



(12)发明专利

(10)授权公告号 CN 108460427 B

(45)授权公告日 2020.05.19

(21)申请号 201810272116.9

审查员 邹琴

(22)申请日 2018.03.29

(65)同一申请的已公布的文献号

申请公布号 CN 108460427 A

(43)申请公布日 2018.08.28

(73)专利权人 国信优易数据有限公司

地址 100070 北京市丰台区南四环西路188号总部广场31号楼

(72)发明人 孙源良 夏虎 李长升 刘萌

(74)专利代理机构 北京超凡志成知识产权代理事务所(普通合伙) 11371

代理人 戈丰

(51)Int.Cl.

G06K 9/62(2006.01)

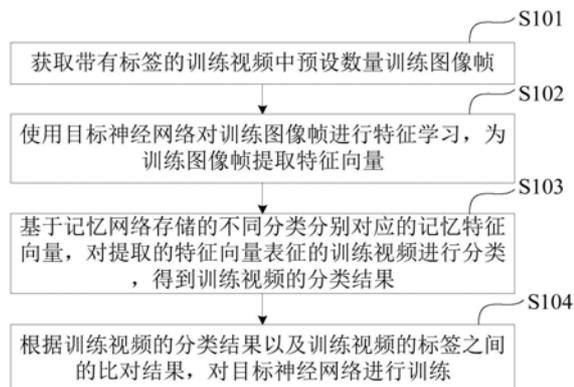
权利要求书2页 说明书10页 附图5页

(54)发明名称

一种分类模型训练方法、装置以及分类方法及装置

(57)摘要

本申请提供一种分类模型训练方法、装置以及分类方法及装置,分类模型训练方法包括:获取带有标签的训练视频中预设数量训练图像帧;使用目标神经网络对训练图像帧进行特征学习,为训练图像帧提取特征向量;基于记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的训练视频进行分类,得到训练视频的分类结果;根据训练视频的分类结果以及训练视频的标签之间的比对结果,对目标神经网络进行训练。记忆网络具有强大的记忆能力,能够弥补神经网络本身存在的记忆能力差的问题,不论对于任何长度的输入序列,记忆网络都能够对其特征进行很好的记忆,因而不论输入序列为何,都不会由于神经网络记忆能力差而影响分类模型的识别精度。



1. 一种分类模型训练方法,其特征在于,该方法包括:
 - 获取带有标签的训练视频中预设数量训练图像帧;
 - 使用目标神经网络对所述训练图像帧进行特征学习,为训练图像帧提取特征向量;
 - 基于记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的所述训练视频进行分类,得到所述训练视频的分类结果;
 - 根据所述训练视频的分类结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练。
2. 根据权利要求1所述的方法,其特征在于,所述基于所述记忆网络存储的不同类别分别对应的记忆特征向量,对提取的特征向量表征的所述训练视频进行分类,得到所述训练视频的分类结果,具体包括:
 - 将提取的特征向量分别与所述记忆网络中不同分类所对应的记忆特征向量进行相似度匹配,得到每个分类与提取的特征向量的相似度匹配结果;
 - 将得到相似度匹配结果中匹配程度最高的分类确定为所述训练视频的分类结果。
3. 根据权利要求2所述的方法,其特征在于,将提取的特征向量分别与所述记忆网络中不同分类所对应的记忆特征向量进行相似度匹配,具体包括:
 - 分别对所述记忆网络中各分类对应的记忆特征向量进行聚类;
 - 对提取的特征向量和聚类结果作相似度度量。
4. 根据权利要求3所述的方法,其特征在于,对提取的特征向量和所述聚类结果做相似度度量,具体包括:
 - 根据预设的基准向量确认规则,为所述聚类结果中得到的每个聚类确认基准向量;
 - 针对每个提取的特征向量,计算该特征向量与每个所述聚类对应的基准向量之间的距离;以及
 - 将与该特征向量距离最小的基准向量所在的分类,作为该特征向量所对应训练图像帧的分类;
 - 将对应训练图像帧最多的分类,作为所述训练视频的分类结果。
5. 根据权利要求1-4任意一项所述的方法,其特征在于,根据所述训练视频的分类结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练,具体包括:
 - 执行如下比对操作,直至所述训练视频的分类结果与所述训练视频的标签一致;所述比对操作包括:
 - 将所述训练视频的分类结果与所述训练视频的标签进行比对;
 - 若所述训练视频的分类结果与所述训练视频的标签不一致,则对所述目标神经网络的参数进行调整;
 - 基于调整后的参数,使用所述目标神经网络为训练图像帧提取新的特征向量;
 - 基于所述记忆网络存储的不同分类分别对应的记忆特征向量以及新的特征向量,对所述训练视频进行分类,得到所述训练视频的新的分类结果,并再次执行所述比对操作。
6. 根据权利要求5所述的方法,其特征在于,还包括:
 - 在所述分类结果与所述训练视频的标签一致时,将对应提取的特征向量添加至所述记忆网络对应分类中,对所述记忆网络进行更新。
7. 根据权利要求1-4任意一项所述的方法,其特征在于,获取带有标签的训练视频中预

设数量训练图像帧,包括:

获取带有标签的训练视频中,前预设数量训练图像帧;

基于所述记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的所述训练视频进行分类,得到所述训练视频的分类结果,包括:

基于所述记忆网络存储的不同动作事件分别对应的记忆特征向量,对所述训练视频中,提取的特征向量对应的训练图像帧之后的图像帧包含的动作事件进行预测,得到所述训练视频的预测结果;

根据所述训练视频的分类结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练,包括:

根据所述训练视频的预测结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练。

8. 一种分类方法,其特征在于,该方法包括:

获取待分类视频;

将所述待分类视频输入至通过权利要求1-7任意一项的分类模型训练方法得到的分类模型中,得到所述待分类视频的分类结果;

其中,所述分类模型包括:所述目标神经网络以及所述记忆网络。

9. 一种分类模型训练装置,其特征在于,该装置包括:

获取模块,用于获取带有标签的训练视频中预设数量训练图像帧;

提取模块,用于使用目标神经网络对所述训练图像帧进行特征学习,为训练图像帧提取特征向量;

分类模块,用于基于记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的所述训练视频进行分类,得到所述训练视频的分类结果;

训练模块,用于根据所述训练视频的分类结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练。

10. 一种分类装置,其特征在于,该装置包括:

待分类数据获取模块,用于获取待分类视频;

分类模块,用于将所述待分类视频输入至通过权利要求1-7任意一项的分类模型训练方法得到的分类模型中,得到所述待分类视频的分类结果;

其中,所述分类模型包括:所述目标神经网络以及所述记忆网络。

一种分类模型训练方法、装置以及分类方法及装置

技术领域

[0001] 本申请涉及机器学习技术领域,具体而言,涉及一种分类模型训练方法、装置以及分类方法及装置。

背景技术

[0002] 多层反馈(Recurrent neural Network、RNN)神经网络,又称循环神经网络,是一种节点定向连接成环的人工神经网络,其可以利用内部的记忆来处理任意时序的输入序列,在视频动作预测中得到广泛的应用。

[0003] 但是RNN网络的记忆能力较差,当输入序列较长的时候,无法精确实现对输入序列的分类,无法达到对精度的使用需求。

发明内容

[0004] 有鉴于此,本申请实施例的目的在于提供一种分类模型训练方法、装置以及分类方法及装置,能够提高基于RNN网络的分类模型对视频的识别精度。

[0005] 第一方面,本申请实施例提供了一种分类模型训练方法,该方法包括:

[0006] 获取带有标签的训练视频中预设数量训练图像帧;

[0007] 使用目标神经网络对所述训练图像帧进行特征学习,为训练图像帧提取特征向量;

[0008] 基于所述记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的所述训练视频进行分类,得到所述训练视频的分类结果;

[0009] 根据所述训练视频的分类结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练。

[0010] 第二方面,本申请实施例还提供一种分类方法,该方法包括:

[0011] 获取待分类视频;

[0012] 将所述待分类视频输入至通过第一方面所述的分类模型训练方法得到的分类模型中,得到所述待分类视频的分类结果;

[0013] 其中,所述分类模型包括:所述目标神经网络以及所述记忆网络。

[0014] 第三方面,本申请实施例还提供一种分类模型训练装置,该装置包括:

[0015] 获取模块,用于获取带有标签的训练视频中预设数量训练图像帧;

[0016] 提取模块,用于使用目标神经网络对所述训练图像帧进行特征学习,为训练图像帧提取特征向量;

[0017] 分类模块,用于基于所述记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的所述训练视频进行分类,得到所述训练视频的分类结果;

[0018] 训练模块,用于根据所述训练视频的分类结果以及所述训练视频的标签之间的比对结果,对所述目标神经网络进行训练。

[0019] 第四方面,本申请实施例还提供一种分类装置,该装置包括:

[0020] 待分类数据获取模块,用于获取待分类视频;

[0021] 分类模块,用于将所述待分类视频输入至通过上述第一方面提供的分类模型训练方法得到的分类模型中,得到所述待分类视频的分类结果;

[0022] 其中,所述分类模型包括:所述目标神经网络以及所述记忆网络。

[0023] 本申请实施例,通过目标神经网络和记忆网络共同构成分类模型;在对分类模型进行训练的时候,将带有标签的训练视频中预设数量训练图像帧输入至目标神经网络,会使用目标神经网络为训练图像帧提取特征向量,然后将提取的特征向量输入至记忆网络,在记忆网络中存储了不同分类分别对应的记忆特征向量;当记忆网络接收到提取的特征向量后,能够基于其所存储的不同分类分别对应的记忆特征向量,对提取的特征向量标准横的训练视频进行分类,得到训练视频的分类结果,然后根据训练视频的分类结果以及分类视频的标签之间的比对结果,对目标神经网络进行训练。在该分类模型中,记忆网络具有强大的记忆能力,能够弥补神经网络本身存在的记忆能力差的问题,不论对于任何长度的输入序列,记忆网络都能够对其特征进行很好的记忆,因而不论输入序列为何,都不会由于神经网络记忆能力差而影响分类模型的识别精度。

[0024] 为使本申请的上述目的、特征和优点能更明显易懂,下文特举较佳实施例,并配合所附附图,作详细说明如下。

附图说明

[0025] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,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相关的附图。

[0026] 图1示出了本申请实施例一提供的一种分类模型训练方法的流程图;

[0027] 图2示出了本申请实施例二提供的基于记忆网络存储的不同类别分别对应的记忆特征向量,对提取的特征向量表征的训练视频进行分类,得到训练视频的分类结果的具体方法的流程图;

[0028] 图3示出了申请实施例三提供的将提取的特征向量分别与记忆网络中不同分类所对应的记忆特征向量进行相似度匹配的具体方法流程图;

[0029] 图4示出了申请实施例四提供的对提取的特征向量和聚类结果做相似度度量的具体方法流程图;

[0030] 图5示出了申请实施例五提供的比对操作的方法流程图;

[0031] 图6示出了申请实施例六提供的分类模型训练方法流程图;

[0032] 图7示出了申请实施例七提供的分类模型训练装置的结构示意图;

[0033] 图8示出了申请实施例八提供的分类方法流程图;

[0034] 图9示出了申请实施例九提供的分类装置的结构示意图;

[0035] 图10示出了本申请实施例十提供的一种计算机设备的结构示意图。

具体实施方式

[0036] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例

中附图,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实施例,而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以以各种不同的配置来布置和设计。因此,以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围,而是仅仅表示本申请的选定实施例。基于本申请的实施例,本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0037] RNN网络记忆能力差,导致其在想要对视频动作进行精度较高的预测,需要大量的训练数据输入对RNN网络进行训练;同时,对于一些分类较为接近的数据,RNN无法得到精确的分类结果;例如在将RNN网络应用于自动驾驶场景的时候,当前方车辆在现动作比较相似的场景下,如在汽车拐弯、并线以及掉头三种场景下,RNN网络是无法对这三个动作进行精确的识别的,因而无法达到对精度的使用需求。基于此,本申请实施例提供一种分类模型训练方法、装置,分类方法以及装置,可以提高基于RNN神经网络的分类模型对视频的识别精度。

[0038] 为便于对本实施例进行理解,首先对本申请实施例所公开的一种分类模型训练方法进行详细介绍。通过本申请实施例所提供的分类模型训练方法得到的分类模型包括:目标神经网络以及记忆网络。在将该分类模型进行具体应用时,神经网络用于为待分类数据提取特征向量;记忆网络用于根据其所存储的不同分别对应的记忆特征向量,基于为分类数据提取的特征向量,为特征向量保证的待分类数据进行分类,得到待分类数据的分类结果。

[0039] 参见图1所示,本申请实施例一提供的分类模型训练方法包括:

[0040] S101:获取带有标签的训练视频中预设数量训练图像帧。

[0041] 在具体实现的时候,训练视频通常是一段比较长的视频,其一般包括一个完整的动作,方便对其进行标签的标注。使用整个的训练视频对分类模型进行训练,通常都会存在由于输入的数据量较大而造成模型收敛速度降低,训练过程需要耗费的时间长,资源多等问题。因此,为了加快模型收敛,减少模型训练过程中需要耗费的时间和资源,需要从整个的训练视频中获取预设数量的训练图像帧。

[0042] 对应不同的场景,训练图像帧的获取方式也有所不同。

[0043] 例如将该分类模型应用在对视频进行标签标注场景下时,可以采用按照预设的频率对训练视频进行采样,将采样获得的帧作为训练图像帧,然后基于该训练图像帧对神经网络进行训练。

[0044] 在将该分类模型应用在无人机、自动驾驶场景下时,其应用场景决定了在外界发生动作的时候,就要能够根据外界发生的动作及时的做出反应,因而通常情况下都不能在获得一段包括了完整动作的视频后,基于该包括了完整动作的视频对视频进行分类,而是需要在外界动作刚刚发生,或者发生到一定阶段的时候,就要对外界动作有明确的判断,因此,在此种场景下,通常会采用获取训练视频前预设数量的图像作为训练图像帧,以能够在所获得的视频并未包括一个完整的动作,能够对该动作进行精准判断。

[0045] 在对分类模型进行训练的时候,可以将视频文件中前2%-10%的图像作为训练图像帧。具体地,图像的数量,可以根据训练视频的训练难度进行具体的设定;训练难度较大的,可以取较多的图像作为训练图像帧,例如车辆在执行拐弯和并线两个动作的时候,车头

以及车轮都会有一定的偏转,但车辆在拐弯时,车头以及车轮的偏转角度一般都会大于并线,因此若训练图像帧的数量过少,在无法明确判断出车头以及车轮偏转角度的时候,就会造成模型对这两个动作的分类不准确;此时就需要相应增加训练图像帧的数量,使得分类模型能够在训练过程中更多的学习到各训练图像帧的内在联系,进而能够对这两个动作进行更精确的分类。

[0046] S102:使用目标神经网络对训练图像帧进行特征学习,为训练图像帧提取特征向量。

[0047] 在具体实现的时候,目标神经网络可以采用RNN神经网络。在使用目标神经网络对训练图像帧进行特征学习,为训练图像帧提取特征向量的时候,可以对所有获得的训练图像帧提取特征向量,也可以对按照一定的规则从中选取部分训练图像帧提取特征向量。

[0048] S103:基于记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的训练视频进行分类,得到训练视频的分类结果。

[0049] 在具体实现的时候,记忆网络具有记忆能力,是指记忆网络能够存储与不同分类分别对应的记忆特征向量。该记忆特征向量有两个来源,其一,使用辅助神经网络为辅助训练视频中辅助训练图像帧提取的特征向量;其二,使用目标神经网络为训练图像帧提取的特征向量,如,在对目标神经网络训练的过程中,可以对记忆网络中的各记忆单元进行动态更新,在训练视频的分类结果以及训练视频的标签一致时,将此时目标神经网络为该训练视频的训练图像帧提取的特征向量作为该分类下的记忆特征向量,添加至相应的记忆单元中,对记忆网络进行更新。

[0050] 此处,辅助神经网络可以是一个通过辅助训练视频训练得到的分类模型,该辅助神经网络能够实现对辅助训练视频的分类。其可以和目标神经网络为同一个神经网络,也可以是不同的神经网络。辅助训练视频和训练视频来源于同一个视频集,两者可以相同,也可以不同;当两者不同时,辅助训练视频中的图像和训练视频中的图像具有相同或者相似的数据分布。

[0051] 具体地,在分类模型开始训练之前,在记忆网络中与不同分类分别对应的记忆特征向量为空时,首先要在记忆网络中生成记忆单元,每一个记忆单元对应一个分类,每一个记忆单元中存储有与该记忆单元对应的分类所对应的记忆特征向量。

[0052] 在记忆网络中生成记忆单元的时候,首先将带有标签的辅助训练视频中的预设数量辅助训练图像帧输入至辅助神经网络,使用该辅助神经网络为每一个训练视频的辅助训练图像帧提取特征向量,并将得到的特征向量传输至记忆网络;记忆网络根据辅助训练视频对应的标签,将该特征向量作为记忆特征向量,保存至与标签对应的分类下,生成记忆单元。

[0053] 在生成各分类对应的记忆单元后,要对分类模型进行训练。

[0054] 针对辅助神经网络和目标神经网络为同一神经网络的情况,将训练视频中预设数量的训练图像帧输入至该辅助神经网络中,使用辅助神经网络对训练图像帧进行特征学习,为训练图像帧提取特征向量。

[0055] 针对辅助神经网络和目标神经网络并非同一神经网络的情况,则首先要构造目标神经网络,并对目标神经网络的参数进行初始赋值;然后将训练视频中预设数量的训练图像帧输入至进行初始赋值的目标神经网络中,使用该目标神经网络对训练图像帧进行特征

学习,为训练图像帧提取特征向量。

[0056] 在为训练图像帧提取了特征向量之后,将特征向量传输至记忆网络,基于记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的训练视频进行分类,从而得到训练视频的分类结果。

[0057] 参见图2所示,本申请实施例二提供一种基于记忆网络存储的不同类别分别对应的记忆特征向量,对提取的特征向量表征的训练视频进行分类,得到训练视频的分类结果的具体方法,该方法包括:

[0058] S201:将提取的特征向量分别与记忆网络中不同分类所对应的记忆特征向量进行相似度匹配,得到每个分类与提取的特征向量的相似度匹配结果。

[0059] 在具体实现的时候,可以使用每个分类对应的记忆单元中所有记忆特征向量与提取的特征向量进行相似度匹配;可选地,由于每个分类对应的记忆单元中都存储了多个记忆特征向量,而每一个训练视频中的训练图像帧的数量是有限的,且随着目标神经网络的训练过程,每一个记忆单元中的记忆特征向量也会越来越多。在通常情况下,如果将为训练视频中训练图像帧提取特征向量依次与每个记忆单元中的每个记忆特征向量进行相似度匹配,会导致计算量够大,且会耗费过多的计算时间。因此,一般地,可以从每个分类对应的记忆单元中随机抽取预设数量的记忆特征向量与提取的特征向量进行相似度匹配。

[0060] S202:将得到相似度匹配结果中匹配程度最高的分类确定为训练视频的分类结果。

[0061] 参见图3所示,本申请实施例三还提供一种将提取的特征向量分别与记忆网络中不同分类所对应的记忆特征向量进行相似度匹配的具体方法,该方法包括:

[0062] S301:分别对记忆网络中各分类对应的记忆特征向量进行聚类。

[0063] S302:对提取的特征向量和聚类结果作相似度度量。

[0064] 在具体实现的时候,可以分别将为训练图像帧提取的特征向量,以及记忆特征向量看作是映射到高维空间中的点,根据点与点之间的距离,分别对这些点进行聚类。将聚类在预设阈值以内的点划分到同一个类中。

[0065] 这里,由于在记忆单元中的记忆特征向量表征的辅助训练图像帧都属于同一分类,因此在理论上,每一个记忆单元中所包括的全部记忆特征向量在聚类后,都应当聚类为同一个类。然后对提取的特征向量和对每个记忆单元的聚类结果做相似度度量,能够得到提取的特征向量与每个记忆单元中特征向量的相似度匹配结果。

[0066] 参见图4所示,本申请实施例四还提供一种对提取的特征向量和聚类结果做相似度度量的具体方法,包括:

[0067] S401:根据预设的基准向量确认规则,为聚类结果中得到的每个聚类确认基准向量。

[0068] 在具体实现的时候,基准向量确认规则可以根据实际的需要进行确定,例如可以将每个记忆单元中记忆特征向量聚类结果中离中心点最近的记忆特征向量作为基准向量,或者随机抽取任意一个点对应的记忆特征向量作为基准特征向量,或者将聚类结果中离中心点最远的记忆特征向量作为基准向量。

[0069] S402:针对每个提取的特征向量,计算该特征向量与每个分类对应的基准向量之间的距离。

[0070] 此处,提取的特征向量与每个分类对应的基准向量之间的距离,可以是欧式距离、曼哈顿距离、切比雪夫距离、闵可夫斯基距离、标准化欧式距离、马氏距离、夹角余弦、汉明距离、杰卡德距离、相关距离以及信息熵中任意一种。

[0071] S403:将与该特征向量距离最小的基准向量所在的分类,作为该特征向量所对应训练图像帧的分类;

[0072] S404:将对应有训练图像帧最多的分类,作为训练视频的分类结果。

[0073] 在具体实现时,假如记忆网络中的分类有A、B、C、D、E五种,训练视频中包括10张训练图像帧,其中,有三张训练图像帧的特征向量与分类A中的基准向量之间的距离最小;有五张训练图像帧的特征向量与分类C中基准向量之间的距离最小;有两张训练图像帧的特征向量与分类D中基准向量之间的距离最小,则最终确定训练视频的分类结果为C。

[0074] 在得到训练视频的分类结果后,要基于该分类结果对目标神经网络记忆训练。也即:

[0075] S104:根据训练视频的分类结果以及训练视频的标签之间的比对结果,对目标神经网络进行训练。

[0076] 在具体实现的时候,对目标神经网络进行训练的过程,实际上是使用目标神经网络为训练图像帧提取的特征向量越来越趋向于正确分类中记忆特征向量的过程。

[0077] 本申请实施例五还提供一种根据训练视频的分类结果以及训练视频的标签之间的比对结果,对目标神经网络进行训练的具体方法,包括:执行如下比对操作,直至训练视频的分类结果与训练视频的标签一致;

[0078] 参见图5所示,该比对操作包括:

[0079] S501:比对训练视频的分类结果与训练视频的标签是否一致;若是,则跳转至502;若否,则跳转至S503。

[0080] S502:完成对目标神经网络的本轮训练。本流程结束。

[0081] S503:对目标神经网络的参数进行调整,并基于调整后的参数,使用目标神经网络为训练图像帧提取新的特征向量。

[0082] S504:基于记忆网络存储的不同分类分别对应的记忆特征向量以及新的特征向量,对训练视频进行分类,得到训练视频的新的分类结果,并再次执行S501。

[0083] 具体地,要保证为使用目标神经网络为训练图像帧提取的特征向量越来越接近正确分类中记忆特征向量,那么就要在对训练视频的分类结果与训练视频的标签不一致时,调整神经网络的参数,使得使用调整了参数后的神经网络为训练图像帧提取的特征向量后,较之调整参数之前,要更接近于正确分类中记忆特征向量。

[0084] 本申请实施例提供的分类模型训练方法,通过目标神经网络和记忆网络共同构成分类模型;在对分类模型进行训练的时候,将电邮标签的训练视频中预设数量训练图像帧输入至目标神经网络,会使用目标神经网络为训练图像帧提取特征向量,然后将提取的特征向量输入至记忆网络,在记忆网络中存储了不同分类分别对应的记忆特征向量;当记忆网络接收到提取的特征向量后,能够基于其所存储的不同分类分别对应的记忆特征向量,对提取的特征向量标准横的训练视频进行分类,得到训练视频的分类结果,然后根据训练视频的分类结果以及分类视频的标签之间的比对结果,对目标神经网络进行训练。在该分类模型中,记忆网络具有强大的记忆能力,能够弥补神经网络本身存在的记忆能力差的问题。

题,不论对于任何长度的输入序列,记忆网络都能够对其特征进行很好的记忆,因而不论输入序列为何,都不会由于神经网络记忆能力差而影响分类模型的识别精度。

[0085] 参见图6所示,本申请实施例六还提供另外一种分类模型训练方法,该方法主要用于对视频中动作事件的预测。该方法包括:

[0086] S601:获取带有标签的训练视频中,前预设数量训练图像帧。

[0087] S602:使用目标神经网络对S601中的训练图像帧进行特征学习,为训练图像帧提取特征向量。

[0088] S603:基于记忆网络存储的不同动作事件分别对应的记忆特征向量,对训练视频中,提取的特征向量对应的训练图像帧之后的图像帧包含的动作事件进行预测,得到训练视频的预测结果;

[0089] S604:根据训练视频的预测结果以及训练视频的标签之间的比对结果,对目标神经网络进行训练。

[0090] 其中,上述S601-S604可以参见上述S101-S104的描述,在此不再赘述。

[0091] 考虑如下场景:汽车在自动驾驶时,需要根据当前获取的图像判断图像中移动物体下一步动作,从而判断自身应该做出的动作反应。针对这种情况,可以基于本发明实施例提供的分类模型训练方法训练动作事件预测模型。

[0092] 在训练过程中,将动作事件视频作为训练样本,获取带有标签的训练视频中前预设数量训练图像帧(例如:前20%的训练图像帧),即动作事件发生之前的起始动作,并且可以将动作视频中该起始动作之后引发的动作事件作为分类,不同的起始动作对应的不同的动作事件对应不同的分类,通过上述实施例所提供的分类模型训练方法训练得到的分类模型,能够根据起始动作确定出对应的分类,即该起始动作引发后续动作事件对应的分类,从而实现了通过起始动作对后续动作事件的预测。

[0093] 在需要进行动作事件预测的场景中,例如:自动驾驶、无人机等,使用本发明实施例提供的分类模型时,可以根据当前拍摄的画面内容中包含的起始动作对该起始动作引发的后续动作事件进行准确预测,从而做出相应的反应。

[0094] 基于同一申请构思,本申请实施例中还提供了与分类模型训练方法对应的分类模型训练装置,由于本申请实施例中的装置解决问题的原理与本申请实施例上述分类模型训练方法相似,因此装置的实施可以参见方法的实施,重复之处不再赘述。

[0095] 参见图7所示,本申请实施例七提供的分类模型训练装置,包括:

[0096] 获取模块71,用于获取带有标签的训练视频中预设数量训练图像帧;

[0097] 提取模块72,用于使用目标神经网络对训练图像帧进行特征学习,为训练图像帧提取特征向量;

[0098] 分类模块73,用于基于记忆网络存储的不同分类分别对应的记忆特征向量,对提取的特征向量表征的训练视频进行分类,得到训练视频的分类结果;

[0099] 训练模块74,用于根据训练视频的分类结果以及训练视频的标签之间的比对结果,对目标神经网络进行训练。

[0100] 本申请实施例提供的分类模型训练装置,通过目标神经网络和记忆网络共同构成分类模型;在对分类模型进行训练的时候,将电邮标签的训练视频中预设数量训练图像帧输入至目标神经网络,会使用目标神经网络为训练图像帧提取特征向量,然后将提取的特

征向量输入至记忆网络,在记忆网络中存储了不同分类分别对应的记忆特征向量;当记忆网络接收到提取的特征向量后,能够基于其所存储的不同分类分别对应的记忆特征向量,对提取的特征向量标准横的训练视频进行分类,得到训练视频的分类结果,然后根据训练视频的分类结果以及分类视频的标签之间的比对结果,对目标神经网络进行训练。在该分类模型中,记忆网络具有强大的记忆能力,能够弥补神经网络本身存在的记忆能力差的问题,不论对于任何长度的输入序列,记忆网络都能够对其特征进行很好的记忆,因而不论输入序列为何,都不会由于神经网络记忆能力差而影响分类模型的识别精度。

[0101] 可选地,提取模块73,具体用于将提取的特征向量分别与记忆网络中不同分类所对应的记忆特征向量进行相似度匹配,得到每个分类与提取的特征向量的相似度匹配结果;

[0102] 将得到相似度匹配结果中匹配程度最高的分类确定为训练视频的分类结果。

[0103] 可选地,提取模块73,具体用于通过如下步骤将提取的特征向量分别与记忆网络中不同分类所对应的记忆特征向量进行相似度匹配:

[0104] 分别对记忆网络中各分类对应的记忆特征向量进行聚类;

[0105] 对提取的特征向量和聚类结果作相似度度量。

[0106] 可选地,提取模块73,具体用于通过如下步骤对提取的特征向量和聚类结果做相似度度量:

[0107] 根据预设的基准向量确认规则,为聚类结果中得到的每个聚类确认基准向量;

[0108] 针对每个提取的特征向量,计算该特征向量与每个分类对应的基准向量之间的距离;以及

[0109] 将与该特征向量距离最小的基准向量所在的分类,作为该特征向量所对应训练图像帧的分类;

[0110] 将对应有训练图像帧最多的分类,作为训练视频的分类结果。

[0111] 可选地,训练模块74,具体用于:执行如下比对操作,直至训练视频的分类结果与训练视频的标签一致;比对操作包括:

[0112] 将训练视频的分类结果与训练视频的标签进行比对;

[0113] 若训练视频的分类结果与训练视频的标签不一致,则对目标神经网络的参数进行调整;

[0114] 基于调整后的参数,使用目标神经网络为训练图像帧提取新的特征向量;

[0115] 基于记忆网络存储的不同分类分别对应的记忆特征向量以及新的特征向量,对训练视频进行分类,得到训练视频的新的分类结果,并再次执行比对操作。

[0116] 可选地,训练模块74,还用于在分类结果与训练视频的标签一致时,将对应提取的特征向量添加至记忆网络对应分类中,对记忆网络进行更新。

[0117] 可选地,提取模块72,具体用于获取带有标签的训练视频中,前预设数量训练图像帧;

[0118] 分类模块73,具体用于基于记忆网络存储的不同动作事件分别对应的记忆特征向量,对训练视频中,提取的特征向量对应的训练图像帧之后的图像帧包含的动作事件进行预测,得到训练视频的预测结果;

[0119] 训练模块74,具体用于根据训练视频的预测结果以及训练视频的标签之间的比对

结果,对目标神经网络进行训练。

[0120] 参见图8所示,本申请实施例八还提供一种分类方法,该方法包括:

[0121] S801:获取待分类视频。

[0122] S802:将待分类视频输入至通过本申请实施例提供的分类模型训练方法得到的分类模型中,得到待分类视频的分类结果;

[0123] 其中,分类模型包括:目标神经网络以及记忆网络。

[0124] 采用本申请实施例提供的分类模型训练方法得到的分类模型,对待分类视频进行分类,能够得到更加精确的分类结果。

[0125] 参见图9所示,本申请实施例九还提供一种分类装置,包括:

[0126] 待分类数据获取模块91,用于获取待分类视频;

[0127] 分类模块92,用于将待分类视频输入至通过本申请实施例提供的分类模型训练方法得到的分类模型中,得到待分类视频的分类结果;

[0128] 其中,分类模型包括:目标神经网络以及记忆网络。

[0129] 对应于图1中的分类模型训练方法,本申请实施例还提供了一种计算机设备,如图10所示,该设备包括存储器1000、处理器2000及存储在该存储器1000上并可在该处理器2000上运行的计算机程序,其中,上述处理器2000执行上述计算机程序时实现上述分类模型训练方法的步骤。

[0130] 具体地,上述存储器1000和处理器2000能够为通用的存储器和处理器,这里不做具体限定,当处理器2000运行存储器1000存储的计算机程序时,能够执行上述分类模型训练方法,从而解决RNN神经网络记忆能力差,无法对较长输入序列进行高精度分类的问题,进而达到提高基于RNN网络的分类模型对视频的识别精度的效果。

[0131] 对应于图1中的分类模型训练方法,本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质上存储有计算机程序,该计算机程序被处理器运行时执行上述分类模型训练方法的步骤。

[0132] 具体地,该存储介质能够为通用的存储介质,如移动磁盘、硬盘等,该存储介质上的计算机程序被运行时,能够执行上述分类模型训练方法,从而解决RNN神经网络记忆能力差,无法对较长输入序列进行高精度分类的问题,进而达到提高基于RNN网络的分类模型对视频的识别精度的效果。

[0133] 本申请实施例所提供的分类模型训练方法、装置以及分类方法及装置的计算机程序产品,包括存储了程序代码的计算机可读存储介质,程序代码包括的指令可用于执行前面方法实施例中的方法,具体实现可参见方法实施例,在此不再赘述。

[0134] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统 and 装置的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0135] 功能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存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本申请各个实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器

(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0136] 以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应所述以权利要求的保护范围为准。

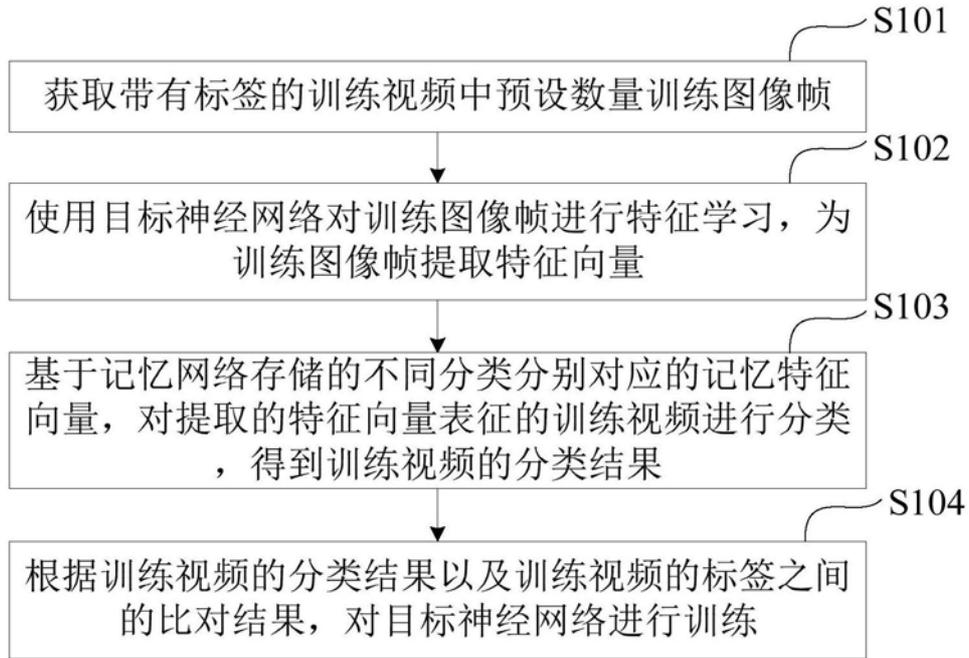


图1

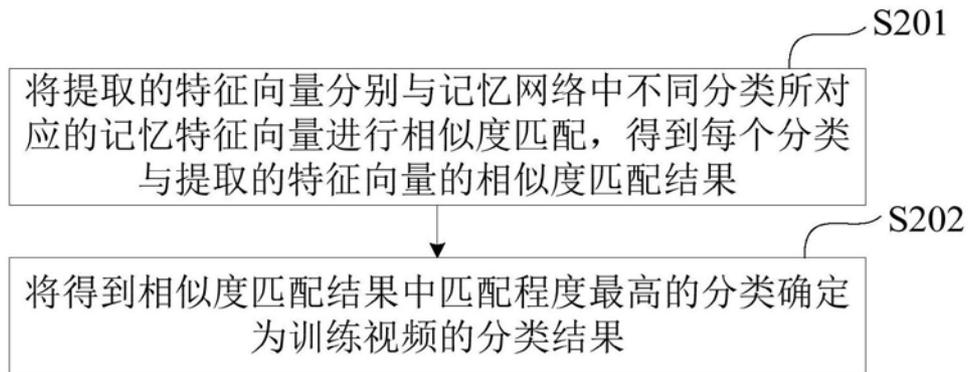


图2

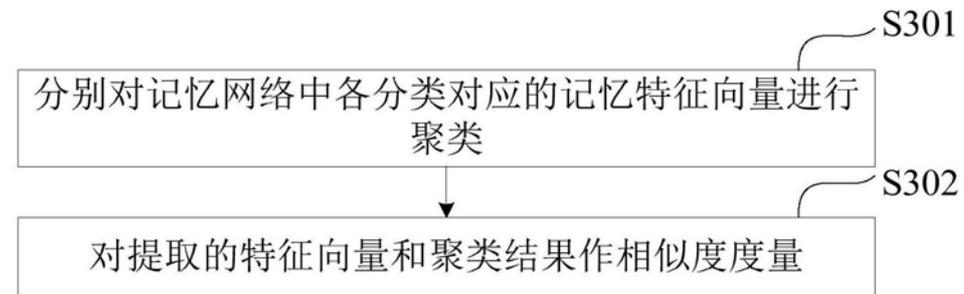


图3

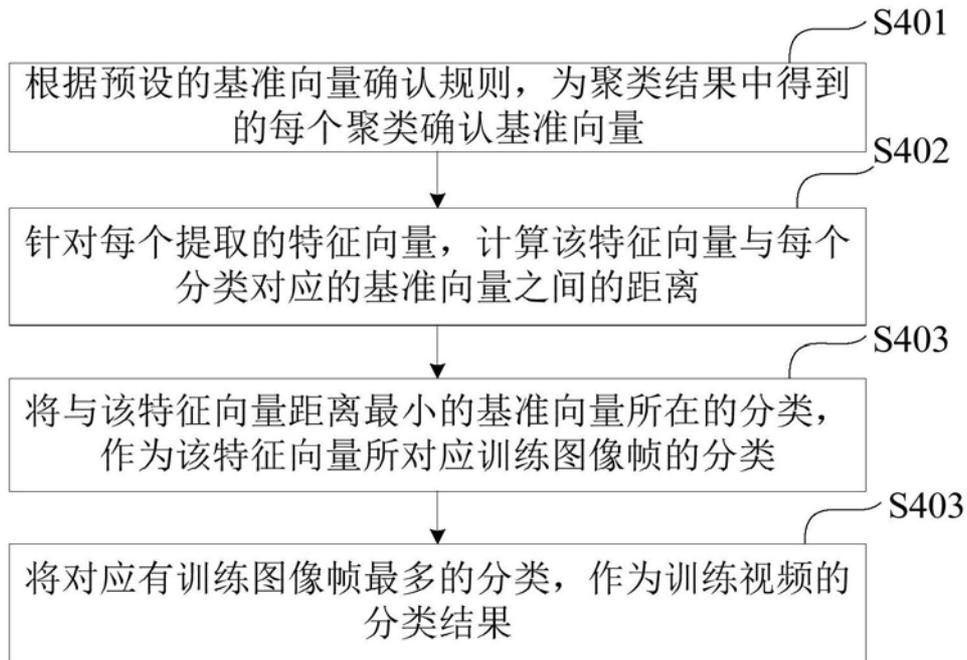


图4

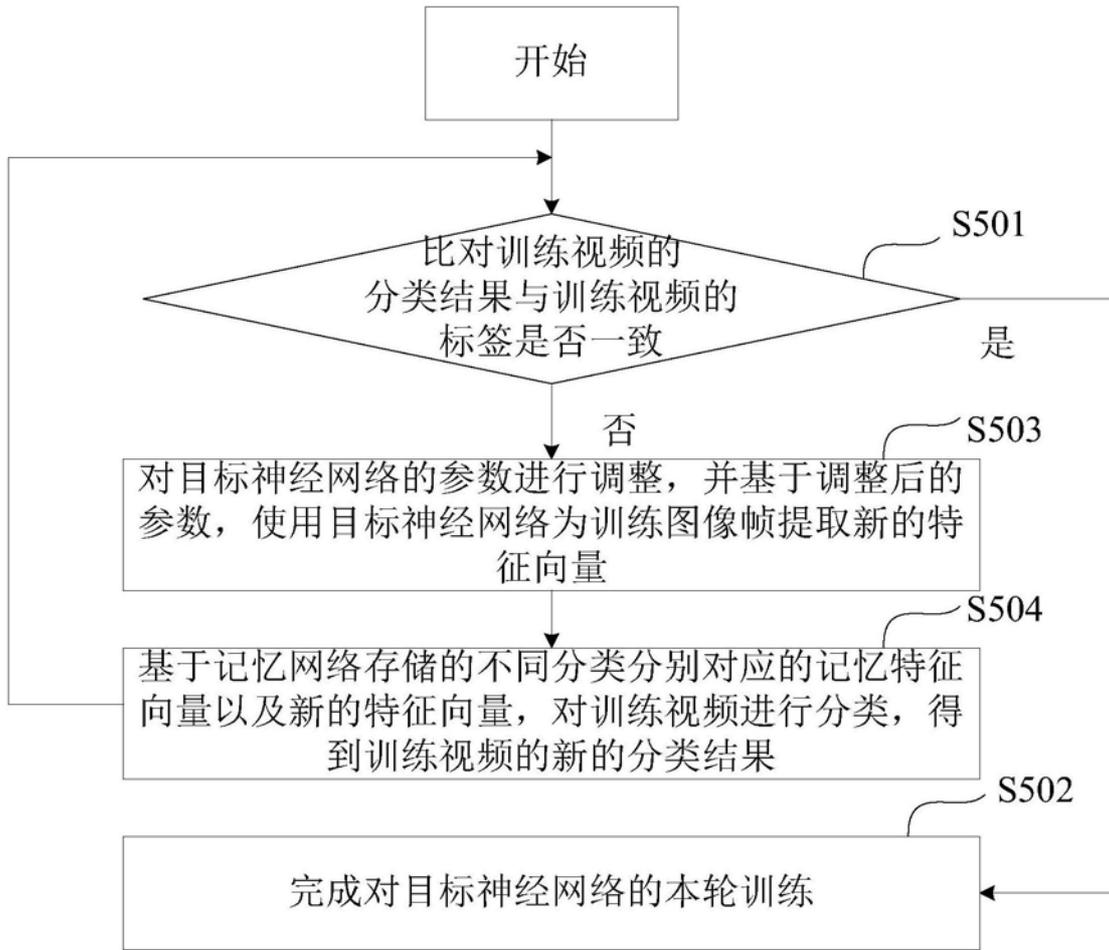


图5

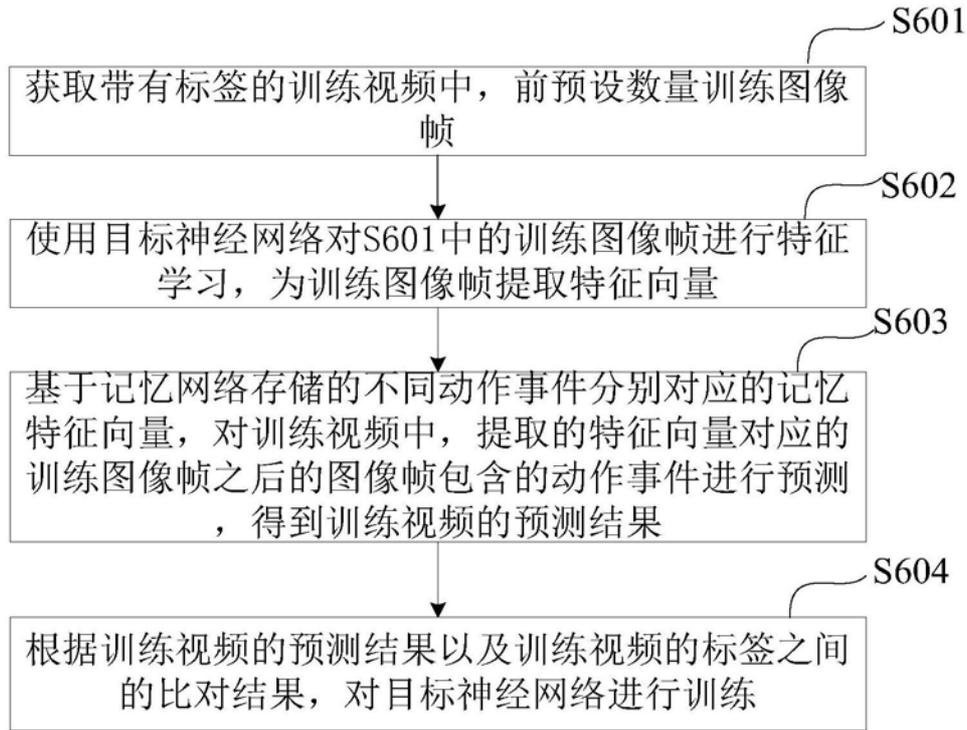


图6

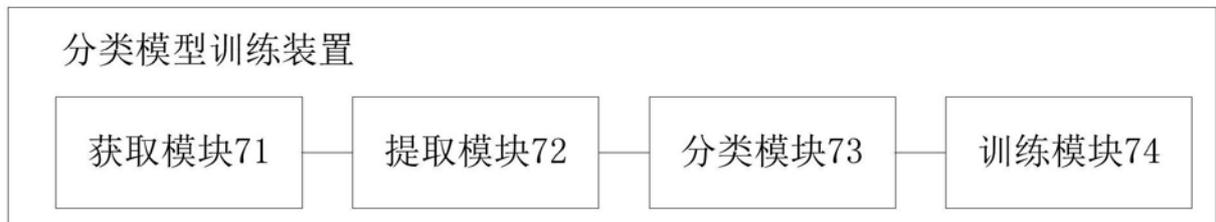


图7

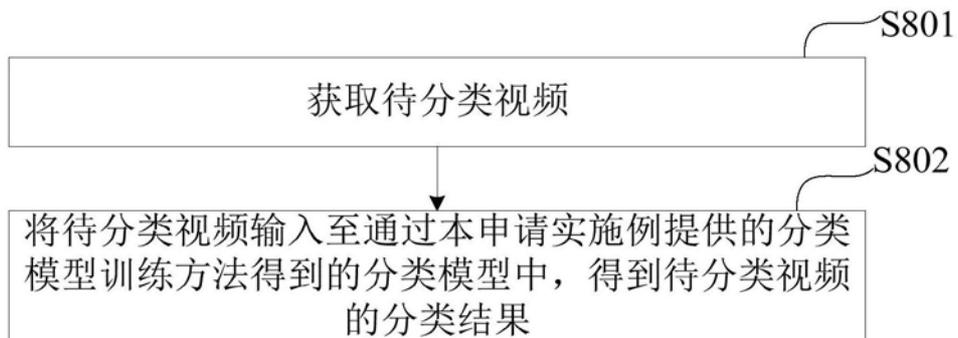


图8

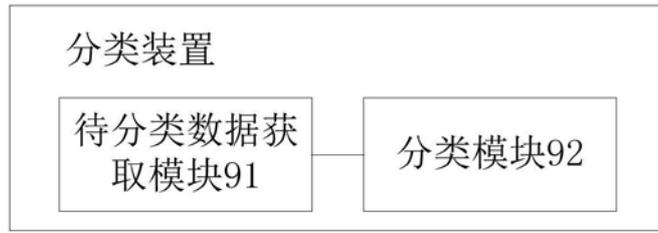


图9

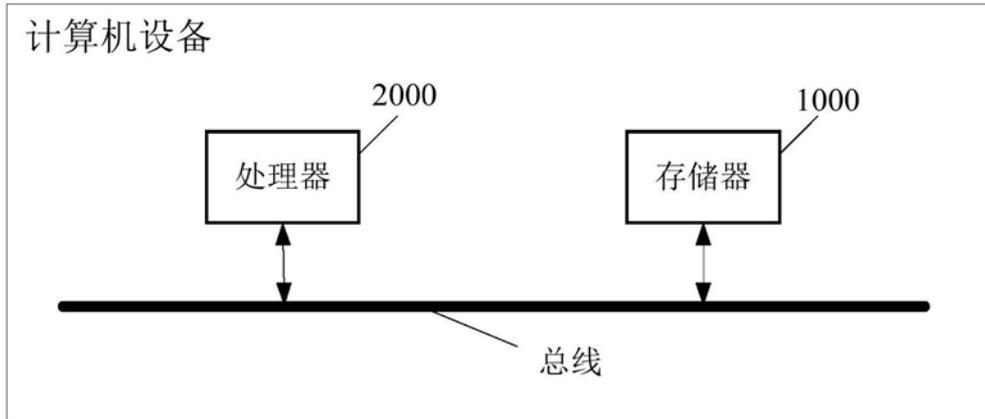


图10