开实施例中涉及的名词和术语进行说明,本公开实施例中涉及的名词和术语适用于如下的解释。

[0038] 1) CNN(Convolutional Neural Networks),指卷积神经网络,是一种带有卷积结构的深度神经网络;它本质上是一种输入到输出的映射,能够学习大量的输入与输出之间的映射关系,用已知的模式对卷积网络加以训练,网络就具有输入输出对之间的映射能力;通常用在图像识别领域。

[0039] 2) CAM(Class Activation Map),指类别激活热力图,简称类激活热力图或类激活图,是一个二维图,实质是与特定输出类别相关的二维特征分数网络,网络的每个位置表示对于该类别的重要程度。对于一张输入到CNN模型且被分类成“狗”的图片,可以以热力图形式呈现图片中每个位置与“狗”类的相似程度。类激活图有助于了解一张图片的哪个部分使得卷积神经网络做出最终的决策。

[0040] 下面说明本公开实施例提供的电子设备的示例性应用,本公开实施例提供的电子设备可以实施为AR眼镜、笔记本电脑,平板电脑,台式计算机,机顶盒,移动设备(例如,移动电话,便携式音乐播放器,个人数字助理,专用消息设备,便携式游戏设备)等各种类型的用

户终端(以下简称终端),也可以实施为服务器。

[0041] 下面,将说明电子设备实施为终端时的示例性应用。图1是本公开实施例提供的行为识别方法的一个可选的流程示意图,将结合图1示出的步骤进行说明。

[0042] S101、对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与至少一个图像帧对应的识别结果;重要区域根据至少一个图像帧中目标对象所在的检测框区域确定。

[0043] 电子设备可以在得到图像帧序列的至少一个图像帧中各个图像帧的重要区域的情况下,将这至少一个图像帧中各个图像帧的重要区域同时输入行为识别模型中进行分类识别,得到与这至少一个图像帧对应的识别结果;每个图像帧的重要区域是根据这至少一个图像帧中的所有目标对象所在的检测框区域而确定出的;例如,电子设备在每得到图像帧序列的8个图像帧中各个图像帧的重要区域的情况下,便可将这8个图像帧中各个图像帧的重要区域同时输入行为识别模型中进行分类识别,从而得到与这8个图像帧对应的一个识别结果。

[0044] 这里,电子设备可以将这至少一个图像帧中各个图像帧的重要区域裁剪出来,并对裁剪出的这些重要区域进行分类识别。

[0045] 本公开实施例中,检测框用于对图像帧中的目标对象的位置进行标记,每个目标对象对应一个检测框,例如下述图3中的存在与两个目标对象一一对应的两个检测框。

[0046] 在一些实施例中,电子设备可以从其他设备处直接获得图像帧序列中的这至少一个图像帧和每个图像帧的重要区域,以根据这至少一个图像帧中各个图像帧的重要区域进行分类识别,从而得到与这至少一个图像帧对应的识别结果。

[0047] 本公开实施例中,行为识别模型是预先训练好的神经网络模型;例如,Two-Stream模型、3D(3-dimensional)ConvNets模型和2D(2-dimensional)ConvNets等,本公开实施例对此不作限定。电子设备通过对这至少一个图像帧中各个图像帧的重要区域进行分类识别,可以确定这至少一个图像帧中是否存在打架、聚众闹事、交通事故等异常事件。

[0048] 在一些实施例中,图像帧序列可以是设置于特定场景的摄像头等采集设备所采集的视频流,这至少一个图像帧则可以是视频流中的连续帧;特定场景可以是街道、商场、景区等等场景,本公开实施例对此不作限定。

[0049] S102、在识别结果表征存在预设异常事件的情况下,确定每个图像帧的与识别结果对应的类激活图的激活中心,得到至少一个图像帧的至少一个激活中心;一个激活中心表征对应的一个图像帧的类激活图中的异常位置。

[0050] 电子设备在识别到这至少一个图像帧中存在预设异常事件(即识别到了预设异常事件这一类别)的情况下,可以确定出至少一个图像帧中每个图像帧对应的类激活图的激活中心,从而得到这至少一个图像帧对应的至少一个激活中心,其中,类激活图是与识别结果对应的类别相关的类激活图,一个激活中心表征对应的一个图像帧的类激活图中的异常位置。例如,如图2所示,在根据 y_1 、 y_2 、 y_3 、 y_4 、 y_5 、 y_6 、 y_7 和 y_8 这8个连续图像帧,识别出存在打架这一预设异常事件的情况下,可以生成与 y_1 、 y_2 、 y_3 、 y_4 、 y_5 、 y_6 、 y_7 和 y_8 这8个连续图像帧一一对应的、且与打架事件对应的8个类激活图 w_1 、 w_2 、 w_3 、 w_4 、 w_5 、 w_6 、 w_7 和 w_8 ,其中,类激活图 w_1 、 w_2 、 w_3 、 w_4 、 w_5 、 w_6 、 w_7 和 w_8 中,对应标示出了区域 r_1 、 r_2 、 r_3 、 r_4 、 r_5 、 r_6 、 r_7 和 r_8 ,分别代表了各自的图像帧中对于模型识别出预设异常事件这一类别时提供的信息最多的区域,

例如,类激活图w1中的区域r1代表了图像帧y1中对于模型识别出打架事件这一类别时提供的信息最多的区域;以及,还确定出w1、w2、w3、w4、w5、w6、w7和w8这8个类激活图的激活中心,如图2中的j1、j2、j3、j4、j5、j6、j7和j8(由于这8个类激活图的尺寸与对应的8个图像帧的尺寸一致,所以图2中直接将类激活图的激活中心标注在了对应的图像帧中,且如图2所示,激活中心为标注有五角心形的位置)。

[0051] 本公开实施例中,预设异常事件例如可以是打架、聚众闹事、交通事故等等,本公开实施例对此不作限定。

[0052] S103、基于至少一个激活中心,识别出各个图像帧的异常目标对象。

[0053] 本公开实施例中,电子设备在确定出这至少一个图像帧中每个图像帧的激活中心的情况下,对于每个图像帧而言,可以根据该图像帧的激活中心来确定出该图像帧中是否存在异常目标对象,以及在存在异常目标对象的情况下识别出存在的异常目标对象。

[0054] 在一些实施例中,由于每个激活中心均对应了一个图像帧,对于每个激活中心而言,电子设备可以确定该激活中心在对应的一个图像帧中的位置,并在该位置位于该图像帧中的任一检测框区域的情况下,将该任一检测框区域对应的目标对象,确定为该图像帧中的异常目标对象;从而对于这至少一个激活中心而言,可以得到与这至少一个激活中心对应的至少一个图像帧中各个图像帧的异常目标对象。例如,如图3所示,对于图像帧a而言,在激活中心31位于两个不同的人体检测框32和33的情况下,可以确定人体检测框32和33对应的两个行人为图像帧a中的异常目标对象。

[0055] 这里,对于每个激活中心而言,在该激活中心在对应的一个图像帧中的位置不属于该图像帧中的任一检测框区域的情况下,可以确定该图像帧中不存在异常目标对象。

[0056] 本公开实施例中,由于是对确定出的图像帧的重要区域进行分类识别,所以,相比于根据每个图像帧的全图的数据进行识别而言,有效缩小了识别时的搜索范围,同时降低了对识别的干扰,从而提高了识别准确度;另一方面,由于是通过确定每个图像帧的与识别结果对应的类激活图中的异常位置,来定位每个图像帧的异常中心(激活中心),并根据得到的异常中心识别出对应的图像帧中的异常目标对象,并且,与识别结果对应的类激活图能够表征出图像帧中与识别结果相关的目标对象,所以,可以使后续识别出的每个图像帧的异常目标对象更准确,从而提高了对异常目标对象的识别准确度,同时,使得在识别时缩小模型的搜索范围也不会对后续的异常目标对象的识别产生影响;因此,本公开实施例的技术方案能够提高识别准确度。

[0057] 图4是本公开实施例提供的行为识别方法的一个可选的流程示意图,如图4所示,上述S102可以通过S1021-S1023实现,将结合图4进行说明。

[0058] S1021、在识别结果表征存在预设异常事件的情况下,确定每个图像帧的与识别结果对应的一组激活值;该组激活值用于生成该图像帧的类激活图,且类激活图中的每个像素位置与该组激活值中的一个激活值一一对应。

[0059] 电子设备在通过行为识别模型识别到这至少一个图像帧中存在预设异常事件的情况下,对于每张图像帧,电子设备可以通过类激活映射方法,确定出该图像帧的与预设异常事件对应的一组激活值,其中该组激活值用于绘制出该图像帧的与预设异常事件对应的类激活图,并且该类激活图中的每个像素位置与该组激活值中的一个像素位置一一对应;在根据该组激活值绘制该图像帧的类激活图的过程中,例如,在一个激活值表征了与该激

活值对应的一个像素点的颜色为红色的情况下,便在与该激活值对应的像素位置生成红色,通过此方法绘制出与该组激活值对应的类激活图,即在类激活图中用不同深浅的颜色表示不同像素点对此识别结果的贡献。

[0060] 本公开实施例中,根据该组激活值绘制出的该图像帧的类激活图的尺寸与该图像帧的尺寸大小一致,如此便于后续根据该图像帧的类激活图像中的异常位置,对原始的该图像帧中的异常目标对象进行准确定位。

[0061] 在一些实施例中,图5是本公开实施例提供的行为识别方法的一个可选的流程示意图,如图5所示,上述S1021可以通过S201-S203实现,将结合图5进行说明。

[0062] S201、生成每个图像帧的与识别结果对应的一组初始类激活映射值。

[0063] 对于每个图像帧,电子设备在得到该图像帧的识别结果的情况下,可以生成该图像帧的、且与识别结果对应的一组初始类激活映射值。

[0064] 在一些实施例中,上述S201可以通过下述方式实现:对于每个图像帧,基于与识别结果对应的分类层的权重,以及行为识别模型中的卷积层输出的特征图,生成该图像帧的与识别结果对应的一组初始类激活映射值;其中,识别结果是采用行为识别模型对重要区域进行分类识别得到的;行为识别模型包括:分类层和多个卷积层。在一些实施例中,电子设备可以采用包括分类层和多个卷积层的行为识别模型,对这至少一个图像帧中各个图像帧的重要区域进行分类识别,以识别出预设异常事件这一事件类别。对于每个图像帧,行为识别模型中的最后一卷积层能够生成该图像帧的重要区域的 y 张特征图,每个特征图主要提取了与 y 个类别中的某一类别相关的特征,例如为, y 张特征图可以表示为 A^1, A^2, \dots, A^y ;分类层中对应包括了 y 个神经元,一个神经元对应一个类别,且每个神经元对应有 y 个权重值,例如, w^1, w^2, \dots, w^y ;在预设异常事件这一事件类别为 c 类的情况下,该图像帧对应的一组初始类激活映射值,可以采用下述公式(1)计算得到:

$$[0065] \quad L_{CAM}^c = \sum_{i=1}^y w_i^c A^i \quad (1)$$

[0066] 其中, i 表示第 i 个神经元, w_i^c 表示第 i 个神经元对应的 c 类的权重值, A^i 表示第 i 个神经元对应的特征图, L_{CAM}^c 表征一组初始类激活映射值。

[0067] S202、对初始类激活映射值进行上采样处理,得到一组中间激活值;一组中间激活值用于生成与该图像帧的重要区域对应的类激活图。

[0068] S203、将中间激活值与该图像帧的原始像素值进行融合处理,得到一组激活值。

[0069] 这里,对于一个图像帧而言,得到的这一组初始类激活映射值所能绘制出的类激活图与最后一个卷积层生成的特征图的尺寸是一致的,所以,电子设备可以对这一组初始类激活映射值进行上采样操作,以生成一组中间激活值,且生成的这一组中间激活值所能绘制出的类激活图的大小与该图像帧的重要区域的大小一致,即得到了该图像帧的重要区域对应的类激活图;在此基础上,电子设备还可以将这一组中间激活值与该图像帧的一组原始像素值进行融合处理,得到该图像帧对应的一组激活值,且得到的这一组激活值所能绘制出的类激活图的大小与该图像帧的原图的大小一致,从而实现将该图像帧的原图与该图像帧的重要区域对应的类激活图进行叠加,最终得到与该图像帧对应的、且与该图像帧的尺寸一致的类激活图。例如,图6示出了根据 $y_{10}, y_{20}, y_{30}, y_{40}, y_{50}, y_{60}, y_{70}$ 和 y_{80} (图6中

未示出)这8个连续图像帧,识别出打架事件的情况下,生成的与y10、y20、y30、y40、y50、y60、y70和y80这8个连续图像帧一一对应的、且与打架事件对应的8个类激活图w10、w20、w30、w40、w50、w60、w70和w80,其中,类激活图w10、w20、w30、w40、w50、w60、w70和w80中,对应标示出了区域r10、r20、r30、r40、r50、r60、r70和r80,分别代表了各自的图像帧中对于模型识别出预设异常事件这一类别时提供的信息最多的区域,例如,类激活图w10中的区域r10代表了图像帧y10中对于模型识别出打架事件这一类别时提供的信息最多的区域。

[0070] 以下继续对上述图4中的S1022~S1023进行解释说明:

[0071] S1022、确定每个图像帧的一组激活值中最大的激活值,并将最大的激活值所对应的像素位置确定为异常位置。

[0072] 对于每个图像帧,电子设备可以确定出与该图像帧对应的一组激活值中的最大激活值,并将该最大激活值对应的像素位置确定为该图像帧的异常位置。由于类激活图是二维图像,所以,异常位置(像素位置)可以采用二维坐标进行表示,例如, $X(i, j)$,以表征该异常位置在类激活图或在该图像帧中的位置。

[0073] S1023、基于该图像帧的异常位置,得到该图像帧的激活中心,对于至少一个图像帧,对应得到至少一个激活中心。

[0074] 对于每个图像帧,电子设备可以将该图像帧的类激活图中的异常位置,确定为该图像帧的类激活图的激活中心,从而对于这至少一个图像帧,可以对应得到至少一个激活中心。

[0075] 图7A是本公开实施例提供的行为识别方法的一个可选的流程示意图,如图7A所示,上述S103可以通过S1031-S1032实现,将以图1为例,结合图7A进行说明。

[0076] S1031、根据至少一个激活中心,确定出目标激活中心。

[0077] 由于不同类激活图中的激活中心的位置可能会存在偏移,所以电子设备在确定出这至少一个图像帧对应的所有激活中心的情况下,可以根据得到的激活中心确定出一个目标激活中心;如此,可以缓解激活中心的漂移现象,从而使得后续采用目标激活中心确定出的这至少一个图像帧中每个图像帧的异常目标对象更准确,从而有利于提高对异常目标对象的识别准确度。

[0078] 在一些实施例中,电子设备可以计算出这至少一个图像帧对应的所有激活中心的平均值,并将平均值作为目标激活中心。这里,由于每个激活中心实际为一个像素点的二维坐标,所以,电子设备可以通过计算这些二维坐标的坐标平均值,得到目标激活中心,例如,可以分别计算出x坐标的平均值和y坐标的平均值。

[0079] S1032、根据目标激活中心,识别出各个图像帧中的异常目标对象。

[0080] 电子设备在确定出这至少一个图像帧对应的一个目标激活中心的情况下,对于每个图像帧而言,均可以根据该目标激活中心来确定出该图像帧中是否存在异常目标对象,以及在确定存在异常目标对象的情况下,识别出存在的异常目标对象。

[0081] 这里,对于每个图像帧而言,电子设备可以确定该目标激活中心在该图像帧中的位置,并在该位置位于该图像帧中的任一检测框区域的情况下,将该任一检测框区域对应的目标对象,确定为该图像帧中的异常目标对象;从而对于这至少一个图像帧而言,可以得到这至少一个图像帧中各个图像帧的异常目标对象。例如,如上述图3中的32和33为图像帧a中的异常目标对象。

[0082] 图7B是本公开实施例提供的行为识别方法的一个可选的流程示意图,如图7B所示,上述S103还可以通过S1033-S1034实现,将以图1为例,结合图7B进行说明。

[0083] S1033、通过无监督聚类方式对至少一个激活中心进行聚类,得到目标激活中心。

[0084] S1034、根据目标激活中心,识别出各个图像帧中的异常目标对象。

[0085] 电子设备在确定出与至少一个图像帧——对应的至少一个激活中心的情况下,可以采用无监督聚类方式根据得到的这至少一个激活中心,确定出一个目标激活中心;如此,可以缓解激活中心的漂移现象,从而使得后续采用目标激活中心确定出的这至少一个图像帧中每个图像帧的异常目标对象更准确,从而有利于提高对异常目标对象的识别准确度。

[0086] 在一些实施例中,电子设备可以采用K-Means的聚类方法,对这至少一个图像帧对应的所有激活中心进行聚类,如此可以实现通过无监督聚类的方式剔除漂移点,使得后续采用目标激活中心确定出的这至少一个图像帧中每个图像帧的异常目标对象更准确,有利于提高对异常目标对象的识别准确度。

[0087] 这里,无监督聚类方式还可以是其他的聚类方法,本公开实施例对此不作具体限定。

[0088] 在一些实施例中,上述S1031还可以通过上述的S1033实现。

[0089] 在一些实施例中,上述S1033可以通过S301-S302实现:

[0090] S301、通过无监督聚类方式对至少一个激活中心进行聚类,得到包含的激活中心数量最多的中心簇。

[0091] 电子设备可以通过无监督聚类方式对这至少一个图像帧对应的所有激活中心进行聚类,从而得到多个簇(cluster),其中每个簇中包含至少一个激活中心,之后,可以从这多个簇中选出包含的激活中心的数量最多的簇,从而得到了中心簇。

[0092] S302、确定中心簇中包含的激活中心的平均值,并将平均值确定为目标激活中心。

[0093] 电子设备在确定出中心簇的情况下,可以计算出中心簇中包含的所有激活中心的平均值,并将该平均值作为目标激活中心。同样的,由于每个激活中心实际为一个像素点的二维坐标,所以,电子设备可以通过计算这些二维坐标的坐标平均值,得到目标激活中心,例如,得到目标激活中心 $X(i, j)$ 。图8示出了在根据 y_{11} 、 y_{22} 、 y_{33} 、 y_{44} 、 y_{55} 、 y_{66} 、 y_{77} 和 y_{88} 这8个连续图像帧,识别出存在打架这一预设异常事件的情况下,可以生成与 y_{11} 、 y_{22} 、 y_{33} 、 y_{44} 、 y_{55} 、 y_{66} 、 y_{77} 和 y_{88} 这8个连续图像帧——对应的、且与打架事件对应的8个类激活图 w_{11} 、 w_{22} 、 w_{33} 、 w_{44} 、 w_{55} 、 w_{66} 、 w_{77} 和 w_{88} ,其中,类激活图 w_{11} 、 w_{22} 、 w_{33} 、 w_{44} 、 w_{55} 、 w_{66} 、 w_{77} 和 w_{88} 中,对应标示出了区域 r_{11} 、 r_{22} 、 r_{33} 、 r_{44} 、 r_{55} 、 r_{66} 、 r_{77} 和 r_{88} ,分别代表了各自的图像帧中对于模型识别出预设异常事件这一类别时提供的信息最多的区域,例如,类激活图 w_{11} 中的区域 r_{11} 代表了图像帧 y_{11} 中对于模型识别出打架事件这一类别时提供的信息最多的区域;以及,还确定出 w_{11} 、 w_{22} 、 w_{33} 、 w_{44} 、 w_{55} 、 w_{66} 、 w_{77} 和 w_{88} 这8个类激活图的目标激活中心,如图8中的 j_{11} 、 j_{22} 、 j_{33} 、 j_{44} 、 j_{55} 、 j_{66} 、 j_{77} 和 j_{88} 所示(由于这8个类激活图的尺寸与对应的8个图像帧的尺寸一致,所以图8中直接将类激活图的目标激活中心标注在了对应的图像帧中,且如图8所示目标激活中心为标注有五角心形的位置)。

[0094] 在一些实施例中,对于上述S101中的至少一个图像帧,以及这至少一个图像中各个图像帧的重要区域,电子设备也可以通过一些方法自己获得,例如图9所示的方法。图9是本公开实施例提供的行为识别方法的一个可选的流程示意图,如图9所示,在上述S101之

前,还可以执行S001-S003,将以图1为例结合图9进行说明。

[0095] S001、对图像帧序列的当前图像帧进行目标检测,确定存在的至少一个目标对象对应的至少一个检测框区域。

[0096] 电子设备可以从设置于特定场景的摄像头等采集设备处得到实时采集的特定场景的视频流,并对实时采集的视频流以图像帧为单位进行实时目标检测,在目标检测过程中,对于当前图像帧而言,电子设备可以采用预先训练好的目标检测网络,检测当前图像帧中是否存在目标对象,并在存在目标对象的情况下,用检测框将当前图像帧中所有的目标对象标注出来,从而得到当前图像帧中的所有目标对象对应的检测框区域;例如,上述图3中的图像帧a中存在与两个目标对象一一对应的两个检测框区域32和33。

[0097] 本公开实施例中,目标检测网络可以是CNN网络,也可以是其他网络等等,本公开实施例对此不作限定。

[0098] 对于一个图像帧而言,可能存在目标对象也可能不存在目标对象,本公开实施例中针对图像帧中存在目标对象的情况进行说明。

[0099] S002、确定至少一个检测框区域中的中心检测框区域。

[0100] S003、在已检测至少一个图像帧的情况下,基于至少一个图像帧对应的至少一个中心检测框区域,确定至少一个图像帧中各个图像帧的重要区域。

[0101] 对于每个图像帧,电子设备在获得了该图像帧中的所有检测框区域的情况下,可以从中确定出一个检测框区域作为中心检测框区域,从而在电子设备获得了至少一个图像帧的情况下,可以对应得到至少一个中心检测框区域,并根据这至少一个中心检测框区域,确定这至少一个图像帧中各个图像帧的重要区域。

[0102] 本公开实施例中,电子设备可以在确定满足预设条件的情况下,确定已检测了至少一个图像帧。示例性地,电子设备可以在确定出已检测时长达到预设时长的情况下,确定满足预设条件,并获得达到预设时长时已检测的图像帧,预设时长可以根据实际需要设定;例如,预设时长可以为6秒,则电子设备可以将每6秒内检测过的图像帧作为已检测的图像帧;也就是说,电子设备每隔6秒钟会得到至少一个图像帧,便会确定出这至少一个图像帧中各个图像帧的重要区域;示例性地,电子设备还可以在确定已检测过的图像帧的数量达到预设数量的情况下,确定满足预设条件,并获得已检测的图像帧,预设数量可以根据实际需要设定;例如,在预设数量为8的情况下,电子设备每检测8个图像帧,便将这8个图像帧作为已检测的图像帧;也就是说,电子设备每检测8个图像帧,便会确定出这8个图像帧中各个图像帧的重要区域。

[0103] 在一些实施例中,图10是本公开实施例提供的行为识别方法的一个可选的流程示意图,如图10所示,上述S002可以通过S401-S404实现,将结合图10进行说明。

[0104] S401、将至少一个检测框区域扩大第一预设比例,得到与至少一个检测框区域对应的至少一个第一区域。

[0105] 对于每个图像帧中的所有检测框区域,电子设备可以将所有检测框区域的面积扩大第一预设比例,从而得到与所有检测框区域对应的第一区域;例如,在每个图像帧中存在N个检测框区域的情况下,电子设备可以将这N个检测框区域的面积扩大第一预设比例,从而得到N个第一区域,其中,N为大于0的整数。需要说明的是,将一个检测框区域的面积扩大第一预设比例是指将该检测框区域包含的范围扩大第一预设比例,从而得到包含的范围更

大的第一区域。

[0106] 本公开实施例中,第一预设比例可以根据实际需要设定,例如,第一预设比例可以是1.5,本公开实施例对此不作限定。

[0107] S402、基于至少一个第一区域中任意两个第一区域之间的面积交并比,确定任意两个第一区域之间的匹配结果;匹配结果表征任意两个第一区域是否匹配。

[0108] 对于每个图像帧中的所有检测框区域对应的第一区域,电子设备可以计算任意两个第一区域之间的面积交并比,并根据得到的面积交并比与预设阈值之间的大小关系来确定这任意两个第一区域之间是否匹配。例如,在一个图像帧对应了3个第一区域(第一区域1、第一区域2和第一区域3)的情况下,电子设备可以计算第一区域1与第一区域2之间的面积交并比、第一区域1与第一区域3之间的面积交并比,以及第一区域2与第一区域3之间的面积交并比,并在第一区域1与第一区域2之间的面积交并比小于预设阈值的情况下,确定第一区域1与第一区域2之间不匹配,而在第一区域1与第一区域3之间的面积交并比大于或等于预设阈值的情况下,确定第一区域1与第一区域3相匹配,以及在第一区域2与第一区域3之间的面积交并比大于或等于预设阈值的情况下,确定第一区域2与第一区域3相匹配。

[0109] 本公开实施例中,预设阈值可以根据实际需要设定,例如,可以为0,对此本公开实施例不作限定。

[0110] S403、基于匹配结果,确定每个第一区域的匹配次数,并确定匹配次数最大的第一区域。

[0111] 对于每个图像帧,电子设备在确定出了该图像帧的所有第一区域之间的匹配结果的情况下,可以根据这些匹配结果来确定出每个第一区域所匹配的其他第一区域的个数,并将该个数作为匹配次数,并对所有第一区域的匹配次数进行降序或增序排列,根据排序结果从中确定出匹配次数最大的一个或多个第一区域。继续以上举例进行说明,由于通过面积交并比确定出了第一区域1与第一区域2之间不匹配,第一区域1与第一区域3相匹配,以及第一区域2与第一区域3相匹配,所以,可以确定第一区域1所匹配的其他第一区域是第一区域3,第一区域2所匹配的其他第一区域是第一区域3,以及第一区域3所匹配的其他第一区域是第一区域1和第一区域2;由此可以得到第一区域1的匹配次数为1、第一区域2的匹配次数为1,第一区域3的匹配次数为2,并且,第一区域3是匹配次数最大的第一区域。

[0112] S404、根据匹配次数最大的第一区域确定密集中心,并将密集中心对应的检测框区域,确定为中心检测框区域。

[0113] 对于每个图像帧,电子设备在得到了该图像帧的匹配次数最大的第一区域的情况下,可以根据该匹配次数最大的第一区域确定出该图像帧的密集中心,并将密集中心对应的检测框区域,确定为该图像帧的中心检测框区域。例如,以上述图3为例,在密集中心对应的检测框区域为33的情况下,可以将检测框区域为33作为图像帧a的中心检测框区域。

[0114] 在一些实施例中,上述S404中的根据匹配次数最大的第一区域确定密集中心,可以通过以下方式实现:在存在一个匹配次数最大的第一区域的情况下,将匹配次数最大的第一区域确定为密集中心;在存在两个或两个以上的匹配次数最大的第一区域的情况下,将两个或两个以上的匹配次数最大的第一区域中面积最大的第一区域,确定为密集中心;如此,可以提高得到的密集中心的准确性。

[0115] 在一些实施例中,图11是本公开实施例提供的行为识别方法的一个可选的流程示

意图,如图11所示,上述S003可以通过S501-S503实现,将结合图11进行说明。

[0116] S501、在已检测至少一个图像帧的情况下,将至少一个图像帧对应的至少一个中心检测框区域扩大第二预设比例,得到与至少一个检测框区域对应的至少一个第二区域。

[0117] 对于这至少一个图像帧中每个图像帧的所有检测框区域,电子设备可以将所有检测框区域的面积扩大第二预设比例,从而得到与所有检测框区域对应的第二区域;例如,在每个图像帧中存在N个检测框区域的情况下,电子设备可以将这N个检测框区域的面积扩大第二预设比例,从而得到N个第二区域,其中,N为大于0的整数。需要说明的是,将一个检测框区域的面积扩大第二预设比例是指将该检测框区域包含的范围扩大第二预设比例,从而得到包含的范围更大的第二区域。

[0118] 本公开实施例中,第二预设比例可以根据实际需要设定,例如,第二预设比例可以是2,本公开实施例对此不作限定。

[0119] S502、根据至少一个第二区域中各个第二区域在所属的图像帧中的位置,确定出概括区域位置。

[0120] 在一些实施例中,对于得到的这至少一个图像帧对应的所有第二区域,电子设备可以根据每个第二区域在对应图像帧中的位置信息,来确定出这至少一个图像帧对应的所有第二区域共同对应的一个概括区域位置,并使得该概括区域位置所对应的区域,可以包含这至少一个图像帧对应的所有第二区域中的任意一个第二区域。

[0121] 这里,概括区域位置可以为区域坐标,对于这至少一个图像帧对应的所有第二区域而言,概括区域位置可以是这至少一个图像帧对应的所有第二区域的区域坐标的并集,例如,可以是最小的并集。举例说明,在获得了2个图像帧对应的2个第二区域,且第一个第二区域m11在第1个图像帧中的区域坐标为 (x_{11}, y_{11}) 、 (x_{12}, y_{12}) ,第二个第二区域m12在第2个图像帧中的区域坐标为 (x_{21}, y_{21}) 、 (x_{22}, y_{22}) ,且 $x_{11} < x_{12} < x_{21} < x_{22}$, $y_{21} < y_{22} < y_{11} < y_{12}$ 的情况下,这两个第二区域的区域坐标的最小坐标并集为: (x_{11}, y_{21}) 、 (x_{22}, y_{12}) ;其中,该最小的坐标并集在该第1个图像帧中所对应的区域,包含了第一个第二区域m11,并且,该最小的坐标并集在该第2个图像帧中所对应的区域,包含了第二个第二区域m12。例如,图12示出了第1个图像帧中的第二区域m11的区域坐标 (x_{11}, y_{11}) 、 (x_{12}, y_{12}) 与第2个图像帧中的第二区域m12的区域坐标为 (x_{21}, y_{21}) 、 (x_{22}, y_{22}) ,以及这两个第二区域之间的最小坐标并集所对应的区域111,其中,区域111包含了区域坐标 (x_{11}, y_{11}) 、 (x_{12}, y_{12}) 对应的区域112,以及包含了区域坐标为 (x_{21}, y_{21}) 、 (x_{22}, y_{22}) 对应的区域113。

[0122] S503、对于每个图像帧,将概括区域位置在该图像帧中对应的区域,确定为该图像帧的重要区域;每个图像帧的重要区域包含该图像帧对应的第二区域。

[0123] 在得到了这至少一个图像帧的所有关注区域共同对应的一个概括区域位置的情况下,对于这至少一个图像帧中的每个图像帧而言,电子设备可以确定出概括区域位置在该图像帧中的区域,并将确定出的该区域作为该图像帧的重要区域,其中,该图像帧的重要区域中包含了该图像帧的第二区域。

[0124] 本公开还提供一种行为识别装置,图13为本公开实施例提供的行为识别装置的结构示意图;如图13所示,行为识别装置1包括:识别单元11,用于对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与至少一个图像帧对应的识别结果;所述重要区域根据所述至少一个图像帧中目标对象所在的检测框区域确定;确定单元12,

用于在所述识别结果表征存在预设异常事件的情况下,确定每个图像帧的与所述识别结果对应的类激活图的激活中心,得到所述至少一个图像帧的至少一个激活中心;一个激活中心表征对应的一个图像帧的所述类激活图中的异常位置;所述识别单元11,还用于基于所述至少一个激活中心,识别出所述各个图像帧的异常目标对象。

[0125] 在本公开的一些实施例中,所述确定单元12,还用于确定每个图像帧的与所述识别结果对应的一组激活值;所述一组激活值用于生成所述每个图像帧的类激活图,且所述类激活图中的每个像素位置与所述一组激活值中的一个激活值一一对应;确定所述每个图像帧的一组激活值中最大的激活值,并将所述最大的激活值所对应的像素位置确定为所述异常位置;基于所述每个图像帧的所述异常位置,得到所述每个图像帧的激活中心,对于所述至少一个图像帧,对应得到所述至少一个激活中心。

[0126] 在本公开的一些实施例中,所述识别单元11,还用于确定每个激活中心在对应的一个图像帧中的位置;在所述位置位于所述图像帧中的任一检测框区域的情况下,将所述任一检测框区域对应的目标对象,确定为所述图像帧中的异常目标对象,对于所述至少一个激活中心,对应得到所述各个图像帧的异常目标对象。

[0127] 在本公开的一些实施例中,所述识别单元11,还用于根据所述至少一个激活中心,确定出目标激活中心;根据所述目标激活中心,识别出所述各个图像帧中的异常目标对象。

[0128] 在本公开的一些实施例中,所述识别单元11,还用于通过无监督聚类方式对所述至少一个激活中心进行聚类,得到目标激活中心;根据所述目标激活中心,识别出所述各个图像帧中的异常目标对象。

[0129] 在本公开的一些实施例中,所述识别单元11,还用于通过无监督聚类方式对所述至少一个激活中心进行聚类,得到包含的激活中心数量最多的中心簇;确定所述中心簇中包含的激活中心的平均值,并将所述平均值确定为所述目标激活中心。

[0130] 在本公开的一些实施例中,所述确定单元12,还用于生成所述每个图像帧的与所述识别结果对应的一组初始类激活映射值;对所述初始类激活映射值进行上采样处理,得到一组中间激活值;所述一组中间激活值用于生成与所述每个图像帧的重要区域对应的类激活图;将所述一组中间激活值与所述每个图像帧的原始像素值进行融合处理,得到所述一组激活值。

[0131] 在本公开的一些实施例中,所述行为识别装置1还包括检测单元13,所述检测单元13,用于在所述对获得的图像帧序列的至少一个图像帧中,各个图像帧的重要区域分类识别,得到与至少一个图像帧对应的识别结果之前,对所述图像帧序列中的当前图像帧进行目标检测,确定存在的至少一个目标对象对应的至少一个检测框区域;所述确定单元12,还用于确定所述至少一个检测框区域中的中心检测框区域;在已检测所述至少一个图像帧的情况下,基于所述至少一个图像帧对应的至少一个中心检测框区域,确定所述至少一个图像帧中各个图像帧的所述重要区域。

[0132] 在本公开的一些实施例中,所述确定单元12,还用于将所述至少一个检测框区域扩大第一预设比例,得到与所述至少一个检测框区域对应的至少一个第一区域;基于所述至少一个第一区域中任意两个第一区域之间的面积交并比,确定所述任意两个第一区域之间的匹配结果;所述匹配结果表征所述任意两个第一区域是否匹配;基于所述匹配结果,确定每个第一区域的匹配次数,并确定匹配次数最大的第一区域;根据所述匹配次数最大的

第一区域,确定密集中心,并将所述密集中心对应的检测框区域,确定为所述中心检测框区域。

[0133] 在本公开的一些实施例中,所述确定单元12,还用于将所述至少一个图像帧对应的至少一个中心检测框区域扩大第二预设比例,得到与所述至少一个检测框区域对应的至少一个第二区域;根据所述至少一个第二区域中各个第二区域在所属的图像帧中的位置,确定出概括区域位置;对于每个图像帧,将所述概括区域位置在所述每个图像帧中对应的区域,确定为所述每个图像帧的所述重要区域;所述每个图像帧的重要区域包含所述每个图像帧对应的第二区域。

[0134] 在本公开的一些实施例中,所述确定单元12,还用于在存在一个所述匹配次数最大的第一区域的情况下,将所述匹配次数最大的第一区域确定为所述密集中心;在存在两个或两个以上的所述匹配次数最大的第一区域的情况下,将所述两个或两个以上的所述匹配次数最大的第一区域中面积最大的第一区域,确定为所述密集中心。

[0135] 本公开实施例还提供一种电子设备,图14为本公开实施例提供的电子设备的结构示意图,如图14所示,电子设备2包括:存储器21和处理器22,其中,存储器21和处理器22通过总线23连接;存储器21,用于存储可执行计算机程序;处理器22,用于执行存储器21中存储的可执行计算机程序时,实现本公开实施例提供的方法,例如,本公开实施例提供的行为识别方法。

[0136] 本公开实施例提供一种计算机可读存储介质,存储有计算机程序,用于引起处理器22执行时,实现本公开实施例提供的方法,例如,本公开实施例提供的行为识别方法。

[0137] 在本公开的一些实施例中,存储介质可以是FRAM、ROM、PROM、EPROM、EEPROM、闪存、磁表面存储器、光盘、或CD-ROM等存储器;也可以是包括上述存储器之一或任意组合的各种设备。

[0138] 在本公开的一些实施例中,计算机程序可以采用软件、软件模块、脚本或代码的形式,按任意形式的编程语言(包括编译或解释语言,或者声明性或过程性语言)来编写,并且其可按任意形式部署,包括被部署为独立的程序或者被部署为模块、组件、子例程或者适合在计算环境中使用的其它单元。

[0139] 作为示例,计算机程序可以但不一定对应于文件系统中的文件,可以可被存储在保存其它程序或数据的文件的一部分,例如,存储在超文本标记语言(HTML,Hyper Text Markup Language)文档中的一个或多个脚本中,存储在专用于所讨论的程序的单个文件中,或者,存储在多个协同文件(例如,存储一个或多个模块、子程序或代码部分的文件)中。

[0140] 作为示例,计算机程序可被部署为在一个计算设备上执行,或者在位于一个地点的多个计算设备上执行,又或者,在分布在多个地点且通过通信网络互连的多个计算设备上执行。

[0141] 综上所述,采用本技术实现方案,一方面,由于是对确定出的图像帧的重要区域进行分类识别,所以,相比于根据每个图像帧的全图的数据增强进行识别而言,有效缩小了识别时的搜索范围,同时降低了对识别时的干扰,从而提高了识别准确度;另一方面,由于是通过确定每个图像帧的与识别结果对应的类激活图中的异常位置,来定位每个图像帧的异常中心(激活中心),并根据得到的异常中心识别出对应的图像帧中的异常目标对象,并且,与识别结果对应的类激活图能够表征出图像帧中与识别结果相关的目标对象,所以,使得

识别出的每个图像帧的异常目标对象更准确,从而提高了对异常目标对象的识别准确度;因此,本公开实施例提供的行为识别方法能够提高识别准确度。

[0142] 以上所述,仅为本公开的实施例而已,并非用于限定本公开的保护范围。凡在本公开的精神和范围之内所作的任何修改、等同替换和改进等,均包含在本公开的保护范围之内。

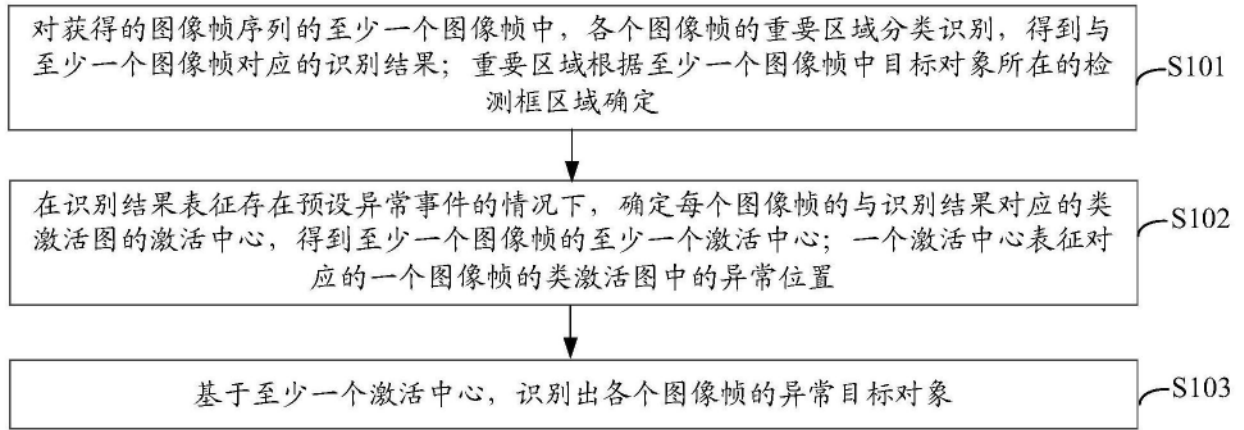


图1

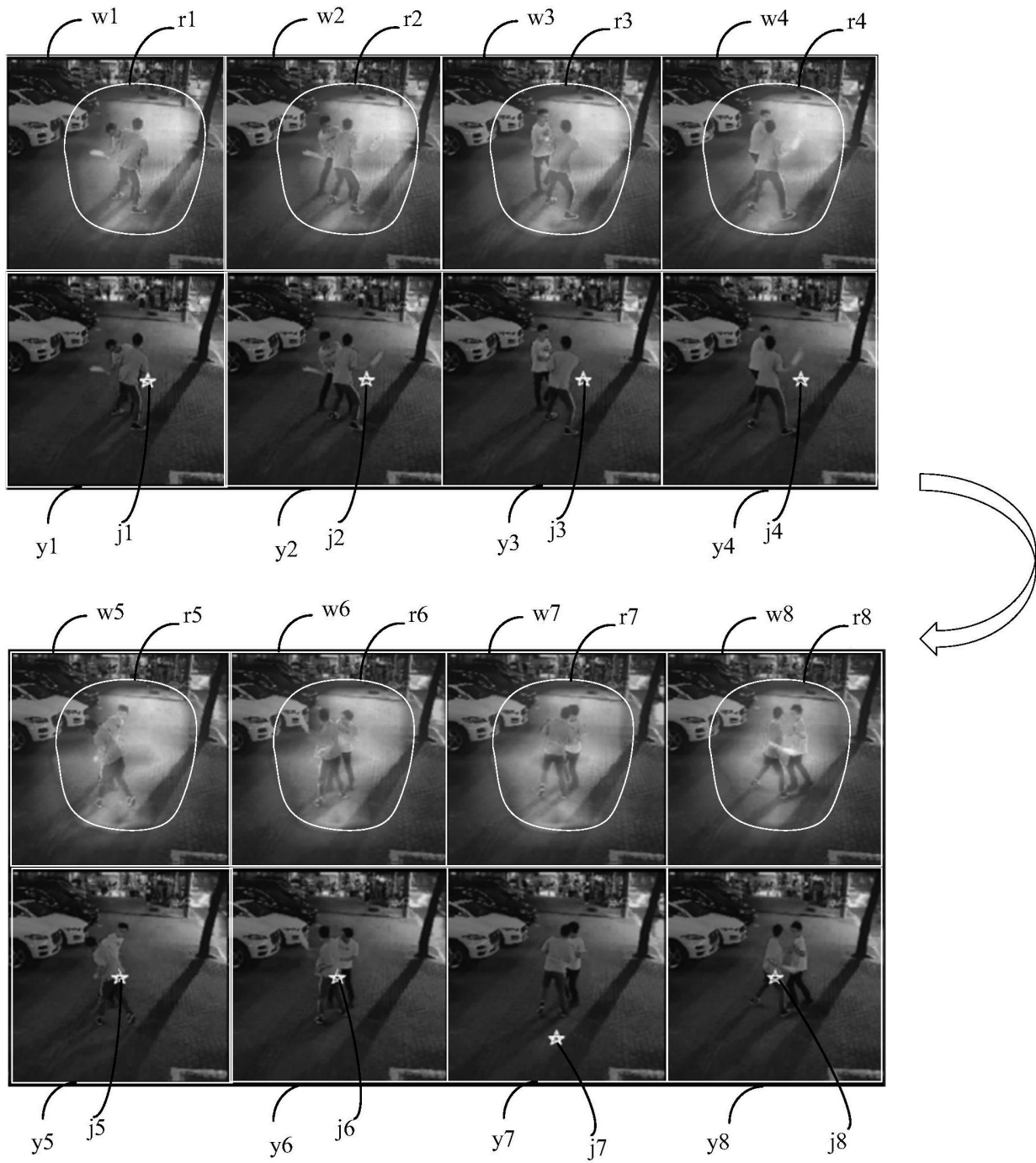


图2



图像帧a

图3

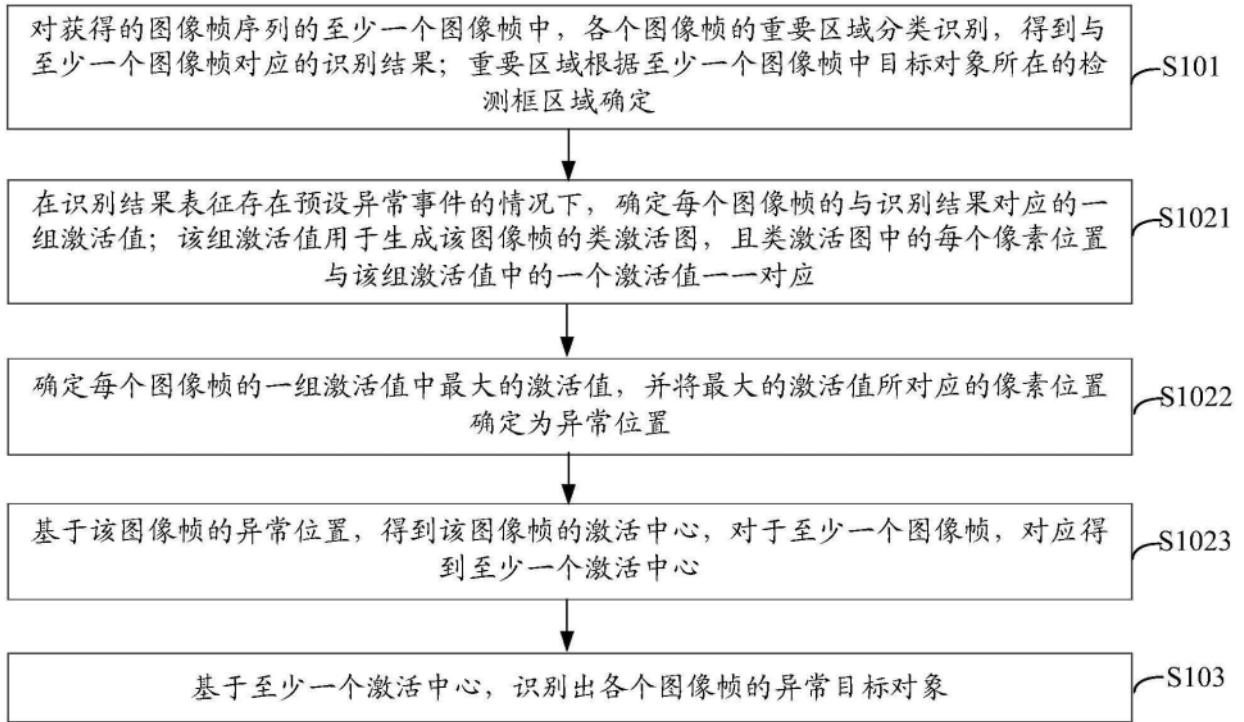


图4

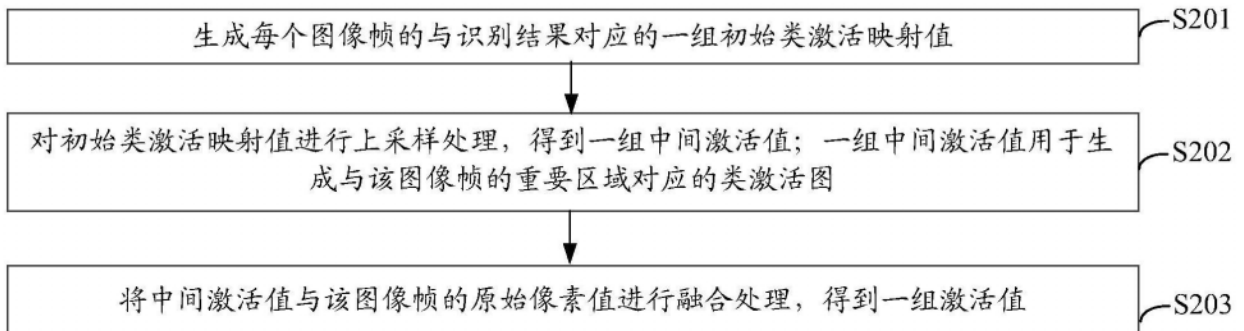


图5

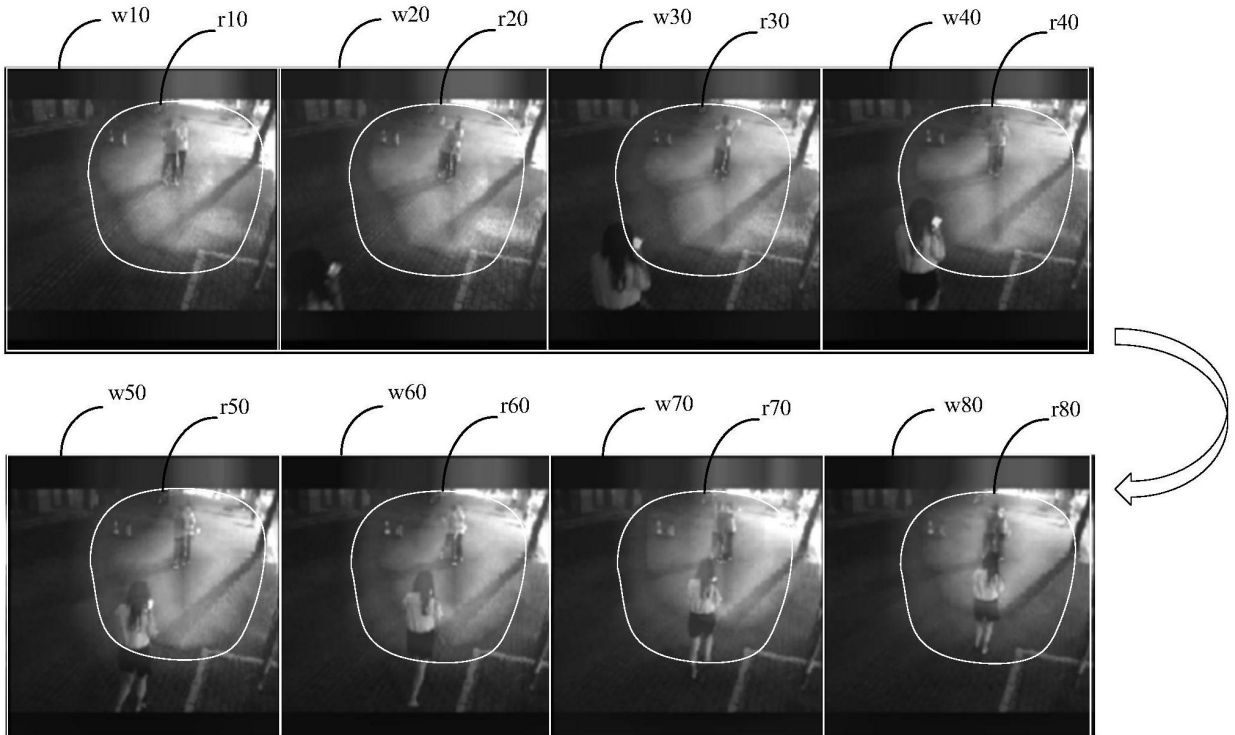


图6

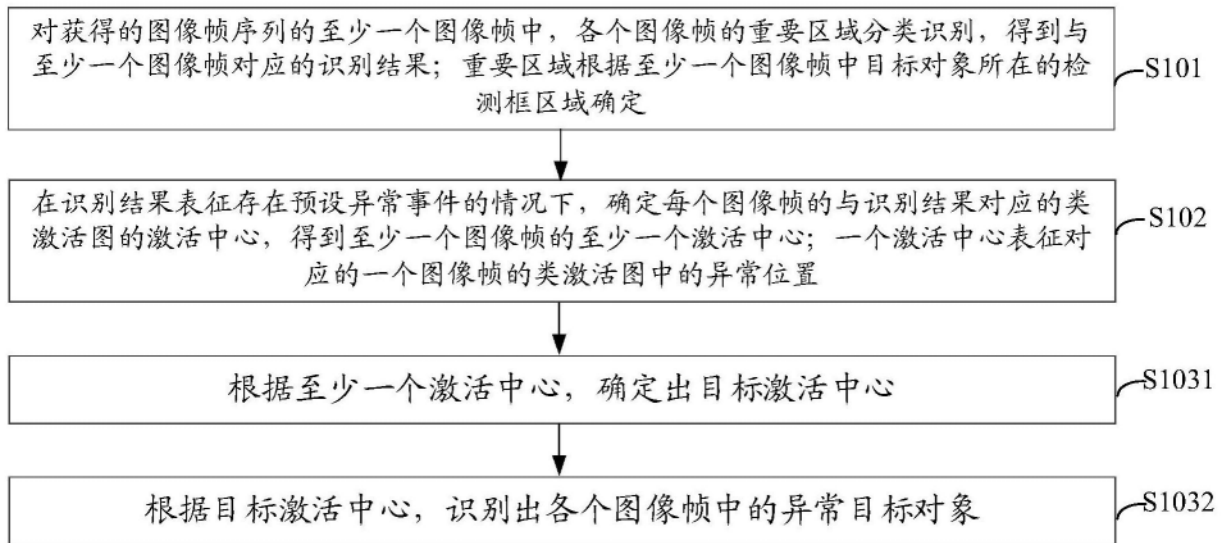


图7A

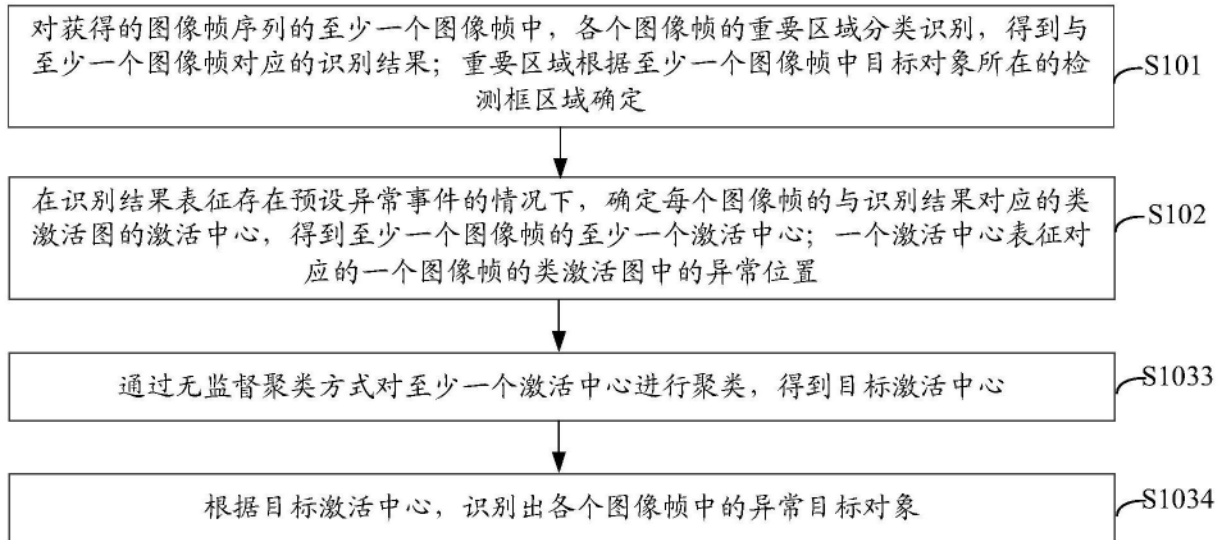


图7B

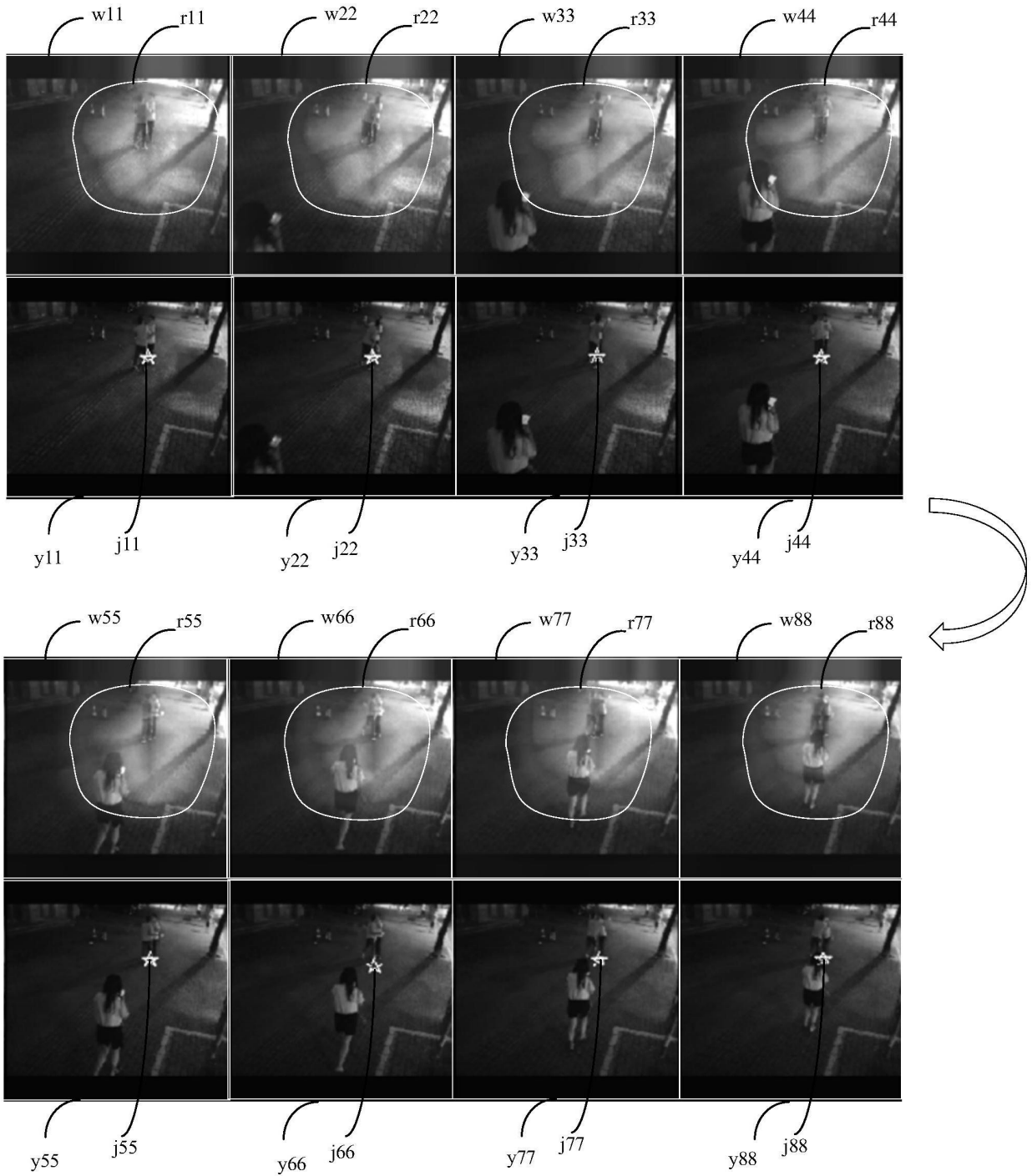


图8

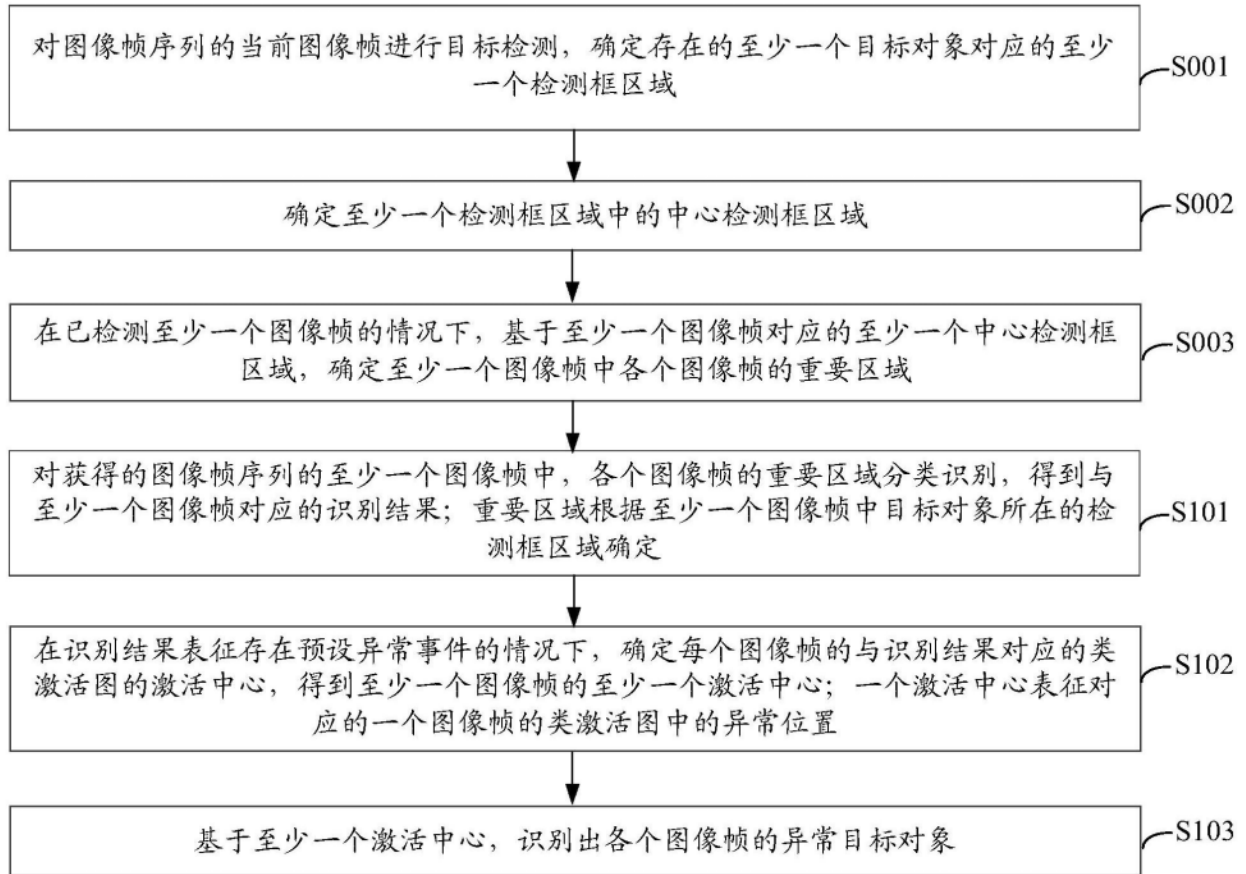


图9

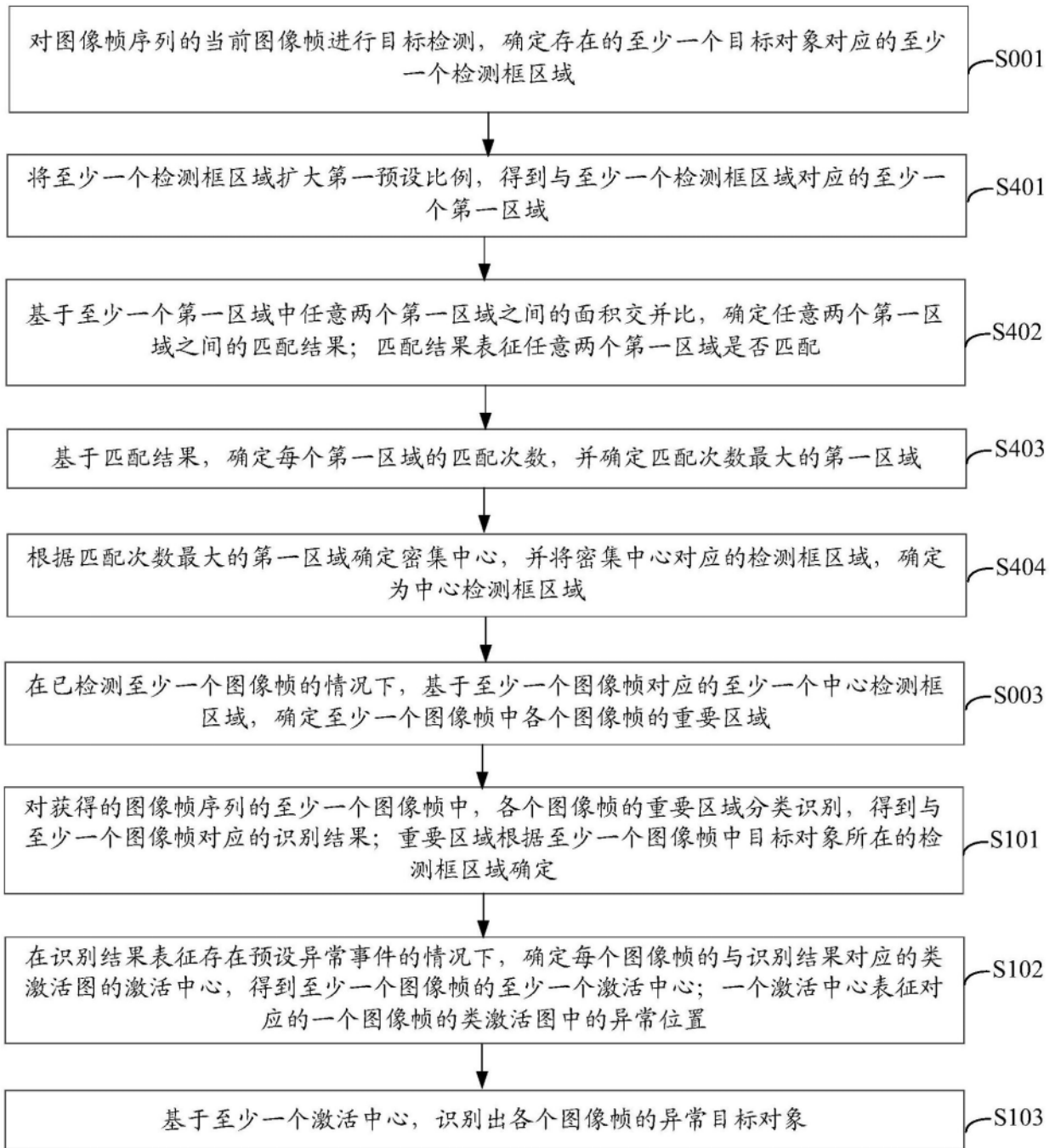


图10

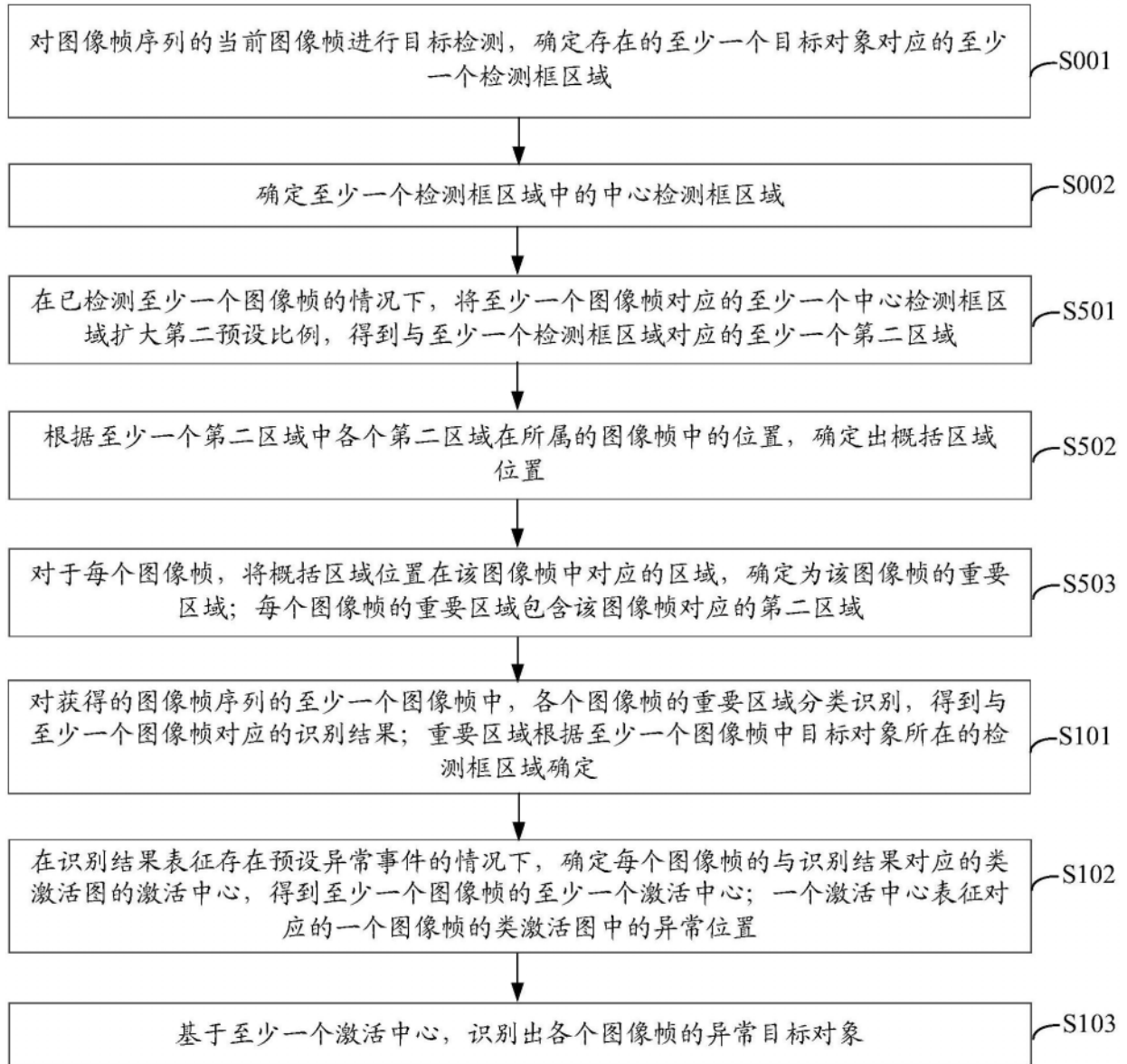


图11

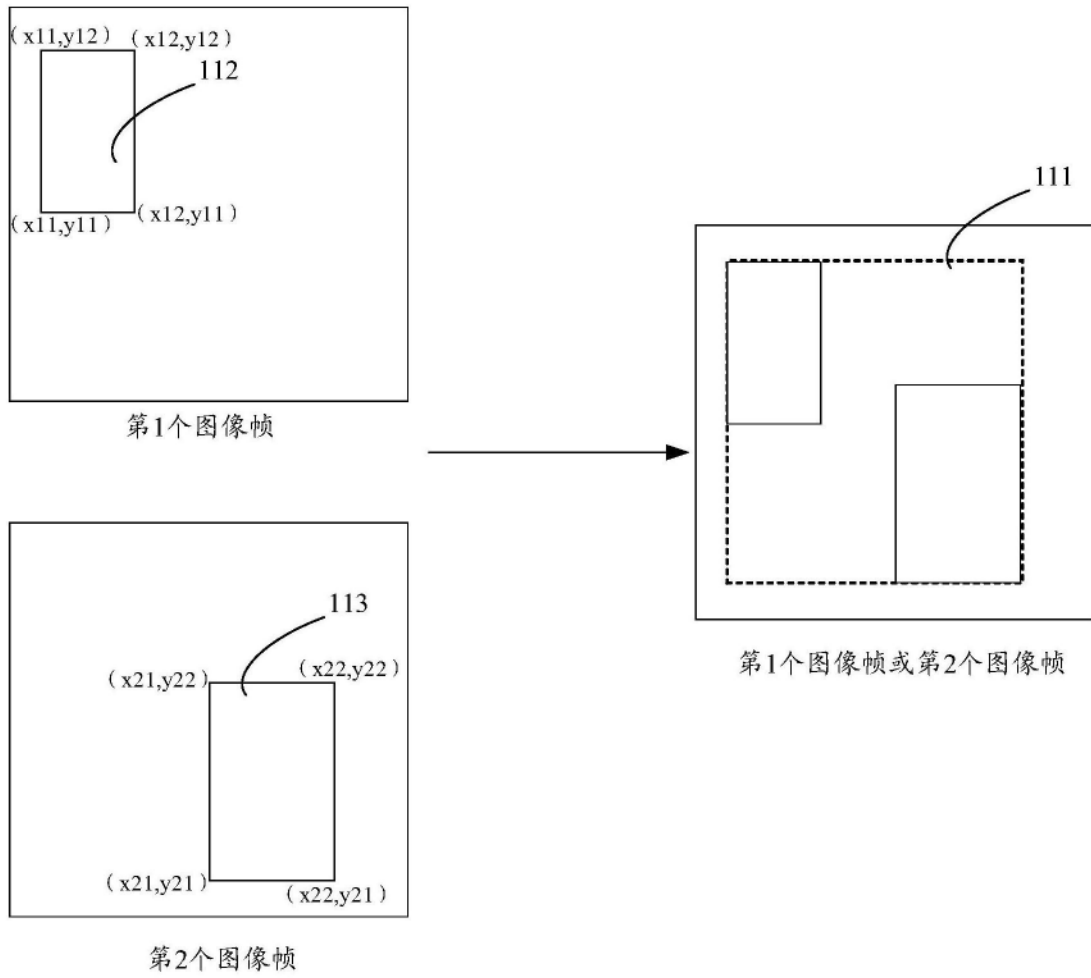


图12

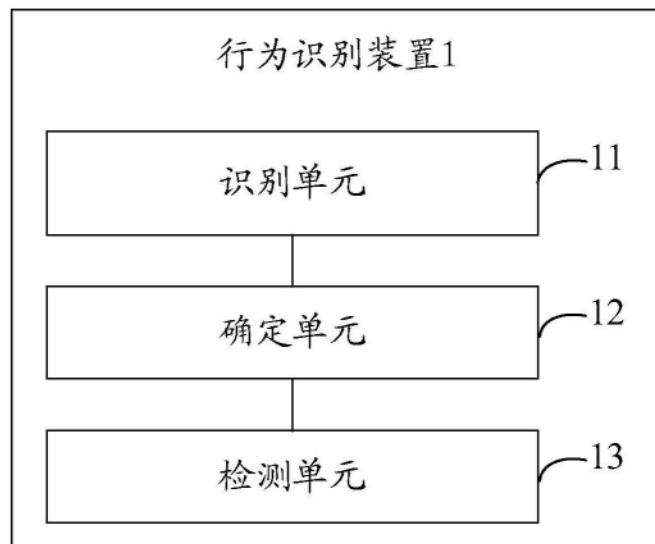


图13

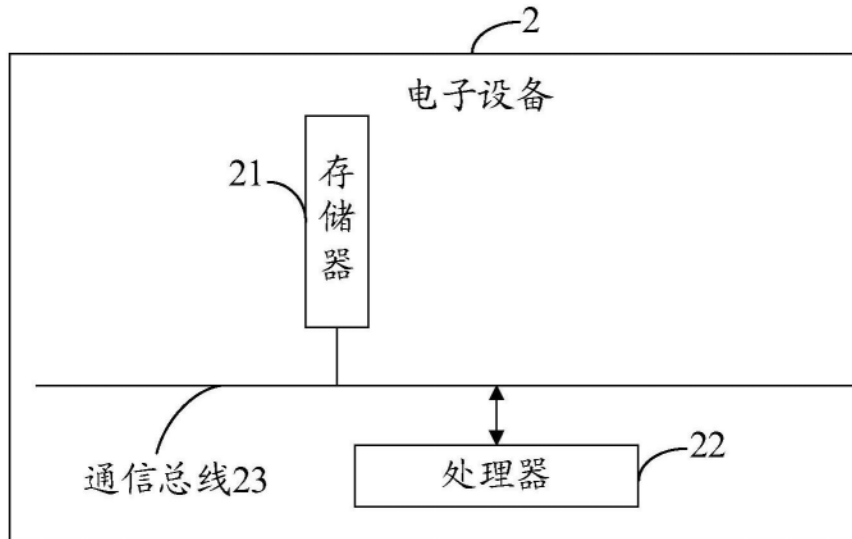


图14