



(12) 发明专利申请

(10) 申请公布号 CN 114494784 A

(43) 申请公布日 2022. 05. 13

(21) 申请号 202210110222.3

(22) 申请日 2022.01.28

(71) 申请人 北京百度网讯科技有限公司
地址 100085 北京市海淀区上地十街10号
百度大厦2层

(72) 发明人 张婉平

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021
专利代理师 鄢功军

(51) Int. Cl.
G06V 10/764 (2022.01)
G06V 10/74 (2022.01)
G06V 10/82 (2022.01)
G06N 3/08 (2006.01)

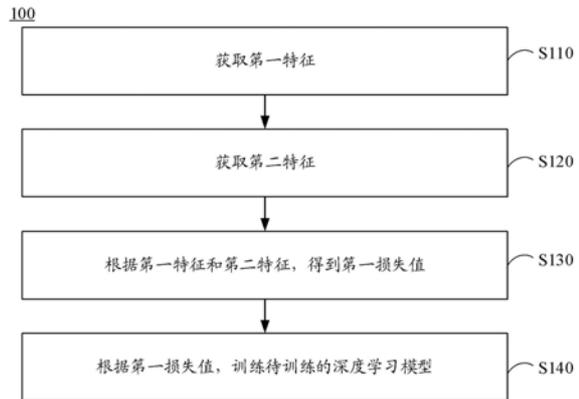
权利要求书3页 说明书10页 附图6页

(54) 发明名称

深度学习模型的训练方法、图像处理方法和对象识别方法

(57) 摘要

本公开提供了一种深度学习模型的训练方法,涉及人工智能技术领域,尤其涉及深度学习技术和计算机视觉技术。具体实现方案为:获取第一特征,其中,第一特征是利用参考深度学习模型处理第一目标图像得到的;获取第二特征,其中,第二特征是利用参考深度学习模型处理第二目标图像得到的;根据第一特征和第二特征,得到第一损失值;以及根据第一损失值,训练待训练的深度学习模型。本公开还提供了一种图像处理方法、对象识别方法、装置、电子设备和存储介质。



1. 一种深度学习模型的训练方法,包括:

获取第一特征,其中,所述第一特征是利用参考深度学习模型处理第一目标图像得到的;

获取第二特征,其中,所述第二特征是利用所述参考深度学习模型处理第二目标图像得到的;

根据所述第一特征和所述第二特征,得到第一损失值;以及

根据所述第一损失值,训练待训练的深度学习模型。

2. 根据权利要求1所述的方法,其中,所述根据所述第一特征和所述第二特征,得到第一损失值包括:

计算所述第一特征和所述第二特征之间的相似度;

根据所述相似度和预设相似度阈值,得到第一差异值;以及

根据所述第一差异值,得到所述第一损失值。

3. 根据权利要求2所述的方法,其中,所述根据所述第一差异值,得到所述第一损失值包括:

将所述第一目标图像输入所述待训练的深度学习模型,得到第三特征;

将所述第二目标图像输入所述待训练的深度学习模型,得到第四特征;

根据所述第三特征和所述第四特征,得到第二差异值;以及

根据所述第一差异值和所述第二差异值,得到所述第一损失值。

4. 根据权利要求3所述的方法,其中,所述根据所述第一差异值和所述第二差异值,得到所述第一损失值包括:

根据所述第一差异值和所述第二差异值,通过以下公式进行计算,得到所述第一损失值:

$$Loss_1 = \begin{cases} \exp(Diff_1) \times (Diff_2)^2, & y = 1 \\ \exp(-Diff_1) \times \max(\text{margin} - Diff_2, 0)^2, & y = 0 \end{cases}$$

其中, $Loss_1$ 为第一损失值, $Diff_1$ 为第一差异值, $Diff_2$ 为第二差异值, margin 为所述待训练的深度学习模型的预设类别阈值;在所述第一目标图像的标签表征的类别和所述第二目标图像的标签表征的类别相同的情况下, $y = 1$;在所述第一目标图像的标签表征的类别和所述第二目标图像的标签表征的类别不相同的情况下, $y = 0$ 。

5. 根据权利要求1所述的方法,其中,所述根据所述第一损失值,训练待训练的深度学习模型包括:

根据所述第一目标图像和所述第二目标图像,确定输入图像;

获取第一输出特征,其中,所述第一输出特征是利用所述参考深度学习模型处理所述输入图像得到的;

获取第二输出特征,其中,所述第二输出特征是利用所述待训练的深度学习模型处理所述输入图像得到的;

根据所述第一输出特征和所述第二输出特征,得到第二损失值;以及

根据所述第一损失值和所述第二损失值,训练所述待训练的深度学习模型。

6. 根据权利要求5所述的方法,其中,所述根据所述第一损失值和所述第二损失值,训

练所述待训练的深度学习模型包括：

根据预设权重、所述第一损失值和所述第二损失值，得到加权损失值；以及根据所述加权损失值，训练所述待训练的深度学习模型。

7. 根据权利要求1至6任一项所述的方法，其中，所述参考深度学习模型的参数多于所述待训练的深度学习模型的参数。

8. 一种图像处理方法，包括：

将图像集中的目标图像输入深度学习模型，得到所述目标图像中对象的分类类别；根据所述分类类别和所述目标图像的标签，得到第三差异值；以及响应于第三差异值大于或等于预设差异阈值，从所述图像集中删除所述目标图像，其中，所述深度学习模型是根据权利要求1至7任一项所述的方法训练的。

9. 一种对象识别方法，包括：

将目标图像输入深度学习模型，得到所述目标图像中对象的分类类别，其中，所述深度学习模型是根据权利要求1至7任一项所述的方法训练的。

10. 一种深度学习模型的训练装置，包括：

第一获取模块，用于获取第一特征，其中，所述第一特征是利用参考深度学习模型处理第一目标图像得到的；

第二获取模块，用于获取第二特征，其中，所述第二特征是利用所述参考深度学习模型处理第二目标图像得到的；

第一获得模块，用于根据所述第一特征和所述第二特征，得到第一损失值；以及训练模块，用于根据所述第一损失值，训练待训练的深度学习模型。

11. 根据权利要求10所述的装置，其中，所述第一获得模块包括：

计算子模块，用于计算所述第一特征和所述第二特征之间的相似度；

第一获得子模块，用于根据所述相似度和预设相似度阈值，得到第一差异值；以及

第二获得子模块，用于根据所述第一差异值，得到所述第一损失值。

12. 根据权利要求11所述的装置，其中，所述第二获得子模块包括：

第一输入单元，用于将所述第一目标图像输入所述待训练的深度学习模型，得到第三特征；

第二输入单元，用于将所述第二目标图像输入所述待训练的深度学习模型，得到第四特征；

第一获得单元，用于根据所述第三特征和所述第四特征，得到第二差异值；以及

第二获得单元，用于根据所述第一差异值和所述第二差异值，得到所述第一损失值。

13. 根据权利要求12所述的装置，其中，所述第二获得单元还用于：

根据所述第一差异值和所述第二差异值，通过以下公式进行计算，得到所述第一损失值：

$$Loss_1 = \begin{cases} \exp(Diff_1) \times (Diff_2)^2 & , y = 1 \\ \exp(-Diff_1) \times \max(\text{margin} - Diff_2, 0)^2 & , y = 0 \end{cases}$$

其中，Loss_1为第一损失值，Diff为第一差异值，d为第二差异值，margin为所述待训练的深度学习模型的预设类别阈值；在所述第一目标图像的标签表征的类别和所述第二目标图像的标签表征的类别相同的情况下， $y=1$ ；在所述第一目标图像的标签表征的类别和所

述第二目标图像的标签表征的类别不相同的情况下, $y=0$ 。

14. 根据权利要求10所述的装置, 其中, 所述训练模块包括:

确定子模块, 用于根据所述第一目标图像和所述第二目标图像, 确定输入图像;

第一获取子模块, 用于获取第一输出特征, 其中, 所述第一输出特征是利用所述参考深度学习模型处理所述输入图像得到的;

第二获取子模块, 用于获取第二输出特征, 其中, 所述第二输出特征是利用所述待训练的深度学习模型处理所述输入图像得到的;

第三获得子模块, 用于根据所述第一输出特征和所述第二输出特征, 得到第二损失值; 以及

训练子模块, 用于根据所述第一损失值和所述第二损失值, 训练所述待训练的深度学习模型。

15. 根据权利要求14所述的装置, 其中, 所述训练子模块包括:

第三获得单元, 用于根据预设权重、所述第一损失值和所述第二损失值, 得到加权损失值; 以及

训练单元, 用于根据所述加权损失值, 训练所述待训练的深度学习模型。

16. 根据权利要求10至15任一项所述的装置, 其中, 所述参考深度学习模型的参数多于所述待训练的深度学习模型的参数。

17. 一种图像处理装置, 包括:

第二获得模块, 用于将图像集中的目标图像输入深度学习模型, 得到所述目标图像中对象的分类类别;

第三获得模块, 用于根据所述分类类别和所述目标图像的标签, 得到第三差异值; 以及

删除模块, 用于响应于第三差异值大于或等于预设差异阈值, 从所述图像集中删除所述目标图像,

其中, 所述深度学习模型是根据权利要求10至16任一项所述的装置训练的。

18. 一种对象识别装置, 包括:

第四获得模块, 用于将目标图像输入深度学习模型, 得到所述目标图像中对象的分类类别,

其中, 所述深度学习模型是根据权利要求10至16任一项所述的装置训练的。

19. 一种电子设备, 包括:

至少一个处理器; 以及

与所述至少一个处理器通信连接的存储器; 其中,

所述存储器存储有可被所述至少一个处理器执行的指令, 所述指令被所述至少一个处理器执行, 以使所述至少一个处理器能够执行权利要求1至9中任一项所述的方法。

20. 一种存储有计算机指令的非瞬时计算机可读存储介质, 其中, 所述计算机指令用于使所述计算机执行根据权利要求1至9中任一项所述的方法。

21. 一种计算机程序产品, 包括计算机程序, 所述计算机程序在被处理器执行时实现根据权利要求1至9中任一项所述的方法。

深度学习模型的训练方法、图像处理方法和对象识别方法

技术领域

[0001] 本公开涉及人工智能技术领域,尤其涉及深度学习技术和计算机视觉技术。更具体地,本公开提供了一种深度学习模型的训练方法、图像处理方法、对象识别方法、装置、电子设备和存储介质。

背景技术

[0002] 可以利用深度学习模型进行对象识别。而在深度学习模型的训练过程中,需要大量的训练数据。

发明内容

[0003] 本公开提供了一种深度学习模型的训练方法、图像处理方法、对象识别方法、装置、设备以及存储介质。

[0004] 根据第一方面,提供了一种深度学习模型的训练方法,该方法包括:获取第一特征,其中,上述第一特征是利用参考深度学习模型处理第一目标图像得到的;获取第二特征,其中,上述第二特征是利用上述参考深度学习模型处理第二目标图像得到的;根据上述第一特征和上述第二特征,得到第一损失值;以及根据上述第一损失值,训练待训练的深度学习模型。

[0005] 根据第二方面,提供了一种图像处理方法,该方法包括:将图像集中的目标图像输入深度学习模型,得到上述目标图像中对象的分类类别;根据上述分类类别和上述目标图像的标签,得到第三差异值;以及响应于第三差异值大于或等于预设差异阈值,从上述图像集中删除上述目标图像,其中,上述深度学习模型是根据本公开提供的方法训练的。

[0006] 根据第三方面,提供了一种对象识别方法,该方法包括:将目标图像输入深度学习模型,得到上述目标图像中对象的分类类别,其中,上述深度学习模型是根据本公开提供的方法训练的。

[0007] 根据第四方面,提供了一种深度学习模块的训练装置,该装置包括:第一获取模块,用于获取第一特征,其中,上述第一特征是利用参考深度学习模型处理第一目标图像得到的;第二获取模块,用于获取第二特征,其中,上述第二特征是利用上述参考深度学习模型处理第二目标图像得到的;第一获得模块,用于根据上述第一特征和上述第二特征,得到第一损失值;以及训练模块,用于根据上述第一损失值,训练待训练的深度学习模型。

[0008] 根据第五方面,提供了一种图像处理装置,该装置包括:第二获得模块,用于将图像集中的目标图像输入深度学习模型,得到上述目标图像中对象的分类类别;第三获得模块,用于根据上述分类类别和上述目标图像的标签,得到第三差异值;以及删除模块,用于响应于第三差异值大于或等于预设差异阈值,从上述图像集中删除上述目标图像,其中,上述深度学习模型是根据本公开提供的装置训练的。

[0009] 根据第六方面,提供了一种对象识别装置,该装置包括:第四获得模块,用于将目标图像输入深度学习模型,得到上述目标图像中对象的分类类别,其中,上述深度学习模型

是根据本公开提供的装置训练的。

[0010] 根据第七方面,提供了一种电子设备,包括:至少一个处理器;以及与至少一个处理器通信连接的存储器;其中,存储器存储有可被至少一个处理器执行的指令,指令被至少一个处理器执行,以使至少一个处理器能够执行根据本公开提供的方法。

[0011] 根据第八方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,该计算机指令用于使计算机执行根据本公开提供的方法。

[0012] 根据第九方面,提供了一种计算机程序产品,包括计算机程序,上述计算机程序在被处理器执行时实现根据本公开提供的方法。

[0013] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0014] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0015] 图1是根据本公开的一个实施例的深度学习模型的训练方法的流程图;

[0016] 图2是根据本公开的另一个实施例的深度学习模型的训练方法的流程图;

[0017] 图3是根据本公开的另一个实施例的深度学习模型的训练方法的流程图;

[0018] 图4A是根据本公开的一个实施例的图像集中的一个图像的示意图;

[0019] 图4B是根据本公开的一个实施例的图像集中的另一个图像的示意图;

[0020] 图4C是根据本公开的一个实施例的深度学习模型的训练方法的原理图;

[0021] 图5是根据本公开的一个实施例的图像处理方法的流程图;

[0022] 图6是根据本公开的一个实施例的对象识别方法的流程图;

[0023] 图7是根据本公开的一个实施例的深度学习模型的训练装置的框图;

[0024] 图8是根据本公开的一个实施例的图像处理装置的框图;

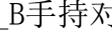
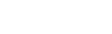
[0025] 图9是根据本公开的一个实施例的对象识别装置的框图;以及

[0026] 图10是根据本公开的一个实施例的可以应用深度学习的训练方法、图像处理方法和/或对象识别方法的电子设备的框图。

具体实施方式

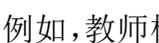
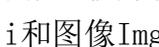
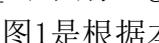
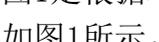
[0027] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0028] 在深度学习模型的训练过程中,需要大量的训练数据。可以通过各种手段获取训练数据。例如,可以从网络上获取某个对象的图像,以获取训练数据。

[0029] 但是,训练数据会存在脏数据。例如,训练数据集为图像集。在图像集中,图像的标签为对象Obj_A,图像的标签也为对象Obj_A。但,图像为对象Obj_B手持对象Obj_A的照片的图像。在此情况下,图像为图像集中的脏数据。

[0030] 教师模型通常是单个复杂网络或若干网络的集合,具有良好的性能和泛化能力。学生模型由于网络规模较小,表达能力有限。因此,可以利用教师模型学习到的知识指导学

生模型训练,使得学生模型具有与教师模型相当的性能。但学生模型的参数量大为降低,从而实现模型压缩和加速。

[0031] 例如,教师模型可以高效区分图像和图像。而学生模型无法高效区分图像和图像。

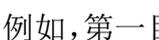
[0032] 图1是根据本公开的一个实施例的深度学习模型的训练方法的流程图。

[0033] 如图1所示,该方法100可以包括操作S110至操作S140。

[0034] 在操作S110,获取第一特征。

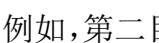
[0035] 例如,第一特征是利用参考深度学习模型处理第一目标图像得到的。

[0036] 例如,参考深度学习模型可以是教师模型。

[0037] 例如,第一目标图像可以是图像集中的图像。

[0038] 在操作S120,获取第二特征。

[0039] 例如,第二特征是利用参考深度学习模型处理第二目标图像得到的。

[0040] 例如,第二目标图像可以是图像集中的图像。

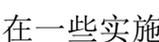
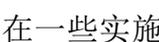
[0041] 在操作S130,根据第一特征和第二特征,得到第一损失值。

[0042] 例如,可以通过各种方式根据第一特征和第二特征,得到第一损失值。在一个示例中,可以根据第一特征和第二特征之间的距离,得到第一损失值。

[0043] 在操作S140,根据第一损失值,训练待训练的深度学习模型。

[0044] 例如,待训练的深度学习模型可以是学生模型。可以根据第一损失值调整待训练的深度学习模型的参数,以训练该深度学习模型。在一个示例中,待训练的深度学习模型可以是ResNet(残差网络)模型。

[0045] 通过本公开实施例,可以利用参考深度学习模型指导待训练的深度学习模型的训练,使得参考深度学习模型的表达能力可以有效地传递给待训练的深度学习模型,使得待训练的深度学习模型可以识别出图像集中的脏数据,也使得待训练的深度学习模型可以更加准确地识别出图像集中各图像的对象类别。

[0046] 在一些实施例中,与方法100不同之处在于,第二目标图像也可以是图像。图像的标签可以为对象Obj_B,即图像的标签可以与不同。

[0047] 在一些实施例,可以根据预设相似度阈值、第一特征和第二特征之间的相似度,得到第一损失值。

[0048] 图2是根据本公开的另一个实施例的深度学习模型的训练方法的流程图。

[0049] 如图2所示,该方法230可以根据第一特征和第二特征,得到第一损失值,下面将结合操作S231至操作S233进行详细说明。

[0050] 在操作S231,计算第一特征和第二特征之间的相似度。

[0051] 例如,本实施例中,第一特征可以是利用参考深度学习模型处理图像得到的,第二特征可以是利用参考深度学习模型处理图像得到的。

[0052] 例如,可以计算第一特征和第二特征之间的余弦相似度,作为二者的相似度S。

[0053] 可以理解,可以通过各种方式计算第一特征和第二特征之间的相似度,本公开在此不做限制。

[0054] 在操作S232,根据相似度和预设相似度阈值,得到第一差异值。

[0055] 例如,预设相似度阈值thred可以是针对参考深度学习模型设置的。可以将相似度

S和预设相似度阈值thred之间的差值,作为第一差异值Diff_1。

[0056] 在一个示例中,可以通过以下公式计算第一差异值:

[0057] $Diff_1 = S - thred$ (公式一)

[0058] 在操作S233,根据第一差异值,得到第一损失值。

[0059] 在本公开实施例中,可以将第一目标图像输入待训练的深度学习模型,得到第三特征。

[0060] 例如,可以将图像Img_i输入待训练的深度学习模型,得到第三特征。

[0061] 在本公开实施例中,可以将第二目标图像输入待训练的深度学习模型,得到第四特征。

[0062] 例如,可以将图像Img_j输入待训练的深度学习模型,得到第四特征。

[0063] 在本公开实施例中,可以根据第三特征和第四特征,得到第二差异值。

[0064] 例如,可以将第三特征和第四特征之间的欧式距离d,作为第二差异值Diff_2。

[0065] 在本公开实施例中,可以根据第一差异值和第二差异值,得到第一损失值。

[0066] 例如,可以通过以下公式进行计算,得到第一损失值Loss_1:

[0067] $Loss_1 = \begin{cases} \exp(Diff_1) \times (Diff_2)^2 & , y = 1 \\ \exp(-Diff_1) \times \max(\text{margin} - Diff_2, 0)^2 & , y = 0 \end{cases}$ (公式二)

[0068] margin为针对待训练的深度学习模型设置的预设类别阈值。在第一目标图像的标签表征的类别和第二目标图像的标签表征的类别相同的情况下, $y=1$;在第一目标图像的标签表征的类别和第二目标图像的标签表征的类别不相同的情况下, $y=0$ 。在一个示例中, $\exp(\cdot)$ 可以是以自然常数e为底的指数函数。 $\max(\cdot)$ 可以是求最大值的函数。

[0069] 在本实施例中,图像Img_i和图像Img_j的标签均为对象Obj_A,二者表征的类别相同。在此情况下, $y=1$ 。

[0070] 通过本公开实施例,在确定第一损失值时,增加了 $\exp(Diff_1)$ 或 $\exp(-Diff_1)$,使得待训练的深度学习模型可以进一步学习到参考深度学习模型的表达能力,提高了待训练的深度学习模型的性能。

[0071] 在一些实施例中,与方法230不同之处在于,第二特征可以是利用参考深度学习模型处理图像Img_k得到的。图像Img_k的标签可以为对象Obj_B。即在操作S233中,图像Img_i和图像Img_k表征的类别不相同。在此情况下,利用例如公式二计算第一损失值时, $y=0$ 。

[0072] 图3是根据本公开的另一个实施例的深度学习模型的训练方法的流程图。

[0073] 如图3所示,该方法340可以根据第一损失值,训练待训练的深度学习模型,下面将结合操作S341至操作S345进行详细说明。

[0074] 在操作S341,根据第一目标图像和第二目标图像,确定输入图像。

[0075] 例如,可以将第一目标图像和第二目标图像中的任意一个,作为输入图像。本实施例中,可以将第一目标图像作为输入图像。

[0076] 在一个示例中,可以将图像Img_i作为输入图像。

[0077] 在操作S342,获取第一输出特征。

[0078] 例如,第一输出特征是利用参考深度学习模型处理输入图像得到的。

[0079] 例如,可以通过各种方式获取第一输出特征。在一个示例中,可以直接获取上文所述的第一特征作为第一输出特征。在另一个示例中,可以将图像Img_i再次输入参考深度学

习模型,以获取第一输出特征。

[0080] 在操作S343,获取第二输出特征。

[0081] 例如,第二输出特征是利用待训练的深度学习模型处理输入图像得到的。

[0082] 例如,可以通过各种方式获取第二输出特征。在一个示例中,可以直接获取上文所述的第三特征作为第二输出特征。在另一个示例中,可以将图像Img_i再次输入待训练的深度学习模型,以获取第二输出特征。

[0083] 在操作S344,根据第一输出特征和第二输出特征,得到第二损失值。

[0084] 例如,可以根据第一输出特征和第二输出特征,通过以下公式进行计算,得到第二损失值:

[0085] $Loss_2 = ||f - F||^2$ (公式三)

[0086] Loss_2为第二损失值,f为第二输出特征,F为第一输出特征。

[0087] 在操作S345,根据第一损失值和第二损失值,训练待训练的深度学习模型。

[0088] 在本公开实施例中,根据预设权重、第一损失值和第二损失值,得到加权损失值。

[0089] 例如,可以通过以下公式进行计算,得到加权损失值:

[0090] $Loss_w = \lambda \times Loss_1 + (1 - \lambda) \times Loss_2$ (公式四)

[0091] Loss_w为加权损失值, λ 为预设权重。

[0092] 在一个示例中, $\lambda = 0.3$ 。

[0093] 在一些实施例中,与方法340不同之处在于,可以将第二目标图像作为输入图像。例如,可以将图像Img_j作为输入图像。

[0094] 在一些实施例中,参考深度学习模型的参数多于待训练的深度学习模型的参数。

[0095] 图4A是根据本公开的一个实施例的图像集中的一个图像的示意图。

[0096] 如图4A所示,图像Img_i 401中只有一个对象Obj_A。在图像集中,图像Img_i 401标签为对象Obj_A。

[0097] 图4B是根据本公开的一个实施例的图像集中的另一个图像的示意图。

[0098] 如图4B所示,图像Img_j 402中包含一个对象Obj_B和一张对象Obj_A的照片。在图像集中,图像Img_j 402标签也可以为对象Obj_A。

[0099] 图4C是根据本公开的一个实施例的深度学习模型的训练方法的原理图。

[0100] 如图4C所示,本实施例中,将例如图4A中的图像Img_i 401作为上文所述的第一目标图像,将例如图4B中的图像Img_j 402作为上文所述的第二目标图像。此外,将第一目标图像,即图像Img_i 401,作为输入图像。

[0101] 参考深度学习模型410可以处理图像Img_i 401,得到第一特征Feat_1 411。参考深度学习模型410也可以处理图像Img_j 402,得到第二特征Feat_2 412。根据第一特征Feat_1 411和第二特征Feat_2 412,可以得到二者的相似度S 413。根据相似度S 413和预设相似度阈值thred 414,可以得到第一差异值Diff_1 415。

[0102] 待训练的深度学习模型420可以处理Img_i 401,得到第三特征Feat_3 421。待训练的深度学习模型420也可以处理图像Img_j 402,得到第四特征Feat_4 422。根据第三特征Feat_3 421和第四特征Feat_4 422,可以得到第二差异值Diff_2 423。

[0103] 根据第一差异值Diff_1 414和第二差异值Diff_2 423,可以得到第一损失值Loss_1 431。

[0104] 接下来,可以将第一特征Feat_1 411作为第一输出特征,将第三特征Feat_3 421作为第二输出特征。根据第一特征Feat_1 411和第三特征Feat_3 421,可以得到第二损失值Loss_2 432。

[0105] 根据第一损失值Loss_1 431和第二损失值Loss_2 432,利用预设权重 λ 进行加权,得到加权损失值Loss_w 433。接下来,可以根据加权损失值Loss_w 433训练待训练的深度学习模型420。

[0106] 在一些实施例中,可以利用随机梯度下降法优化待训练的深度学习模型的参数,经过多批次的训练数据迭代,最终达到预设条件后得到深度学习模型。例如,预设条件可以是训练收敛。

[0107] 图5是根据本公开一个实施例的图像处理方法的流程图。

[0108] 如图5所示,该方法500可以包括操作S510至操作S530。

[0109] 在操作S510,将图像集中的目标图像输入深度学习模型,得到目标图像中对象的分类类别。

[0110] 例如,深度学习模型可以是根据本公开提供的方法训练的。

[0111] 例如,本实施例中,目标图像可以是上文所述的图像Img_j。如上文所述,图像Img_j的标签为对象Obj_A。

[0112] 将图像Img_j输入深度学习模型,可以得到目标图像中对象的分类类别。图像Img_j的分类类别可以是照片Pict_A。

[0113] 在操作S520,根据分类类别和目标图像的标签,得到第三差异值。

[0114] 例如,可以根据各种方式计算标签和分类类别之间的第三差异值,本公开在此不做限制。在一个示例中,可以将对象Obj_A和照片Pict_A分别编码,得到两个向量。计算两个向量之间的距离,作为第三差异值。

[0115] 在操作S530,响应于第三差异值大于或等于预设差异阈值,从图像集中删除目标图像。

[0116] 例如,如上文所述,可以从图像集中删除图像Img_j。

[0117] 通过本公开实施例,利用训练好的深度学习模型,可以准确地识别出脏数据。可以高效地筛选训练数据集中的数据,提高训练数据的质量,以便该训练数据集用于其他模型的训练。

[0118] 图6是根据本公开一个实施例的对象识别方法的流程图。

[0119] 如图6所示,该方法600可以包括操作S610。

[0120] 在操作S610,将目标图像输入深度学习模型,得到目标图像中对象的分类类别。

[0121] 例如,深度学习模型可以是根据本公开提供的方法训练的。

[0122] 例如,以目标图像为图像Img_i为示例。将图像Img_i输入深度学习模型,可以得到图像Img_i中对象的分类类别。该分类类别可以为对象Obj_A,与图像Img_i标签相同。

[0123] 又例如,以目标图像为图像Img_j为示例。将图像Img_j输入深度学习模型,可以得到图像Img_j中对象的分类类别。该分类类别可以为照片Pict_A,与图像Img_j标签不相同。

[0124] 图7是根据本公开的一个实施例的深度学习模型的训练装置的框图。

[0125] 如图7所示,该装置700可以包括第一获取模块710、第二获取模块720、第一获得模块730和训练模块740。

[0126] 第一获取模块710,用于获取第一特征。例如,上述第一特征是利用参考深度学习模型处理第一目标图像得到的。

[0127] 第二获取模块720,用于获取第二特征。例如,上述第二特征是利用上述参考深度学习模型处理第二目标图像得到的;

[0128] 第一获得模块730,用于根据上述第一特征和上述第二特征,得到第一损失值。

[0129] 训练模块740,用于根据上述第一损失值,训练待训练的深度学习模型。

[0130] 在一些实施例中,上述第一获得模块包括:计算子模块,用于计算上述第一特征和上述第二特征之间的相似度;第一获得子模块,用于根据上述相似度和预设相似度阈值,得到第一差异值;以及第二获得子模块,用于根据上述第一差异值,得到上述第一损失值。

[0131] 在一些实施例中,上述第二获得子模块包括:第一输入单元,用于将上述第一目标图像输入上述待训练的深度学习模型,得到第三特征;第二输入单元,用于将上述第二目标图像输入上述待训练的深度学习模型,得到第四特征;第一获得单元,用于根据上述第三特征和上述第四特征,得到第二差异值;以及第二获得单元,用于根据上述第一差异值和上述第二差异值,得到上述第一损失值。

[0132] 在一些实施例中,上述第二获得单元还用于:根据上述第一差异值和上述第二差异值,通过以下公式进行计算,得到上述第一损失值:

$$Loss_1 = \begin{cases} \exp(Diff_1) \times (Diff_2)^2 & , y = 1 \\ \exp(-Diff_1) \times \max(margin - Diff_2, 0)^2 & , y = 0 \end{cases}$$

[0133] 例如,Loss_1为第一损失值,Diff为第一差异值,d为第二差异值,margin为上述待训练的深度学习模型的预设类别阈值;在上述第一目标图像的标签表征的类别和上述第二目标图像的标签表征的类别相同的情况下,y=1;在上述第一目标图像的标签表征的类别和上述第二目标图像的标签表征的类别不相同的情况下,y=0。

[0134] 在一些实施例中,上述训练模块包括:确定子模块,用于根据上述第一目标图像和上述第二目标图像,确定输入图像;第一获取子模块,用于获取第一输出特征,其中,上述第一输出特征是利用上述参考深度学习模型处理上述输入图像得到的;第二获取子模块,用于获取第二输出特征,其中,上述第二输出特征是利用上述待训练的深度学习模型处理上述输入图像得到的;第三获得子模块,用于根据上述第一输出特征和上述第二输出特征,得到第二损失值;以及训练子模块,用于根据上述第一损失值和上述第二损失值,训练上述待训练的深度学习模型。

[0135] 在一些实施例中,上述训练子模块包括:第三获得单元,用于根据预设权重、上述第一损失值和上述第二损失值,得到加权损失值;以及训练单元,用于根据上述加权损失值,训练上述待训练的深度学习模型。

[0136] 在一些实施例中,上述参考深度学习模型的参数多于上述待训练的深度学习模型的参数。

[0137] 图8是根据本公开的一个实施例的图像处理装置的框图。

[0138] 如图8所示,该装置800可以包括第一输入模块810、第二获得模块820和删除模块830。

[0139] 第二获得模块810,用于将图像集中的目标图像输入深度学习模型,得到上述目标图像中对象的分类类别。

[0140] 第三获得模块820,用于根据上述分类类别和上述目标图像的标签,得到第三差异值。

[0141] 删除模块830,用于响应于第三差异值大于或等于预设差异阈值,从上述图像集中删除上述目标图像。

[0142] 例如,上述深度学习模型是根据本公开提供的装置训练的。

[0143] 图9是根据本公开的一个实施例的对象识别装置的框图。

[0144] 如图9所示,该装置900可以包括第二输入模块910。

[0145] 第四获得模块910,用于将目标图像输入深度学习模型,得到上述目标图像中对象的分类类别。

[0146] 例如,上述深度学习模型是根据本公开提供的装置训练的。

[0147] 本公开的技术方案中,所涉及的用户个人信息的收集、存储、使用、加工、传输、提供和公开等处理,均符合相关法律法规的规定,且不违背公序良俗。

[0148] 根据本公开的实施例,本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。

[0149] 图10示出了可以用来实施本公开的实施例的示例电子设备1000的示意性框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0150] 如图10所示,设备1000包括计算单元1001,其可以根据存储在只读存储器(ROM) 1002中的计算机程序或者从存储单元1008加载到随机访问存储器(RAM) 1003中的计算机程序,来执行各种适当的动作和处理。在RAM 1003中,还可存储设备1000操作所需的各种程序和数据。计算单元1001、ROM 1002以及RAM 1003通过总线1004彼此相连。输入/输出(I/O)接口1005也连接至总线1004。

[0151] 设备1000中的多个部件连接至I/O接口1005,包括:输入单元1006,例如键盘、鼠标等;输出单元1007,例如各种类型的显示器、扬声器等;存储单元1008,例如磁盘、光盘等;以及通信单元1009,例如网卡、调制解调器、无线通信收发机等。通信单元1009允许设备1000通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0152] 计算单元1001可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元1001的一些示例包括但不限于中央处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元1001执行上文所描述的各个方法和处理,例如深度学习模块的训练方法、图像处理方法和/或对象识别方法。例如,在一些实施例中,深度学习模块的训练方法、图像处理方法和/或对象识别方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元1008。在一些实施例中,计算机程序的部分或者全部可以经由ROM 1002和/或通信单元1009而被载入和/或安装到设备1000上。当计算机程序加载到RAM 1003并由计算单元1001执行时,可以执行上文描述的深度学习模块的训练方法、图像处理方法和/或对象识别方法的一个或多个步骤。备选地,在其他实施

例中,计算单元1001可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行深度学习模块的训练方法、图像处理方法和/或对象识别方法。

[0153] 本文中以上描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、负载可编程逻辑设备(CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0154] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0155] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0156] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0157] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0158] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务器的关系。

[0159] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0160] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

100

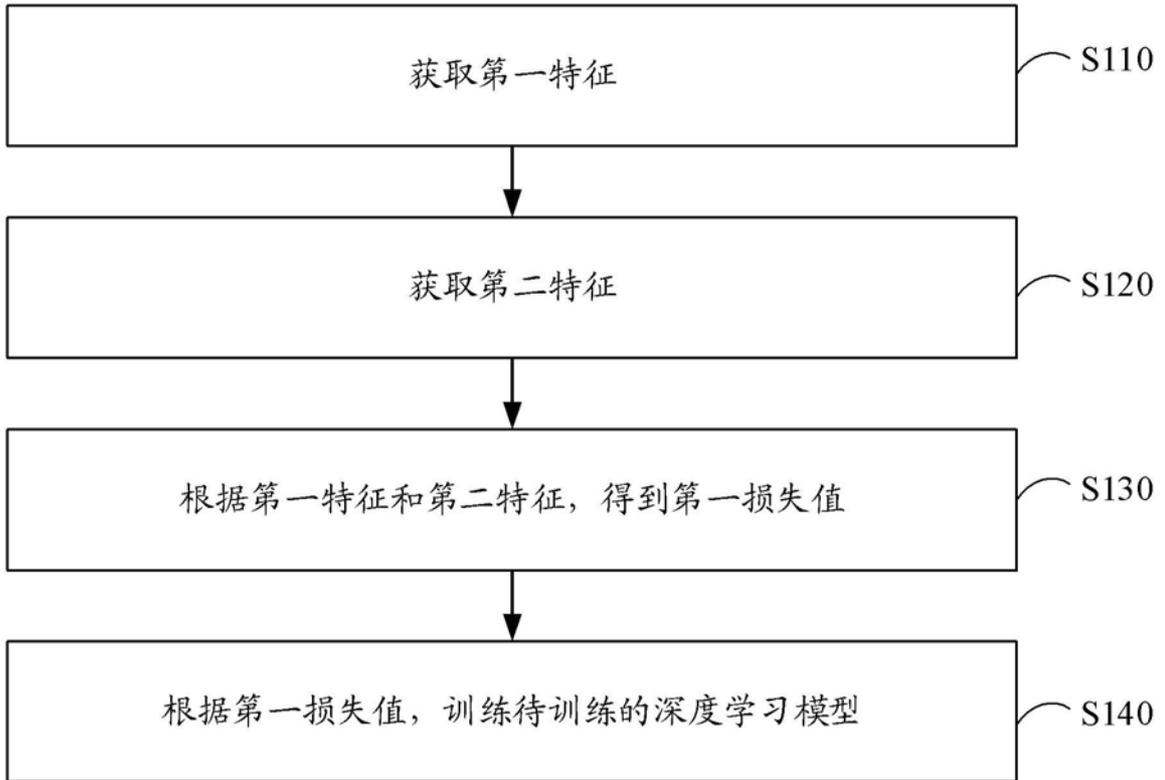


图1

230

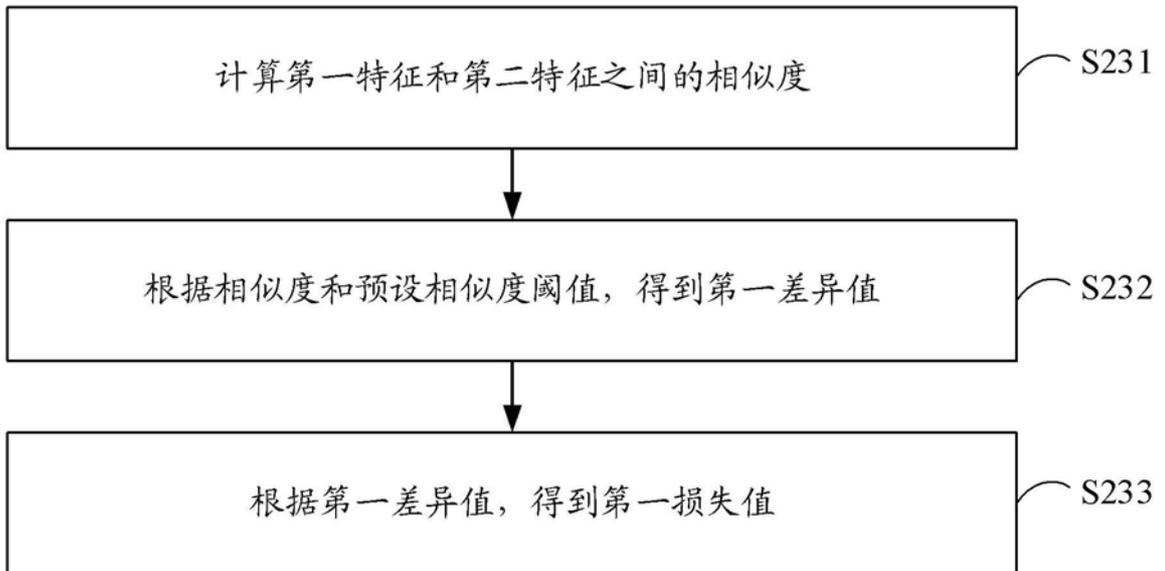


图2

340

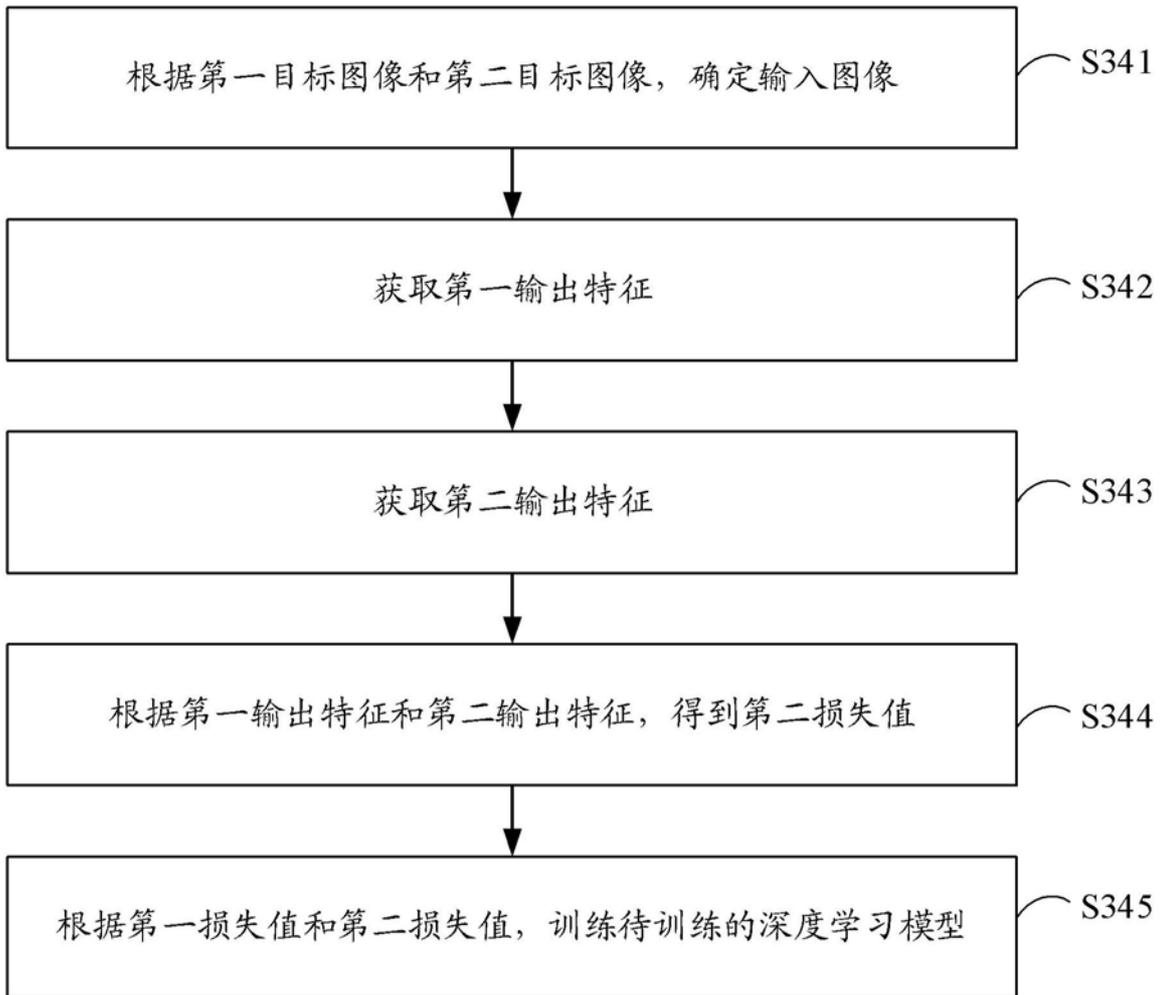


图3

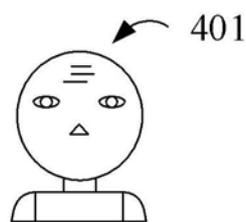


图4A

402

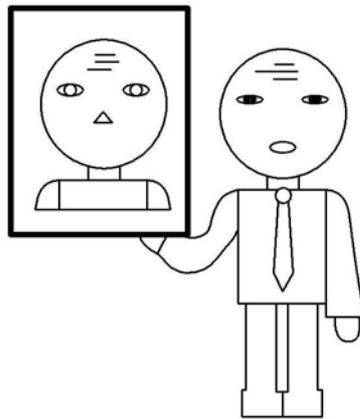


图4B

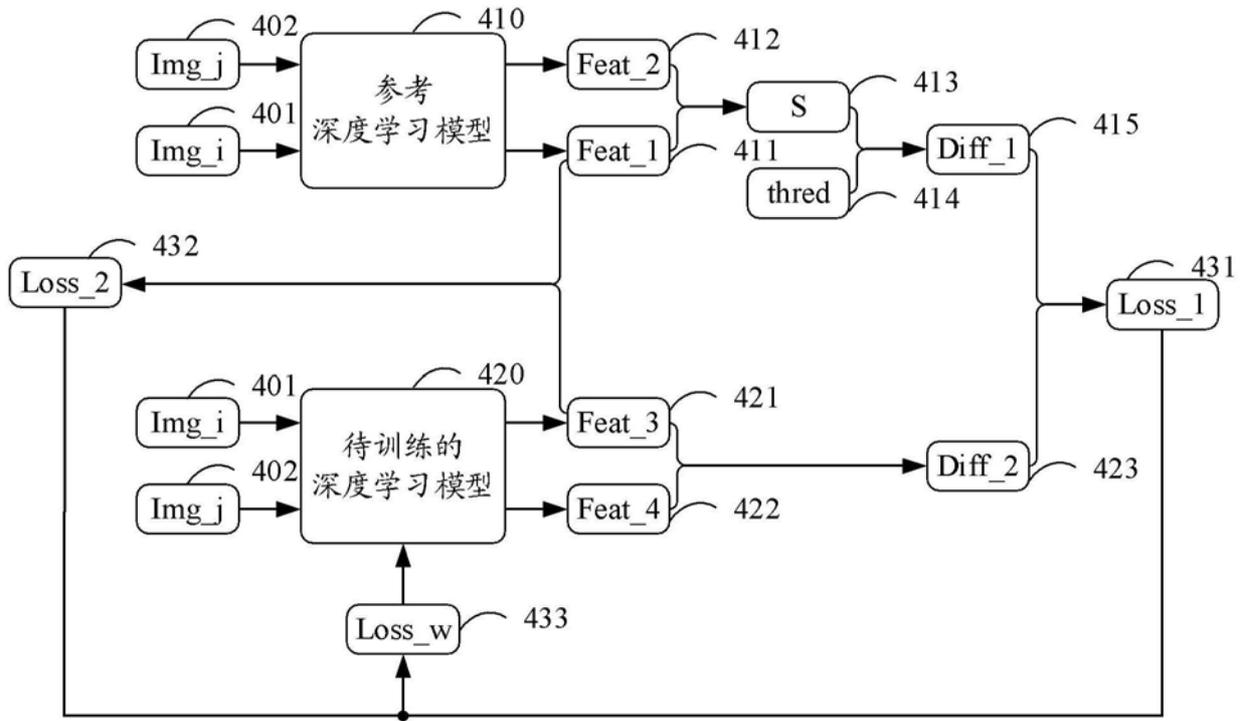


图4C

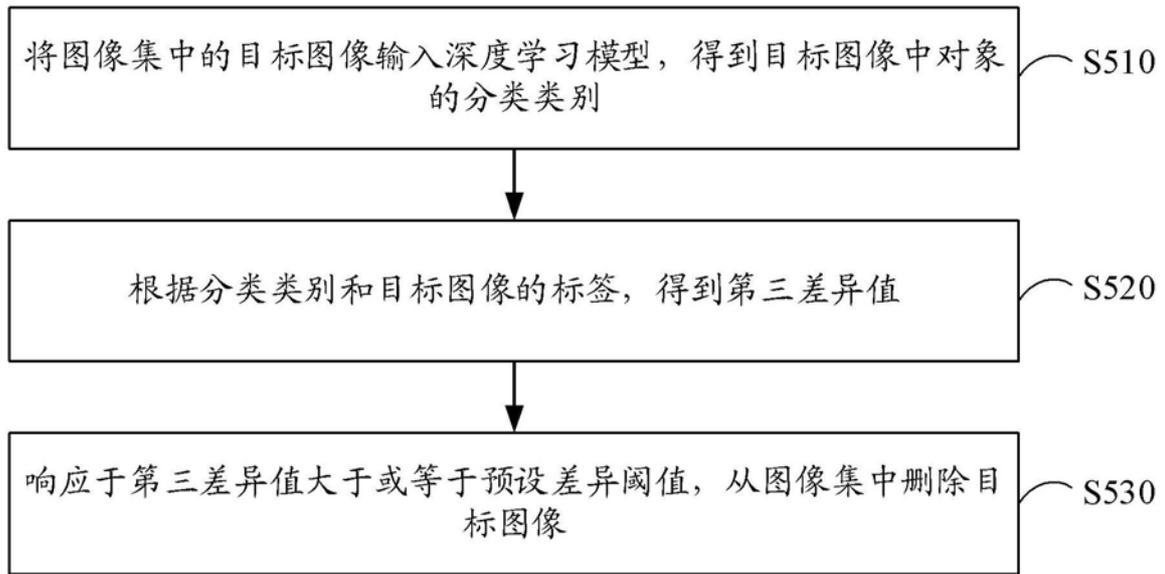
500

图5

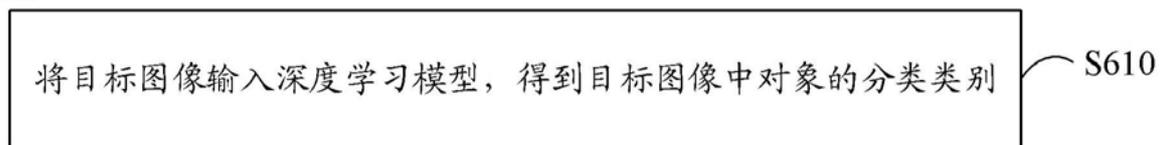
600

图6

700

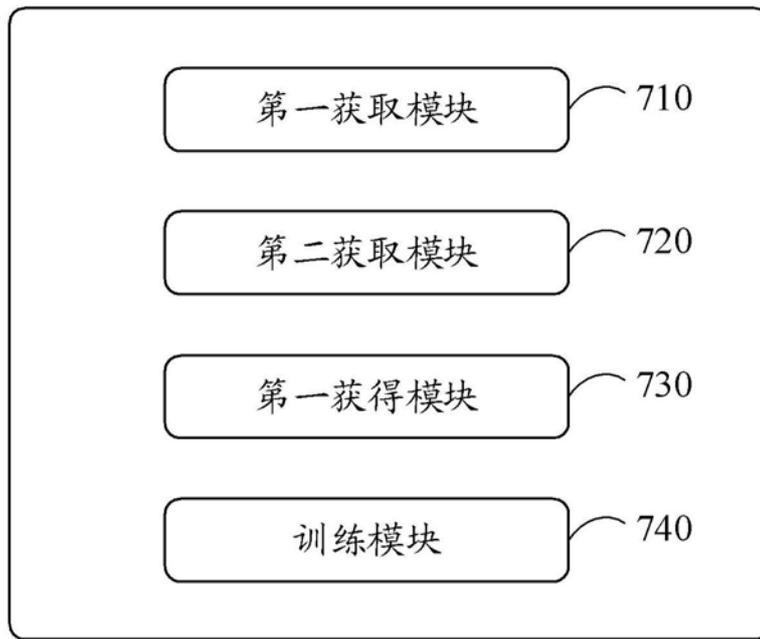


图7

800

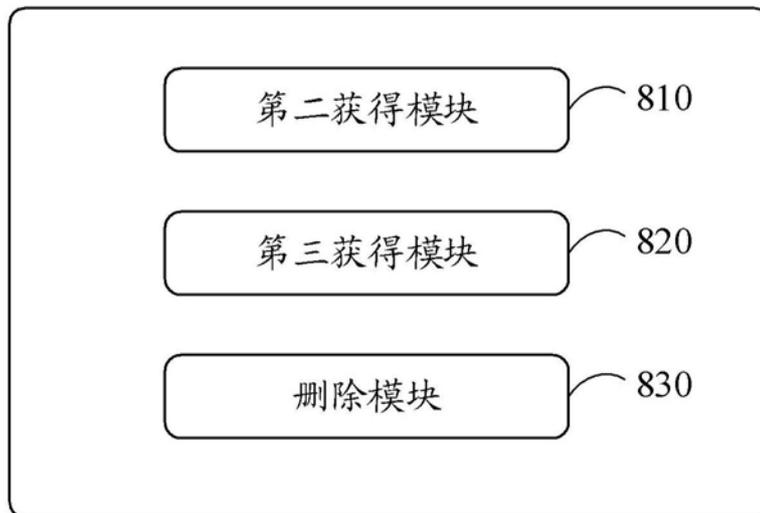


图8

900

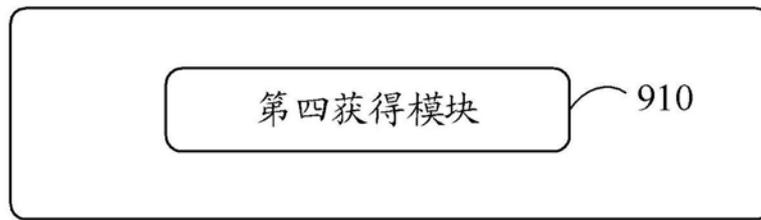


图9

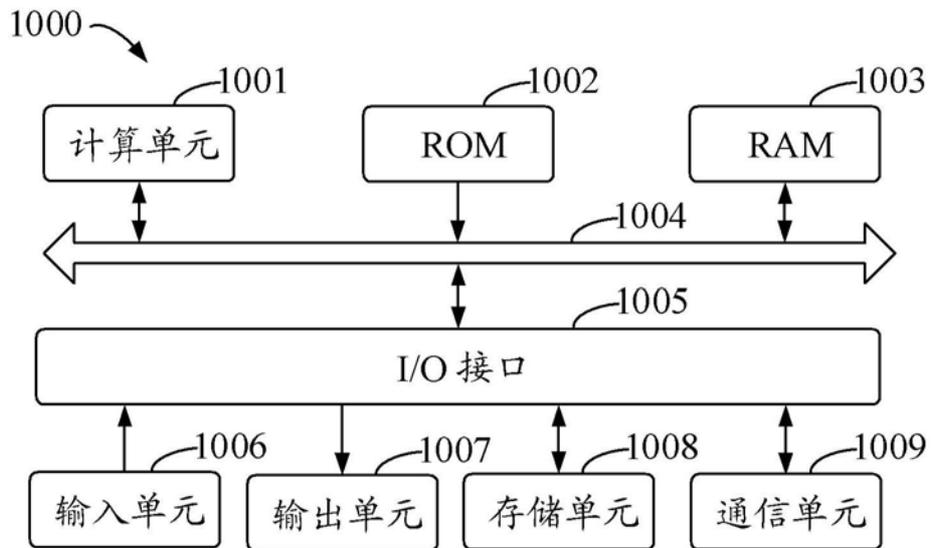


图10