



(12)发明专利申请

(10)申请公布号 CN 111104965 A
(43)申请公布日 2020.05.05

(21)申请号 201911167501.8

(22)申请日 2019.11.25

(71)申请人 河北科技大学

地址 050000 河北省石家庄市裕华区裕翔街26号

(72)发明人 侯卫民 罗欣 许文波 苏佳
贾海涛 王萌 王泉 张建
李卓韬

(74)专利代理机构 石家庄国为知识产权事务所
13120
代理人 付晓娣

(51)Int.Cl.
G06K 9/62(2006.01)

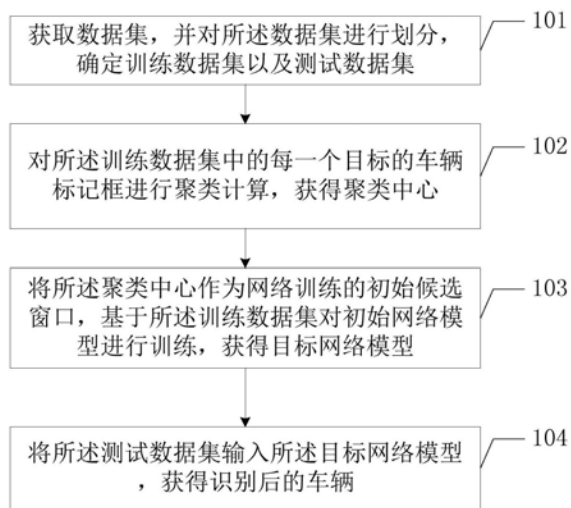
权利要求书2页 说明书11页 附图3页

(54)发明名称

车辆目标识别的方法及装置

(57)摘要

本发明适用于地球视觉目标识别技术领域，提供了一种车辆目标识别的方法及装置，该方法包括：获取数据集，并对所述数据集进行划分，确定训练数据集以及测试数据集；对所述训练数据集中的每一个目标的车辆标记框进行聚类计算，获得聚类中心；将所述聚类中心作为网络训练的初始候选窗口，基于所述训练数据集对初始网络模型进行训练，获得目标网络模型；将所述测试数据集输入所述目标网络模型，获得识别后的车辆，从而可以提高车辆目标的识别效果。



1. 一种车辆目标识别的方法,其特征在于,包括:
获取数据集,并对所述数据集进行划分,确定训练数据集以及测试数据集;
对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心;
将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型;
将所述测试数据集输入所述目标网络模型,获得识别后的车辆。
2. 如权利要求1所述的车辆目标识别的方法,其特征在于,所述获取数据集,包括:
分别获取不同角度拍摄的不同像素的图片,并将所有图片作为数据集。
3. 如权利要求1所述的车辆目标识别的方法,其特征在于,所述对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心,包括:
采用改进的K-means算法对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得预设数量的聚类中心。
4. 如权利要求3所述的车辆目标识别的方法,其特征在于,所述改进的K-means算法聚类采用的距离公式为
$$d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid});$$

其中, $d(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框的距离, $\text{IOU}(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框之间的重叠度的值。
5. 如权利要求4所述的车辆目标识别的方法,其特征在于,所述将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型,包括:
设置所述聚类中心作为网络训练的初始候选窗口,并将所述训练数据集输入初始网络模型;
对所述训练数据集中的所有图像进行特征提取,分别获得各个图像对应的预设像素的特征图;
将所述各个图像根据所述预设像素划分为对应数量的大小均匀的网格;
根据每个网格内的特征点以及对应的初始候选窗口,确定每个目标待匹配的初始候选窗口;
对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测,获得目标标记框,输出各个目标标记框;
根据输出所述各个目标标记框采用的模型参数更新所述初始网络模型,获得更新后的网络模型,并确定多次更新后的网络模型中测试平均精度值最高的网络模型为目标网络模型。
6. 如权利要求5所述的车辆目标识别的方法,其特征在于,所述根据每个网格内的特征点以及对应的初始候选窗口,确定每个目标待匹配的初始候选窗口,包括:
根据每个网格内的特征点以及对应的初始候选窗口,计算各个特征点所在的预测框与对应的实际标记的预测框之间的重叠度值;
确定最大的重叠度值对应的预测框为当前目标对应的待匹配的初始候选窗口。
7. 如权利要求6所述的车辆目标识别的方法,其特征在于,所述对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测,获得目标标记框,包括:

根据所述每个网格内的目标对应的所述待匹配的初始候选窗口,采用损失函数和衰减函数对每个网格内的目标的真实标记框进行目标边界框预测,获得每个目标对应的目标边界框;

对每个目标对应的目标边界框进行去重处理,获得优化后的每个目标对应的目标标记框。

8.如权利要求7所述的车辆目标识别的方法,其特征在于,所述损失函数为

$$Loss = \frac{1}{2} \sum_{i=1}^{10647} \lambda_{obj} * \left((2 - truth_w * truth_h) * \sum_{r \in (x,y,w,h)} (truth_r - predict_r)^2 + \sum_{r=0}^{k-1} ((r == truth_{class}) ? 1 : 0 - predict_{class_r})^2 \right) + (truth_{conf} - predict_{conf})^2 ;$$

其中, λ_{obj} 表示一个特征点存在于一个初始候选窗口内时的取值,Loss表示损失值, $truth_w$ 表示真实标记框的宽, $truth_h$ 表示真实标记框的高, $(truth_r - predict_r)^2$ 表示实际与预测之间真实标记框的差异, $truth_{class}$ 表示真实类别, $predict_{class_r}$ 表示预测的类别, $(truth_{conf} - predict_{conf})^2$ 表示置信度损失;

所述衰减函数为

$$lr = lr_{base} \cdot \gamma^{\lfloor \frac{iter}{stepsize} \rfloor};$$

其中,lr表示衰减值, lr_{base} 表示预设学习率, γ 表示预设系数,iter表示,衰减步长,stepsize表示减少学习率的学习轮数。

9.如权利要求7所述的车辆目标识别的方法,其特征在于,所述对每个目标对应的目标边界框进行去重处理,获得优化后的每个目标对应的目标标记框,包括:

分别计算当前目标对应的所有目标边界框的置信度,并确定置信度最大的目标边界框;

分别计算其他置信度对应的目标边界框与置信度最大的目标边界框的置信度差值;

当所述置信度差值大于预设阈值时,去除对应的置信度对应的目标边界框,获得当前目标对应的目标标记框;

根据获得当前目标对应的目标标记框的方法获得每个目标对应的目标标记框。

10.一种车辆目标识别的装置,其特征在于,包括:

获取模块,用于获取数据集,并对所述数据集进行划分,确定训练数据集以及测试数据集;

计算模块,用于对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心;

训练模块,用于将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型;

识别模块,用于将所述测试数据集输入所述目标网络模型,获得识别后的车辆。

车辆目标识别的方法及装置

技术领域

[0001] 本发明属于地球视觉目标识别技术领域,尤其涉及一种车辆目标识别的方法及装置。

背景技术

[0002] 目标识别一直以来都是计算机视觉里的研究热点与难点。在信息与网络日益普及的今天,尤其是随着移动互联网时代的到来,无人机已经是人们生活中的比较重要的组成部分,无人机与各种各样不同的侦查系统一起构成了完整的新时代的战争预警与侦查系统,因此无人机图像中的车辆目标识别在民用与军用领域都具有极其重要的价值。然而,由于无人机图像中的车辆目标一般比较小,目标经过网络运算以后细节信息丢失严重,导致现有的识别算法对无人机图像中的车辆目标的识别效果较差。

发明内容

[0003] 有鉴于此,本发明实施例提供了一种车辆目标识别的方法及装置,以解决现有技术中对无人机图像中的车辆目标的识别效果较差的问题。

[0004] 本发明实施例的第一方面提供了一种车辆目标识别的方法,包括:

[0005] 获取数据集,并对所述数据集进行划分,确定训练数据集以及测试数据集;

[0006] 对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心;

[0007] 将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型;

[0008] 将所述测试数据集输入所述目标网络模型,获得识别后的车辆。

[0009] 在一实施例中,所述获取数据集,包括:

[0010] 分别获取不同角度拍摄的不同像素的图片,并将所有图片作为数据集。

[0011] 在一实施例中,所述对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心,包括:

[0012] 采用改进的K-means算法对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得预设数量的聚类中心。

[0013] 在一实施例中,所述改进的K-means算法聚类采用的距离公式为

[0014] $d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid})$;

[0015] 其中, $d(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框的距离, $\text{IOU}(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框之间的重叠度的值。

[0016] 在一实施例中,所述将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型,包括:

[0017] 设置所述聚类中心作为网络训练的初始候选窗口,并将所述训练数据集输入初始网络模型;

[0018] 对所述训练数据集中的所有图像进行特征提取,分别获得各个图像对应的预设像

素的特征图；

[0019] 将所述各个图像根据所述预设像素划分为对应数量的大小均匀的网格；

[0020] 根据每个网格内的特征点以及对应的初始候选窗口，确定每个目标待匹配的初始候选窗口；

[0021] 对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测，获得目标标记框，输出各个目标标记框；

[0022] 根据输出所述各个目标标记框采用的模型参数更新所述初始网络模型，获得更新后的网络模型，并确定多次更新后的网络模型中测试平均精度值最高的网络模型为目标网络模型。

[0023] 在一实施例中，所述根据每个网格内的特征点以及对应的初始候选窗口，确定每个目标待匹配的初始候选窗口，包括：

[0024] 根据每个网格内的特征点以及对应的初始候选窗口，计算各个特征点所在的预测框与对应的实际标记的预测框之间的重叠度值；

[0025] 确定最大的重叠度值对应的预测框为当前目标对应的待匹配的初始候选窗口。

[0026] 在一实施例中，所述对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测，获得目标标记框，包括：

[0027] 根据所述每个网格内的目标对应的所述待匹配的初始候选窗口，采用损失函数和衰减函数对每个网格内的目标真实标记框进行目标边界框预测，获得每个目标对应的目标边界框；

[0028] 对每个目标对应的目标边界框进行去重处理，获得优化后的每个目标对应的目标标记框。

[0029] 在一实施例中，所述损失函数为

$$[0030] \quad Loss = \frac{1}{2} \sum_{i=1}^{10647} \lambda_{obj} * \left((2 - truth_w * truth_h) * \sum_{r \in (x, y, w, h)} (truth_r - predict_r)^2 + \sum_{r=0}^{k-1} ((r == truth_{class}) ? 1 : 0 - predict_{class_r})^2 \right) + (truth_{conf} - predict_{conf})^2;$$

[0031] 其中，Loss表示损失值， λ_{obj} 表示一个特征点存在于一个初始候选窗口内时的取值， $truth_w$ 表示真实标记框的宽， $truth_h$ 表示真实标记框的高， $(truth_r - predict_r)^2$ 表示实际与预测之间真实标记框的差异， $truth_{class}$ 表示真实类别， $predict_{class_r}$ 表示预测的类别， $(truth_{conf} - predict_{conf})^2$ 表示置信度损失；

[0032] 所述衰减函数为

$$[0033] \quad lr = lr_{base} \cdot \gamma^{\left\lfloor \frac{iter}{stepsize} \right\rfloor};$$

[0034] 其中，lr表示衰减值， lr_{base} 表示预设学习率， γ 表示预设系数，iter表示，衰减步长，stepsize表示减少学习率的学习轮数。

[0035] 在一实施例中，所述对每个目标对应的目标边界框进行去重处理，获得优化后的每个目标对应的目标标记框，包括：

[0036] 分别计算当前目标对应的所有目标边界框的置信度，并确定置信度最大的目标边

界框；

[0037] 分别计算其他置信度对应的目标边界框与置信度最大的目标边界框的置信度差值；

[0038] 当所述置信度差值大于预设阈值时，去除对应的置信度对应的目标边界框，获得当前目标对应的目标标记框；

[0039] 根据获得当前目标对应的目标标记框的方法获得每个目标对应的目标标记框。

[0040] 本发明实施例的第二方面提供了一种车辆目标识别的装置，包括：

[0041] 获取模块，用于获取数据集，并对所述数据集进行划分，确定训练数据集以及测试数据集；

[0042] 计算模块，用于对所述训练数据集中的每一个目标的车辆标记框进行聚类计算，获得聚类中心；

[0043] 训练模块，用于将所述聚类中心作为网络训练的初始候选窗口，基于所述训练数据集对初始网络模型进行训练，获得目标网络模型；

[0044] 识别模块，用于将所述测试数据集输入所述目标网络模型，获得识别后的车辆。

[0045] 本发明实施例与现有技术相比存在的有益效果是：通过获取数据集，并对所述数据集进行划分，确定训练数据集以及测试数据集；对所述训练数据集中的每一个目标的车辆标记框进行聚类计算，获得聚类中心；将所述聚类中心作为网络训练的初始候选窗口，基于所述训练数据集对初始网络模型进行训练，获得目标网络模型；将所述测试数据集输入所述目标网络模型，获得识别后的车辆，从而可以提高车辆目标的识别效果。

附图说明

[0046] 为了更清楚地说明本发明实施例中的技术方案，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发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0047] 图1是本发明实施例提供的车辆目标识别的方法的实现流程示意图；

[0048] 图2是本发明实施例提供的获得目标网络模型的示意图；

[0049] 图3是本发明实施例提供的车辆目标识别的装置的示例图；

[0050] 图4是本发明实施例提供的终端设备的示意图。

具体实施方式

[0051] 以下描述中，为了说明而不是为了限定，提出了诸如特定系统结构、技术之类的具体细节，以便透彻理解本发明实施例。然而，本领域的技术人员应当清楚，在没有这些具体细节的其它实施例中也可以实现本发明。在其它情况中，省略对众所周知的系统、装置、电路以及方法的详细说明，以免不必要的细节妨碍本发明的描述。

[0052] 为了说明本发明所述的技术方案，下面通过具体实施例来进行说明。

[0053] 图1为本发明实施例提供的车辆目标识别的方法的实现流程示意图，详述如下。

[0054] 步骤101，获取数据集，并对所述数据集进行划分，确定训练数据集以及测试数据集。

[0055] 可选的,本步骤可以包括分别获取不同角度拍摄的不同像素的图片,并将所有图片作为数据集。可选的,可以获取高层建筑上拍摄的图像,因此在一定程度上类似于无人机的图像视图,例如,获取PKLot的子数据集PUCPR,其中PUCPR中的图像的大小可以为 1280×720 像素,还可以获取无人机拍摄的停车场数据集,例如PKCAR数据集,其中PKCAR数据集的图像的大小可以为 1280×720 像素。还可以为自己拍摄的图像的大小为 4608×3456 像素的图片。将上述获取的所有图像合在一起,构成本实施例中的数据集。

[0056] 可选的,对数据集按照一定比例进行划分,为了训练出合格的网络模型,可以将训练数据集中包括的图像数量大于测试数据集中包括的图像数据,例如按照7:3的比例划分数据集,70%的图像构成训练数据集,30%的图像构成测试数据集。当然还可以采用其它的比例划分数据集,例如6:4、8:2、9:1等等都可以。

[0057] 步骤102,对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心。

[0058] 可选的,现有技术中对车辆目标识别通常采用YOLO V3网络识别方法,但是原有的YOLO V3在多尺度上预测目标位置与分类时,采用的初始候选窗口anchor都是在VOC数据集上聚类而来,而且聚类算法比较依赖初始值的选取。因此在本实施例中采用改进的K-means++算法获取聚类中心。

[0059] 可选的,本步骤包括采用改进的K-means算法对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得预设数量的聚类中心。可选的,由于YOLO V3网络的预测框有九个,则可以采用改进的K-means算法对所述训练数据集中的每一个目标的车辆标记框进行 $K=9$ 的聚类计算,这样可以将计算的九个聚类中心作为网络训练的初始候选窗口。

[0060] 可选的,聚类过程中使用的距离度量是一个值得考量的问题,假如使用欧式距离来计算的话,就会造成大的anchor会比小的anchor产生更多的错误,因此不采用欧式距离度量公式。聚类的目的是anchor boxes和临近的真实标记框ground truth有更大的重叠度值,这和anchor box的尺寸没有直接关系。聚类使用的距离公式为

[0061] $d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid})$;

[0062] 其中, $d(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框的距离, $\text{IOU}(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框之间的重叠度的值。

[0063] 当前特征点到聚类中心的距离越小越好,但IOU值是越大越好,所以使用 $(1 - \text{IOU})$ 就可以保证距离越小,IOU值越大。聚类后计算出来的九个聚类中心作为网络训练的初始的anchor值,这样改进以后既可以提高模型的收敛速度,减小发散的风险,也可以提高模型的准确率。

[0064] 步骤103,将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型。

[0065] 可选的,如图2所示,本步骤可以包括以下步骤。

[0066] 步骤201,设置所述聚类中心作为网络训练的初始候选窗口,并将所述训练数据集输入初始网络模型。

[0067] 可选的,通过使用改进的K-means++算法对训练数据集的真实标记框进行聚类,得到的预测框的尺度可以为 $(12, 18)$ 、 $(14, 36)$ 、 $(20, 36)$ 、 $(50, 62)$ 、 $(70, 92)$ 、 $(80, 96)$ 、 $(56, 97)$ 、 $(72, 132)$ 、 $(110, 226)$ 。

[0068] 步骤202,对所述训练数据集中的所有图像进行特征提取,分别获得各个图像对应的预设像素的特征图。

[0069] 可选的,采用特征提取网络结构Darknet53对所述训练数据集中的所有图像进行特征提取,输入的图像经过Darknet53网络进行提取特征之后,得到一个一定大小的特征图 feature map。可选的,预设像素可以为 13×13 。

[0070] 步骤203,将所述各个图像根据所述预设像素划分为对应数量的大小均匀的网格。

[0071] 可选的,将各个图像划分为 13×13 个大小均匀的网格。

[0072] 步骤204,根据每个网格内的特征点以及对应的初始候选窗口,确定每个目标待匹配的初始候选窗口。

[0073] 可选的,如果图像中某个目标的真实标记框的中心坐标坐落到网格中,那么该网格负责预测该物体。图像中的每个小网格都会生成固定数量的预测框,即初始候选窗口。在本实施例中每个网格对应的初始候选窗口为九个,每个初始候选窗口的尺度不同,在实际预测的时候,应该选择哪一个预测框来对该目标进行预测,通过该预测框与网中实际标记的预测框之间的重叠度的值IOU来决定,重叠度的值最大的预测框就作为该目标的预测框。

[0074] 可选的,本步骤可以包括根据每个网格内的特征点以及对应的初始候选窗口,计算各个特征点所在的预测框与对应的实际标记的预测框之间的重叠度值;确定最大的重叠度值对应的预测框为当前目标对应的待匹配的初始候选窗口。

[0075] 步骤205,对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测,获得目标标记框,输出各个目标标记框。

[0076] 可选的,本步骤可以包括:根据所述每个网格内的目标对应的所述待匹配的初始候选窗口,采用损失函数和衰减函数对每个网格内的目标的真实标记框进行目标边界框预测,获得每个目标对应的目标边界框;对每个目标对应的目标边界框进行去重处理,获得优化后的每个目标对应的目标标记框。

[0077] 可选的,所述损失函数为

$$[0078] \quad Loss = \frac{1}{2} \sum_{i=1}^{10647} \lambda_{obj} * \left((2 - truth_w * truth_h) * \sum_{r \in (x,y,w,h)} (truth_r - predict_r)^2 \right) + (truth_{conf} - predict_{conf})^2; \\ + \sum_{r=0}^{k-1} ((r = truth_{class}) ? 1 : 0 - predict_{class_r})^2$$

[0079] 其中,Loss表示损失值,即坐标位置的损失、置信度的损失和分类损失之和, λ_{obj} 表示一个特征点存在于一个初始候选窗口内时的取值, $truth_w$ 表示真实标记框的宽, $truth_h$ 表示真实标记框的高, $(truth_r - predict_r)^2$ 表示实际与预测之间真实标记框的差异, $truth_{class}$ 表示真实类别, $predict_{class_r}$ 表示预测的类别, $(truth_{conf} - predict_{conf})^2$ 表示置信度损失;

[0080] 损失函数主要包括三个大的方面:坐标位置的损失、置信度的损失和分类损失。 λ_{obj} 的值可以为1或0,当一个特征点存在一个检测物体的时候,它的值就取1,否则就取值为0。Loss表示的是所有的损失函数之和。使用这种方式主要是出于网络的预测的特殊机制的考虑,训练数据集中的数据正负样本不均衡,特别是在计算置信度损失函数的那一部分,假如图像中只有一个目标,那么置信度的正负样本比就达到1:10646。假如依然使用平均损失函数的话,这样损失函数的值就接近于0,这样网络就没有办法继续计算下去,造成梯度消失。

[0081] 所述衰减函数为 $lr = lr_{base} \cdot \gamma^{\lfloor \frac{iter}{stepsize} \rfloor}$;

[0082] 其中, lr表示衰减值, lr_{base}表示预设学习率, γ表示预设系数, iter表示衰减步长, stepsize表示减少学习率的学习轮数。

[0083] 可选的, 在对网络进行训练的时候, 使用Darknet-53框架在训练数据集上采用随机梯度下降法SGD训练1万次, 在训练过程中, 网络每次输入的是64张图片组成的一个矩阵, 开始的时候以0.0003的初始学习率开始训练网络, 选择均匀分步速率衰减策略, 步长设置为在4000次与6000次时, 学习率衰减为前一步的0.1倍。

[0084] 可选的, 在获得每个目标对应的目标边界框之后, 可以一个目标对应多个目标边界框, 这样需要将重叠的目标边界框去掉, 以便获得准确的目标边界框。现有技术中通常采用非极大值抑制的算法进行去重处理, 然而由于停车场里的车辆因为每个目标在原图中相互之间的距离很近, 尤其对于那些斜着拍摄的车辆, 当对其进行标记的时候, 两个车辆目标之间的框重叠度会很高, 所以直接用非极大值抑制的算法很可能会造成某一些重叠度高的目标的漏检。本实施例中采用改进的非极大值抑制的算法进行去重处理。

[0085] 可选的, 对每个目标对应的目标边界框进行去重处理, 获得优化后的每个目标对应的目标标记框, 可以包括: 分别计算当前目标对应的所有目标边界框的置信度, 并确定置信度最大的目标边界框; 分别计算其他置信度对应的目标边界框与置信度最大的目标边界框的置信度差值; 当所述置信度差值大于预设阈值时, 去除对应的置信度对应的目标边界框, 获得当前目标对应的目标标记框; 根据获得当前目标对应的目标标记框的方法获得每个目标对应的目标标记框。这样的做法可以在某种程度上减小密集目标漏检的问题, 从而提高模型的准确率。

[0086] 步骤206, 根据输出所述各个目标标记框采用的模型参数更新所述初始网络模型, 获得更新后的网络模型, 并确定多次更新后的网络模型中测试平均精度值最高的网络模型为目标网络模型。

[0087] 可选的, 采用YOLO V3网络在不同的训练次数获得对应的测试测试平均精度(Average Precision, AP)值, 如表一所示。

[0088] 表一

次数	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
AP (%)	83.2	86.7	89.1	92.0	91.8	91.8	90.2	90.8	91.5	91.5

[0090] 从上表可以看出YOLO V3网络在网络训练到4000次的时候, 达到最优解, 其在测试数据集上的AP为92.01%, 使用网络在4000次迭代的时候的参数对应的目标网络模型在测试数据集上进行测试, 得到的测试样例图如表二所示。

[0091] 表二

AP	召回率	准确率	误检率	漏检率
92.01%	94.23%	99.74%	0.26%	5.77%

[0093] 采用本实施例中通过使用改进的K-means算法对训练数据集的Ground Truth进行

聚类,运用新得到的anchor值对网络进行训练,训练的策略与初始网络的训练策略相同,得到的测试样例图如表三所示。

[0094] 表三

[0095]	AP	召回率	准确率	误检率	漏检率
	96.32%	97.70%	99.76%	2.30%	0.24%

[0096] 在网络使用改进的K-means聚类的基础上,使用改进的NMS算法进行优化,改进后的网络在不同的训练次数对应的测试AP值如表四所示。

[0097] 表四

	次数	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
[0098]	AP (%)	86.25	89.6	93.5	96.3	97.2	97.4	96.5	97.12	93.1	94.51
			2	6	8	0	9	5		2	

[0099] 从上表可以看出采用改进的网络在网络训练到6000次的时候,达到最优解,其在测试数据集上的AP为97.49%,使用网络在6000次迭代的时候的参数对应的目标网络模型在测试数据集上进行测试,得到的测试样例图如表五所示。

[0100] 表五

	AP	召回率	准确率	误检率	漏检率
[0101]	97.49%	98.74%	99.66%	1.26%	0.34%

[0102] 从上表可以看出,使用改进的K-means聚类算法后,网络的准确率与召回率有较明显的上升,对不完整目标的检测效果更好。使用改进后的NMS算法以后,网络的综合性能稍有提升,主要是因为训练数据集中的目标会被NMS错误的抑制掉的只是很少一部分。改进后的网络对不完整的车辆目标与NMS错误抑制的车辆目标的检测效果都有改善。

[0103] 步骤104,将所述测试数据集输入所述目标网络模型,获得识别后的车辆。

[0104] 上述车辆目标识别的方法,通过获取数据集,并对所述数据集进行划分,确定训练数据集以及测试数据集;对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心;将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型;将所述测试数据集输入所述目标网络模型,获得识别后的车辆,从而可以提高车辆目标的识别效果,并且可以达到实时检测的目的。

[0105] 应理解,上述实施例中各步骤的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不应对本发明实施例的实施过程构成任何限定。

[0106] 对应于上文实施例所述的车辆目标识别的方法,图3示出了本发明实施例提供的车辆目标识别的装置的示例图。如图3所示,该装置可以包括:获取模块301、计算模块302、训练模块303以及识别模块304。

[0107] 获取模块301,用于获取数据集,并对所述数据集进行划分,确定训练数据集以及测试数据集;

[0108] 计算模块302,用于对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心;

[0109] 训练模块303,用于将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型;

[0110] 识别模块304,用于将所述测试数据集输入所述目标网络模型,获得识别后的车辆。

[0111] 可选的,所述获取模块301获取数据集时,可以用于:

[0112] 分别获取不同角度拍摄的不同像素的图片,并将所有图片作为数据集。

[0113] 可选的,所述计算模块302对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心时,可以用于:

[0114] 采用改进的K-means算法对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得预设数量的聚类中心。

[0115] 可选的,所述改进的K-means算法聚类采用的距离公式为

[0116] $d(\text{box}, \text{centroid}) = 1 - \text{IOU}(\text{box}, \text{centroid})$;

[0117] 其中, $d(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框的距离, $\text{IOU}(\text{box}, \text{centroid})$ 表示当前预测框与实际标记的预测框之间的重叠度的值。

[0118] 可选的,所述训练模块303将所述聚类中心作为网络训练的初始候选窗口,基于所述训练数据集对初始网络模型进行训练,获得目标网络模型时,可以用于:

[0119] 设置所述聚类中心作为网络训练的初始候选窗口,并将所述训练数据集输入初始网络模型;

[0120] 对所述训练数据集中的所有图像进行特征提取,分别获得各个图像对应的预设像素的特征图;

[0121] 将所述各个图像根据所述预设像素划分为对应数量的大小均匀的网格;

[0122] 根据每个网格内的特征点以及对应的初始候选窗口,确定每个目标待匹配的初始候选窗口;

[0123] 对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测,获得目标标记框,输出各个目标标记框;

[0124] 根据输出所述各个目标标记框采用的模型参数更新所述初始网络模型,获得更新后的网络模型,并确定多次更新后的网络模型中测试平均精度值最高的网络模型为目标网络模型。

[0125] 可选的,所述训练模块303根据每个网格内的特征点以及对应的初始候选窗口,确定每个目标待匹配的初始候选窗口时,可以用于:

[0126] 根据每个网格内的特征点以及对应的初始候选窗口,计算各个特征点所在的预测框与对应的实际标记的预测框之间的重叠度值;

[0127] 确定最大的重叠度值对应的预测框为当前目标对应的待匹配的初始候选窗口。

[0128] 可选的,所述训练模块303对所述每个网格内的目标对应的所述待匹配的初始候选窗口进行目标边界框预测,获得目标标记框时,可以用于:

[0129] 根据所述每个网格内的目标对应的所述待匹配的初始候选窗口,采用损失函数和衰减函数对每个网格内的目标的真实标记框进行目标边界框预测,获得每个目标对应的目标边界框;

[0130] 对每个目标对应的目标边界框进行去重处理,获得优化后的每个目标对应的目标

标记框。

[0131] 可选的,所述损失函数为

$$[0132] \quad Loss = \frac{1}{2} \sum_{i=1}^{10647} \lambda_{obj} * \left((2 - truth_w * truth_h) * \sum_{r \in (x,y,w,h)} (truth_r - predict_r)^2 + \sum_{r=0}^{k-1} ((r == truth_{class}) ? 1 : 0 - predict_{class_r})^2 \right) + (truth_{conf} - predict_{conf})^2;$$

[0133] 其中, Loss表示损失值, λ_{obj} 表示一个特征点存在于一个初始候选窗口内时的取值, $truth_w$ 表示真实标记框的宽, $truth_h$ 表示真实标记框的高, $(truth_r - predict_r)^2$ 表示实际与预测之间真实标记框的差异, $truth_{class}$ 表示真实类别, $predict_{class_r}$ 表示预测的类别, $(truth_{conf} - predict_{conf})^2$ 表示置信度损失;

[0134] 所述衰减函数为

$$[0135] \quad lr = lr_{base} \cdot \gamma^{\left\lfloor \frac{iter}{stepsize} \right\rfloor};$$

[0136] 其中, lr表示衰减值, lr_{base} 表示预设学习率, γ 表示预设系数, iter表示, 衰减步长, stepsize表示减少学习率的学习轮数。

[0137] 可选的,所述训练模块303对每个目标对应的目标边界框进行去重处理,获得优化后的每个目标对应的目标标记框时,可以用于:分别计算当前目标对应的所有目标边界框的置信度,并确定置信度最大的目标边界框;分别计算其他置信度对应的目标边界框与置信度最大的目标边界框的置信度差值;当所述置信度差值大于预设阈值时,去除对应的置信度对应的目标边界框,获得当前目标对应的目标标记框;根据获得当前目标对应的目标标记框的方法获得每个目标对应的目标标记框。

[0138] 上述车辆目标识别的装置,通过获取模块获取数据集,并对所述数据集进行划分,确定训练数据集以及测试数据集;计算模块对所述训练数据集中的每一个目标的车辆标记框进行聚类计算,获得聚类中心;将所述聚类中心作为网络训练的初始候选窗口,训练模块基于所述训练数据集对初始网络模型进行训练,获得目标网络模型;识别模块将所述测试数据集输入所述目标网络模型,获得识别后的车辆,从而可以提高车辆目标的识别效果,并且可以达到实时检测的目的。

[0139] 图4是本发明一实施例提供的终端设备的示意图。如图4所示,该实施例的终端设备400包括:处理器401、存储器402以及存储在所述存储器402中并可在所述处理器401上运行的计算机程序403,例如车辆目标识别的程序。所述处理器401执行所述计算机程序403时实现上述车辆目标识别的方法实施例中的步骤,例如图1所示的步骤101至104,或者图2所示的步骤201至步骤206,所述处理器401执行所述计算机程序403时实现上述各装置实施例中各模块的功能,例如图3所示模块301至304的功能。

[0140] 示例性的,所述计算机程序403可以被分割成一个或多个程序模块,所述一个或者多个程序模块被存储在所述存储器402中,并由所述处理器401执行,以完成本发明。所述一个或多个程序模块可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述所述计算机程序403在所述车辆目标识别的装置或者终端设备400中的执行过程。例如,所述计算机程序403可以被分割成获取模块301、计算模块302、训练模块303以及识别模块304,各模块具体功能如图3所示,在此不再一一赘述。

[0141] 所述终端设备400可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述终端设备可包括,但不仅限于,处理器401、存储器402。本领域技术人员可以理解,图4仅仅是终端设备400的示例,并不构成对终端设备400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,例如所述终端设备还可以包括输入输出设备、网络接入设备、总线等。

[0142] 所称处理器401可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0143] 所述存储器402可以是所述终端设备400的内部存储单元,例如终端设备400的硬盘或内存。所述存储器402也可以是所述终端设备400的外部存储设备,例如所述终端设备400上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。进一步地,所述存储器402还可以既包括所述终端设备400的内部存储单元也包括外部存储设备。所述存储器402用于存储所述计算机程序以及所述终端设备400所需的其他程序和数据。所述存储器402还可以用于暂时地存储已经输出或者将要输出的数据。

[0144] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能单元、模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能单元、模块完成,即将所述装置的内部结构划分成不同的功能单元或模块,以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能单元、模块的具体名称也只是为了便于相互区分,并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0145] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述或记载的部分,可以参见其它实施例的相关描述。

[0146] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0147] 在本发明所提供的实施例中,应该理解到,所揭露的装置/终端设备和方法,可以通过其它的方式实现。例如,以上所描述的装置/终端设备实施例仅仅是示意性的,例如,所述模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以通过一些接口,装置或单元的间接耦合或通讯连接,可以是电性,机械或其它的形式。

[0148] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0149] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0150] 所述集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明实现上述实施例方法中的全部或部分流程,也可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一计算机可读存储介质中,该计算机程序在被处理器执行时,可实现上述各个方法实施例的步骤。其中,所述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。需要说明的是,所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减,例如在某些司法管辖区,根据立法和专利实践,计算机可读介质不包括电载波信号和电信信号。

[0151] 以上所述实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围,均应包含在本发明的保护范围之内。

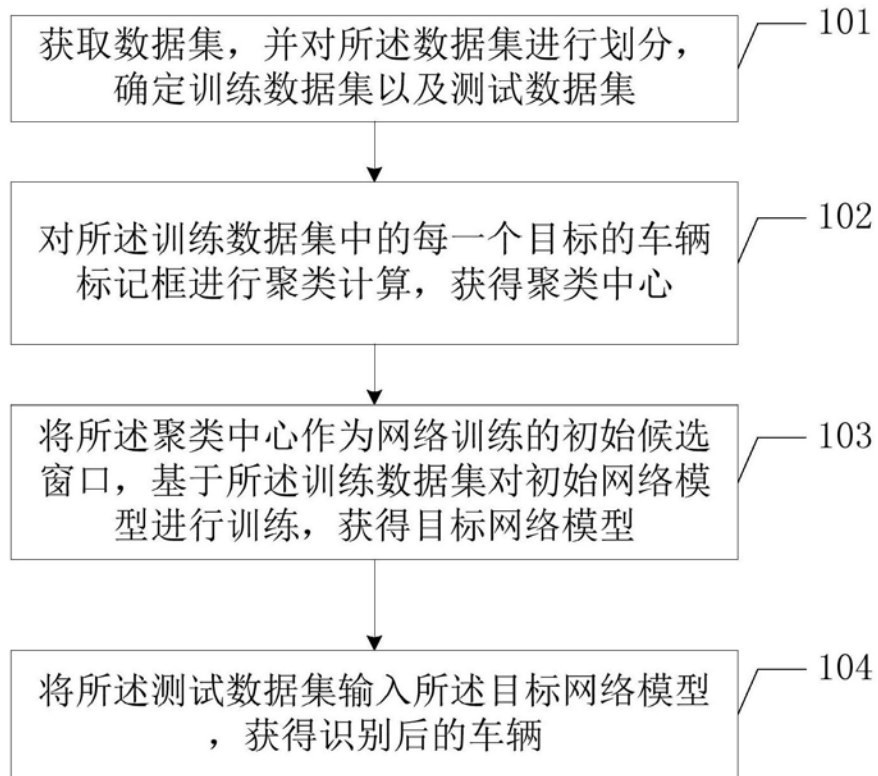


图1

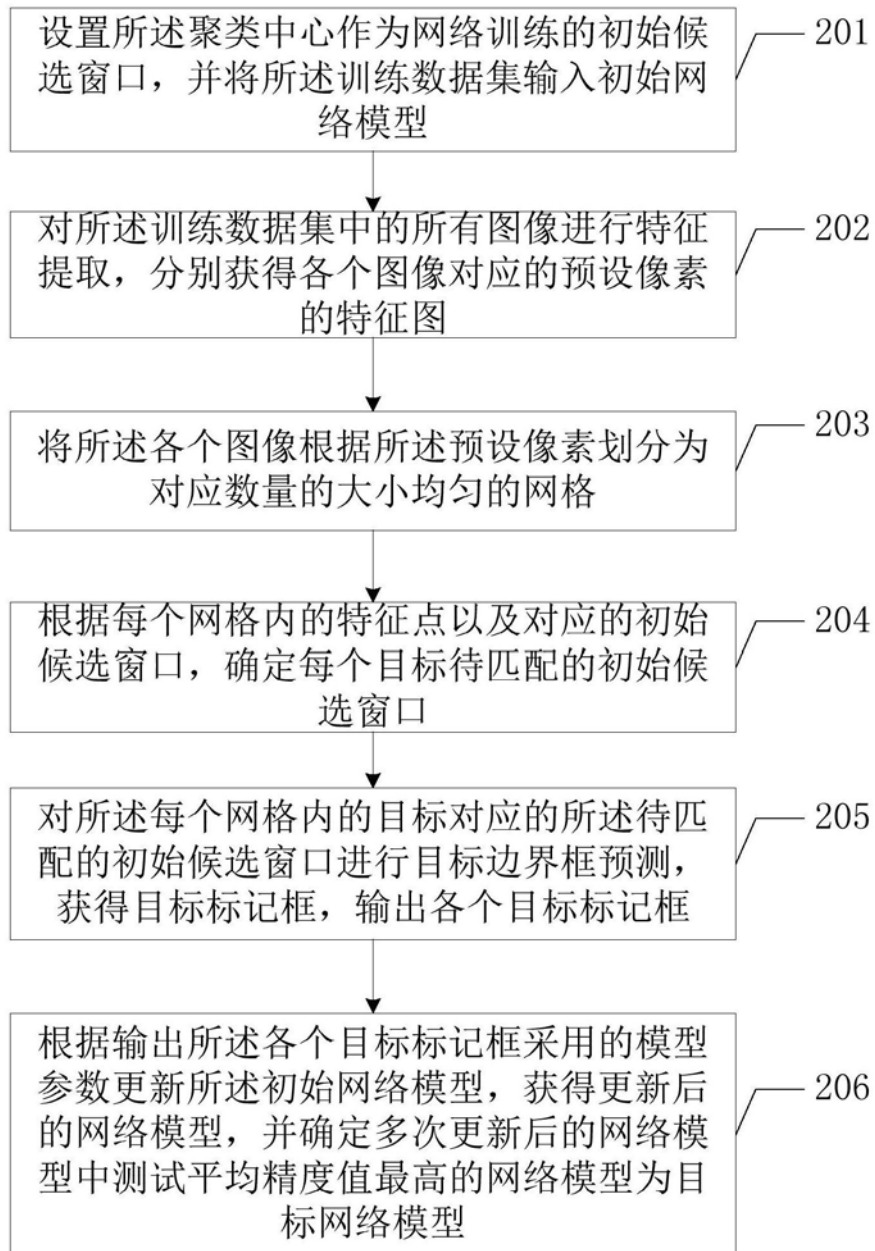


图2

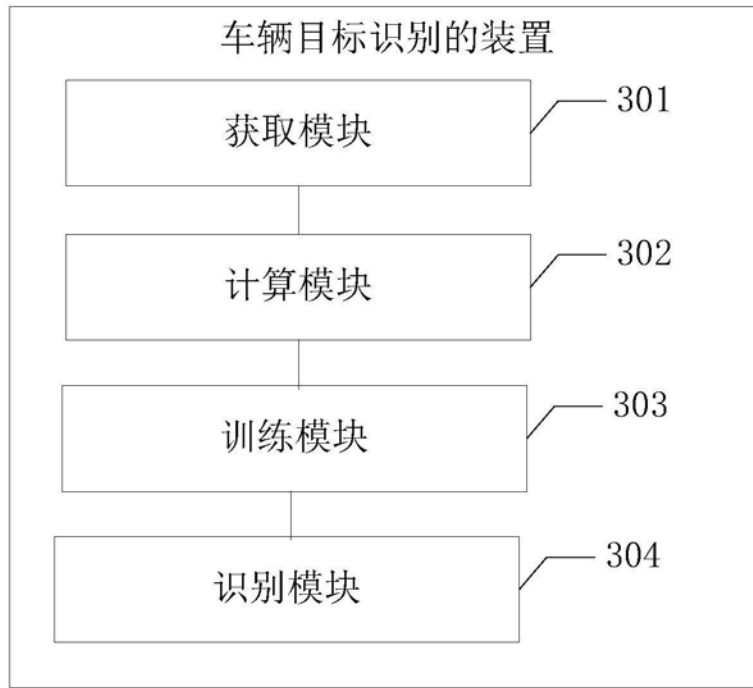


图3

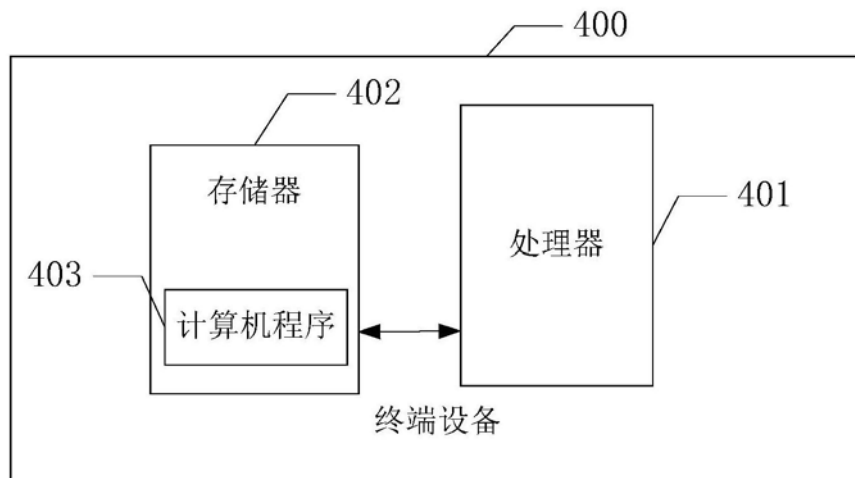


图4