



(12)发明专利申请

(10)申请公布号 CN 113569600 A

(43)申请公布日 2021. 10. 29

(21)申请号 202010359731.0

(22)申请日 2020.04.29

(71)申请人 深圳云天励飞技术有限公司  
地址 518000 广东省深圳市龙岗区园山街  
道龙岗大道8288号深圳大运软件小镇  
17栋1楼

(72)发明人 曾卓熙 胡文泽 王孝宇

(74)专利代理机构 广州三环专利商标代理有限  
公司 44202

代理人 熊永强

(51)Int.Cl.  
G06K 9/00(2006.01)  
G06K 9/32(2006.01)  
G06K 9/62(2006.01)  
G06N 3/04(2006.01)

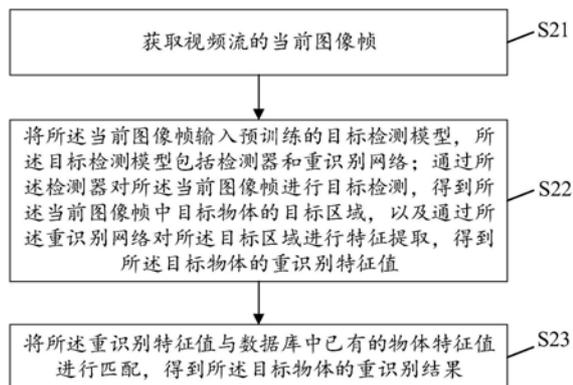
权利要求书2页 说明书11页 附图6页

(54)发明名称

物体重识别的方法、装置、电子设备及存储  
介质

(57)摘要

本申请提供了一种物体重识别的方法、装  
置、电子设备及存储介质,其中,该物体重识别的  
方法包括:获取视频流的当前图像帧;将当前图  
像帧输入预训练的目标检测模型,目标检测模型  
包括检测器和重识别网络;通过所述检测器对所  
述当前图像帧进行目标检测,得到所述当前图像  
帧中目标物体的目标区域,以及通过所述重识别  
网络对所述目标区域进行特征提取,得到所述目  
标物体的重识别特征值;将所述重识别特征值与  
数据库中已有的物体特征值进行匹配,得到所述  
目标物体的重识别结果。采用本申请实施例有利  
于提高视频流中物体重识别的效率。



1. 一种物体重识别的方法,其特征在于,所述方法包括:

获取视频流的当前图像帧;

将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;

将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

2. 根据权利要求1所述的方法,其特征在于,所述检测器为基于锚点的检测器;所述对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,包括:

通过所述检测器对所述当前图像帧进行特征提取,得到目标特征图;

基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,将所述目标检测框所框定的区域确定为所述目标区域。

3. 根据权利要求2所述的方法,其特征在于,所述基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,包括:

根据所述目标特征图的尺寸将所述当前图像帧划分为多个栅格;

若所述目标物体的中心落入所述栅格内,则利用所述栅格预设的所述锚框对所述目标物体进行预测,得到所述目标检测框。

4. 根据权利要求2所述的方法,其特征在于,所述锚框的预测值包括所述锚框的中心点坐标偏移值,所述锚框的宽、高的缩放值,所述得到所述目标检测框,包括:

根据所述中心点坐标偏移值计算得到所述目标检测框的中心点坐标;

根据所述宽、高的缩放值计算得到所述目标检测框的宽和高;

由所述目标检测框的中心点坐标、所述目标检测框的宽和高确定所述目标检测框的位置。

5. 根据权利要求1-4任一项所述的方法,其特征在于,所述目标检测模型的训练过程,包括:

在所述检测器原有的训练损失上加上所述重识别网络的训练损失,得到相加后的训练损失;

将由多个物体构成的训练数据集输入所述目标检测模型进行迭代,根据所述相加后的训练损失的值调整所述目标检测模型的网络参数,并拟合所述重识别网络的输出;所述多个物体中包括所述目标物体。

6. 一种物体重识别的装置,其特征在于,所述装置包括:

图像帧获取模块,用于获取视频流的当前图像帧;

目标检测和重识别模块,用于将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;

重识别特征值匹配模块,用于将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

7. 根据权利要求6所述的装置, 其特征在于, 在对所述当前图像帧进行目标检测, 得到所述当前图像帧中目标物体的目标区域方面, 所述目标检测和重识别模块, 具体用于:

通过所述检测器对所述当前图像帧进行特征提取, 得到目标特征图;

基于所述目标特征图及预设的锚框对所述目标物体进行预测, 得到所述目标物体的目标检测框, 将所述目标检测框所框定的区域确定为所述目标区域。

8. 根据权利要求7所述的装置, 其特征在于, 在基于所述目标特征图及预设的锚框对所述目标物体进行预测, 得到所述目标物体的目标检测框方面, 所述目标检测和重识别模块, 具体用于:

根据所述目标特征图的尺寸将所述当前图像帧划分为多个栅格;

若所述目标物体的中心落入所述栅格内, 则利用所述栅格预设的所述锚框对所述目标物体进行预测, 得到所述目标检测框。

9. 一种电子设备, 包括输入设备和输出设备, 其特征在于, 还包括:

处理器, 适于实现一条或多条指令; 以及,

计算机存储介质, 所述计算机存储介质存储有一条或多条指令, 所述一条或多条指令适于由所述处理器加载并执行如权利要求1-5任一项所述的方法。

10. 一种计算机存储介质, 其特征在于, 所述计算机存储介质存储有一条或多条指令, 所述一条或多条指令适于由处理器加载并执行如权利要求1-5任一项所述的方法。

## 物体重识别的方法、装置、电子设备及存储介质

### 技术领域

[0001] 本申请涉及机器学习技术领域,尤其涉及一种物体重识别的方法、装置、电子设备及存储介质。

### 背景技术

[0002] 随着计算机视觉技术的发展,视频、图像处理领域有了更多的研究方向,而物体重识别就是其中的热门之一。视频流中,物体重识别就是将帧与帧之间可能相同的物体识别出来,在智能视频监控、自动驾驶、目标跟踪等方面有着不可或缺的作用。现有物体重识别通常采用检测器对当前视频图像帧中的物体进行识别,对于识别出的物体用单独的特征提取网络进行特征提取,然后与数据库中已有的特征进行匹配,整个过程需要大量资源,物体重识别效率较低。

### 发明内容

[0003] 针对上述问题,本申请提供了一种物体重识别的方法、装置、电子设备及存储介质,有利于提高对视频流中的物体进行重识别的效率。

[0004] 为实现上述目的,本申请实施例第一方面提供了一种物体重识别的方法,该方法包括:

[0005] 获取视频流的当前图像帧;

[0006] 将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;

[0007] 将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

[0008] 结合第一方面,在一种可能的实施方式中,所述检测器为基于锚点的检测器;所述对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,包括:

[0009] 通过所述检测器对所述当前图像帧进行特征提取,得到目标特征图;

[0010] 基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,将所述目标检测框所框定的区域确定为所述目标区域。

[0011] 结合第一方面,在一种可能的实施方式中,所述基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,包括:

[0012] 根据所述目标特征图的尺寸将所述当前图像帧划分为多个栅格;

[0013] 若所述目标物体的中心落入所述栅格内,则利用所述栅格预设的所述锚框对所述目标物体进行预测,得到所述目标检测框。

[0014] 结合第一方面,在一种可能的实施方式中,所述锚框的预测值包括所述锚框的中心点坐标偏移值,所述锚框的宽、高的缩放值,所述得到所述目标检测框,包括:

- [0015] 根据所述中心点坐标偏移值计算得到所述目标检测框的中心点坐标；
- [0016] 根据所述宽、高的缩放值计算得到所述目标检测框的宽和高；
- [0017] 由所述目标检测框的中心点坐标、所述目标检测框的宽和高确定所述目标检测框的位置。
- [0018] 结合第一方面,在一种可能的实施方式中,所述目标检测模型的训练过程,包括:
- [0019] 在所述检测器原有的训练损失上加上所述重识别网络的训练损失,得到相加后的训练损失;
- [0020] 将由多个物体构成的训练数据集输入所述目标检测模型进行迭代,根据所述相加后的训练损失的值调整所述目标检测模型的网络参数,并拟合所述重识别网络的输出;所述多个物体中包括所述目标物体。
- [0021] 本申请实施例第二方面提供了一种物体重识别的装置,该装置包括:
- [0022] 图像帧获取模块,用于获取视频流的当前图像帧;
- [0023] 目标检测和重识别模块,用于将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;
- [0024] 重识别特征值匹配模块,用于将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。
- [0025] 本申请实施例第三方面提供了一种电子设备,该电子设备包括输入设备和输出设备,还包括处理器,适于实现一条或多条指令;以及,计算机存储介质,所述计算机存储介质存储有一条或多条指令,所述一条或多条指令适于由所述处理器加载并执行如下步骤:
- [0026] 获取视频流的当前图像帧;
- [0027] 将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;
- [0028] 将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。
- [0029] 本申请实施例第四方面提供了一种计算机存储介质,所述计算机存储介质存储有一条或多条指令,所述一条或多条指令适于由处理器加载并执行如下步骤:
- [0030] 获取视频流的当前图像帧;
- [0031] 将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;
- [0032] 将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。
- [0033] 本申请的上述方案至少包括以下有益效果:与现有技术相比,本申请实施例通过获取视频流的当前图像帧;然后将当前图像帧输入预训练的目标检测模型,所述目标检测

模型包括检测器和重识别网络；通过所述检测器对所述当前图像帧进行目标检测，得到所述当前图像帧中目标物体的目标区域，以及通过所述重识别网络对所述目标区域进行特征提取，得到所述目标物体的重识别特征值；最后将所述重识别特征值与数据库中已有的物体特征值进行匹配，得到所述目标物体的重识别结果。这样将用于重识别特征值提取的重识别网络并与检测器构成目标检测模型，目标检测任务与重识别特征值提取任务在同一个网络模型中进行，在原有检测器输出结果的基础上，还同时输出重识别特征值，而不需要对当前图像帧中的每个目标物体单独使用特征提取网络进行重识别特征值的提取，从而有利于提高目标物体重识别的效率。

### 附图说明

[0034] 为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述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申请的一些实施例，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。

[0035] 图1为本申请实施例提供的一种应用环境的示意图；

[0036] 图2为本申请实施例提供的一种物体重识别的方法的流程示意图；

[0037] 图3为本申请实施例提供的另一种物体重识别的方法的流程示意图；

[0038] 图4为本申请实施例提供的一种目标检测模型的结构示意图；

[0039] 图5为本申请实施例提供的一种目标检测模型输入到输出的示例图；

[0040] 图6为本申请实施例提供的一种物体重识别的装置的结构示意图；

[0041] 图7为本申请实施例提供的另一种物体重识别的装置的结构示意图；

[0042] 图8为本申请实施例提供的一种电子设备的结构示意图。

### 具体实施方式

[0043] 为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施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申请一部分的实施例，而不是全部的实施例。基于本申请中的实施例，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他实施例，都应当属于本申请保护的范围。

[0044] 本申请说明书、权利要求书和附图中出现的术语“包括”和“具有”以及它们任何变形，意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元，而是可选地还包括没有列出的步骤或单元，或可选地还包括对于这些过程、方法、产品或设备固有的其它步骤或单元。此外，术语“第一”、“第二”和“第三”等是用于区别不同的对象，而并非用于描述特定的顺序。

[0045] 本申请实施例提出一种对视频流中的物体进行重识别的方案，以提高物体重识别的效率，相比现有单独使用训练数据集训练特征提取网络、将目标检测任务与重识别特征值提取任务分开进行的方案，我们将重识别网络加入到目标检测器中，训练目标检测器时，在目标检测器训练损失的基础上加上重识别网络的训练损失，用蒸馏的方式去训练重识别网络，实现一个模型同时端到端输出目标检测结果和物体的重识别特征值，减少了超参数，

且重识别网络与目标检测器共享100%的计算,整体计算复杂度并未增加。

[0046] 具体的,首先结合相关附图来举例介绍下本申请实施例的方案可能的应用环境。请参见图1,图1为本申请实施例提供的一种应用环境的示意图,如图1所示,包括视频采集区域和视频处理中心两大部分,在一些情况下,视频处理中心可以是理解为视频监控中心,其中,视频处理中心采用VCN(Video Cloud Node,视频云节点)视频管理服务,具备实时视频查看、视频转发、视频回放等功能。视频处理中心以服务器为执行主体,用户可通过终端与服务器进行交互,例如:通过终端在服务器上部署加入有重识别网络的目标检测模型,服务器对视频文件的处理结果可在终端进行显示。视频处理中心中,数据库可以是本地数据库(例如:服务器中的数据库),也可以是独立于服务器的数据库(例如:云端数据库、第三方数据库),可用于结构化存储视频文件、物体的样本图像、重识别特征值等,当然,其存储的数据还可供服务器或终端调用。视频采集区域为图像采集设备所覆盖的区域,可以是工业园区、街道、小区入口等区域,其内布置的图像采集设备可将采集的视频传输至服务器,由服务器执行一系列重识别处理,在一些情况下,图像采集设备采集的视频也可直接传输至数据库进行存储。视频采集区域和视频处理中心各部分之间的通信网络包括但不限于虚拟专用网络、局域网络、城域网络,本申请提出的物体重识别方案可以在图1所示的应用环境下实施,当然,图1仅仅是一种示例,并不对本申请实施例造成限定。

[0047] 基于图1所示的应用环境,以下结合其他附图对本申请实施例提供的物体重识别的方法进行详细阐述。

[0048] 请参加图2,图2为本申请实施例提供的一种物体重识别的方法的流程示意图,该物体重识别的方法由服务器执行,如图2所示,包括步骤S21-S23:

[0049] S21,获取视频流的当前图像帧;

[0050] 本申请具体实施例中,视频流可以是视频采集区域的图像采集设备上传的视频,也可以是服务器通过接口获取的数据库中或其他平台的视频,当前图像帧即当前时间点传输并显示的图像帧。

[0051] S22,将当前图像帧输入预训练的目标检测模型,目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;

[0052] 本申请具体实施例中,当前图像帧中的目标物体可以是行人、车辆等任何物体,对目标检测模型的检测器并不作限定,基于anchor(锚点或锚框)的检测器均可执行对当前图像帧进行目标检测的任务,例如:Faster R-CNN(Faster Region-Convolutional Neural Networks,更快速的候选区域卷积神经网络检测器)、YOLO(You Only Look Once,一瞥目标检测器)-V2、YOLO-V3、SSD(Single Shot MultiBox Detector,单发多盒检测器)等等。对于选定的目标检测器,其整体网络结构基本保持不变,在输出层中加入重识别网络,然后采用训练数据集进行迭代训练,使得到的目标检测模型在原本输出结果的基础上,还能多输出当前图像帧中目标物体的重识别特征值。

[0053] 对于获取到的当前图像帧,将其输入目标检测模型的检测器由全卷积网络进行特征提取,不同的目标检测模型对提取出的特征图会进行不同的处理,例如:Faster R-CNN中特征图的一个分支进入RPN(Region Proposal Network,区域生成网络)生成锚框与边界框

回归的偏移值,然后计算出当前图像帧中目标对象的候选目标检测框,而ROI Pooling (Region of interest Pooling,感兴趣区域池化)层从特征图中提取出目标特征图,将其输入全连接和softmax网络进行目标物体的分类和边界框精细回归。而YOLOV3对于提取出的特征图分别进行32倍、16倍和8倍下采样得到三种不同尺寸的特征图,三种不同尺寸的特征图分别采用不同尺寸的锚框来生成偏移值,以计算目标对象的目标检测框,采用logistic预测目标对象的类别。以上仅是用两种目标检测模型进行说明,虽然不同的目标检测模型的处理方式不同,但最终都将得到当前图像帧中所有目标对象的目标检测框,目标检测框框定的区域即重识别网络进行特征提取的目标区域,重识别网络对目标区域进行卷积操作输出目标物体的重识别特征值,而不需要将不同目标对象的目标区域输入单独的特征提取网络进行重识别特征值的提取,例如:假设当前图像帧中有8个目标物体,现有的做法是使用检测器进行目标检测,然后对每个目标物体单独进行重识别特征提取,需要跑8次特征提取网络,而本申请只需跑一个加入有重识别网络的检测器,即可实现目标检测和重识别特征提取。

[0054] S23,将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

[0055] 本申请具体实施例中,数据库中已有的物体特征值即当前图像帧的上一图像帧或之前的图像帧中出现的物体的特征值,将步骤S22中提取出的重识别特征值与其进行匹配,可以采用矩阵运算实施,快速找到匹配度较高的已有物体的特征值,并可在终端的显示界面突出显示该目标物体,例如:可对上一图像帧和当前图像帧中均出现的行人进行标定,在未找到匹配度较高的已有物体特征值的情况下,说明当前图像帧中的目标物体是第一次出现,可直接将目标物体的重识别特征值存储于数据库中,或者在一些的实施例中,可向终端返回未找到相同目标物体的提示。可选的,将重识别特征值与数据库中已有的物体特征值进行匹配还可通过计算余弦相似度、欧氏距离、搬土距离等进行实施。

[0056] 可以看出,本申请实施例通过获取视频流的当前图像帧;然后将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;最后将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。这样将用于重识别特征值提取的重识别网络并与检测器构成目标检测模型,目标检测任务与重识别特征值提取任务在同一个网络模型中进行,在原有检测器输出结果的基础上,还同时输出重识别特征值,而不需要对当前图像帧中的每个目标物体单独使用特征提取网络进行重识别特征值的提取,从而有利于提高目标物体重识别的效率。

[0057] 请参见图3,图3为本申请实施例提供的另一种物体重识别的方法的流程示意图,同样可基于图1所示的应用环境实施,如图3所示,包括步骤S31-S35:

[0058] S31,获取视频流的当前图像帧;

[0059] S32,将所述当前图像帧输入预训练的目标检测模型的检测器进行特征提取,得到目标特征图;所述检测器为基于锚点的检测器;

[0060] S33,基于所述目标特征图及预设的锚框对所述当前图像帧中的目标物体进行预测,得到所述目标物体的目标检测框,将所述目标检测框所框定的区域确定为目标区域;

[0061] S34,通过所述目标检测模型的重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;

[0062] 本申请具体实施例中,以基于anchor的YOLO-V3检测器为目标检测模型的检测器进行说明,如图4所示,以视频流的当前图像帧为输入图像,采用53个卷积层进行基本特征图的提取,在79层后经过几个卷积层进行32倍下采样得到13\*13的目标特征图,然后对其进行上采样,将上采样得到的特征图与第62层特征图进行融合得到第91层的特征图,对第91层的特征图进行16倍下采样得到26\*26的目标特征图,再对第91层的特征图进行上采样,将本次上采样得到的特征图与第36层的特征图进行融合,最后对其进行8倍下采样得到52\*52的目标特征图。

[0063] 在一种可能的实施方式中,上述基于所述目标特征图及预设的锚框对所述当前图像帧中的目标物体进行预测,得到所述目标物体的目标检测框,包括:

[0064] 根据目标特征图的尺寸将输入的当前图像帧划分为多个栅格。例如:使用13\*13的目标特征图进行目标物体定位时,则将当前图像帧划分为13\*13的栅格,使用26\*26的目标特征图进行目标物体定位时,则将当前图像帧划分为26\*26的栅格,52\*52的目标特征图亦是如此。

[0065] 若目标物体的中心落入某个栅格内,则利用该栅格预设的锚框对目标物体进行预测,得到目标物体的目标检测框。YOLO-V3中对于每个尺度的目标特征图应用了三种不同尺寸的锚框(或者称为先验框),划分出的每个栅格均预设了三个锚框,假设行人A的中心正好落在13\*13的最中间的栅格,则由最中间的这个栅格预测出锚框的中心点坐标偏移值 $t_x$ 、 $t_y$ ,锚框的宽、高的缩放值 $t_w$ 、 $t_h$ ,根据中心点坐标偏移值 $t_x$ 、 $t_y$ ,采用公式: $b_x = \sigma t_x + c_x$ , $b_y = \sigma t_y + c_y$ ,计算得到目标检测框的中心点坐标( $b_x$ , $b_y$ ),其中, $c_x$ 和 $c_y$ 为目标特征图中栅格的左上角坐标,根据宽、高的缩放值 $t_w$ 、 $t_h$ ,采用公式: $b_w = p_w e^{t_w}$ , $b_h = p_h e^{t_h}$ ,计算得到目标检测框的宽 $b_w$ 和高 $b_h$ ,其中, $p_w$ 和 $p_h$ 是锚框映射到目标特征图上的宽和高,由 $b_x$ , $b_y$ , $b_w$ 和 $b_h$ 便能确定目标检测框的位置,目标检测框所框定的区域即目标对象所在的目标区域,采用重识别网络对该目标区域进行卷积,在YOLO-V3原有输出结果的基础上,还多输出512维的重识别特征值。如图5所示,对于输入的416\*416\*3的当前图像帧,经过全卷积网络的特征提取得到三个尺度的目标特征图,YOLO-V3原来在三个目标特征图上进行预测的输出为:13\*13\*3\*(定位坐标(4)+置信度(1)+物体识别概率(C)),26\*26\*3\*(定位坐标(4)+置信度(1)+物体识别概率(C)),52\*52\*3\*(定位坐标(4)+置信度(1)+物体识别概率(C)),由于现在输出层并入了重识别网络(卷积层)对目标区域进行了重识别特征值的提取,那么现在的输出即为:13\*13\*3\*(定位坐标(4)+置信度(1)+物体识别概率(C)+重识别特征值(512)),26\*26\*3\*(定位坐标(4)+置信度(1)+物体识别概率(C)+重识别特征值(512)),52\*52\*3\*(定位坐标(4)+置信度(1)+物体识别概率(C)+重识别特征值(512))。

[0066] S35,将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

[0067] 其中,图3所示的部分步骤在图2所示的实施例中已有相关说明,且能达到相同或相似的有益效果,为避免重复,此处不再赘述。

[0068] 在一种可能的实施方式中,上述目标检测模型的训练过程,包括:

[0069] 在所述检测器原有的训练损失上加上所述重识别网络的训练损失,得到相加后的

训练损失；

[0070] 将由多个物体构成的训练数据集输入所述目标检测模型进行迭代，根据所述相加后的训练损失的值调整所述目标检测模型的网络参数，并拟合所述重识别网络的输出；所述多个物体中包括所述目标物体。

[0071] 本申请具体实施例中，训练数据集为多个物体的图像集，可以采用梯度下降法对训练数据集进行迭代，也可以采用投影的方法对训练数据集进行迭代（迭代投影算法），每次迭代输出时，观察相加后得到的训练损失的值，根据该值调整目标检测模型的网络参数（例如：隐藏层神经元个数等）。优选地，此处采用蒸馏的方式，将训练数据集输入目标检测模型进行迭代，用回归损失函数Huber Loss拟合重识别网络的输出。

[0072] 该实施方式中，将重识别网络并入检测器构成目标检测模型，采用蒸馏的方式去训练重识别网络，在检测器原有的训练损失上加上重识别网络的训练损失即可，并不需要采用复杂的损失函数来单独训练用于重识别特征值提取的网络，能够节省大量显存资源。

[0073] 可以看出，本申请实施例通过获取视频流的当前图像帧；将当前图像帧输入预训练的目标检测模型的检测器进行特征提取，得到目标特征图；基于目标特征图及预设的锚框对当前图像帧中的目标物体进行预测，得到目标物体的目标检测框，将目标检测框所框定的区域确定为目标区域；通过目标检测模型的重识别网络对目标区域进行特征提取，得到目标物体的重识别特征值；最后将重识别特征值与数据库中已有的物体特征值进行匹配，得到目标物体的重识别结果。这样将用于重识别特征值提取的重识别网络并与检测器构成目标检测模型，目标检测任务与重识别特征值提取任务在同一个网络模型中进行，在原有检测器输出结果的基础上，还同时输出重识别特征值，而不需要对当前图像帧中的每个目标物体单独使用特征提取网络进行重识别特征值的提取，从而有利于提高目标物体重识别的效率。

[0074] 请参见图6，图6为本申请实施例提供的一种物体重识别的装置的结构示意图，如图6所示，该装置包括：

[0075] 图像帧获取模块61，用于获取视频流的当前图像帧；

[0076] 目标检测和重识别模块62，用于将所述当前图像帧输入预训练的目标检测模型，所述目标检测模型包括检测器和重识别网络；通过所述检测器对所述当前图像帧进行目标检测，得到所述当前图像帧中目标物体的目标区域，以及通过所述重识别网络对所述目标区域进行特征提取，得到所述目标物体的重识别特征值；

[0077] 重识别特征值匹配模块63，用于将所述重识别特征值与数据库中已有的物体特征值进行匹配，得到所述目标物体的重识别结果。

[0078] 在一种可能的实施方式中，在对所述当前图像帧进行目标检测，得到所述当前图像帧中目标物体的目标区域方面，所述目标检测和重识别模块62，具体用于：

[0079] 通过所述检测器对所述当前图像帧进行特征提取，得到目标特征图；

[0080] 基于所述目标特征图及预设的锚框对所述目标物体进行预测，得到所述目标物体的目标检测框，将所述目标检测框所框定的区域确定为所述目标区域。

[0081] 在一种可能的实施方式中，在基于所述目标特征图及预设的锚框对所述目标物体进行预测，得到所述目标物体的目标检测框方面，所述目标检测和重识别模块62，具体用于：

[0082] 根据所述目标特征图的尺寸将所述当前图像帧划分为多个栅格；

[0083] 若所述目标物体的中心落入所述栅格内，则利用所述栅格预设的所述锚框对所述目标物体进行预测，得到所述目标检测框。

[0084] 在一种可能的实施方式中，所述锚框的预测值包括所述锚框的中心点坐标偏移值，所述锚框的宽、高的缩放值，在得到所述目标检测框方面，所述目标检测和重识别模块62，具体用于：

[0085] 根据所述中心点坐标偏移值计算得到所述目标检测框的中心点坐标；

[0086] 根据所述宽、高的缩放值计算得到所述目标检测框的宽和高；

[0087] 由所述目标检测框的中心点坐标、所述目标检测框的宽和高确定所述目标检测框的位置。

[0088] 在一种可能的实施方式中，如图7所示，所述装置还包括网络训练模块64，在所述目标检测模型的训练方面，网络训练模块64，具体用于：

[0089] 在所述检测器原有的训练损失上加上所述重识别网络的训练损失，得到相加后的训练损失；

[0090] 将由多个物体构成的训练数据集输入所述目标检测模型进行迭代，根据所述相加后的训练损失的值调整所述目标检测模型的网络参数，并拟合所述重识别网络的输出；所述多个物体中包括所述目标物体。

[0091] 根据本申请的一个实施例，图6和图7所示的物体重识别的装置中的各个单元可以分别或全部合并为一个或若干个另外的单元来构成，或者其中的某个(些)单元还可以再拆分为功能上更小的多个单元来构成，这可以实现同样的操作，而不影响本申请的实施例的技术效果的实现。上述单元是基于逻辑功能划分的，在实际应用中，一个单元的功能也可以由多个单元来实现，或者多个单元的功能由一个单元实现。在本申请的其它实施例中，基于物体重识别的装置也可以包括其它单元，在实际应用中，这些功能也可以由其它单元协助实现，并且可以由多个单元协作实现。需要说明的是，本申请实施例提供的物体重识别的装置能够应用在目标检测、物体重识别、目标跟踪等场景中。

[0092] 根据本申请的另一个实施例，可以通过在包括中央处理单元(CPU)、随机存取存储介质(RAM)、只读存储介质(ROM)等处理元件和存储元件的例如计算机的通用计算设备上运行能够执行如图2或图3中所示的相应方法所涉及的各步骤的计算机程序(包括程序代码)，来构造如图6或图7中所示的物体重识别的装置设备，以及来实现本申请实施例的物体重识别的方法。所述计算机程序可以记载于例如计算机可读记录介质上，并通过计算机可读记录介质装载于上述计算设备中，并在其中运行。

[0093] 基于上述方法实施例和装置实施例的描述，本申请实施例还提供一种电子设备。请参见图8，该电子设备至少包括处理器81、输入设备82、输出设备83以及计算机存储介质84。其中，电子设备内的处理器81、输入设备82、输出设备83以及计算机存储介质84可通过总线或其他方式连接。

[0094] 计算机存储介质84可以存储在电子设备的存储器中，所述计算机存储介质84用于存储计算机程序，所述计算机程序包括程序指令，所述处理器81用于执行所述计算机存储介质84存储的程序指令。处理器81(或称CPU(Central Processing Unit, 中央处理器))是电子设备的计算核心以及控制核心，其适于实现一条或多条指令，具体适于加载并执行一

条或多条指令从而实现相应方法流程或相应功能。

[0095] 在一个实施例中,本申请实施例提供的电子设备的处理器81可以用于进行一系物体重识别的处理:获取视频流的当前图像帧;将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

[0096] 在一种实施方式中,所述检测器为基于锚点的检测器;处理器81执行所述对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,包括:

[0097] 通过所述检测器对所述当前图像帧进行特征提取,得到目标特征图;

[0098] 基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,将所述目标检测框所框定的区域确定为所述目标区域。

[0099] 再一种实施方式中,处理器81用于执行所述基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,包括:

[0100] 根据所述目标特征图的尺寸将所述当前图像帧划分为多个栅格;

[0101] 若所述目标物体的中心落入所述栅格内,则利用所述栅格预设的所述锚框对所述目标物体进行预测,得到所述目标检测框。

[0102] 再一种实施方式中,所述锚框的预测值包括所述锚框的中心点坐标偏移值,所述锚框的宽、高的缩放值,处理器81执行所述得到所述目标检测框,包括:

[0103] 根据所述中心点坐标偏移值计算得到所述目标检测框的中心点坐标;

[0104] 根据所述宽、高的缩放值计算得到所述目标检测框的宽和高;

[0105] 由所述目标检测框的中心点坐标、所述目标检测框的宽和高确定所述目标检测框的位置。

[0106] 再一种实施方式中,处理器81执行所述目标检测模型的训练过程,包括:

[0107] 在所述检测器原有的训练损失上加上所述重识别网络的训练损失,得到相加后的训练损失;

[0108] 将由多个物体构成的训练数据集输入所述目标检测模型进行迭代,根据所述相加后的训练损失的值调整所述目标检测模型的网络参数,并拟合所述重识别网络的输出;所述多个物体中包括所述目标物体。

[0109] 示例性的,上述电子设备可以是计算机、服务器、电脑主机等设备,电子设备可包括但不限于处理器81、输入设备82、输出设备83以及计算机存储介质84。本领域技术人员可以理解,所述示意图仅仅是电子设备的示例,并不构成对电子设备的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件。

[0110] 需要说明的是,由于电子设备的处理器81执行计算机程序时实现上述的物体重识别的方法中的步骤,因此上述物体重识别的方法的实施例均适用于该电子设备,且均能达到相同或相似的有益效果。

[0111] 本申请实施例还提供了一种计算机存储介质(Memory),所述计算机存储介质是电子设备中的记忆设备,用于存放程序和数据。可以理解的是,此处的计算机存储介质既可以包括终端中的内置存储介质,当然也可以包括终端所支持的扩展存储介质。计算机存储介

质提供存储空间,该存储空间存储了终端的操作系统。并且,在该存储空间中还存放了适于被处理器81加载并执行的一条或多条的指令,这些指令可以是一个或一个以上的计算机程序(包括程序代码)。需要说明的是,此处的计算机存储介质可以是高速RAM存储器,也可以是非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器;可选的,还可以是至少一个位于远离前述处理器81的计算机存储介质。在一个实施例中,可由处理器81加载并执行计算机存储介质中存放的一条或多条指令,以实现上述有关物体重识别的方法的相应步骤;具体实现中,计算机存储介质中的一条或多条指令由处理器81加载并执行如下步骤:

[0112] 获取视频流的当前图像帧;

[0113] 将所述当前图像帧输入预训练的目标检测模型,所述目标检测模型包括检测器和重识别网络;通过所述检测器对所述当前图像帧进行目标检测,得到所述当前图像帧中目标物体的目标区域,以及通过所述重识别网络对所述目标区域进行特征提取,得到所述目标物体的重识别特征值;

[0114] 将所述重识别特征值与数据库中已有的物体特征值进行匹配,得到所述目标物体的重识别结果。

[0115] 在一种示例中,计算机存储介质中的一条或多条指令由处理器81加载时还执行如下步骤:

[0116] 通过所述检测器对所述当前图像帧进行特征提取,得到目标特征图;

[0117] 基于所述目标特征图及预设的锚框对所述目标物体进行预测,得到所述目标物体的目标检测框,将所述目标检测框所框定的区域确定为所述目标区域。

[0118] 在一种示例中,计算机存储介质中的一条或多条指令由处理器81加载时还执行如下步骤:

[0119] 根据所述目标特征图的尺寸将所述当前图像帧划分为多个栅格;

[0120] 若所述目标物体的中心落入所述栅格内,则利用所述栅格预设的所述锚框对所述目标物体进行预测,得到所述目标检测框。

[0121] 在一种示例中,计算机存储介质中的一条或多条指令由处理器81加载时还执行如下步骤:

[0122] 根据锚框中心点坐标偏移值计算得到所述目标检测框的中心点坐标;

[0123] 根据锚框宽、高的缩放值计算得到所述目标检测框的宽和高;

[0124] 由所述目标检测框的中心点坐标、所述目标检测框的宽和高确定所述目标检测框的位置。

[0125] 在一种示例中,计算机存储介质中的一条或多条指令由处理器81加载时还执行如下步骤:

[0126] 在所述检测器原有的训练损失上加上所述重识别网络的训练损失,得到相加后的训练损失;

[0127] 将由多个物体构成的训练数据集输入所述目标检测模型进行迭代,根据所述相加后的训练损失的值调整所述目标检测模型的网络参数,并拟合所述重识别网络的输出;所述多个物体中包括所述目标物体。

[0128] 示例性的,计算机存储介质的计算机程序包括计算机程序代码,所述计算机程序

代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器 (ROM, Read-Only Memory)、随机存取存储器 (RAM, Random Access Memory)、电载波信号、电信信号以及软件分发介质等。

[0129] 需要说明的是，由于计算机存储介质的计算机程序被处理器执行时实现上述的物体识别的方法中的步骤，因此上述物体识别的方法的所有实施例均适用于该计算机存储介质，且均能达到相同或相似的有益效果。

[0130] 以上对本申请实施例进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

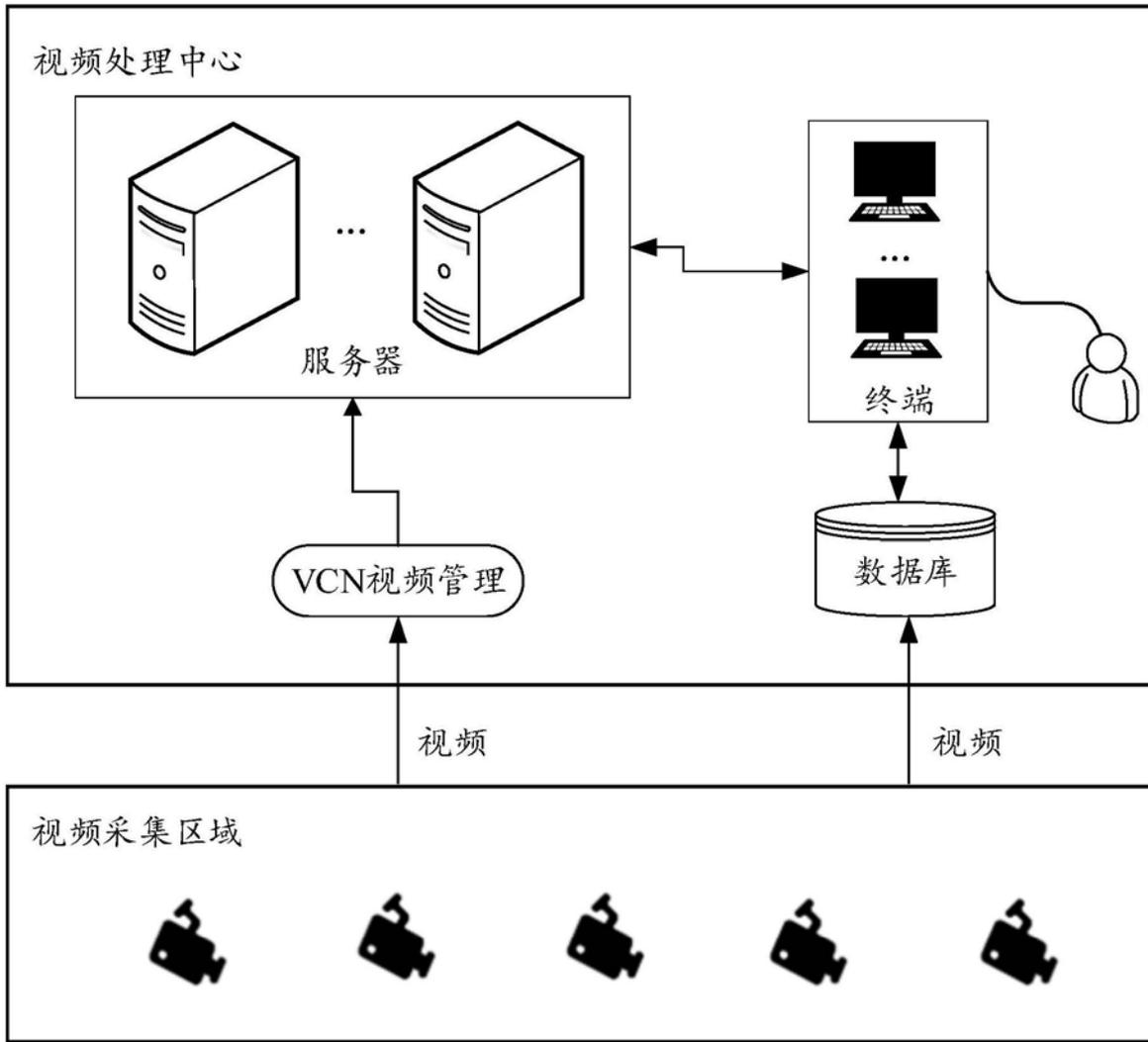


图1

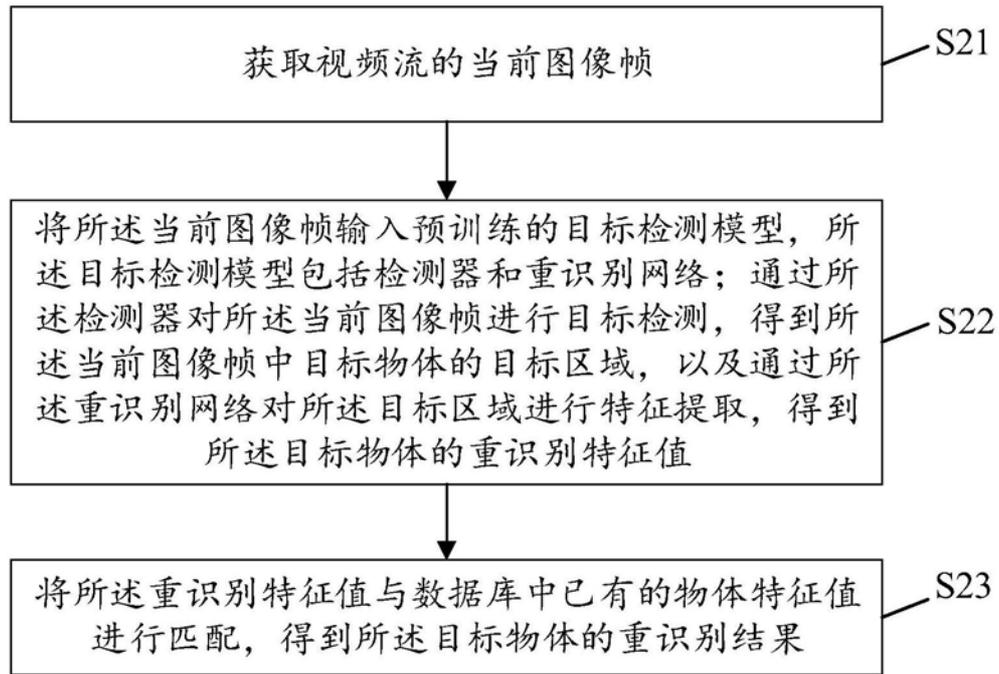


图2

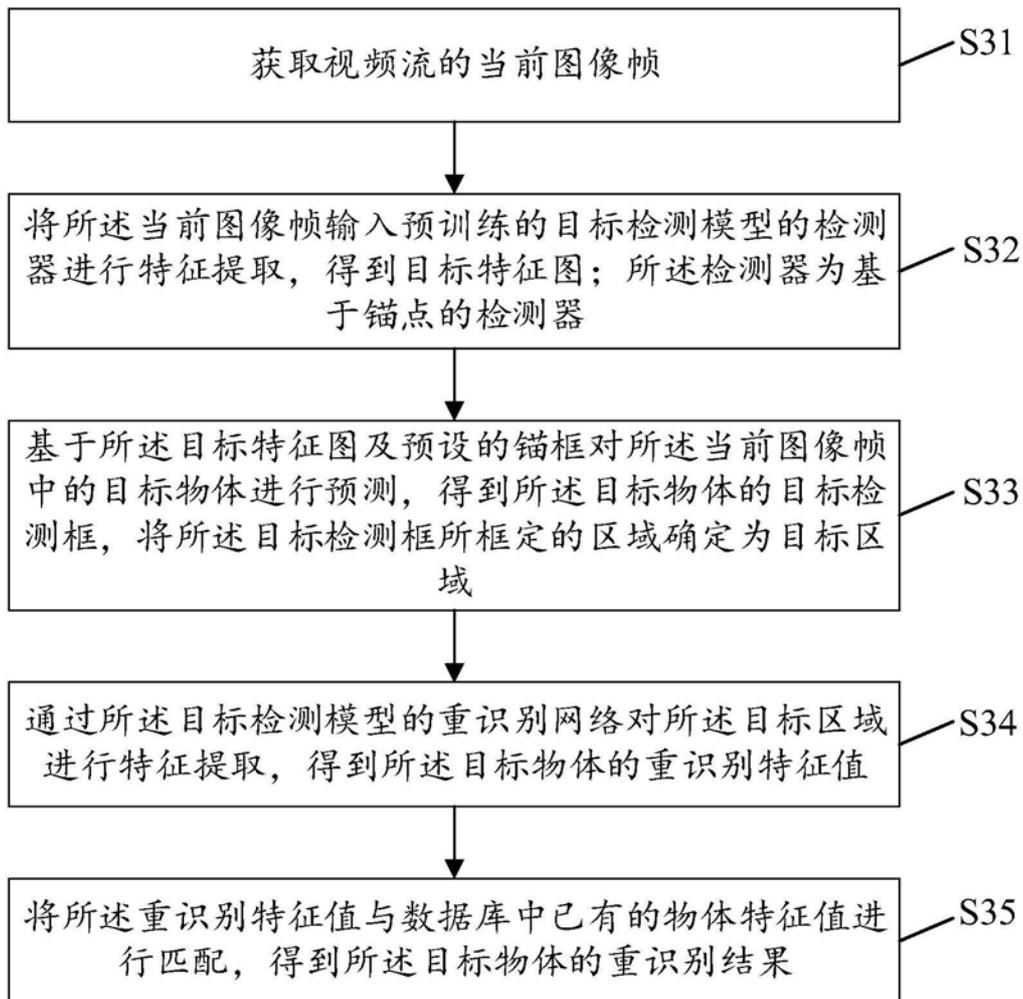


图3

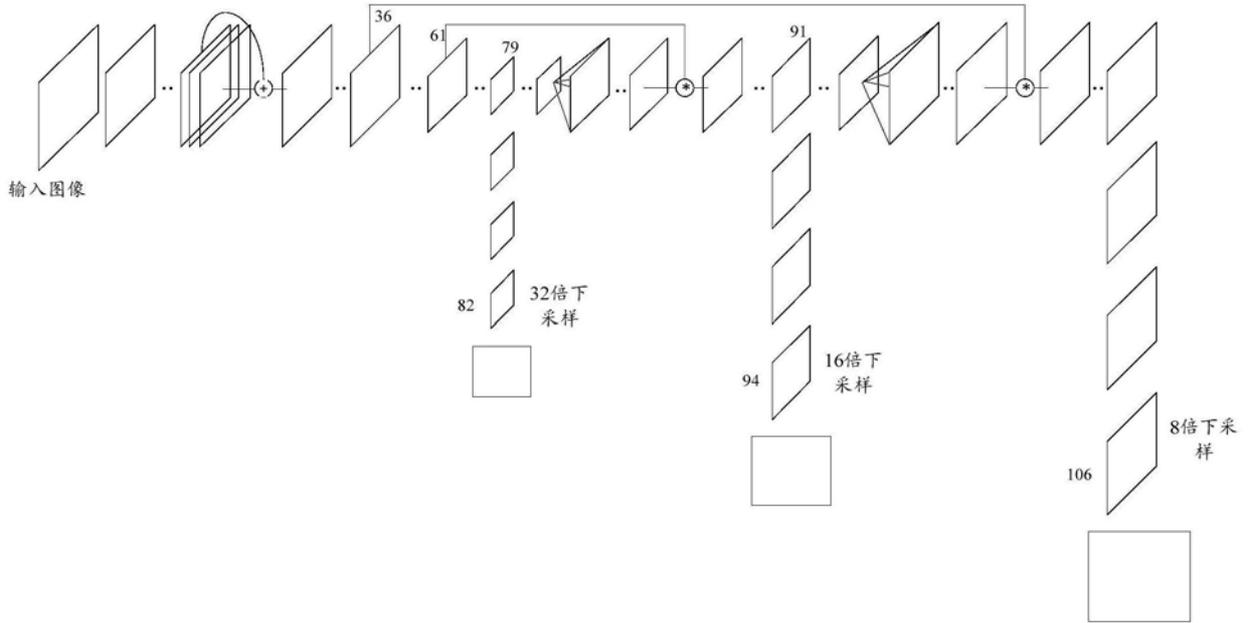


图4

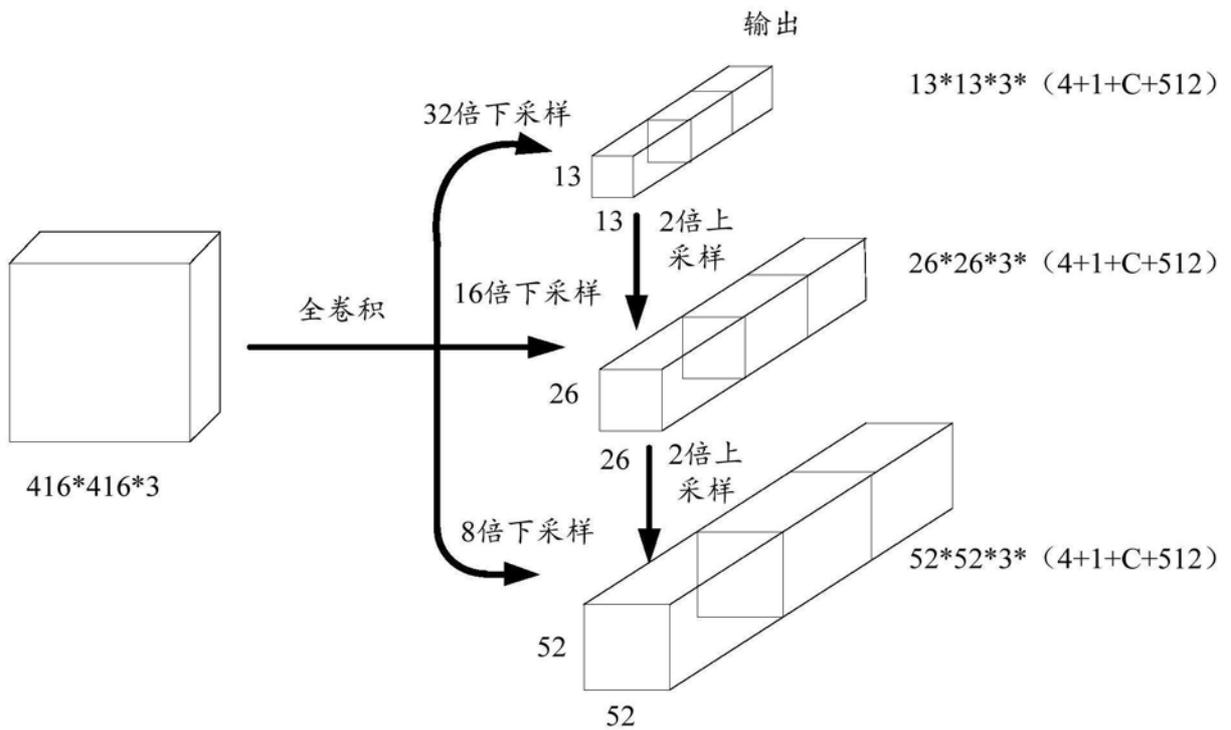


图5

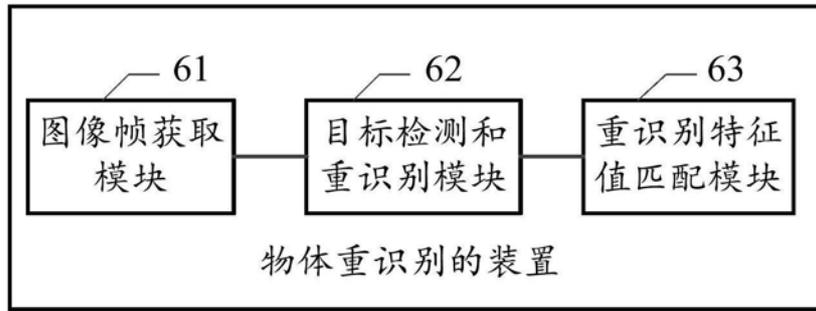


图6

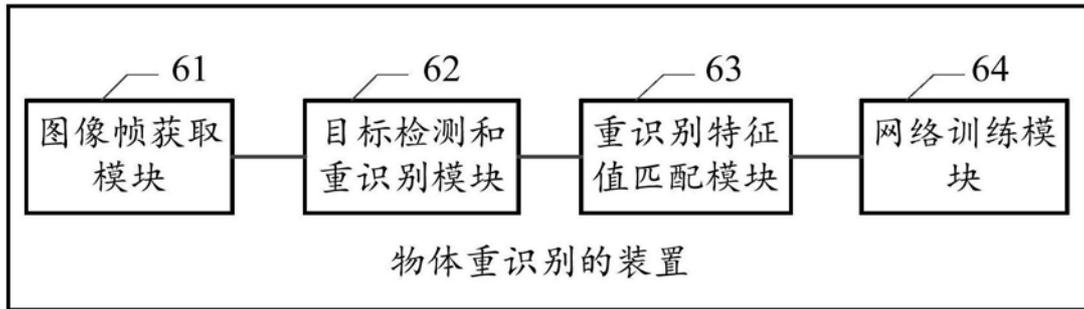


图7

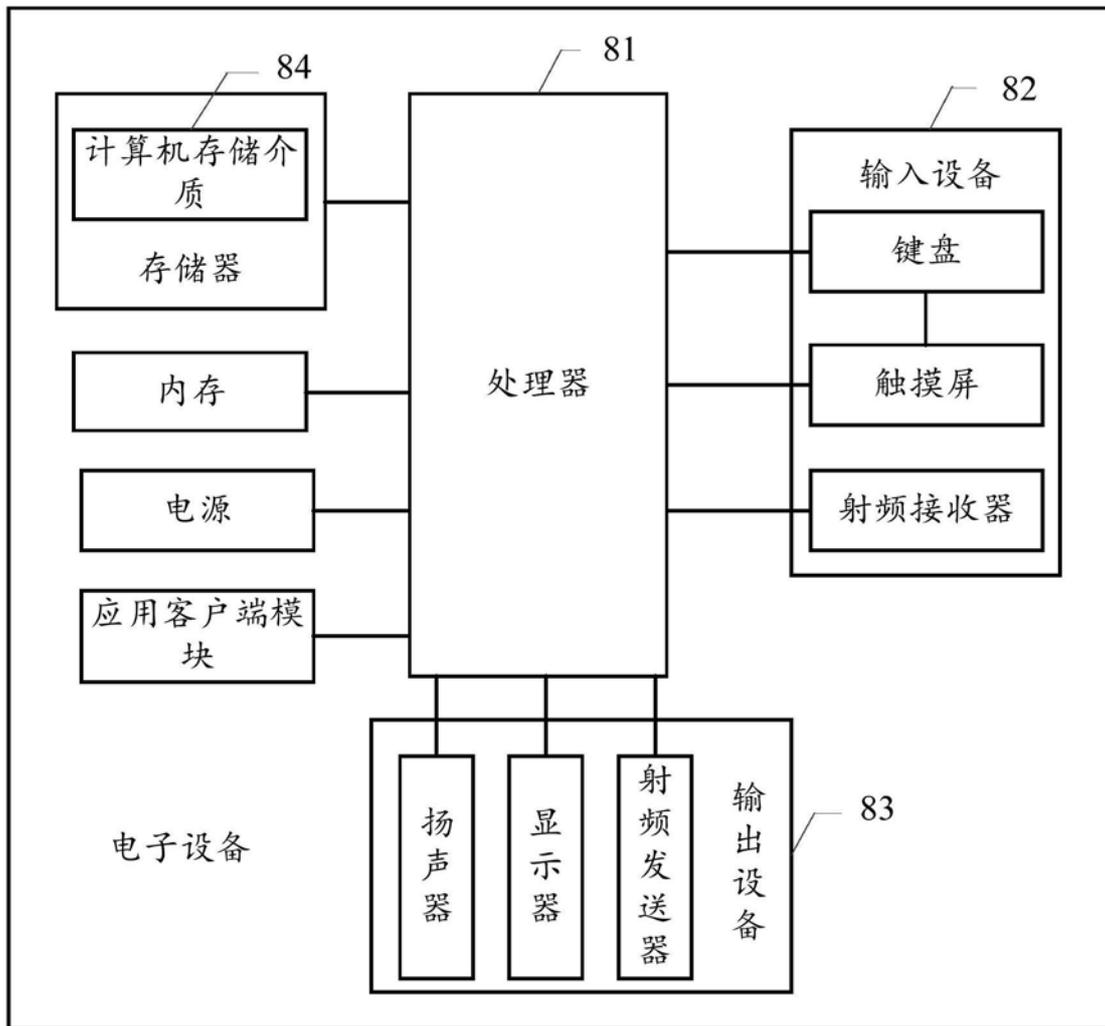


图8