(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 113537462 B
(45)授权公告日 2025.04.04

(21)申请号 202110742803.4
(22)申请日 2021.06.30
(65)同一申请的已公布的文献号
申请公布号 CN 113537462 A
(43)申请公布日 2021.10.22
(73)专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼
(72)发明人 刘传建 韩凯 王云鹤
(74)专利代理机构 广州三环专利商标代理有限公司 44202
专利代理师 熊永强 李稷芳
(51)Int.Cl.
G06N 3/0464(2023.01)
G06N 3/047(2023.01)
G06N 3/048(2023.01)
G06N 3/084(2023.01)
(56)对比文件
CN 110663048 A,2020.01.07
CN 111095301 A,2020.05.01
审查员 王闪

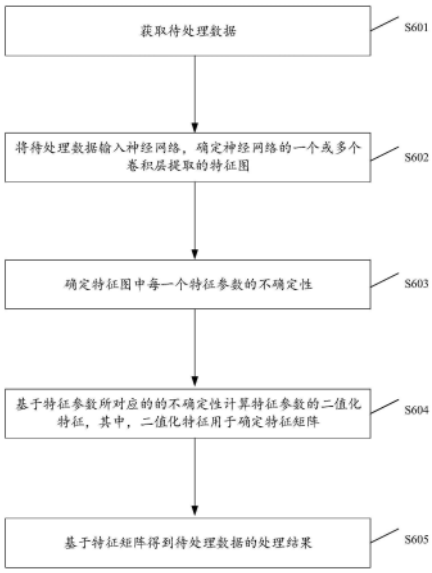
权利要求书6页 说明书39页 附图14页

(54)发明名称

数据处理方法、神经网络的量化方法及相关装置

(57)摘要

本申请实施例提供一种数据处理方法、神经网络的量化方法及相关装置，该方法包括：获取待处理数据；将待处理数据输入神经网络，确定神经网络的一个或多个卷积层提取的特征图，特征图包含m*n个特征参数，m和n为正整数；确定特征图中每一个特征参数的不确定性，其中，特征参数的不确定性用于表征特征参数在二值化过程中，接近于零的特征参数的符号的波动性；基于特征参数所对应的不确定性计算特征参数的二值化特征，二值化特征用于确定特征矩阵，特征矩阵中包含的m*n个二值化特征与m*n个特征参数一一对应；基于特征矩阵得到待处理数据的处理结果。采用本申请实施例，能够减小内存开销，提高运算速度。



CN 113537462 B

1.一种数据处理方法,其特征在于,所述方法包括:

获取待处理数据;所述数据包括图形、或图像、或语音、或文本、或物联网数据;所述物联网数据包括感知数据,所述感知数据包括力、或位移、或液位、或温度、或湿度的感知数据;

将所述待处理数据输入神经网络,确定所述神经网络的一个或多个卷积层提取的特征图,所述特征图包含m*n个特征参数,m和n为正整数;

确定所述特征图中每一个特征参数的不确定性,其中,所述特征参数的不确定性用于表征所述特征参数在二值化过程中,接近于零的特征参数的符号的波动性;

基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,所述二值化特征用于确定特征矩阵,所述特征矩阵中包含的m*n个二值化特征与所述m*n个特征参数一一对应;

基于所述特征矩阵得到所述待处理数据的处理结果。

2.根据权利要求1所述的方法,其特征在于,所述神经网络为基于二值化权重训练得到的,所述二值化权重为根据所述二值化权重对应的权重参数的不确定性对所述权重参数进行二值化处理得到的,所述权重参数的不确定性用于表征所述权重参数在二值化过程中,接近于零的权重参数的符号的波动性。

3.根据权利要求1或2所述的方法,其特征在于,所述确定所述特征图中每一个特征参数的不确定性,包括:

根据不确定性函数计算所述特征图中每一个特征参数的不确定性,其中,在所述不确定性函数的自变量越接近于0时,所述不确定性函数的值越大;在所述不确定性函数的自变量的绝对值越大时,所述不确定性函数的值越小。

4.根据权利要求3所述的方法,其特征在于,所述根据不确定性函数计算特征图中每一个特征参数的不确定性的公式为:

$$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与所述目标特征参数相邻的a个特征参数的不确定性,所述目标参数为所述特征图上的任意一个参数,i,j,a均为正整数。

5.根据权利要求4所述的方法,其特征在于,所述基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,包括:

在所述目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对所述目标特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

6.根据权利要求4所述的方法,其特征在于,所述基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,包括:

在所述目标特征参数的不确定性大于第二预设阈值时,通过符号函数对平均池化后的与所述目标特征参数相邻的一个或多个特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

7.一种神经网络的量化方法,其特征在于,包括:

获取第一权重矩阵,所述第一权重矩阵中包含神经网络中用于提取特征的参数,所述

第一权重矩阵包含s*k个权重参数,s和k为正整数;所述神经网络输入的数据包括图形、或图像、或语音、或文本、或物联网数据;所述物联网数据包括感知数据,所述感知数据包括力、或位移、或液位、或温度、或湿度的感知数据;

计算所述第一权重矩阵中每一个权重参数的不确定性,其中,所述权重参数的不确定性用于表征所述权重参数在二值化过程中,接近于零的权重参数的符号的波动性;

基于所述权重参数所对应的不确定性计算所述权重参数的二值化权重,所述二值化权重用于确定第二权重矩阵,所述第二权重矩阵中包含的s*k个二值化权重与所述s*k个权重参数一一对应。

8.根据权利要求7所述的方法,其特征在于,所述计算所述第一权重矩阵中每一个权重参数的不确定性,包括:

根据不确定性函数计算所述第一权重矩阵中每一个权重参数的不确定性,其中,在所述不确定性函数的自变量越接近于0时,所述不确定性函数的值越大;在所述不确定性函数的自变量的绝对值越大时,所述不确定性函数的值越小。

9.根据权利要求8所述的方法,其特征在于,所述根据不确定性函数计算所述第一权重矩阵中每一个权重参数的不确定性,包括:

在当前迭代次数小于或等于预设迭代次数时,根据所述不确定性函数计算所述当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

10.根据权利要求9所述的方法,其特征在于,所述根据不确定性函数计算所述第一权重矩阵中每一个权重参数的不确定性,包括:

在所述当前迭代次数大于所述预设迭代次数时,根据在参考迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算所述当前迭代次数所对应的第一权重矩阵的每一个权重参数的不确定性,其中,所述参考迭代次数为最接近所述当前迭代次数的预设迭代次数。

11.根据权利要求7至10任一项所述的方法,其特征在于,所述基于所述权重参数所对应的不确定性计算所述权重参数的二值化权重,包括:

在当前迭代次数所对应的所述第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,所述第一值为所述当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,所述目标权重参数为所述第一权重矩阵中的任意一个参数。

12.根据权利要求7至10任一项所述的方法,其特征在于,所述基于所述权重参数所对应的不确定性计算所述权重参数的二值化权重,包括:

在当前迭代次数所对应的所述第一权重矩阵中的目标权重参数的不确定性大于第一值时,将所述当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为所述当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,其中,所述第一值为所述当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,所述目标权重参数为所述第一权重矩阵中的任意一个参数。

13.根据权利要求7至10任一项所述的方法,其特征在于,还包括:获取训练数据;

将所述训练数据输入神经网络,确定所述神经网络的一个或多个卷积层提取的特征

图;

计算所述特征图中每一个特征参数的二值化特征,其中,所述特征图包含m*n个特征参数, m和n为正整数,所述特征图为在所述神经网络的一个或多个卷积层中提取的训练数据的特征。

14.根据权利要求13所述的方法,其特征在于,所述计算所述特征图中每一个特征参数的二值化特征,包括:

确定所述特征图中每一个特征参数的不确定性,其中,所述特征参数的不确定性用于表征所述特征参数在二值化过程中,接近于零的特征参数的符号的波动性;

基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,所述二值化特征用于确定特征矩阵,所述特征矩阵中包含的m*n个二值化特征与所述m*n个特征参数一一对应。

15.根据权利要求14所述的方法,其特征在于,所述确定所述特征图中每一个特征参数的不确定性,包括:

根据不确定性函数计算所述特征图中特征参数的不确定性,其中,在所述不确定性函数的自变量越接近于0时,所述不确定性函数的值越大;在所述不确定性函数的自变量的绝对值越大时,所述不确定性函数的值越小。

16.根据权利要求15所述的方法,其特征在于,所述不确定性函数公式为:

$$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与所述目标特征参数相邻的a个特征参数的不确定性,i,j,a均为自然数。

17.根据权利要求16所述的方法,其特征在于,所述基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,包括:

在所述目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对所述目标特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

18.根据权利要求16所述的方法,其特征在于,所述基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,包括:

在所述目标特征参数的目标不确定性大于第二预设阈值时,通过符号函数对平均池化后的与所述目标特征参数相邻的一个或多个特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

19.一种数据处理装置,其特征在于,所述装置,包括:

获取单元,用于获取待处理数据;所述数据包括图形、或图像、或语音、或文本、或物联网数据;所述物联网数据包括感知数据,所述感知数据包括力、或位移、或液位、或温度、或湿度的感知数据;

输入单元,用于将所述待处理数据输入神经网络,确定所述神经网络的一个或多个卷积层提取的特征图,所述特征图包含m*n个特征参数,m和n为正整数;

计算单元,用于确定所述特征图中每一个特征参数的不确定性,其中,所述特征参数的不确定性用于表征所述特征参数在二值化过程中,接近于零的特征参数的符号的波动性;

量化单元,用于基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,所述二值化特征用于确定特征矩阵,所述特征矩阵中包含的m*n个二值化特征与所述m*n个特征参数一一对应;

所述计算单元,还用于基于所述特征矩阵得到所述待处理数据的处理结果。

20.根据权利要求19所述的装置,其特征在于,所述神经网络为基于二值化权重训练得到的,所述二值化权重为根据所述二值化权重对应的权重参数的不确定性对所述权重参数进行二值化处理得到的,所述权重参数的不确定性用于表征所述权重参数在二值化过程中,接近于零的权重参数的符号的波动性。

21.根据权利要求19或20所述的装置,其特征在于,所述计算单元,具体用于:

根据不确定性函数计算所述特征图中每一个特征参数的不确定性,其中,在所述不确定性函数的自变量越接近于0时,所述不确定性函数的值越大;在所述不确定性函数的自变量的绝对值越大时,所述不确定性函数的值越小。

22.根据权利要求21所述的装置,其特征在于所述根据不确定性函数计算特征图中每一个特征参数的不确定性的公式为:

$$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与所述目标特征参数相邻的a个特征参数的不确定性,所述目标参数为所述特征图上的任意一个参数,i,j,a均为正整数。

23.根据权利要求22所述的装置,所述量化单元,具体用于:

在所述目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对所述目标特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

24.根据权利要求22所述的装置,所述量化单元,具体用于:

在所述目标特征参数的不确定性大于第二预设阈值时,通过符号函数对平均池化后的与所述目标特征参数相邻的一个或多个特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

25.一种神经网络的量化装置,其特征在于,所述装置包括:

获取单元,用于获取第一权重矩阵,所述第一权重矩阵中包含神经网络中用于提取特征的参数,所述第一权重矩阵包含s*k个权重参数,s和k为正整数;所述神经网络输入的数据包括图形、或图像、或语音、或文本、或物联网数据;所述物联网数据包括感知数据,所述感知数据包括力、或位移、或液位、或温度、或湿度的感知数据;

计算单元,用于计算所述第一权重矩阵中每一个权重参数的不确定性,其中,所述权重参数为所述神经网络的权重中的任意一个权重,所述权重参数的不确定性用于表征所述权重参数在二值化过程中,接近于零的权重参数的符号的波动性;

量化单元,用于基于所述权重参数所对应的不确定性计算所述权重参数的二值化权重,所述二值化权重用于确定第二权重矩阵,所述第二权重矩阵中包含的s*k个二值化权重与所述s*k个权重参数一一对应。

26.根据权利要求25所述的装置,其特征在于,所述计算单元,具体用于:

根据不确定性函数计算所述第一权重矩阵中每一个权重参数的不确定性,其中,在所

述不确定性函数的自变量越接近于0时,所述不确定性函数的值越大;在所述不确定性函数的自变量的绝对值越大时,所述不确定性函数的值越小。

27.根据权利要求25所述的装置,其特征在于,所述计算单元,具体用于:

在当前迭代次数小于或等于预设迭代次数时,通过不确定性函数计算所述当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

28.根据权利要求25所述的装置,其特征在于,所述计算单元,具体用于:

在当前迭代次数大于预设迭代次数时,根据在参考迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算所述当前迭代次数所对应的第一权重矩阵的每一个权重参数的不确定性,其中,所述参考迭代次数为最接近所述当前迭代次数的预设迭代次数。

29.根据权利要求25至28任一项所述的装置,其特征在于,所述量化单元,具体用于:

在当前迭代次数所对应的所述第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,所述第一值为所述当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,所述目标权重参数为所述第一权重矩阵中的任意一个参数。

30.根据权利要求25至28任一项所述的装置,其特征在于,所述量化单元,具体用于:

在当前迭代次数所对应的所述第一权重矩阵中的目标权重参数的不确定性大于第一值时,将所述当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为所述当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,其中,所述第一值为所述当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,所述目标权重参数为所述第一权重矩阵中的任意一个参数。

31.根据权利要求25至28任一项所述的装置,其特征在于,还包括输入单元,

所述获取单元,还用于获取训练数据;

所述输入单元,用于将所述训练数据输入神经网络,确定所述神经网络的一个或多个卷积层提取的特征图;

所述量化单元,用于计算所述特征图中每一个特征参数的二值化特征,其中,所述特征图包含m*n个特征参数, m和n为正整数,所述特征图为在所述神经网络的一个或多个卷积层中提取的训练数据的特征。

32.根据权利要求31所述的装置,其特征在于,所述量化单元,具体用于:

确定所述特征图中每一个特征参数的不确定性,其中,所述特征参数的不确定性用于表征所述特征参数在二值化过程中,接近于零的特征参数的符号的波动性;

基于所述特征参数所对应的不确定性计算所述特征参数的二值化特征,所述二值化特征用于确定特征矩阵,所述特征矩阵中包含的m*n个二值化特征与所述m*n个特征参数一一对应。

33.根据权利要求32所述的装置,其特征在于,所述量化单元,具体用于:

根据不确定性函数计算所述特征图中特征参数的不确定性,其中,在所述不确定性函数的自变量越接近于0时,所述不确定性函数的值越大;在所述不确定性函数的自变量的绝对值越大时,所述不确定性函数的值越小。

34.根据权利要求33所述的装置,其特征在于,所述不确定性函数公式为:

$$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与所述目标特征参数相邻的a个特征参数的不确定性,i,j,a均为自然数。

35.根据权利要求34所述的装置,其特征在于,所述量化单元,具体用于:

在所述目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对所述目标特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

36.根据权利要求34所述的装置,其特征在于,所述量化单元,具体用于:

在所述目标特征参数的目标不确定性大于第二预设阈值时,通过符号函数对平均池化后的与所述目标特征参数相邻的一个或多个特征参数进行二值化处理,得到所述目标特征参数的二值化特征。

37.一种数据处理设备,其特征在于,包括:存储器和处理器,所述存储器用于程序,所述处理器执行所述存储器存储的程序,当存储器存储的程序被执行时,所述处理器用于执行如权利要求1至6任一项所述的数据处理方法。

38.一种神经网络的量化设备,其特征在于,包括:存储器和处理器,所述存储器用于程序,所述处理器执行所述存储器存储的程序,当存储器存储的程序被执行时,所述处理器用于执行如权利要求7至18任一项所述的神经网络的量化方法。

39.一种计算机可读存储介质,其特征在于,所述计算机可读介质存储用于电子设备执行的程序代码,所述程序代码包括如权利要求1至6或权利要求7至18任一项所述的方法。

40.一种包含指令的计算机程序产品,其特征在于,所述计算机程序产品在电子设备上运行时,使得所述电子设备执行如权利要求1至6或权利要求7至18任一项所述的方法。

数据处理方法、神经网络的量化方法及相关装置

技术领域

[0001] 本申请涉及人工智能技术领域,尤其涉及一种数据处理方法、神经网络的量化方法及相关装置。

背景技术

[0002] 二值神经网络(Binary Neural Network,BNN)可以将权重和/或特征量等神经网络的参数量化到单个比特,使得模型的参数可以占用更小的存储空间。另外,相比于全精度神经网络中使用浮点数的乘法和累加实现卷积操作来说,二值神经网络可以通按位异或非来实现卷积操作。因此,二值神经网络可以降低模型的计算量,加快模型的推断过程,在很大程度上方便了模型在资源受限设备上的部署。

[0003] 但是,二值化会不可避免地带来信息损失,其量化函数不连续性也给网络的优化带来了困难。其中,二值神经网络中的权值优化是导致网络性能下降的主要原因之一。为了解决上述问题,现有技术提供了直接量化的朴素二值化方法,以及使用最小量化误差、改善网络损失函数和减小梯度误差等技术的改进二值化方法。但是,上述方法都强调了权值的梯度大小,而忽略了权值的梯度方向,而权值的梯度方向可以确定权值的优化方向。但是,不稳定的优化方向可能会导致神经网络的收敛速度缓慢和不稳定,因此,如何降低提高优化方向的稳定性是亟需解决的技术问题。

发明内容

[0004] 本申请实施例提高了一种数据处理方法、神经网络的量化方法及相关装置,能够减小内存开销,提高运算速度。

[0005] 第一方面,本申请实施例提供了一种数据处理方法,该方法可以包括:获取待处理数据;将待处理数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图,特征图包含 m*n个特征参数,m和n为正整数;确定特征图中每一个特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;基于特征参数所对应的不确定性计算特征参数的二值化特征,二值化特征用于确定特征矩阵,特征矩阵中包含的m*n个二值化特征与m*n个特征参数一一对应;基于特征矩阵得到待处理数据的处理结果。

[0006] 实施本申请实施例,可以考虑到神经网络中特征参数的二值化所带来的不确定性影响,从而通过不确定性函数来定量计算特征参数的不确定性。并且,为了减少推理过程中的不确定性,提高推理的稳定性,通过计算得到的特征参数的不确定性对特征参数进行二值化处理。这样,可以提高神经网络的运算速度和稳定性。

[0007] 在一种可能的实现方式中,神经网络为基于二值化权重训练得到的,二值化权重为根据二值化权重对应的权重参数的不确定性对权重参数进行二值化处理得到的,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性。可以理解的是,在对模型的训练过程中考虑到神经网络中权重参数以及特征数值的二值化

所带来的不确定性影响,从而通过不确定性函数来定量计算不确定性。通过计算得到的不确定性对神经网络参数(比如说权重参数)进行二值化。这样,可以提高神经网络的收敛速度和稳定性。

[0008] 在一种可能的实现方式中,确定特征图中每一个特征参数的不确定性,包括:根据不确定性函数计算特征图中每一个特征参数的不确定性,其中,在不确定函数的自变量越接近于 0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。其中,不确定性函数可以计算特征图中特征参数在推理过程中的不确定性,提升网络性能。

[0009] 在一种可能的实现方式中,不确定性函数的公式为:

[0010] $$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

[0011] 其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与目标特征参数相邻的a个特征参数的不确定性,目标参数为特征图上的任意一个参数,i,j,a均为正整数。为了提高特征参数的不确定性的稳定性,可以对特征图中一个或多个特征参数的不确定性进行联合考虑,来综合计算目标特征参数的不确定性。

[0012] 在一种可能的实现方式中,基于特征参数所对应的不确定性计算特征参数的二值化特征,包括:在目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0013] 在一种可能的实现方式中,基于特征参数所对应的不确定性计算特征参数的二值化特征,包括:在目标特征参数的不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标特征参数相邻的一个或多个特征参数进行二值化处理,得到目标位置点的二值化特征。

[0014] 基于不确定性的相关计算对特征参数进行二值化处理,对于不确定性较大(也即目标位置点的不确定性大于第二预设阈值)的特征参数采用平均池化并且引入sign函数,并从空间维度对目标位置点的特征参数进行二值化。

[0015] 第二方面,本申请实施例提供了一种神经网络的量化方法,该方法可以包括:获取第一权重矩阵,第一权重矩阵中包含神经网络中用于提取特征的参数,第一权重矩阵包含s*k个权重参数,s和k为正整数;计算第一权重矩阵中每一个权重参数的不确定性,其中,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性;基于权重参数的不确定性计算权重参数的二值化权重,二值化权重用于确定第二权重矩阵,第二权重矩阵中包含的s*k个二值化权重与s*k个权重参数一一对应。

[0016] 实施本申请实施例,可以考虑到神经网络中权重参数的二值化所带来的不确定性影响,从而通过不确定性函数来定量计算不确定性。并且,为了减少训练过程中的不确定性,提高训练的稳定性,可以通过计算得到的不确定性对权重参数进行二值化处理。这样,可以提高神经网络的收敛速度和稳定性。

[0017] 在一种可能的实现方式中,计算第一权重矩阵中每一个权重参数的不确定性,包括:根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大

时,不确定性函数的值越小。

[0018] 其中,不确定性函数可以确定权重参数在迭代更新过程中权重参数的不确定性,从而完善参数量化机制,提升网络性能。

[0019] 在一种可能的实现方式中,根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,包括:在当前迭代次数小于或等于预设迭代次数时,通过不确定性函数计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0020] 在一种可能的实现方式中,根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,包括:在当前迭代次数大于预设迭代次数时,根据在参考迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性,其中,参考迭代次数为最接近当前迭代次数的预设迭代次数。

[0021] 为了使得神经网络的不确定性最小,可以通过与当前迭代次数最接近的预设迭代次数内的权重参数的不确定性来计算当前迭代次数的权重参数的不确定性。

[0022] 在一种可能的实现方式中,基于权重参数所对应的不确定性计算权重参数的二值化权重,包括:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0023] 在一种可能的实现方式中,基于权重参数所对应的不确定性计算权重参数的二值化权重,包括:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性大于第一值时,将当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0024] 为了使得神经网络的不确定性最小,将当前迭代次数的不确定性与其他值(比如说前一迭代次数的不确定性或者第一预设阈值)进行比较,在满足条件的情况下,将前一迭代次数的二值化权重作为当前迭代次数的二值化权重。

[0025] 在一种可能的实现方式中,该方法还可以包括:还包括:获取训练数据;将训练数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图;计算特征图中每一个特征参数的二值化特征,其中,特征图包含m*n个特征参数,m和n为正整数,特征图为在神经网络的一个或多个卷积层中提取的训练数据的特征。

[0026] 考虑到神经网络中特征参数的二值化所带来的不确定性影响,从而通过不确定性函数来定量计算不确定性。并且,为了减少推理过程中的不确定性,提高推理的稳定性,通过计算得到的不确定性对特征参数进行二值化处理。这样,可以提高神经网络的运算速度和稳定性。

[0027] 在一种可能的实现方式中,计算特征图中每一个特征参数的二值化特征,包括:确定特征图中每一个特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;基于特征参数所对应的不确定性计

算特征参数的二值化特征,二值化特征用于确定特征矩阵,特征矩阵中包含的m*n个二值化特征与 m*n个特征参数一一对应。

[0028] 在一种可能的实现方式中,确定特征图中每一个特征参数的不确定性,包括:根据不确定性函数计算特征图中特征参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0029] 其中,不确定性函数可以计算特征参数在迭代更新过程中的不确定性,从而完善参数量化机制,提升网络性能。

[0030] 在一种可能的实现方式中,不确定性函数公式为:

[0031] $$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

[0032] 其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与目标特征参数相邻的a个特征参数的不确定性,i,j,a均为自然数。对训练数据中一个或多个位置点的特征参数的不确定性进行联合考虑,来综合计算目标位置点(或者位置点)的特征图的不确定性。

[0033] 在一种可能的实现方式中,基于特征参数所对应的不确定性计算特征参数的二值化特征,包括:在目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0034] 在一种可能的实现方式中,基于特征参数所对应的不确定性计算特征参数的二值化特征,包括:在目标特征参数的目标不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标特征参数相邻的一个或多个特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0035] 基于不确定性的相关计算对特征参数进行二值化处理,对于不确定性较大(也即目标位置点的不确定性大于第二预设阈值)的特征参数采用平均池化并且引入sign函数,并从空间维度对目标位置点的特征参数进行二值化。

[0036] 第三方面,本申请实施例提供了一种神经网络的训练方法,该方法可以包括:获取第一权重矩阵和特征图,第一权重矩阵中包含神经网络中用于提取特征的参数,第一权重矩阵包含s*k个权重参数,特征图为在神经网络的一个或多个卷积层中提取的训练数据的特征,特征图包含m*n个特征参数,s,k,m和n均为正整数;计算第一权重矩阵中每一个权重参数的不确定性,其中,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性;基于权重参数所对应的不确定性计算权重参数的二值化权重,二值化权重用于确定第二权重矩阵,第二权重矩阵中包含的s*k个二值化权重与s*k个权重参数一一对应;计算特征图中每一个特征参数的二值化特征,其中,特征图包含m*n个特征参数, m和n为正整数,特征图为在神经网络的一个或多个卷积层中提取的训练数据的特征;基于第二权重矩阵和特征矩阵对神经网络进行训练。实施本申请实施例,可以考虑到神经网络中权重以及特征参数的二值化所带来的不确定性影响,从而通过不确定性函数来定量计算不确定性。并且,为了减少训练过程中的不确定性,提高训练的稳定性,通过计算得到的不确定性来进行二值化。这样,可以提高神经网络的收敛速度和稳定性。

[0037] 在一种可能的实现方式中,计算第一权重矩阵中每一个权重参数的不确定性,包

括:根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0038] 其中,不确定性函数可以计算权重参数在迭代更新过程中的不确定性,从而完善参数量化机制,提升网络性能。

[0039] 在一种可能的实现方式中,根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,包括:在当前迭代次数小于或等于预设迭代次数时,根据不确定性函数计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0040] 在一种可能的实现方式中,根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,包括:在当前迭代次数大于预设迭代次数时,根据在参考迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算当前迭代次数所对应的第一权重矩阵的每一个权重参数的不确定性,其中,参考迭代次数为最接近当前迭代次数的预设迭代次数。

[0041] 为了使得神经网络的不确定性最小,可以通过预设迭代次数内的权重参数的不确定性来计算当前迭代次数的权重参数的不确定性。

[0042] 在一种可能的实现方式中,基于权重参数所对应的不确定性计算权重参数的二值化权重,包括:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0043] 在一种可能的实现方式中,基于权重参数所对应的不确定性计算权重参数的二值化权重,包括:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性大于第一值时,将当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0044] 为了使得神经网络的不确定性最小,将当前迭代次数的不确定性与其他值(比如说前一迭代次数的不确定性或者第一预设阈值)进行比较,在满足条件的情况下,将前一迭代次数的二值化权重作为当前迭代次数的二值化权重。

[0045] 在一种可能的实现方式中,计算特征图中每一个特征参数的二值化特征,包括:

[0046] 确定特征图中每一个特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;基于特征参数所对应的不确定性计算特征参数的二值化特征,二值化特征用于确定特征矩阵,特征矩阵中包含的m*n个二值化特征与m*n个特征参数一一对应。

[0047] 基于特征参数的不确定性,为了使得神经网络的不确定性最小,根据不确定性来计算特征参数的二值化特征。

[0048] 在一种可能的实现方式中,确定特征图中每一个特征参数的不确定性,包括:根据不确定性函数计算特征图中特征参数的不确定性,其中,在不确定函数的自变量越接近于0

时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0049] 其中,不确定性函数可以计算特征参数在迭代更新过程中的不确定性,从而完善参数量化机制,提升网络性能。

[0050] 在一种可能的实现方式中,不确定性函数公式为:

[0051] $$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

[0052] 其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与目标特征参数相邻的a个特征参数的不确定性,i,j,a均为自然数。

[0053] 在一种可能的实现方式中,基于特征参数所对应的不确定性计算特征参数的二值化特征,包括:在目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标特征参数进行二值化处理,得到目标特征参数的二值化特征。对特征图中一个或多个位置点的特征参数的不确定性进行联合考虑,来综合计算目标位置点(或者位置点)的特征参数的不确定性。

[0054] 在一种可能的实现方式中,基于特征参数所对应的不确定性计算特征参数的二值化特征,包括:在目标特征参数的目标不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标特征参数相邻的一个或多个特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0055] 基于不确定性的相关计算对特征参数进行二值化处理,对于不确定性较大(也即目标位置点的不确定性大于第二预设阈值)的特征参数采用平均池化并且引入sign函数,并从空间维度对目标位置点的特征参数进行二值化。

[0056] 在一种可能的实现方式中,基于第二权重矩阵和特征矩阵对神经网络进行训练,包括:对第二权重矩阵和特征矩阵进行二维卷积,得到神经网络中输出层的输出结果;根据输出结果得到损失函数;通过损失函数计算第一权重矩阵中权重参数的梯度;根据权重参数的梯度更新权重参数来对神经网络进行训练。其中,基于不确定性计算得到的二值化特征和二值化权重可以较小优化方向的不稳定性,提高神经网络的收敛速度。

[0057] 第四方面,本申请实施例提供了一种数据处理设备,该数据处理设备可以包括:获取单元,用于获取待处理数据;输入单元,用于将待处理数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图;计算单元,用于计算特征图中特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;量化单元,用于根据特征参数的不确定性计算特征参数的二值化特征;计算单元,还用于基于二值化特征得到待处理数据的处理结果。

[0058] 在一种可能的实现方式中,神经网络为根据二值化权重训练得到的,二值化权重为根据神经网络中权重参数的不确定性所得到的,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的一个或多个权重参数的符号的波动性。

[0059] 在一种可能的实现方式中,计算单元,具体用于:根据不确定性函数计算特征图中特征参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0060] 在一种可能的实现方式中,计算单元,具体用于:根据不确定性函数计算特征图上的位置点所对应的特征参数不确定性;根据与目标位置点相邻的一个或多个位置点的特征参数的不确定性,计算目标位置点的不确定性,目标位置点特征图上的任意一个位置点。

[0061] 在一种可能的实现方式中,量化单元,具体用于:在目标位置点的特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标位置点的特征参数进行二值化处理,得到目标位置点的二值化特征。

[0062] 在一种可能的实现方式中,量化单元,具体用于:在目标位置点的特征参数的不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标位置点相邻的一个或多个位置点的特征参数进行二值化处理,得到目标位置点的二值化特征。

[0063] 第五方面,本申请实施例提供了一种神经网络的量化装置,该装置可以包括输入单元,

[0064] 获取单元,还用于获取第一权重矩阵,第一权重矩阵中包含神经网络中用于提取特征的参数,第一权重矩阵包含s*k个权重参数,s和k为正整数;计算单元,用于计算第一权重矩阵中每一个权重参数的不确定性,其中,权重参数为神经网络的权重中的任意一个权重,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性;量化单元,用于基于权重参数所对应的不确定性计算权重参数的二值化权重,二值化权重用于确定第二权重矩阵,第二权重矩阵中包含的s*k个二值化权重与s*k个权重参数一一对应。

[0065] 在一种可能的实现方式中,计算单元,具体用于:根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0066] 在一种可能的实现方式中,计算单元,具体用于:在当前迭代次数小于或等于预设迭代次数时,通过不确定性函数计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0067] 在一种可能的实现方式中,计算单元,具体用于:在当前迭代次数大于预设迭代次数时,根据在参考迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算当前迭代次数所对应的第一权重矩阵的每一个权重参数的不确定性,其中,参考迭代次数为最接近当前迭代次数的预设迭代次数。

[0068] 在一种可能的实现方式中,量化单元,具体用于:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0069] 在一种可能的实现方式中,量化单元,具体用于:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性大于第一值时,将当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0070] 在一种可能的实现方式中,该装置还可以包括:获取单元,用于获取训练数据;输入单元,用于将训练数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图;量化单元,用于计算特征图中每一个特征参数的二值化特征,其中,特征图包含m*n个特征参数, m和n为正整数,特征图为在神经网络的一个或多个卷积层中提取的训练数据的特征。在一种可能的实现方式中,量化单元,具体用于:确定特征图中每一个特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;基于特征参数所对应的不确定性计算特征参数的二值化特征,二值化特征用于确定特征矩阵,特征矩阵中包含的m*n个二值化特征与m*n个特征参数一一对应。

[0071] 在一种可能的实现方式中,量化单元,具体用于:根据不确定性函数计算特征图中特征参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0072] 在一种可能的实现方式中,不确定性函数公式为:

[0073] $$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

[0074] 其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与目标特征参数相邻的a个特征参数的不确定性,i,j,a均为自然数。

[0075] 在一种可能的实现方式中,量化单元,具体用于:在目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0076] 在一种可能的实现方式中,量化单元,具体用于:在目标特征参数的目标不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标特征参数相邻的一个或多个特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0077] 第六方面,本申请实施例提供了一种数据处理设备,该数据处理设备可以包括存储器和处理器,存储器用于存储支持训练设备执行上述方法的计算机程序,计算机程序包括程序指令,处理器被配置用于调用程序指令,执行上述第一方面的方法。

[0078] 第七方面,本申请实施例提供了一种量化设备,该量化设备可以包括存储器和处理器,存储器用于存储支持数据处理装置执行上述方法的计算机程序,计算机程序包括程序指令,处理器被配置用于调用程序指令,执行上述第二方面的方法。

[0079] 第八方面,本申请实施例还提供一种计算机可读存储介质,计算机存储介质存储有计算机程序,计算机程序包括程序指令,程序指令当被处理器执行时使处理器执行上述第一方面的方法。

[0080] 第九方面,本申请实施例还提供一种计算机可读存储介质,计算机存储介质存储有计算机程序,计算机程序包括程序指令,程序指令当被处理器执行时使处理器执行上述第二方面的方法。

[0081] 第十方面,本申请实施例还提供了一种计算机程序,计算机程序包括计算机软件指令,计算机软件指令当被计算机执行时使计算机执行如第一方面、第二方面或者第三方面的任一种方法。

[0082] 第十一方面,本申请实施例还提供了一种包含指令的计算机程序产品,计算机程

序产品在电子设备上运行时,使得电子设备执行如第一方面、第二方面或者第三方面的任一种方法。

附图说明

[0083] 以下对本申请实施例用到的附图进行介绍。

[0084] 图1A为本申请实施例提供的一种通过全精度神经网络模型处理数据所需要的计算时间的示意图;

[0085] 图1B为本申请实施例提供的一种高阶近似方法对符号函数进行逼近的示意图;

[0086] 图1C为本申请实施例提供的一种在二值化过程中接近于零的一个或多个权重参数的符号波动性的示意图;

[0087] 图2为本申请实施例提供的一种人工智能主体框架的一种结构示意图;

[0088] 图3为本申请实施例提供的一种系统架构100的示意图;

[0089] 图4A为本申请实施例提供的一种卷积神经网络的结构示意图;

[0090] 图4B为本申请实施例提供的另一种卷积神经网络的结构示意图;

[0091] 图4C为本申请实施例提供的一种全连接网络的结构示意图;

[0092] 图5为本申请实施例提供的一种芯片的硬件结构示意图;

[0093] 图6为本申请实施例提供的一种数据处理方法的流程示意图;

[0094] 图7为本申请实施例提供的一种数据处理方法的网络架构示意图;

[0095] 图8A为本申请实施例提供的一种神经网络的量化方法的流程示意图;

[0096] 图8B为本申请实施例提供的一种不确定性函数的示意图;

[0097] 图8C为本申请实施例提供的一种二值化结果的示意图;

[0098] 图9为本申请实施例提供的一种神经网络的量化方法的流程示意图;

[0099] 图10为本申请实施例提供的一种神经网络的训练方法的流程示意图;

[0100] 图11为本申请实施例提供的一种数据处理装置的结构示意图;

[0101] 图12为本申请实施例提供的一种神经网络的量化装置的结构示意图;

[0102] 图13为本申请实施例提供的一种数据处理设备的结构示意图;

[0103] 图14为本申请实施例提供的一种神经网络的量化设备的结构示意图。

具体实施方式

[0104] 下面结合附图对本申请实施例中的技术方案进行清楚、完整的描述,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。

[0105] 本申请的说明书以及附图中的术语“第一”和“第二”等是用于区分不同的对象,或者用于区别对同一对象的不同处理,而不是用于描述对象的特定顺序。此外,本申请的描述中所提到的术语“包括”和“具有”以及它们的任何变形,意图在于覆盖不排他的包含。例如包含了一些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,而是可选地还包括其他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他步骤或单元。需要说明的是,本申请实施例中,“示例性地”或者“例如”等词用于表示作例子、例证或说明。本申请实施例中被描述为“示例性地”或者“例如”的任何实施例或设计方法不应被解释为比其他实施例或设计方案更优地或更具优势。确切而

言,使用“示例性地”或者“例如”等词旨在以具体方式呈现相关概念。在本申请实施例中,“A和/或B”表示A和 B,A或B两个含义。“A,和/或B,和/或C”表示A、B、C中的任一个,或者,表示A、B、 C中的任两个,或者,表示A和B和C。下面将结合附图,对本申请中的技术方案进行描述。

[0106] 随着人工智能技术(Artificial Intelligence,AI)的发展,算法的准确率往往依赖于海量数据的训练分析,因此需要把大量的数据输入到模型中来完成对模型的训练。由于基于海量数据的训练分析会占用较大的存储量和计算量,而神经网络二值化能够最大程序地降低全精度神经网络模型的存储占用和模型的计算量,将神经网络中原本32浮点数的参数量化到1bit 整型的参数,降低了模型部署的存储资源消耗,同时极大加速了神经网络的推断过程。如表 1所示为全精度神经网络模型所需要的计算量和存储量。其中,视觉几何组(visual geometry group,VGG)模型是一种卷积神经网络模型;残差神经网络(residual networks,ResNets) 指的是在传统卷积神经网络中加入了残差学习(residual learning)的思想;SENet(Squeeze-and-Excitation Networks)是一种图像识别结构。

[0107] 表1全精度神经网络模型所需要的计算量和存储量

[0108]

模型 (Model)	每秒所执行的浮点运算次数 (floating-point operations per second, FLOPs)	内存使用量 (Memory Usage)
VGG-16	15.74 千兆字节 (Gigabyte, GB)	11.49GB
ResNet-50	7.7GB	819.45 兆字节(Megabytes, MB)
SENet-154	21.21GB	440.34MB

[0109] 如图1A所示为本申请实施例提供的一种通过全精度神经网络模型处理数据所需要的计算时间的示意图。图1A中的(a)为VGG16模型在不同的批尺寸(batch size)下所需要的计算时间,图1A中的(b)为GoogleNet模型在不同的批尺寸(batch size)下所需要的计算时间。其中,batch size为一次训练所选取的样本数。从图1A可以看出,全精度神经网络模型所需要的计算时间较长,难以满足实时性计算要求。

[0110] 但是,二值化会不可避免地导致严重的信息损失,其量化函数不连续性也给深度网络的优化带来了困难。其中,二值神经网络中离散权值优化是导致性能下降的主要原因之一,该优化通常使用一个非平滑符号函数执行,除零点外,其函数的导数处处为零,将零周围的点称为“敏感点”。

[0111] 为了处理这些“敏感点”,现有的二值神经网络倾向于近似符号函数的导数或者符号函数本身。通过直通估计器(straight through estimator,STE)提出用二值神经网络的恒等式来近似符号函数的导数,但是可能会带来较大的误差。受STE的启发,引入了更精确的近似,包括一阶近似和指数多项式近似,来代替导数从而减小梯度误差。可以看出,上述这些方法提供了“敏感点”的梯度,如图1B所示为高阶近似方法对符号函数进行逼近的示意图。

[0112] 其中,图1B中的(a)为符号函数sign(x)和sign(x)的梯度$\frac{\partial \text{sign}(x)}{\partial x}$的示意图。图1B的(b) 为裁剪函数Clip(-1,x,1)和Clip(-1,x,1)的梯度$\frac{\partial \text{Clip}(-1,\ x,\ 1)}{\partial x}$的示意图。图1B中的(c)为自定义的符号函数Approxsign(x)和Approxsign的梯度$\frac{\partial \text{Approxsign}(x)}{\partial x}$的示意图。图1B中的(d) 为自定义的符号函数3-oder-Approxsign(x)和3-oder-Approxsign(x)的梯度

$\frac{\partial 3-order-Approxsign(x)}{\partial x}$的示意图。

[0113] 从图1B可以看出,除直接导数近似外,在对模型的训练中还可以采用具有自适应参数或者可学习参数的符号函数,比如说采用渐进方法来估计符号函数。而随着对模型训练的进行,这些类符号函数扩大了“敏感点”的梯度,使梯度大到足以可以改变“敏感点”的二值权值。

[0114] 另外,本申请实施例还提供了以下方案来处理这些“敏感点”。

[0115] 方案一:BONN基于贝叶斯方法最小量化误差,并将实权值重新分配为双峰分布。利用贝叶斯学习的有效性以端到端的方式构建二值神经网络(1-Bit Convolutional Neural Networks,1-Bit CNNs)。特别地,引入了两个新的贝叶斯损失,在此基础上优化1-Bit CNNs,从而可以提高效率和稳定性。在统一的理论框架下,这些贝叶斯损失不仅考虑了1-Bit CNNs的核与权重的分布,而且还监督了特征分布。基于贝叶斯的核损失改善了每个卷积层的逐层核分布,而基于贝叶斯特征损失引入了类紧性来减轻量化过程所带来的干扰。需要注意的是,贝叶斯特征损失仅适用于全连接层。

[0116] 方案一引入的贝叶斯损失增加了模型训练的复杂度,在实际应用过程中稳定性不强,反向传播过程中存在梯度爆炸的风险。

[0117] 方案二:抗锯齿卷积神经网络(Anti-aliasing CNN,AA-CNN)。传统信息处理领域对于抗锯齿的技术,一般是采用增大采样频率。但是由于图像处理任务一般都需要降采样,所以还可以采用图像模糊(bluring)技术。根据奈奎斯特采样理论,先给定采样频率,通过降低原始信号的频率来使得信号可以被重构处理。对于模糊化处理和未处理的原图像进行下采样,模糊化处理的原图像下采样的图像可以看出一些轮廓,而未处理的原图像下采样的图像就显得更加混乱。通过将抗锯齿和降采样结合到一起,面对图像损坏,模型具有鲁棒性。

[0118] 二值神经网络的目的是将特征和权重进行1bit量化,而方案二的抗锯齿操作虽然能够有效应对图片损坏,但是无法解决模型二值化产生的信息丢失问题。量化不确定度的问题依然存在,全精度模型中零附近的参数依然不具有量化鲁棒性。

[0119] 方案三:信息保持网络(information preserving network,IR-Net)从统一信息的角度研究了二值神经网络的前向和后向传播过程,引入了信息熵损失,同时对其进行了量化误差优化。网络在前向传播过程中可以携带足够的信息和反向传播中得到的梯度能够为网络优化提高正确的信息是网络保持高性能的关键之一。IR-Net在前向传播中引入了一种称为Libra Parameter Binarization(Libra)二值化的平衡标准量化方法,来最大化量化参数的信息熵和最小化量化误差,通过整数移位标量扩展了权重的表示能力。在反向传播过程中采用误差衰减估计器来计算梯度,保证训练开始时的充分更新和训练结束时的精确梯度。

[0120] 方案三中IR-Net无法最大化权重参数的差异,造成量化不准确。对“敏感点”的量化能力较差,从而导致量化信息丢失。

[0121] 综上所述,对网络模式进行1bit量化,可能会由于量化误差而造成信息丢失。为了减少信息丢失,提出了许多解决方法,如最小化实值权值与二值化权值之间的距离或调整参数的分布以减少量化误差。然而,二值神经网络和全精度神经网络之间仍然存在着不小

的精度差距。现有的方法都无法解决关于“敏感点”的量化问题。比如说现有导数或符号函数逼近方法都强调了“敏感点”的梯度大小,而忽略了“敏感点”的优化方法。由于“敏感点”的不稳定性。该符号函数可能会导致不稳定的优化方向。如图1C所示为在二值化过程中接近于零的一个或多个权重参数的符号波动性的示意图。从图1C可以看出,在二值化过程中,接近于量的权重更不稳定。比如说假设权重为0.001时,该权重的符号为+1;当权重进行了微小的变化,比如说权重从0.001更新到-0.001时,该权重的符号为-1。因此,权重在零附近的波动会由于频繁越过零点而导致权值优化方向频繁跳变,方向性不确定的学习可能会导致神经网络收敛缓慢和不稳定。

[0122] 因此,本申请实施提出了对二值化的不确定性进行建模,并更加不确定性来确定优化方向,从而改善二值化神经网络的相关性能(比如说精度,收敛速度等)。

[0123] 首先对人工智能系统总体工作流程进行描述,请参见图2,图2示出的为人工智能主体框架的一种结构示意图,下面从“智能信息链”(水平轴)和“IT价值链”(垂直轴)两个维度对上述人工智能主体框架进行阐述。其中,“智能信息链”反映从数据的获取到处理的一系列过程。举例来说,可以是智能信息感知、智能信息表示与形成、智能推理、智能决策、智能执行与输出的一般过程。在这个过程中,数据经历了“数据—信息—知识—智慧”的凝练过程。“IT价值链”从人智能的底层基础设施、信息(提供和处理技术实现)到系统的产业生态过程,反映人工智能为信息技术产业带来的价值。

[0124] (1)基础设施

[0125] 基础设施为人工智能系统提供计算能力支持,实现与外部世界的沟通,并通过基础平台实现支撑。通过传感器与外部沟通;计算能力由智能芯片(CPU、GPU、NPU、ASIC、FPGA等硬件加速芯片)提供;基础平台包括分布式计算框架及网络等相关的平台保障和支持,可以包括云存储和计算、互联互通网等。举例来说,传感器和外部沟通获取数据,这些数据提供给基础平台提供的分布式计算系统中的智能芯片进行计算。

[0126] (2)数据

[0127] 基础设施的上一层的数据用于表示人工智能领域的数据来源。数据涉及到图形、图像、语音、文本,还涉及到传统设备的物联网数据,包括已有系统的业务数据以及力、位移、液位、温度、湿度等感知数据。

[0128] (3)数据处理

[0129] 数据处理通常包括数据训练,机器学习,深度学习,搜索,推理,决策等方式。

[0130] 其中,机器学习和深度学习可以对数据进行符号化和形式化的智能信息建模、抽取、预处理、训练等。

[0131] 推理是指在计算机或智能系统中,模拟人类的智能推理方式,依据推理控制策略,利用形式化的信息进行机器思维和求解问题的过程,典型的功能是搜索与匹配。

[0132] 决策是指智能信息经过推理后进行决策的过程,通常提供分类、排序、预测等功能。

[0133] (4)通用能力

[0134] 对数据经过上面提到的数据处理后,进一步基于数据处理的结果可以形成一些通用的能力,比如可以是算法或者一个通用系统,例如,翻译,文本的分析,计算机视觉的处理,语音识别,图像的识别等等。

[0135] (5)智能产品及行业应用

[0136] 智能产品及行业应用指人工智能系统在各领域的产品和应用,是对人工智能整体解决方案的封装,将智能信息决策产品化、实现落地应用,其应用领域主要包括:智能终端、智能交通、智能医疗、自动驾驶、平安城市等。

[0137] 本申请实施例主要应用在驾驶辅助、自动驾驶、手机终端等领域。

[0138] 下面介绍几种应用场景:

[0139] 应用场景1:高级驾驶辅助系统(Advanced Driver Assistance System,ADAS)/自动驾驶解决方案(Autonomous Driving Solution,ADS)

[0140] 在ADAS和ADS中,需要实时进行多类型的2D目标检测,包括:动态障碍物(行人(Pedestrian)、骑行者(Cyclist)、三轮车(Tricycle)、轿车(Car)、卡车(Truck)、公交车(Bus)),静态障碍物(交通锥标(TrafficCone)、交通棍标(TrafficStick)、消防栓(FireHydrant)、摩托车 (Motocycle)、自行车(Bicycle)),交通标志(TrafficSign、导向标志(GuideSign)、广告牌(Billboard)、红色交通灯(TrafficLight_Red)/黄色交通灯(TrafficLight_Yellow)/绿色交通灯 (TrafficLight_Green)/黑色交通灯(TrafficLight_Black)、路标(RoadSign))。另外,为了准确获取动态障碍物的在3维空间所占的区域,还需要对动态障碍物进行3D估计,输出3D框。为了与激光雷达的数据进行融合,需要获取动态障碍物的Mask,从而把打到动态障碍物上的激光点云筛选出来;为了进行精确的泊车位,需要同时检测出泊车位的4个关键点;为了进行构图定位,需要检测出静态目标的关键点。这是一个语义分割问题。自动驾驶车辆的摄像头捕捉到道路画面,需要对画面进行分割,分出路面、路基、车辆、行人等不同物体,从而保持车辆行驶在正确的区域。对于安全型要求极高的自动驾驶需要实时对画面进行理解,能够实时运行的进行语义分割的卷积神经网络至关重要。

[0141] 应用场景2:图像分类场景

[0142] 物体识别装置在获取待分类图像后,通过基于本申请实施例的神经网络的量化方法所训练得到的分类模型对待分类图像中的物体进行处理,得到待分类图像的类别,然后可根据待分类图像中物体的物体类别对待分类图像进行分类。对于摄影师来说,每天会拍很多照片,有动物的,有人物,有植物的。采用本申请的方法可以快速地将照片按照照片中的内容进行分类,可分成包含动物的照片、包含人物的照片和包含植物的照片。

[0143] 对于图像数量比较庞大的情况,人工分类的方式效率比较低下,并且人在长时间处理同一件事情时很容易产生疲劳感,此时分类的结果会有很大的误差。

[0144] 应用场景3:商品分类

[0145] 物体识别装置获取商品的图像后,通过基于本申请实施例的神经网络的量化方法所训练得到的分类模型商品的图像进行处理,得到商品的图像中商品的类别,然后根据商品的类别对商品进行分类。对于大型商场或超市中种类繁多的商品,采用本申请的物体识别方法可以快速完成商品的分类,降低了时间开销和人工成本。

[0146] 应用场景4:入口闸机人脸验证

[0147] 这是一个图像相似度比对问题。在高铁、机场等入口的闸机上,乘客进行人脸认证时,摄像头会拍摄人脸图像,使用卷积神经网络抽取特征,和存储在系统中的身份证件的图像特征进行相似度计算,如果相似度高就验证成功。其中,卷积神经网络抽取特征是最耗时

的,要快速进行人脸验证,需要高效的卷积神经网络进行特征提取。

[0148] 应用场景5:翻译机同声传译

[0149] 这是一个语音识别和机器翻译问题。在语音识别和机器翻译问题上,卷积神经网络也是常有的一种识别模型。在需要同声传译的场景,必须达到实时语音识别并进行翻译,高效的卷积神经网络可以给翻译机带来更好的体验。

[0150] 本申请实施例训练出的神经网络模型可以实现上述功能。

[0151] 本申请实施例提供的神经网络的量化方法,可以涉及计算机视觉的处理或自然语言的处理等等,具体可以应用于数据训练、机器学习、深度学习等数据处理方法,对训练数据进行符号化和形式化的智能信息建模、抽取、预处理、训练等,最终得到训练好的神经网络模型 (也即:目标模型/规则)。并且,本申请实施例提供的数据处理方法可以运用上述训练好的神经网络模型中,得到输出数据(如:图片的识别结果)。需要说明的是,本申请实施例提供的神经网络的训练方法和数据处理方法是基于同一个构思产生的发明。

[0152] 由于本申请实施例涉及大量神经网络的应用,为了便于理解,下面先对本申请实施例涉及的相关术语及神经网络等相关概念进行介绍。

[0153] (1)神经网络

[0154] 神经网络可以是由神经单元组成的,神经单元可以是指以x_s和截距b为输入的运算单元,该运算单元的输出可以为:

[0155] $$h_{w,b}(x)=f(W^T x)=f\left(\sum_{s=1}^{n}W_s x_s+b\right) \quad (1\text{-}1)$$

[0156] 其中,s=1、2、……n,n为大于1的自然数,w_s为x_s的权重,b为神经单元的偏置。f 为神经单元的激活函数(activation functions),用于将非线性特性引入神经网络中,来将神经单元中的输入信号转换为输出信号。该激活函数的输出信号可以作为下一层卷积层的输入。激活函数可以是sigmoid函数。神经网络是将许多个上述单一的神经单元联结在一起形成的网络,即一个神经单元的输出可以是另一个神经单元的输入。每个神经单元的输入可以与前一层的局部接受域相连,来提取局部接受域的特征,局部接受域可以是由若干个神经单元组成的区域。

[0157] (2)深度神经网络

[0158] 深度神经网络(deep neural network,DNN),也称多层神经网络,可以理解为具有很多层隐含层的神经网络,这里的“很多”并没有特别的度量标准。从DNN按不同层的位置划分, DNN内部的神经网络可以分为三类:输入层,隐含层,输出层。一般来说第一层是输入层,最后一层是输出层,中间的层数都是隐含层。层与层之间是全连接的,也就是说,第i层的任意一个神经元一定与第i+1层的任意一个神经元相连。虽然DNN看起来很复杂,但是就每一层的工作来说,其实并不复杂,简单来说就是如下线性关系表达式:$\vec{y}=\alpha(W\vec{x}+\vec{b})$,其中,$\vec{x}$是输入向量,$\vec{y}$是输出向量,b是偏移向量,W是权重矩阵(也称系数),α()是激活函数。每一层仅仅是对输入向量α经过如此简单的操作得到输出向量$\vec{y}$。由于DNN层数多,则系数 W和偏移向量b的数量也就很多了。这些参数在DNN中的定义如下所述:以系数W为例:假设在一个三层的DNN中,第二层的第4个神经元到第三层的第2个神经元的线性系数定义为W_{24}^3。上标3代表系数W所在的层数,而下标对应的是输出的第三层索引2和输入的第二层索引4。

总结就是:第L-1层的第k个神经元到第L层的第j个神经元的系数定义为W_{jk}^{L}。需要注意的是,输入层是没有W参数的。在深度神经网络中,更多的隐含层让网络更能够刻画现实世界中的复杂情形。理论上而言,参数越多的模型复杂度越高,“容量”也就越大,也就意味着它能完成更复杂的学习任务。训练深度神经网络的也就是学习权重矩阵的过程,其最终目的是得到训练好的深度神经网络的所有层的权重矩阵(由很多层的向量w形成的权重矩阵)。

[0159] (3)卷积神经网络

[0160] 卷积神经网络(convolutional neuron network,CNN)是一种带有卷积结构的深度神经网络。卷积神经网络包含了一个由卷积层和子采样层构成的特征抽取器。该特征抽取器可以看作是滤波器,卷积过程可以看作是使用一个可训练的滤波器与一个输入的数据(如图像数据,以图像数据为例描述)或者卷积特征平面(feature map)做卷积。卷积层是指卷积神经网络中对输入信号进行卷积处理的神经元层。在卷积神经网络的卷积层中,一个神经元可以只与部分邻层神经元连接。一个卷积层中,通常包含若干个特征平面,每个特征平面可以由一些矩形排列的神经单元组成。同一特征平面的神经单元共享权重,这里共享的权重就是卷积核。共享权重可以理解为提取图像信息的方式与位置无关。这其中隐含的原理是:图像的某一部分的统计信息与其他部分是一样的。即意味着在某一部分学习的图像信息也能用在另一部分上。所以对于图像上的所有位置,都能使用同样的学习得到的图像信息。在同一卷积层中,可以使用多个卷积核来提取不同的图像信息,一般地,卷积核数量越多,卷积操作反映的图像信息越丰富。

[0161] 卷积核可以以随机大小的矩阵的形式初始化,在卷积神经网络的训练过程中卷积核可以通过学习得到合理的权重。另外,共享权重带来的直接好处是减少卷积神经网络各层之间的连接,同时又降低了过拟合的风险。

[0162] (4)循环神经网络

[0163] 循环神经网络(recurrent neural networks,RNN)是用来处理序列数据的。在传统的神经网络模型中,是从输入层到隐含层再到输出层,层与层之间是全连接的,而对于每一层层内之间的各个节点是无连接的。这种普通的神经网络虽然解决了很多难题,但是却仍然对很多问题却无能无力。例如,你要预测句子的下一个单词是什么,一般需要用到前面的单词,因为一个句子中前后单词并不是独立的。RNN之所以称为循环神经网路,即一个序列当前的输出与前面的输出也有关。具体的表现形式为网络会对前面的信息进行记忆并应用于当前输出的计算中,即隐含层本层之间的节点不再无连接而是有连接的,并且隐含层的输入不仅包括输入层的输出还包括上一时刻隐含层的输出。理论上,RNN能够对任何长度的序列数据进行处理。对于RNN的训练和对传统的CNN或DNN的训练一样。同样使用误差反向传播算法,不过有一点区别:即,如果将RNN进行网络展开,那么其中的参数,如W,是共享的;而如上举例上述的传统神经网络却不是这样。并且在使用梯度下降算法中,每一步的输出不仅依赖当前步的网络,还依赖前面若干步网络的状态。该学习算法称为基于时间的反向传播算法Back propagation Through Time(也即:BPTT)。

[0164] 既然已经有了卷积神经网络,为什么还要循环神经网络?原因很简单,在卷积神经网络中,有一个前提假设是:元素之间是相互独立的,输入与输出也是独立的,比如猫和狗。但现实世界中,很多元素都是相互连接的,比如股票随时间的变化,再比如一个人说了:我喜欢旅游,其中最喜欢的地方是云南,以后有机会一定要去(__)。这里填空,人类应该都知

道是填“云南”。因为人类会根据上下文的内容进行推断,但如何让机器做到这一步?RNN就应运而生了。RNN旨在让机器像人一样拥有记忆的能力。因此,RNN的输出就需要依赖当前的输入信息和历史的记忆信息。

[0165] (5)损失函数

[0166] 在训练深度神经网络的过程中,因为希望深度神经网络的输出尽可能的接近真正想要预测的值,所以可以通过比较当前网络的预测值和真正想要的目标值,再根据两者之间的差异情况来更新每一层神经网络的权重向量(当然,在第一次更新之前通常会有过程,即为深度神经网络中的各层预先配置参数),比如,如果网络的预测值高了,就调整权重向量让它预测低一些,不断的调整,直到深度神经网络能够预测出真正想要的目标值或与真正想要的目标值非常接近的值。因此,就需要预先定义“如何比较预测值和目标值之间的差异”,这便是损失函数(loss function)或目标函数(objective function),它们是用于衡量预测值和目标值的差异的重要方程。其中,以损失函数举例,损失函数的输出值(loss)越高表示差异越大,那么深度神经网络的训练就变成了尽可能缩小这个loss的过程。

[0167] (6)反向传播算法

[0168] 卷积神经网络可以采用误差反向传播(back propagation,BP)算法在训练过程中修正初始模型中参数的大小,使得初始模型的重建误差损失越来越小。具体地,前向传递输入信号直至输出会产生误差损失,通过反向传播误差损失信息来更新初始模型中参数,从而使误差损失收敛。反向传播算法是以误差损失为主导的反向传播运动,旨在得到最优的目标模型的参数,例如权重矩阵。

[0169] (7)模型量化

[0170] 模型量化(model quantization)是通用的深度学习优化的手段之一,一方面模型量化可以降低内存和存储的开销,另一方面还可以加快模型的收敛速度,以及提高模型的推理效率。在本申请实施例中,量化是将一组原始值域范围内的数,通过一个数学变换将原始值域映射到另一个目标值域范围的过程。例如,将神经网络的模型参数由浮点数转换为整形数。

[0171] (8)二值神经网络

[0172] 二值神经网络(Binary Neural Network,BNN)是指在全精度神经网络(参数为32为浮点数的网络)的基础上,将全精度神经网络中的参数值进行二值化处理得到的神经网络。也即,将32为浮点数的参数二值化为1bit整型(1或者-1)。通过二值化处理,可以使得参数占用更小的存储空间(内存消耗理论上减少为原来的1/32倍,从float32到1bit),同时利用位操作来代替网络中的乘加运算,可以降低运算时间。

[0173] 下面介绍本申请实施例提供的系统架构。

[0174] 参见图3,本申请实施例提供的一种系统架构100。如所述系统架构100所示,数据采集设备160用于采集或生成训练数据,本申请实施例中训练数据包括:带标签的多张图像或者多个语音片段等;并将训练数据存入数据库130,训练设备120可确定神经网络中一个或多个卷积层提取的特征图,特征图包含m*n个特征参数,m和n为正整数。然后,训练设备120 可以确定特征图中每一个特征参数的不确定性,基于特征参数所对应的不确定性计算特征参数的二值化特征。其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性。二值化特征用于确定特征矩阵,所述特征矩阵中包含

的m*n个二值化特征与所述m*n个特征参数一一对应。

[0175] 训练设备120可以获取第一权重矩阵,第一权重矩阵中包含神经网络中用于提取特征的参数,第一权重矩阵包含s*k个权重参数,s和k为正整数。然后,训练设备120可以计算第一权重矩阵中每一个权重参数的不确定性,基于权重参数所对应的不确定性计算权重参数的二值化权重。其中,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性。二值化权重用于确定第二权重矩阵,第二权重矩阵中包含的s*k个二值化权重与所述s*k个权重参数一一对应。

[0176] 最后,训练设备120可以基于上述二值化权重和上述二值化特征对神经网络进行训练。具体来说,该训练好的神经网络模型能够用于实现本申请实施例提供的数据处理方法

[0177] 需要说明的是,在实际应用中,数据库130中维护的训练数据不一定都来自于数据采集设备160的采集,也有可能是从其他设备接收得到的。另外需要说明的是,训练设备120也不一定完全基于数据库130维护的训练数据进行神经网络模型的训练,也有可能从云端或其他地方获取训练数据进行模型训练,上述描述不应该作为对本申请实施例的限定。

[0178] 根据训练设备120训练得到的目标模型/规则101可以应用于不同的系统或设备中,如应用于图3所示的执行设备110,所述执行设备110可以是终端,如手机终端,平板电脑,笔记本电脑,增强现实(augmented reality,AR)AR/虚拟现实(virtual reality,VR),车载终端等,还可以是服务器或者云端等。在图3中,执行设备110配置输入/输出(input/output,I/O)接口 112,用于与外部设备进行数据交互,用户可以通过客户设备140向I/O接口112输入数据,所述输入数据在本申请实施例中可以包括:待识别图像、视频或待识别的语音片段。

[0179] 在执行设备120对输入数据进行预处理,或者在执行设备120的计算模块111执行计算等相关的处理(比如进行本申请中神经网络的功能实现)过程中,执行设备120可以调用数据存储系统150中的数据、代码等以用于相应的处理,也可以将相应处理得到的数据、指令等存入数据存储系统150中。

[0180] 最后,I/O接口112将处理结果,如图像、视频或语音的识别结果或分类结果返回给客户设备140,从而客户设备140可以提供给用户设备150。该用户设备150可以是需要使用目标模型/规则101的轻量级终端,如手机终端、笔记本电脑、AR/VR终端或车载终端等,以用于响应于终端用户的相应需求,如对终端用户输入的图像进行图像识别输出识别结果给该终端用户,或对终端用户输入的文本进行文本分类输出分类结果给该终端用户等。

[0181] 值得说明的是,训练设备120可以针对不同的目标或称不同的任务,基于不同的训练数据生成相应的目标模型101,该相应的目标模型101即可以用于实现上述目标或完成上述任务,从而为用户提供所需的结果。

[0182] 在图3中所示情况下,用户可以手动给定输入数据,该手动给定可以通过I/O接口112 提供的界面进行操作。另一种情况下,客户设备140可以自动地向I/O接口112发送输入数据,如果要求客户设备140自动发送输入数据需要获得用户的授权,则用户可以在客户设备 140中设置相应权限。用户可以在客户设备140查看执行设备110输出的结果,具体的呈现形式可以是显示、声音、动作等具体方式。客户设备140也可以作为数据采集端,采集如图所示输入I/O接口112的输入数据及输出I/O接口112的输出结果作为新的样本数据,并存入

数据库130。当然,也可以不经过客户设备140进行采集,而是由I/O接口112直接将如图所示输入I/O接口112的输入数据及输出I/O接口112的输出结果,作为新的样本数据存入数据库130。

[0183] 客户设备140在接收到输出结果后,可以将结果传输给用户设备150,用户设备150可以是终端,如手机终端,平板电脑,笔记本电脑,AR/VR,车载终端等。在其中一个示例中,用户设备150可以运行目标模型/规则101,以实现特定的功能。

[0184] 值得注意的是,图3仅是本申请实施例提供的一种系统架构的示意图,图中所示设备、器件、模块等之间的位置关系不构成任何限制,例如,在图3中,数据存储系统150相对执行设备110是外部存储器,在其它情况下,也可以将数据存储系统150置于执行设备110中。

[0185] 如图3所示,根据训练设备120训练得到目标模型/规则101,该目标模型/规则101可以是应用场景2和应用场景3中的分类模型,应用场景4中的图像识别模型,应用场景5中的语音识别模型,具体的,本申请实施例提供的目标模型/规则101,例如,图像识别模型;又例如,语音识别模型等等,在实际应用中,图像识别模型、语音识别模型都可以是卷积神经网络模型。

[0186] 为了便于理解以及出于阐述的便利,在本申请实施例中,神经网络模型可以包括卷积神经网络、全连接网络等。如前文的基础概念介绍所述,卷积神经网络是一种带有卷积结构的深度神经网络,是一种深度学习(deep learning)架构,深度学习架构是指通过机器学习的算法,在不同的抽象层级上进行多个层次的学习。作为一种深度学习架构,CNN是一种前馈 (feed-forward)人工神经网络,该前馈人工神经网络中的各个神经元可以对输入其中的图像作出响应。

[0187] 在一些可能的实现方式中,如图4A所示的卷积神经网络的结构示意图,卷积神经网络 (CNN)200可以包括输入层210,卷积层/池化层220(其中池化层为可选的),以及神经网络层230。其中,输入层210可以获取待处理数据,并将获取到的待处理数据交由卷积层/池化层220以及后面的神经网络层230进行处理,可以得到图像的处理结果。下面对图4A中的 CNN 200中内部的层结构进行详细的介绍。

[0188] 卷积层/池化层220:

[0189] 卷积层:

[0190] 如图4A所示卷积层/池化层220可以包括如示例221-226层,举例来说:在一种实现中,221层为卷积层,222层为池化层,223层为卷积层,224层为池化层,225为卷积层,226为池化层;在另一种实现方式中,221、222为卷积层,223为池化层,224、225为卷积层,226 为池化层。即卷积层的输出可以作为随后的池化层的输入,也可以作为另一个卷积层的输入以继续进行卷积操作。

[0191] 下面将以卷积层221为例,介绍一层卷积层的内部工作原理。

[0192] 卷积层221可以包括很多个卷积算子,卷积算子也称为核,其在图像处理中的作用相当于一个从输入图像矩阵中提取特定信息的过滤器,卷积算子本质上可以是一个权重矩阵,这个权重矩阵通常被预先定义,在对图像进行卷积操作的过程中,权重矩阵通常在输入图像上沿着水平方向一个像素接着一个像素(或两个像素接着两个像素……这取决于步长stride的取值)的进行处理,从而完成从图像中提取特定特征的工作。该权重矩阵的大小应该与图像的大小相关,需要注意的是,权重矩阵的纵深维度(depth dimension)和输入图像

的纵深维度是相同的,在进行卷积运算的过程中,权重矩阵会延伸到输入图像的整个深度。因此,和一个单一的权重矩阵进行卷积会产生一个单一纵深维度的卷积化输出,但是大多数情况下不使用单一权重矩阵,而是应用多个尺寸(行×列)相同的权重矩阵,即多个同型矩阵。每个权重矩阵的输出被堆叠起来形成卷积图像的纵深维度,这里的维度可以理解为由上面所述的“多个”来决定。不同的权重矩阵可以用来提取图像中不同的特征,例如一个权重矩阵用来提取图像边缘信息,另一个权重矩阵用来提取图像的特定颜色,又一个权重矩阵用来对图像中不需要的噪点进行模糊化等。该多个权重矩阵尺寸(行×列)相同,经过该多个尺寸相同的权重矩阵提取后的卷积特征图的尺寸也相同,再将提取到的多个尺寸相同的卷积特征图合并形成卷积运算的输出。

[0193] 这些权重矩阵中的权重值在实际应用中需要经过大量的训练得到,通过训练得到的权重值形成的各个权重矩阵可以用来从输入图像中提取信息,从而使得卷积神经网络200进行正确的预测。

[0194] 当卷积神经网络200有多个卷积层的时候,初始的卷积层(例如221)往往提取较多的一般特征,该一般特征也可以称之为低级别的特征;随着卷积神经网络200深度的加深,越往后的卷积层(例如226)提取到的特征越来越复杂,比如高级别的语义之类的特征,语义越高的特征越适用于待解决的问题。

[0195] 池化层:

[0196] 由于常常需要减少训练参数的数量,因此卷积层之后常常需要周期性的引入池化层,在如图4A中220所示例的221-226各层,可以是一层卷积层后面跟一层池化层,也可以是多层卷积层后面接一层或多层池化层。具体来说,池化层,用于对数据进行采样,降低数据的数量。例如,以数据为图像数据为例,在图像处理过程中,通过池化层可以减少图像的空间大小。一般情况下,池化层可以包括平均池化算子和/或最大池化算子,以用于对输入图像进行采样得到较小尺寸的图像。平均池化算子可以在特定范围内对图像中的像素值进行计算产生平均值作为平均池化的结果。最大池化算子可以在特定范围内取该范围内值最大的像素作为最大池化的结果。另外,就像卷积层中用权重矩阵的大小应该与图像尺寸相关一样,池化层中的运算符也应该与图像的大小相关。通过池化层处理后输出的图像尺寸可以小于输入池化层的图像的尺寸,池化层输出的图像中每个像素点表示输入池化层的图像的对应子区域的平均值或最大值。

[0197] 神经网络层230:

[0198] 在经过卷积层/池化层220的处理后,卷积神经网络200还不足以输出所需要的输出信息。因为如前所述,卷积层/池化层220只会提取特征,并减少输入图像带来的参数。然而为了生成最终的输出信息(所需要的类信息或其他相关信息),卷积神经网络200需要利用神经网络层230来生成一个或者一组所需要的类的数量的输出。因此,在神经网络层230中可以包括多层隐含层(如图4A所示的231、232至23n)以及输出层240,该多层隐含层中所包含的参数可以根据具体的任务类型的相关训练数据进行预先训练得到,例如该任务类型可以包括图像识别,图像分类,图像超分辨率重建等等。

[0199] 在神经网络层230中的多层隐含层之后,也就是整个卷积神经网络200的最后层为输出层240,该输出层240具有类似分类交叉熵的损失函数,具体用于计算预测误差,一旦整个卷积神经网络200的前向传播(如图4A由210至240方向的传播为前向传播)完成,反向传

播 (如图4A由240至210方向的传播为反向传播)就会开始更新前面提到的各层的权重值以及偏差,以减少卷积神经网络200的损失,及卷积神经网络200通过输出层输出的结果和理想结果之间的误差。

[0200] 需要说明的是,如图4A所示的卷积神经网络200仅作为一种卷积神经网络的示例,在具体的应用中,卷积神经网络还可以以其他网络模型的形式存在。例如,如图4B所示的另一种卷积神经网络的结构示意图,图4B所示的卷积神经网络(CNN)300可以包括输入层310,卷积层/池化层320(其中池化层为可选的),以及神经网络层130。与图4A相比,图 4B中的卷积层/池化层320中的多个卷积层/池化层并行,将分别提取的特征均输入给神经网络层330进行处理。

[0201] 又例如,上述神经网络模型为全连接网络。全连接网络是指对n-1层和n层而言,n-1层的任意一个节点(又称为神经元),都和n层的所有节点有连接。具体地,参见图4C,是本申请实施例提供的一种全连接层的结构示意图,如图4C所示,该神经网络包括输入层、隐含层以及输出层,其中,输入层到隐含层之间的这一全连接层的二维参数矩阵为(3,4),该二维参数矩阵(3,4)表示在输入层到隐含层之间的全连接层结构中,输入神经元的个数为3,输出神经元的个数为4,权值数量为12。可以理解的是,神经元与神经元之间均具有连接关系。

[0202] 下面介绍本申请实施例提供的一种芯片硬件结构。

[0203] 图5为本申请实施例提供额一种芯片硬件结构,该芯片包括人工智能处理器50。该芯片可以被设置在如图3所示的执行设备110中,用以完成计算模块111的计算工作。该芯片也可以被设置在如图3所示的训练设备120中,用以完成训练设备120的训练工作并输出目标模型/规则101。如图4A和图4B所示的卷积神经网络中各层的算法均可在如图5所示的芯片中得以实现。

[0204] 人工智能处理器50可以是神经网络处理器(network processing unit,NPU),张量处理器 (tensor processing unit,TPU)或者图形处理器(graphics processing unit,GPU)等一切适合用于大规模异或运算处理的处理器。以NPU为例:NPU可以作为协处理器挂载到主CPU(Host CPU)上,由主CPU为其分配任务。NPU的核心部分为运算电路503,通过控制器504控制运算电路503提取存储器(权重存储器或输入存储器)中的数据并进行运算。

[0205] 在一些实现中,运算电路503内部包括多个处理单元(process engine,PE)。在一些实现中,运算电路503是二维脉动阵列。运算电路503还可以是一维脉动阵列或者能够执行例如乘法和加法这样的数字运算的其他电子线路。在一些实现中,运算电路503是通用的矩阵处理器。

[0206] 举例来说,假设有输入矩阵A,权重矩阵B,输出矩阵C。运算电路503从权重存储器503中取矩阵B相应的数据,并缓存在运算电路503中的每一个PE上。运算电路503从输入存储器501中取矩阵A的输入数据,根据矩阵A的输入数据与矩阵B的权重数据进行矩阵运算,得到的矩阵的部分结果或最终结果,保存在累加器(accumulator)508中。

[0207] 统一存储器506用于存放输入数据以及输出数据。权重数据直接通过存储单元访问控制器(direct memory access controller,DMAC)505被搬运到权重存储器502中。输入数据也通过DMAC被搬运到统一存储器506中。

[0208] 总线接口单元(bus interface unit,BIU)510,用于DMCA和取指存储器

(instruction fetch buffer)509的交互;总线接口单元310还用于取指存储器509从外部存储器获取指令;总线接口单元510还用于存储单元访问控制器505从外部存储器获取输入矩阵A或者权重矩阵B 的原数据。

[0209] DMAC主要用于将外部存储器DDR中的输入数据搬运到统一存储器506中,或将权重数据搬运到权重存储器502中,或将输入数据搬运到输入存储器501中。

[0210] 向量计算单元507可以包括多个运算处理单元,在需要的情况下,对运算电路503的输出做进一步处理,如向量乘,向量加,指数运算,对数运算,大小比较等等。向量计算单元 507主要用于神经网络中非卷积层,或者全连接层(fully connected layers,FC)的计算,具体可以处理:池化(pooling),批归一化(batch normalization),局部响应归一化(local response normalization)等。例如,向量计算单元507可以将非线性函数应用到运算电路 503的输出,例如累加值的向量,用以生成激活值。在一些实现中,向量计算单元507生成归一化的值、合并值,或二者均有

[0211] 在一些实现中,向量计算单元507将经处理的输出的向量存储到统一缓存器506。例如,向量计算单元507可以将非线性函数应用到运算电路503的输出,例如累加值的向量,用以生成激活值。在一些实现中,向量计算单元507生成归一化的值、合并值,或二者均有。在一些实现中,处理过的输出的向量能够用作到运算电路503的激活输入,例如用于在神经网络中的后续层中的使用。

[0212] 与控制器504连接的取指存储器(instruction fetch buffer)509,用于存储控制器504使用的指令。

[0213] 控制器504,用于调用指存储器509中缓存的指令,实现控制该运算加速器的工作过程。

[0214] 一般地,统一存储器506,输入存储器501,权重存储器502以及取指存储器509均为片上(On-Chip)存储器,外部存储器为该NPU外部的存储器,该外部存储器可以为双倍数据率同步动态随机存储器(double data rate synchronous dynamic random access memory, DDR SDRAM)、高带宽存储器(high bandwidth memory,HBM)或其他可读可写的存储器。

[0215] 上文中介绍的图3中的执行设备110能够执行本申请实施例的神经网络的量化方法或者神经网络的量化方法的各个步骤,图4A和图4B的卷积神经网络模型和图5所示的芯片也可以执行本申请实施例的神经网络的量化方法或者神经网络的量化方法的各个步骤。

[0216] 本申请实施例提供了一种系统架构。该系统架构包括一个或多个本地设备、执行设备和数据存储系统。其中,本地设备通过通信网络与执行设备连接。

[0217] 执行设备可以由一个或多个服务器实现。可选的,执行设备可以与其它计算设备配合使用,例如:数据存储器、路由器、负载均衡器等设备。执行设备可以布置在一个物理站点上,或者分布在多个物理站点上。执行设备可以使用数据存储系统中的数据,或者调用数据存储系统中的程序代码来实现本申请实施例的神经网络的量化方法。

[0218] 用户可以操作各自的用户设备(例如一个或多个本地设备)与执行设备进行交互。每个本地设备可以表示任何计算设备,例如个人计算机、计算机工作站、智能手机、平板电脑、智能摄像头、智能汽车或其他类型蜂窝电话、媒体消费设备、可穿戴设备、机顶盒、游戏机等。

[0219] 每个用户的本地设备可以通过任何通信机制/通信标准的通信网络与执行设备进行交互,通信网络可以是广域网、局域网、点对点连接等方式,或它们的任意组合。

[0220] 在一种实现方式中,本地设备从执行设备获取到目标神经网络的相关参数,将目标神经网络部署在本地设备、本地设备上,利用该目标神经网络进行图像分类或者图像处理等等。其中,目标神经网络为根据本申请实施例的神经网络的量化方法训练得到的。

[0221] 在另一种实现中,执行设备上可以直接部署目标神经网络,执行设备通过从本地设备和本地设备获取待处理数据,并根据目标神经网络对待处理数据进行分类或者其他类型的处理。

[0222] 上述执行设备也可以称为云端设备,此时执行设备一般部署在云端。

[0223] 下面详细介绍本申请实施例涉及的方法。图6为本申请实施例提供的一种数据处理方法的流程示意图。该方法可由数据处理设备来执行,数据处理设备具体可以是图3所示的系统架构100中的执行设备120、客户设备140或者用户设备150,该方法包括但不限于如下步骤:

[0224] 步骤S601、获取待处理数据。

[0225] 步骤S602、将待处理数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图。

[0226] 具体地,神经网络是基于二值化权重训练得到的,其中,二值化权重为根据二值化权重所对应的权重参数的不确定性对权重参数进行二值化处理所得到的。其中,权重参数的不确定用于表征权重参数在二值化过程中,接近于零的一个或多个权重值的符号的波动性。

[0227] 容易理解的是,神经网络包含一个或多个卷积层,在一种实现中,一个卷积层的输出(即输出的特征图)可以作为随后的池化层的输入,也可以作为另一个卷积层的输入以继续进行卷积操作。其中,在每个卷积层,数据是以三维形式存在的,可以将其看成是许多个二维数据堆叠在一起,而每一个二维数据可以成为一个特征图。特征图可以包括m*n个特征参数, m和n为正整数。

[0228] 步骤S603、确定特征图中每一个特征参数的不确定性。

[0229] 具体地,为了加快模型的运算速度,一般来说,在神经网络的每一层处理的过程中,可以对特征图中的每一个特征参数进行二值化处理,也即将全精度特征处理为二值化特征。在本申请实施例中,为了提高模型的稳定性,数据处理设备需要基于特征参数所对应的不确定性来对特征参数进行二值化处理。其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性。

[0230] 可以理解的是,通过符号函数sign(x)来计算特征图中每一个特征参数的符号,当x≥0, sign(x)=1;当x<0,sign(x)=-1。因此,当对于在零附近的特征参数进行微小的改变,在零附近的特征参数的符号可能会在1和-1之间跳变,处于不稳定性的状态。

[0231] 为了可以定量测量特征图中每一个特征参数的不确定性,在一种实现方式中,数据处理设备可以根据不确定函数计算特征图中每一个特征参数的不确定性。其中,不确定性函数为通过大量的数据建模得到的,在不确定性函数的自变量x越接近于0时,不确定性函数的值f(x)越大;在不确定性函数的自变量x的绝对值越大时,不确定性函数的值f(x)越小。不确定性可以通过多种函数来表示,在一种实现方式中,不确定性函数可以由高斯函数

来表示。其中,不确定性函数的表达式具体可以如公式(1-2)所示。

[0232] $$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (1\text{-}2)$$

[0233] 其中,σ是一个超参数。在一种实现方式中,σ可以表示为矩阵的元素函数。

[0234] 在本申请实施例中,为了提高模型的处理速度,本申请实施例引入了特征图中连续k个特征参数的不确定性来综合估计当前位置点的不确定性。其中,k的数值为预设位置点的数值,预设位置点为根据人为经验或者历史值所确定的。

[0235] 在一种实现方式中,根据不确定性函数计算特征图中每一个特征参数的不确定性的数学表达式可以为:

[0236] $$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}\left(1-f(x_j)\right) \quad (1\text{-}3)$$

[0237] 其中,i为目标位置点,也可以理解为当前处理的特征参数。若待处理数据为图像,则位置点可以理解为像素点。a为与目标位置点所相邻的位置点,$\hat{f}(x_i)$为特征图上目标特征参数所对应的不确定性,$f(x_j)$为特征图上与目标特征参数相邻的一个或多个特征参数所对应的不确定性,f()为不确定性函数。

[0238] 步骤S604、基于特征参数所对应的不确定性计算特征参数的二值化特征,其中,二值化特征用于确定特征矩阵。

[0239] 具体地,为了加快模型的处理速度和提高模型的稳定性,数据处理设备可以根据特征参数所对应的不确定性对特征参数进行二值化处理,得到特征参数的二值化特征。可以理解的是,在特征图中包含m*n个特征参数,数据处理设备在对特征图进行二值化处理时,实际是对特征图中每一个特征数值进行二值化处理。因此,数据处理设备可以基于每一个特征数值所对应的不确定性计算该特征数值的二值化特征。

[0240] 在一种实现方式中,在目标特征数值的不确定性小于或等于第二预设阈值时,数据处理设备可以通过符号函数对目标特征参数进行二值化处理,得到上述目标特征数值的二值化特征。

[0241] 在一种实现方式中,在目标特征数值的不确定性大于所述第二预设阈值时,也即对于不确定性较大的特征,数据处理设备可以采用平均池化(average pooling)并且引入符号函数,从空间维度上对目标特征参数进行二值化处理。其中,目标数据特征为特征图上的任意一个特征参数。

[0242] 也即,数据处理设备可以通过符号函数对平均池化后的与目标特征数值相邻的一个或多个特征参数进行二值化处理,得到上述目标位置点的二值化特征。

[0243] 在一种实现方式中,对于特征图上的任意一个特征参数,通过该特征参数所对应的不确定性来计算该特征参数的二值化特征的数学表达式可以参见公式(1-4)。

[0244] $$csign(x_i)=\begin{cases} sign(x_i) & \hat{f}\le\Delta \\ sign\left(averagepool\left(x_{i-\frac{a}{2}},x_{i+1-\frac{a}{2}},\ldots,x_i,x_{i+\frac{a}{2}},\right)\right) & \hat{f}>\Delta \end{cases} \quad (1\text{-}4)$$

[0245] 其中,公式(1-4)中x_i为当前进行二值化处理的目标特征参数,

$\left(x_{i-\frac{a}{2}},x_{i+1-\frac{a}{2}},\ldots,x_i,x_{i+\frac{a}{2}},\right)$为与包括目标特征参数在内的与目标特征参数相邻的a个特征参数。

Δ为第二预设阈值,第二阈值为根据经验人为设置的,用于表示神经网络中每一层需要进行二值化处理的特征个数。第二预设阈值可以根据实际需求进行自适应改变,举例来说,根据实际需求需要对从大到小排在前30%的特征参数进行二值化处理,则可以将第一预设阈值设为0.3。这样,当上述特征参数位于排序值的前30%时,数据处理设备可以通过符号函数来更新该特征参数的二值化特征。当上述特征参数没有位于排序值的前40%时,数据处理设备可以通过符号函数对平均池化后的与目标位置点相邻的一个或多个位置点的特征参数进行二值化处理,得到目标位置点的特征图的二值化特征。

[0246] 当基于特征数值所对应的不确定性对特征图中的每一个特征数值进行二值化处理后,可以得到特征矩阵。其中,特征矩阵中包含的m*n个二值化特征与特征图中包含的m*n个特征参数一一对应。

[0247] 步骤S605、基于特征矩阵得到待处理数据的处理结果。

[0248] 具体地,为了确保待处理数据的信息的完整度,在除去第一个卷积层和最后一个卷积层以外的每一个卷积层中,待处理设备可以将提取的特征矩阵与权重矩阵进行二维卷积运算,提取待处理数据中的特征。需要说明的是,若特征矩阵与权重矩阵的尺寸大小一致,则待处理设备可以将权重矩阵上的每个参数与特征矩阵上的参数相乘,最后把计算得到结果作为本次卷积的结果;若特征矩阵与权重矩阵的尺寸大小不一致,则可以将权重矩阵上的每个参数与特征矩阵上的部分参数进行对应相乘,然后待处理设备再将权重矩阵移动一个步长接着与特征矩阵上的其他参数进行下一次卷积,直到遍历完整个特征矩阵中的参数,将遍历完后的结果为本次卷积的结果。

[0249] 其中,不同的权重矩阵可以用来提取待处理数据中不同的特征。在一种实现方式中,一个卷积层的输出可以作为随后的池化层的输入,也可以作为另一个卷积层的输入以继续进行卷积操作。在经过所有卷积层的处理以及其他处理后,待处理设备可以输出待处理结果。需要说明的是,该处理结果的内容依赖于训练好的神经网络的功能,而训练好的神经网络的功能依赖于待训练神经网络的功能,处理结果可以是对图像的分结果、识别结果等。

[0250] 举例来说,请参见图7,图7所示为本申请实施例提供的一种数据处理方法的网络架构示意图。该数据处理方法具体为图像处理方法,包括:数据处理设备获取待处理图像,将待处理图像输入到训练好的神经网络模型中,由神经网络模型中的卷积层/池化层以及后面的神经网络层进行处理,可以得到图像的处理结果。其中,为了保证待处理图像的信息的完整度,在卷积层的第一卷积层和最后一个卷积层中不进行二值化运算,在中间卷积层中进行二值化运算。其中,中间卷积层中的每一层中进行二值化运算的二值化处理模块具体可以是第一量化模块或第二量化模块。其中,每一个卷积层提取出的待处理图像的特征图将作为一下层的输入。

[0251] 在第一量化模块中,数据处理设备对输入的特征图进行归一化处理、二值化处理、二维卷积处理、激活处理后得到输出结果,将输出结果输入到下一层,作为下一层的输入。其中,二值化处理流程可参考本申请实施例提供的神经网络的量化方法,详细描述可参考图8A中的部分内容,此处不再赘述。

[0252] 在第二量化模块中,数据处理设备对输入的特征图进行偏差、二值化处理、二维卷积处理、归一化处理、偏差、激活和偏差后得到输出结果,将输出结果输入到下一层,作为下一层的输入。其中,二值化处理流程可参考本申请实施例提供的神经网络的量化方法,详细描述可参考图8A中的部分内容,此处不再赘述。

[0253] 最后,由最后一个卷积层将输出的特征图输入到全连接层中,由全连接层基于上述输出的特征图得到处理结果。需要说明的是,该处理结果的内容依赖于训练好的神经网络的功能,而训练好的神经网络的功能依赖于待训练神经网络的功能,可以是对图像的分结果。识别结果等。

[0254] 图8A为本申请实施例提供的一种神经网络的量化方法的流程示意图。该方法可由量化来执行,量化设备具体可以是图3所示的系统架构100中的训练设备110,该方法包括但不限于如下步骤:

[0255] 步骤S801、获取第一权重矩阵,第一权重矩阵中包含神经网络中用于提取特征的参数,第一权重矩阵包含s*k个权重参数。

[0256] 具体地,神经网络的量化设备获取的第一权重矩阵可以是初始化权重矩阵,也可以是在迭代更新后的权重矩阵,本申请实施例不做任何限制。其中,s和k为正整数。

[0257] 步骤S802、计算第一权重矩阵中每一个权重参数的不确定性。

[0258] 具体地,为了加快神经网络模型在训练过程中的收敛速度,量化设备可以对神经网络中每一层的第一权重矩阵进行二值化处理。其中,第一权重矩阵中的权重参数为全精度参数。可以理解的是,对第一权重矩阵进行二值化处理,也即对第一权重矩阵中的每一个权重参数进行二值化处理。因此,量化设备可以计算第一权重矩阵中每一个权重参数的不确定性。

[0259] 权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性。在一种实现方式中,通过符号函数sign(x)来计算权重参数的符号,当$x\geq 0$,sign (x)=1;当$x<0$,sign(x)=-1。因此,在零附近的权重参数的符号会频繁在1和-1之间跳变,处于不稳定性的状态。为了可以定量测量权重参数的不确定性,量化设备需要根据不确定函数计算神经网络中权重参数的不确定性。其中,不确定性函数为通过大量的数据建模得到的,在不确定性函数的自变量x越接近于0时,不确定性函数的值f(x)越大;在不确定性函数的自变量x的绝对值越大时,不确定性函数的值f(x)越小。

[0260] 不确定性可以通过多种函数来表示,在一种实现方式中,不确定性函数可以由高斯函数来表示。图8B所示为本申请实施例提供的一种不确定性函数的示意图。从图8A可以看出,不确定性函数的值在0处最大,且随着自变量(也即权重参数)接近+1/-1而逐渐变小。因此,本申请实施例通过预测的连续值x(-1≤x≤1)及其目标(+1和-1),通过高斯函数对不确定函数的建模如公式(1-2)所示。

[0261] 通过由高斯函数得到的不确定性函数可以用来计算权重矩阵中每一个权重参数的不确定性,容易理解的是,不确定性函数的值越高的权重参数的置信度越低,也即,该权重参数的符号被反转的可能性越大。比如说,对该权重参数进行微小的改变,可能导致该权重参数的符号从+1改变为-1。不确定性函数的值越低的权重参数的置信度越高,也即,该权重参数的符号被反转的可能性比较小。比如说,对该权重参数进行微小的改变,不太可能导致该权重参数的符号从+1改变为-1。

[0262] 在本申请实施例中,为了保持一个稳定的训练过程,避免出现符号的波动性较为不稳定的权重参数。本申请实施例引入了神经网络中连续m个迭代次数所对应的第一权重矩阵中权重参数的不确定性,来综合估计当前迭代次数所对应的权重参数的不确定性。其中,m的数值为预设迭代次数的数值,预设迭代次数为根据人为经验或者历史值所确定的。

[0263] 因此,在一种实现方式中,在当前迭代次数小于或等于预设迭代次数时,量化设备可以通过不确定性函数计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0264] 在一种实现方式中,在当前迭代次数小于或等于预设迭代次数时,量化设备可以综合考虑预设迭代次数内计算得到第一权重矩阵中每一个权重参数的不确定性。量化设备可以根据在参考预设迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算当前迭代次数所对应的第一权重矩阵中内一个权重参数的不确定性。其中,参考预设迭代次数为最接近当前迭代次数的预设迭代次数。

[0265] 在一种实现方式中,通过不确定性函数来计算第一权重矩阵中每一个权重参数的数学表达式可以为:

[0266] $$\hat{f}(w_t)=\begin{cases} f(w_t) & t<m \\ 1-\prod_{i=t-m+1}^{t}(1-f(w_i)) & t\geq m \end{cases} \quad (1\text{-}5)$$

[0267] 其中,t为当前迭代次数,也可以理解为当前时间步。m为预设迭代次数,也可以理解为预设时间步。$\hat{f}(w_t)$可以标识当前迭代次数所对应的第一权重矩阵中每一个权重参数不确定,$\hat{f}(w_i)$可以表示参考预设次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,f() 为不确定性函数。

[0268] 当t≥m时,$\hat{f}(w_t)-\hat{f}(w_{t-1})=\prod_{i=t-m}^{t-1}(1-f(w_i))-\prod_{i=t-m+1}^{t}(1-f(w_i))=(f(w_t)-f(w_{t-m}))\prod_{i=t-m}^{t-1}(1-f(w_i))$。由0≤f()≤1可得:$\hat{f}(w_t)\leq\hat{f}(w_{t-1})$。

[0269] 举例来说,假设预设迭代次数为5次,在神经网络的训练过程中,若当前迭代次数为4 次,则量化设备可以通过不确定性函数来计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0270] 若当前迭代次数为5,则量化设备可以根据在前5次计算得到的权重参数的不确定性,来计算第5次迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。也即,量化设备可以根据表达式$\hat{f}(w_6)=1-\prod_{i=1}^{5}(1-f(w_i))$来计算第5次迭代次数所对应的第一权重参数中每一份权重参数的不确定性。

[0271] 若当前迭代次数为9次,则量化设备可以根据与当前迭代次数最接近的5次迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,来计算第9次迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。其中,与当前迭代次数9次最接近的5次迭代次数可以是第5次迭代次数、第6次迭代次数、第7次迭代次数、第8次迭代次数、第9次迭

代次数。因此,量化设备可以根据表达式$\hat{f}(w_9)=1-\prod_{i=5}^{9}(1-f(w_i))$来计算第9次迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0272] 需要说明的是,不确定性有多种函数可表示,高斯函数只是其中的一种表达方式,本申请实施例不做任何限制。

[0273] 步骤S803、基于权重参数所对应的不确定性计算权重参数的二值化权重。

[0274] 具体地,在二值化过程中,为了提高神经网络的收敛速度和稳定性,量化设备可以根据权重参数所对应的不确定性来计算权重参数的二值化权重。也即,在每一次的迭代更新过程中,对当前迭代次数的第一权重矩阵中每一个权重参数进行二值化时,需要考虑当前迭代次数的第一权重矩阵中每一个权重参数的不确定性。

[0275] 在一种实现方式中,在当前迭代次数的所对应的第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,量化设备可以通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0276] 在一种实现方式中,在当前迭代次数所对应的第一权重矩阵中的权重参数的不确定性大于上述第一值时,量化设备可以将当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为当前迭代次数所对应的第一权重矩阵中目标权重参数的二值化权重;其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0277] 在一种实现方式中,基于权重参数所对应的不确定性来计算权重参数的二值化权重的数学表达式可以为:

[0278] $$csign(w_t)=\begin{cases} sign(w_t) & \hat{f}\le \min\left(\hat{f}(w_{t-1}),\ \Delta\right) \\ csign(w_{t-1}) & \hat{f}>\min\left(\hat{f}(w_{t-1}),\ \Delta\right) \end{cases} \quad (1\text{-}6)$$

[0279] 其中,t为当前迭代次数,t-1为当前迭代次数的前一迭代次数。w为目标权重参数,也即第一权重矩阵中的任意权重。w_t为当前迭代次数所对应的第一权重矩阵中的目标权重参数,w_{t-1}为当前迭代次数迭代前一迭代次数所对应的第一权重矩阵中的目标权重参数。Δ为第一预设阈值。csign()为本申请实施例提出的为了使得神经网络的不确定性最小所提出的一种确定符号函数。

[0280] 请参见图8C,图8C是本申请实施例提供的一种二值化结果的示意图。其中,图8C中的(a)是权重参数的实值及其不确定性的示意图,图8C中的(b)是权重参数的实值的二值化结果。从图8C中的(b)可以看出,通过csign函数可以降低权重参数的不确定性。

[0281] 举例来说,假设第一预设阈值为0.3,当前迭代次数所对应的目标权重参数的不确定性为 0.1,当前迭代次数的前一迭代次数所对应的第一权重矩阵中目标权重参数的不确定性为0.2。可以看出,0.1小于min(0.2,0.3),因此量化设备可以通过符号函数来计算当前迭代次数的目标权重参数的二值化权重。容易理解的是,对权重的迭代更新过程是根据反向传播梯度对权重参数进行不断地优化。而当前迭代次数的权重参数的二值化权重是通过

上一次的迭代所更新得到的,根据不确定性函数可以得知,当前迭代次数所对应的目标权重参数比前一迭代次数所对应的目标权重参数大。因此对当前迭代次数的目标权重参数进行微小的改变后,不太可能改变当前迭代次数的目标权重参数的符号。所以,量化设备可以根据符号函数对当前迭代次数所对应的目标权重参数进行二值化处理。

[0282] 假设第一预设阈值为0.3,当前迭代次数所对应的目标权重参数的不确定性为0.4,当前迭代次数的前一迭代次数所对应的目标权重参数的不确定性为0.5。可以看出,0.4大于 min(0.5,0.3),因此量化设备可以将当前迭代次数的前一迭代次数的目标权重参数的二值化权重,作为当前迭代次数的目标权重参数的二值化权重。容易理解的是,对权重的迭代更新过程是根据反向传播梯度对权重参数进行不断地优化。而当前迭代次数的权重参数的二值化权重是通过上一次的迭代所更新得到的,根据不确定性函数可以得知,当前迭代次数所对应的目标权重参数比前一迭代次数所对应的目标权重参数小。因此对当前迭代次数的目标权重参数进行微小的改变后,可能会改变当前迭代次数的目标权重参数的符号。所以,量化设备可以将前一迭代次数的目标权重参数的二值化权重,作为当前迭代次数的目标权重参数的二值化权重。这样,可以使得神经网络的不确定性最小,加快神经网络的收敛速度。

[0283] 其中,目标权重参数为第一权重矩阵中的任意一个权重参数。

[0284] 需要说明的,在本申请实施例中,量化设备通过对当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性与第一值的比较,来确定当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重。而第一值是前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,所以第一预设阈值也影响着当前迭代次数的权重参数的二值化权重。

[0285] 其中,第一预设阈值为根据经验人为设置的,用于表示神经网络中每一层中需要进行二值化处理的目标权重参数的个数。第一预设阈值可以根据实际需求进行自适应改变,举例来说,根据实际需求需要对从大到小排在前30%的目标权重参数进行二值化处理,则可以将第一预设阈值设为0.3。这样,当目标权重参数位于排序值的前30%时,量化设备可以通过符号函数来更新目标权重参数的二值化权重。当目标权重参数没有位于排序值的前40%时,量化设备将不更新目标权重参数的二值化权重,也即将前一迭代次数的目标权重参数的二值化权重作为当前迭代次数的。

[0286] 请参见图9,在执行步骤S801至步骤S802之前,或者在执行步骤S801至步骤S802之后,还可以包括以下步骤:

[0287] 步骤S901、获取训练数据。

[0288] 步骤S902、将训练数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图。

[0289] 具体地,容易理解的是,在每个卷积层,数据都是以三维形式存在的,可以将该三维形式看作许多个二维图片堆叠在一起所形成的。其中每一个称为一个特征图(feature map)。在输入层,如果是灰度图片,那就只有一个feature map;如果是彩色图片,一般就是3个feature map(红绿蓝)。层与层之间会有若干个卷积核(kernel),上一层和每个feature map跟每个卷积核做卷积,都会产生下一层的一个feature map。其中,特征图可以包括m*n个特征参数, m和n为正整数。

[0290] 步骤S903、计算特征图中每一个特征参数的二值化特征。

[0291] 具体地,为了加快神经网络模型在训练过程中的收敛速度,可以对神经网络参数进行二值化处理,比如说将特征图中的每一个特征参数进行二值化处理,得到二值化特征。

[0292] 在本申请实施例中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性。在一种实现方式中,通过符号函数sign(x)来计算特征参数的符号,当x≥0,sign(x)=1;当x<0,sign(x)=-1。因此,对于在零附近的特征参数进行微小的改变,可能会使得在零附近的特征参数的符号在1和-1之间跳变,处于不稳定性的状态。

[0293] 为了可以定量测量特征参数的不确定性,在一种实现方式中,量化设备可以根据不确定函数计算神经网络中特征参数的不确定性。其中,不确定性函数为通过大量的数据建模得到的,在不确定性函数的自变量x越接近于0时,不确定性函数的值f(x)越大;在不确定性函数的自变量x的绝对值越大时,不确定性函数的值f(x)越小。不确定性可以通过多种函数来表示,在一种实现方式中,不确定性函数可以由高斯函数来表示。其中,不确定性函数的表达式具体可以如公式(1-2)所示。

[0294] 在本申请实施例中,为了保持一个稳定的训练过程,避免出现符号的波动较为不稳定的特征图。本申请实施例引入了连续k个相邻位置点的特征参数的不确定性,来综合估计当前位置点的特征参数的不确定性。其中,k的数值为预设位置点的数值,预设位置点为根据人为经验或者历史值所确定的。

[0295] 在一种实现方式中,通过不确定性函数计算特征图中每一个特征参数的不确定性的数学表达式如公式(1-3)所示。其中,i为目标位置点,也可以理解为当前处理的特征参数。若训练数据为图像,则位置点可以理解为像素点。m为与目标位置点所相邻的位置点,x为特征图,x_i为特征图上目标位置点所对应的特征参数,f()为不确定性函数。

[0296] 可以理解的是,为了提高在训练过程中神经网络的收敛速度和稳定性,量化设备可以根据特征参数所对应的不确定性来计算特征参数的二值化特征。可以理解的是,在特征图中包含m*n个特征参数,数据处理设备在对特征图进行二值化处理时,实际是对特征图中每一个特征数值进行二值化处理。因此,数据处理设备可以基于每一个特征数值所对应的不确定性计算该特征数值的二值化特征。

[0297] 在一种实现方式中,在目标特征数值的不确定性小于或等于第二预设阈值时,量化设备可以通过符号函数对目标特征参数进行二值化处理,得到上述目标特征参数的二值化特征。

[0298] 在一种实现方式中,在目标特征参数的不确定性大于所述第二预设阈值时,也即对于不确定性较大的特征,量化设备可以采用平均池化(average pooling)并且引入符号函数,从空间维度上对目标特征参数进行二值化。也即,量化设备可以通过符号函数对平均池化后的与目标特征参数相邻的一个或多个特征参数进行二值化处理,得到上述目标位置点的二值化特征。其中,目标数据特征为特征图上的任意一个特征参数。

[0299] 在一种实现方式中,对应特征图上的任意一个特征参数,通过该特征参数所对应的不确定性来计算特征参数的二值化特征的数学表达式可以如公式(1-4)所示。

[0300] 其中,x_i为当前进行二值化处理的特征参数,$\left(x_{i-\frac{k}{2}}, x_{i+1-\frac{k}{2}}, \ldots, x_i, x_{i+\frac{k}{2}}\right)$为与目标特征

参数在内的与目标特征参数相邻的k个特征参数。Δ为第二预设阈值,第二阈值为根据经验人为设置的,用于表示神经网络中每一层需要进行二值化处理的特征参数的个数。第二预设阈值可以根据实际需求进行自适应改变,举例来说,根据实际需求需要对从大到小排在前30%的特征参数进行二值化处理,则可以将第一预设阈值设为0.3。这样,当上述特征参数位于排序值的前30%时,量化设备可以通过符号函数来更新特征参数的二值化特征。当特征参数没有位于排序值的前40%时,量化设备可以通过符号函数对平均池化后的与目标位置点相邻的一个或多个位置点的特征参数进行二值化处理,得到目标位置点的二值化特征。

[0301] 当基于特征数值所对应的不确定性对特征图中的每一个特征数值进行二值化处理后,可以得到特征矩阵。其中,特征矩阵中包含的m*n个二值化特征与特征图中包含的m*n个特征参数一一对应。

[0302] 需要说明的是,在神经网络的训练过程中,通过图8A或图9所示的神经网络的量化方法对神经网络进行训练后,可以得到如图6所示的数据处理设备,用于执行图6所示的数据处理方法。

[0303] 请参见图10,图10是本申请实施例提供的一种神经网络的训练方法的流程示意图。该方法可由图3所示的系统架构中的训练设备120执行。

[0304] 容易理解的是,在一个L层的卷积神经网络中,将第l层(为L层中的任意一层)的权重参数和特征参数分别表示为W^l和F^l,从而可以将第l层发生的运算表示为:

[0305] $$F^{l+1}=\varphi^l(W^l*F^l) \quad (1-7)$$

[0306] 其中,*表示卷积操作,φ^l表示第l层发生的其余操作,譬如批标准化(Batch Normalization,BN),线性整流函数(Rectified Linear Unit,ReLU)等。在二值神经网络中,W^l和F^l的每一个元素都可以被sign函数投影到{-1,+1}。

[0307] 然而,符号函数舍弃了变量中的幅度信息并且可能会导致较大的量化误差。因此,为了减小这种精度损失,现有技术所提供的尺度因子方法被广泛地应用在二值神经网络中,用以减少精度的损失,增强二值神经网络的表征能力。使用尺度因子的操作可以表示为:

[0308] $$F^{l+1}=\phi^l\left(\alpha^l\cdot\left(W_B^{\ l}\otimes F_B^{\ l}\right)\right) \quad (1-8)$$

[0309] 其中,$W_B^{\ l}$表示为进行二值化处理所得到的二值化权重,$F_B^{\ l}$表示为进行而二值化处理所得到的二值化特征,⊗表示由XNOR和popcount操作组成的二值卷积。这样,可以将公式(1-7)中的实权卷积中的多重累积运算替换为简化卷积的轻权XNOR和popcount运算,以加速二值神经网络的运算和减少存储。

[0310] 从图10可以看出,L为二值神经网络的网络层数。首先,训练设备初始化神经网络模型的超参数和所有层的权重参数。在前向传播过程中,训练设备获取训练数据,将所述待处理数据输入神经网络,确定所述神经网络的一个或多个卷积层提取的特征图。

[0311] 训练设备从第1层到第L层逐层计算权重参数的不确定性和特征图中特征参数的不确定性。其中,计算权重参数的不确定性可参考图8A所示的步骤S801中的相关内容,计算特征图中特征参数的不确定性可参考图9所示的步骤S903中的相关内容,此次不再赘述。

[0312] 训练设备可以根据权重参数的不确定性通过csign函数将权重参数更新为二值化

权重。其中,相关计算可参考图8A所示的步骤S802,此次不再赘述。

[0313] 训练设备可以根据特征图中特征参数的不确定性通过csign函数将特征参数更新为二值化特征。其中,相关计算可参考图9所示的步骤S903,此次不再赘述。

[0314] 在得到二值化权重和二值化特征后,训练设备对二值化权重和二值化特征进行二维卷积操作。在前向传播完成后,训练设备从第L层到1层反向传播计算权重参数的梯度,并逐层更新权重参数W^l,直到训练完成。

[0315] 综上所述,在一次迭代中,权重参数可以根据csign函数进行更新。进一步地,在对模型的训练过程中使用异步更新。在正向传播过程中,量化设备可以根据sign函数对权重参数进行二值化处理,更新为二值化权重;由于一些权重参数的不确定性,该权重参数将不会得到更新,也即不会进行二值化处理。这样,可以保证神经网络中不确定性的减小,从而增强神经网络的稳定性和加快其收敛速度。

[0316] 在本申请实施例中,训练设备根据图8A所示的神经网络的量化方法对神经网络模型进行训练后,还需要通过验证数据对训练得到的模型进行评估,保证训练得到的神经网络模型的具有较好的泛化性。

[0317] 在一种实现方式中,训练设备基于Pytorch深度学习框架,对于CIFAR10/100数据集,使用WideResNet-22(WRN-22)作为神经网络中的主干网络来验证图6所示方式的有效性。其中,神经网络模型中的学习速率初始为0.1,随机梯度下降(Stochastic gradien descent,SGD) 优化器的动量为0.9,应用余弦退火衰减方法。在CIFAR10/100数据集上,所以的网络都可以被训练成200时期(epoch)。其中,WRN-22可以是一个具有22个卷积层的WRN网络。由于WRN是一个以ResNet为原型,其引入了新的深度因子k,通过三个阶段调整特征图的深度扩展,保证特征的空间维度不变。在一种实现方式中,将k设为1.第一级通道数是WRN的一个参数,将其设置为16和64,从而可以分别得到16-16-32-64和64-64-128-256的网络配置。

[0318] 需要说明的是,与其他的方法在CIFAR10/100数据集上的测试结果相比,通过图8A所示的神经网络的量化方法所训练得到的模型具有很好的性能,比如说在CIFAR10和CIFAR101 上使用不同的网络配置分别获得了0.69%,0.51%和0.77,0.49%的改进。详细数据可参见表 2:本申请实施例在CIFA数据集上与其他方法的测试结果对比。

[0319] 表2:本申请实施例在CIFA数据集上与其他方法的测试结果对比

[0320]

模型（Model）	内核步（Kernel-Stage）	方法	参数	W/A	CIFAR10（%）	CIFAR100（%）
WRN22	16-16-32-64	XNOR-Net	0.27M	1/1	81.90	53.17
		Bi Real Net	0.27M	1/1	85.16	57.34
		BONN	0.27M	1/1	87.34	60.19
		UaBNN	0.27M	1/1	88.03	61.68
		FP32	0.27M	32/32	91.66	67.51
WRN22	64-64-128-256	Bi Real Net	4.3M	1/1	90.65	68.51
		PCNN	4.3M	1/1	91.37	69.98
		BONN	4.3M	1/1	92.36	-
		Proxy BNN	4.3M	1/1	92.96	71.57
		UaBNN	4.3M	1/1	93.37	72.01
		FP32	4.3M	32/32	95.75	77.34

[0321] 如表2所示,W/A分别表示权重和激活位宽带,FP表示全精度模型,UaBNN通过图8A所示的神经网络的量化方法所训练得到的神经网络模型。从表2可以看出,与其他方法相比较,UaBNN的模型精度更接近于全精度模型。说明通过本申请实施例得到的神经网络模型不仅可以加快模型的收敛速度,还可以保证模型的精度。

[0322] 在一种实现方式中,训练设备对于ImageNet数据集,使用ResNet18作为神经网络中的主干网络来验证图6所示方式的有效性。其中,神经网络模型中的学习速率初始为0.001, Adma优化器的动量为0.9。学习速率采用线性衰减策略,以线性方式降低学习速率。其中,对于ResNet18,遵循Bi-Real Net中的设备和网络修改,将除第一层和最后一层外的骨干卷积层的特征和内核进行了二值化。

[0323] 需要说明的是,与其他的模型在ImageNet数据集上的测试结果相比,通过图8A所示的神经网络的1量化方法训练得到的模型具有很好的性能,比如说在提高了1.0%的Top-1精度和0.6%的Top-5精度。详细数据可参见表3:本申请实施例在ImageNet数据集上与其他模型的测试结果对比。

[0324] 表3:本申请实施例在CIFA数据集上与其他模型的测试结果对比

[0325]

模型（Model）	W	A	Top-1	Top-5
ResNet18	32	32	69.3	89.2
TBN	1	2	55.6	79.0
BNN	1	1	42.2	67.1
XNOR-Net	1	1	51.2	73.2

[0326]

ABC-Net	1	1	42.7	67.6
Bi-Real Net	1	1	56.4	79.5
PCNN	1	1	57.3	80.0
IR-Net	1	1	58.1	80.0
BONN	1	1	59.3	81.6
RBNN	1	1	59.6	81.6
UaBNN	1	1	60.6	82.2
ReActNet*	1	1	61.4	83.2
UaBNN*	1	1	61.9	83.4

[0327] 如表3所示,W和A分别表示权重和激活位宽带,表3中所有模型的主干是ResNet,UaBNN*为通过图8A所示的神经网络的量化方法所训练得到的神经网络模型。从表3可以看出,与其他方法相比较,UaBNN和UaBNN*的模型精度更接近于全精度模型。说明通过本申请实施例得到的神经网络模型不仅可以加快模型的收敛速度,还可以保证模型的精度。

[0328] 综上,本申请实施例考虑到了神经网络中参数(比如说权重和特征)的不确定性,并建模了一个用于定量计算不确定性的函数,根据计算得到的不确定性来对参数进行二值化处理,从而完成对模型的训练。在CIFA和ImageNet上进行的实验表明,通过本申请实施例提供的方法对WRN和ResNet18进行了有效的增强。

[0329] 图11是本申请实施例中数据处理装置的示意性框架图。如图11所示,数据处理装置110 可以包括:获取单元1101,输入单元1102、计算单元1103和量化单元1104。其中,

[0330] 获取单元1101,用于获取待处理数据;

[0331] 输入单元1102,用于将待处理数据输入神经网络,确定神经网络的一个或多个卷积层提取的特征图,特征图包含m*n个特征参数,m和n为正整数;

[0332] 计算单元1103,用于计算特征图中特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;

[0333] 量化单元1104,用于根据特征参数的不确定性计算特征参数的二值化特征;

[0334] 计算单元1103,还用于基于二值化特征得到待处理数据的处理结果。

[0335] 在一种可能的实现方式中,神经网络为根据二值化权重训练得到的,二值化权重为根据神经网络中权重参数的不确定性所得到的,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的一个或多个权重参数的符号的波动性。

[0336] 在一种可能的实现方式中,计算单元1103,具体用于根据不确定性函数计算特征图中特征参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0337] 在一种可能的实现方式中,计算单元1103,具体用于根据不确定性函数计算特征图上的位置点所对应的特征参数不确定性;根据与目标位置点相邻的一个或多个位置点的特征参数的不确定性,计算目标位置点的不确定性,目标位置点特征图上的任意一个位置点。

[0338] 在一种可能的实现方式中,量化单元1104,具体用于在目标位置点的特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标位置点的特征参数进行二值化处理,得到目标位置点的二值化特征。

[0339] 在一种可能的实现方式中,量化单元1104,具体用于在目标位置点的特征参数的不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标位置点相邻的一个或多个位置点的特征参数进行二值化处理,得到目标位置点的二值化特征。

[0340] 应理解,各个器件的实现还可以对应参照上述实施例中的相应描述,本申请实施例不再赘述。

[0341] 图12是本申请实施例中神经网络的量化装置的示意性框架图。如图12所示,神经网络的量化装置120可以包括:获取单元1201、计算单元1202和量化单元1203。其中,

[0342] 获取单元1201,用于获取第一权重矩阵,第一权重矩阵中包含神经网络中用于提取特征的参数,第一权重矩阵包含s*k个权重参数,s和k为正整数;

[0343] 计算单元1202,用于计算第一权重矩阵中每一个权重参数的不确定性,其中,权重参数为神经网络的权重中的任意一个权重,权重参数的不确定性用于表征权重参数在二值化过程中,接近于零的权重参数的符号的波动性;

[0344] 量化单元1203,用于基于权重参数所对应的不确定性计算权重参数的二值化权重,二值化权重用于确定第二权重矩阵,第二权重矩阵中包含的s*k个二值化权重与s*k个权重参数一一对应。

[0345] 在一种可能的实现方式中,计算单元1202,具体用于:根据不确定性函数计算第一权重矩阵中每一个权重参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0346] 在一种可能的实现方式中,计算单元1202,具体用于:在当前迭代次数小于或等于预设迭代次数时,通过不确定性函数计算当前迭代次数所对应的第一权重矩阵中每一个权重参数的不确定性。

[0347] 在一种可能的实现方式中,量化单元1203,具体用于:在当前迭代次数大于预设迭代次数时,根据在参考迭代次数内计算得到的第一权重矩阵中每一个权重参数的不确定性,计算当前迭代次数所对应的第一权重矩阵的每一个权重参数的不确定性,其中,参考迭代次数为最接近当前迭代次数的预设迭代次数。

[0348] 在一种可能的实现方式中,量化单元1203,具体用于:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性小于或等于第一值时,通过符号函数对当前迭代次数所对应的第一权重矩阵中的目标权重参数进行二值化处理,得到二值化权重;其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0349] 在一种可能的实现方式中,量化单元1203,具体用于:在当前迭代次数所对应的第一权重矩阵中的目标权重参数的不确定性大于第一值时,将当前迭代次数的前一迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,作为当前迭代次数所对应的第一权重矩阵中的目标权重参数的二值化权重,其中,第一值为当前迭代次数的前一迭代次数的权重参数的不确定性和第一预设阈值中的最小值,目标权值参数为第一权重矩阵中的任意一个参数。

[0350] 在一种可能的实现方式中,神经网络的量化装置120还可以包括输入单元1204:

[0351] 获取单元1201,还用于获取训练数据;

[0352] 输入单元1204,用于将训练数据输入神经网络,确定神经网络的一个或多个卷积

层提取的特征图;

[0353] 量化单元1203,用于计算特征图中每一个特征参数的二值化特征,其中,特征图包含m*n个特征参数,m和n为正整数,特征图为在神经网络的一个或多个卷积层中提取的训练数据的特征。在一种可能的实现方式中,量化单元1203,具体用于:确定特征图中每一个特征参数的不确定性,其中,特征参数的不确定性用于表征特征参数在二值化过程中,接近于零的特征参数的符号的波动性;基于特征参数所对应的不确定性计算特征参数的二值化特征,二值化特征用于确定特征矩阵,特征矩阵中包含的m*n个二值化特征与m*n个特征参数一一对应。

[0354] 在一种可能的实现方式中,量化单元1203,具体用于:根据不确定性函数计算特征图中特征参数的不确定性,其中,在不确定函数的自变量越接近于0时,不确定性函数的值越大;在不确定性函数的自变量的绝对值越大时,不确定性函数的值越小。

[0355] 在一种可能的实现方式中,不确定性函数公式为:

[0356] $$\hat{f}(x_i)=1-\prod_{j=i-a/2}^{i+a/2}(1-f(x_j))$$

[0357] 其中,$\hat{f}(x_i)$为目标特征参数的不确定性,$f(x_j)$为与目标特征参数相邻的a个特征参数的不确定性,i,j,a均为自然数。

[0358] 在一种可能的实现方式中,量化单元1203,具体用于:在目标特征参数的不确定性小于或等于第二预设阈值时,通过符号函数对目标特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0359] 在一种可能的实现方式中,量化单元1203,具体用于:在目标特征参数的目标不确定性大于第二预设阈值时,通过符号函数对平均池化后的与目标特征参数相邻的一个或多个特征参数进行二值化处理,得到目标特征参数的二值化特征。

[0360] 应理解,各个器件的实现还可以对应参照上述实施例中的相应描述,本申请实施例不再赘述。

[0361] 如图13所示,本申请实施例提供的一种数据处理设备,该数据处理设备1300可以包括处理器1301、存储器1302、通信总线1303和通信接口1304,处理器1301通过通信总线1303 连接存储器1302和通信接口1304。

[0362] 处理器1301可以采用通用的中央处理器(Central Processing Unit,CPU),微处理器,应用专用集成电路(Application Specific Integrated Circuit,ASIC),图形处理器(Graphics Processing Unit,GPU)、神经网络处理器(Network Processing Unit,NPU)或者一个或多个集成电路,用于执行相关程序,以执行本申请方法实施例的数据处理方法。

[0363] 处理器1301还可以是一种集成电路芯片,具有信号的处理能力。在实现过程中,本申请的神经网络的训练方法的各个步骤可以通过处理器1301中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器1301还可以是通用处理器、数字信号处理器(Digital Signal Processing,DSP)、专用集成电路(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。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。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。结合本申请实施

例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器1302,处理器1301读取存储器1302中的信息,结合其硬件执行本申请方法实施例的数据处理方法。

[0364] 存储器1302可以是只读存储器(Read Only Memory,ROM),静态存储设备,动态存储设备或者随机存取存储器(Random Access Memory,RAM)。存储器1302可以存储程序和数据,例如本申请实施例中神经网络的训练方法的程序等。当存储器1302中存储的程序被处理器1301执行时,处理器1301和通信接口1304用于执行本申请实施例的数据处理方法的各个步骤。

[0365] 例如,本申请实施例中用于实现本申请实施例中数据处理方法的程序等。

[0366] 通信接口1304使用例如但不限于收发器一类的收发装置,来实现第二设备1300与其他设备或通信网络之间的通信。例如,可以通过通信接口1304获取训练好的神经网络,以实现与执行设备、客户设备、用户设备或者终端设备等的信息交互。

[0367] 可选地,该数据处理设备1300还可以包括人工智能处理器1305,人工智能处理器1305 可以是神经网络处理器(Network Processing Unit,NPU),张量处理器(Tensor Processing Unit, TPU),或者图形处理器(Graphics Processing Unit,GPU)等一切适合用于大规模异或运算处理的处理器。人工智能处理器1305可以作为协处理器挂载到主CPU(Host CPU)上,由主 CPU为其分配任务。人工智能处理器1305可以实现上述神经网络的训练方法中涉及的一种或多种运算。例如,以NPU为例,NPU的核心部分为运算电路,通过控制器控制运算电路提取存储器1302中的矩阵数据并进行乘加运算。

[0368] 处理器1301用于调用存储器中的数据和程序代码,执行上述方法实施例中数据处理设备 1300执行的具体操作,在此不再赘述。

[0369] 应理解,各个器件的实现还可以对应参照上述数据处理方法实施例中的相应描述,本申请实施例不再赘述。

[0370] 图14为本申请实施例中一种神经网络的量化设备的结构示意图,如图14所示,该神经网络的量化设备1400可以包括处理器1401、存储器1402、通信总线1403和通信接口1404,处理器1401通过通信总线1403连接存储器1402和通信接口1404。

[0371] 处理器1401可以采用通用的中央处理器(Central Processing Unit,CPU),微处理器,应用专用集成电路(Application Specific Integrated Circuit,ASIC),图形处理器(Graphics Processing Unit,GPU)、神经网络处理器(Network Processing Unit,NPU)或者一个或多个集成电路,用于执行相关程序,以执行本申请方法实施例的神经网络的量化方法。

[0372] 处理器1401还可以是一种集成电路芯片,具有信号的处理能力。在实现过程中,本申请的神经网络的量化方法的各个步骤可以通过处理器1401中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器1401还可以是通用处理器、数字信号处理器(Digital Signal Processing,DSP)、专用集成电路(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。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。

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。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器1402,处理器1401读取存储器1402中的信息,结合其硬件执行本申请方法实施例的神经网络的量化方法。

[0373] 存储器1402可以是只读存储器(Read Only Memory,ROM),静态存储设备,动态存储设备或者随机存取存储器(Random Access Memory,RAM)。存储器1402可以存储程序和数据,例如本申请实施例中神经网络的训练方法的程序等。当存储器1402中存储的程序被处理器1401执行时,处理器1401和通信接口1404用于执行本申请实施例的神经网络的量化方法的各个步骤。

[0374] 例如,本申请实施例中用于实现本申请实施例中神经网络的量化方法的程序等。

[0375] 通信接口1404使用例如但不限于收发器一类的收发装置,来实现神经网络的量化设备 1400与其他设备或通信网络之间的通信。例如,可以通过通信接口1404获取训练好的神经网络,以实现与执行设备、客户设备、用户设备或者终端设备等的信息交互。

[0376] 可选地,该神经网络的量化设备还可以包括人工智能处理器1405,人工智能处理器1405 可以是神经网络处理器(Network Processing Unit,NPU),张量处理器(Tensor Processing Unit, TPU),或者图形处理器(Graphics Processing Unit,GPU)等一切适合用于大规模异或运算处理的处理器。人工智能处理器1405可以作为协处理器挂载到主CPU(Host CPU)上,由主 CPU为其分配任务。人工智能处理器1405可以实现上述神经网络的量化方法中涉及的一种或多种运算。例如,以NPU为例,NPU的核心部分为运算电路,通过控制器控制运算电路提取存储器1402中的矩阵数据并进行乘加运算。

[0377] 处理器1401用于调用存储器中的数据和程序代码,执行上述神经网络的量化方法。

[0378] 应理解,各个器件的实现还可以对应参照上述神经网络的训练方法实施例中的相应描述,本申请实施例不再赘述。

[0379] 本发明实施例还提供了一种计算机存储介质,该计算机可读存储介质中存储有指令,当其在计算机或处理器上运行时,使得计算机或处理器执行上述任一个实施例方法中的一个或多个步骤。上述装置的各组成模块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时,可以存储在计算机可读取存储介质中,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全部或部分可以以软件产品的形式体现出来,该计算机产品存储在计算机可读存储介质中。

[0380] 上述计算机可读存储介质可以是前述实施例的设备的内部存储单元,例如硬盘或内存。上述计算机可读存储介质也可以是上述设备的外部存储设备,例如配备的插接式硬盘,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,上述计算机可读存储介质还可以既包括上述设备的内部存储单元也包括外部存储设备。上述计算机可读存储介质用于存储上述计算机程序以及上述设备所需的其他程序和数据。上述计算机可读存储介质还可以用于暂时地存储已经输出或者将要输出的数据。

[0381] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,可通过计算机程序来指令相关的硬件来完成,该计算机的程序可存储于计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。而前述的存储介质包括:ROM、RAM、磁碟或者光盘等各种可存储程序代码的介质。

[0382] 本申请实施例方法中的步骤可以根据实际需要进行顺序调整、合并和删减。

[0383] 本申请实施例装置中的模块可以根据实际需要进行合并、划分和删减。

[0384] 可以理解,本领域普通技术人员可以意识到,结合本申请各个实施例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申请的范围。

[0385] 本领域技术人员能够领会,结合本申请各个实施例中公开描述的各种说明性逻辑框、模块和算法步骤所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件来实施,那么各种说明性逻辑框、模块、和步骤描述的功能可作为一或多个指令或代码在计算机可读媒体上存储或传输,且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体,其对应于有形媒体,例如数据存储媒体,或包括任何促进将计算机程序从一处传送到另一处的媒体(例如,根据通信协议)的通信媒体。以此方式,计算机可读媒体大体上可对应于(1)非暂时性的有形计算机可读存储媒体,或(2)通信媒体,例如信号或载波。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索用于实施本申请中描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0386] 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、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0387] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,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信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

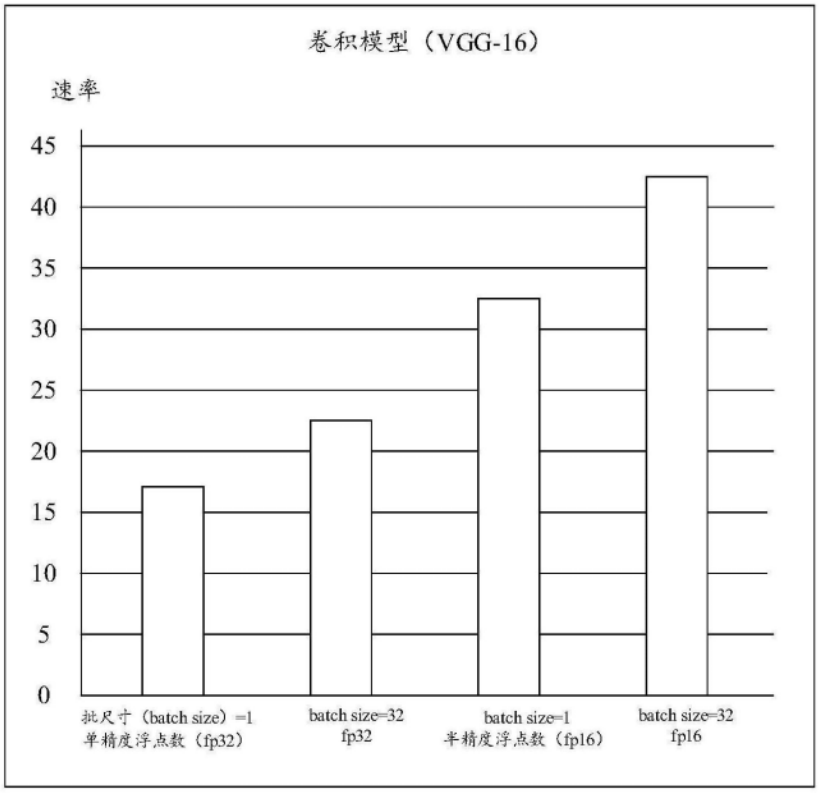
[0388] 作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0389] 另外,在本申请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。

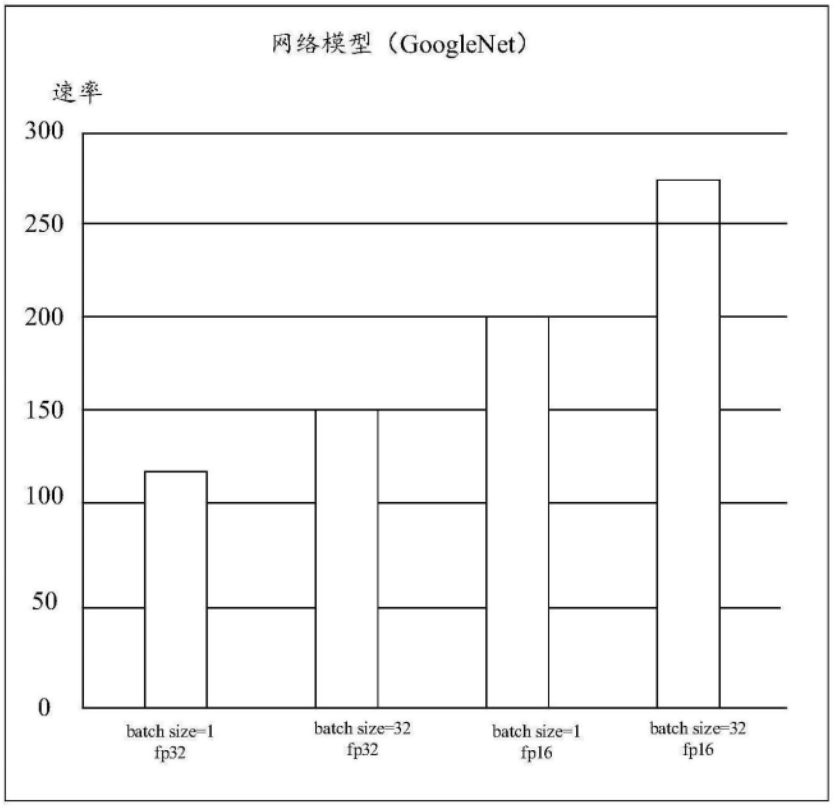
[0390] 功能如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,基站,或者网络设备等)执行本申请各个实施例方法的全部或部分步骤。而前述的存

储介质包括：U盘、移动硬盘、只读存储器(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory,RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0391] 以上,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以所述权利要求的保护范围为准。


卷积模型（VGG-16）
速率
45
40
35
30
25
20
15
10
5
0
批尺寸（batch size）=1
单精度浮点数（fp32）
batch size=32
fp32
batch size=1
半精度浮点数（fp16）
batch size=32
fp16

（a）


网络模型（GoogleNet）
速率
300
250
200
150
100
50
0
batch size=1
fp32
batch size=32
fp32
batch size=1
fp16
batch size=32
fp16

（b）

图1A

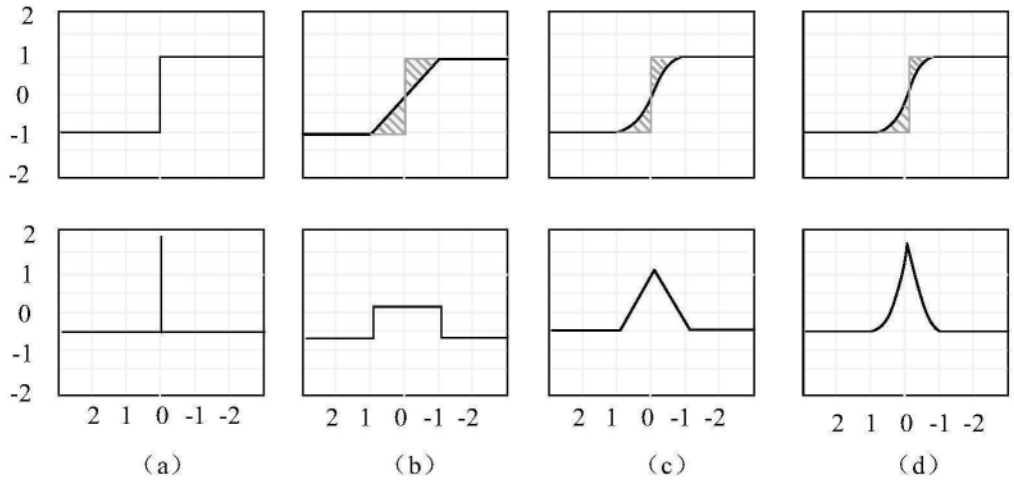


图1B

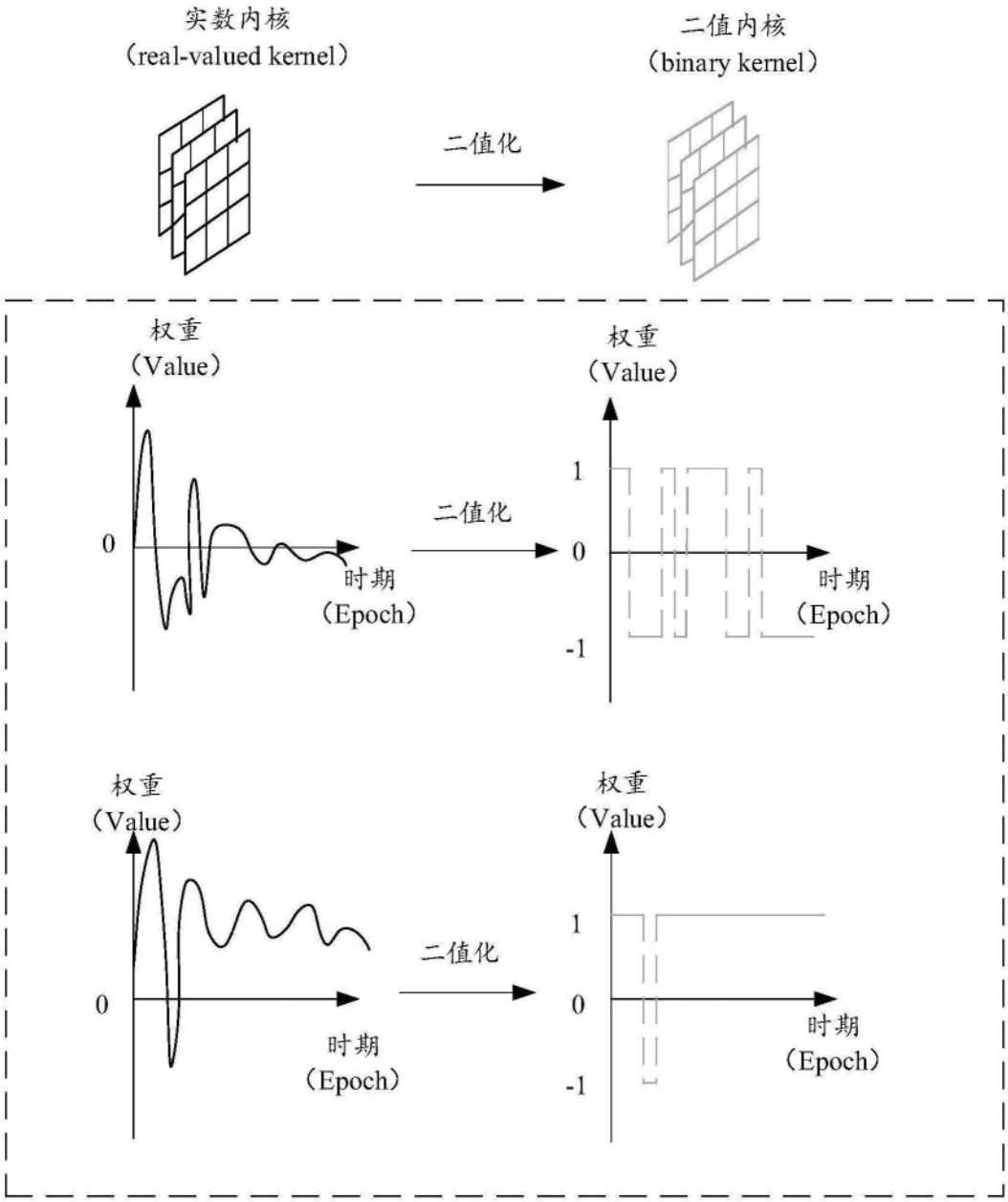

实数内核
(real-valued kernel)
二值内核
(binary kernel)
二值化
权重
(Value)
0
时期
(Epoch)
二值化
权重
(Value)
1
0
-1
时期
(Epoch)
权重
(Value)
0
时期
(Epoch)
二值化
权重
(Value)
1
0
-1
时期
(Epoch)

图1C

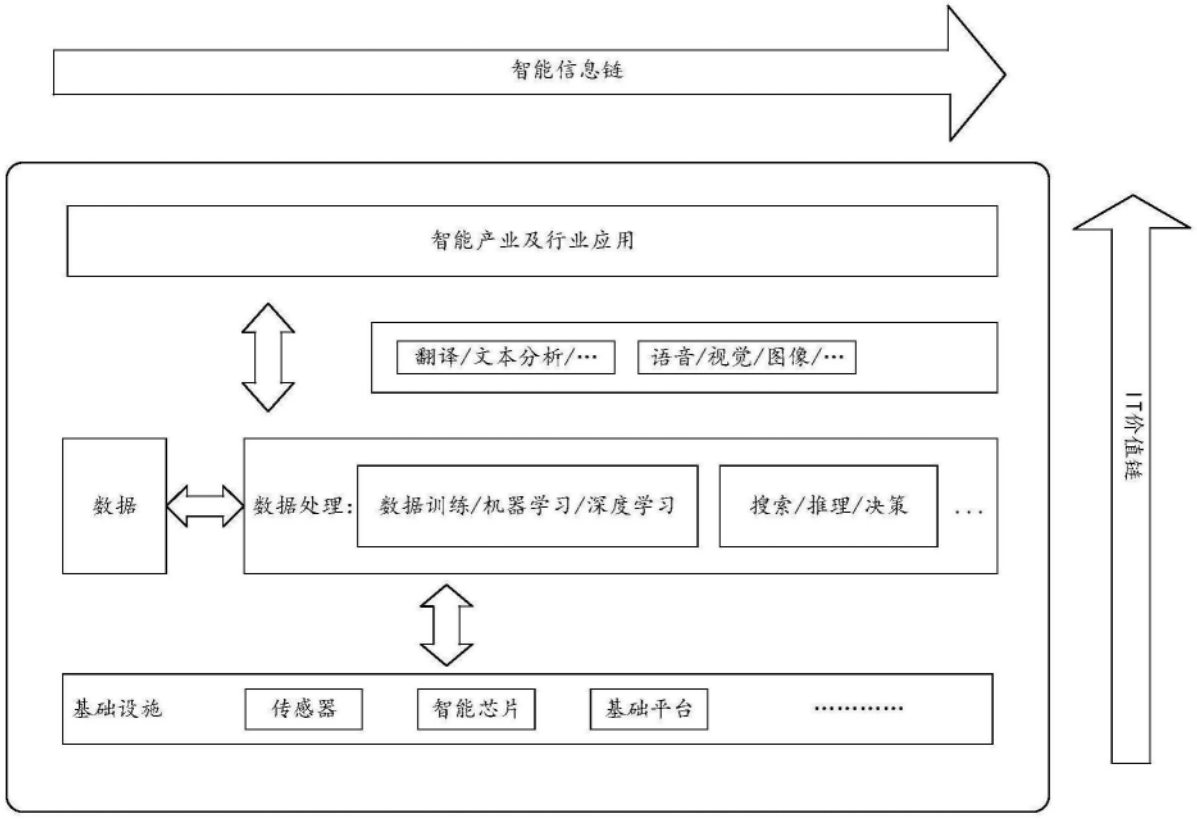

智能信息链
智能产业及行业应用
翻译/文本分析/…
语音/视觉/图像/…
数据
数据处理:
数据训练/机器学习/深度学习
搜索/推理/决策
...
基础设施
传感器
智能芯片
基础平台
……
IT价值链

图2

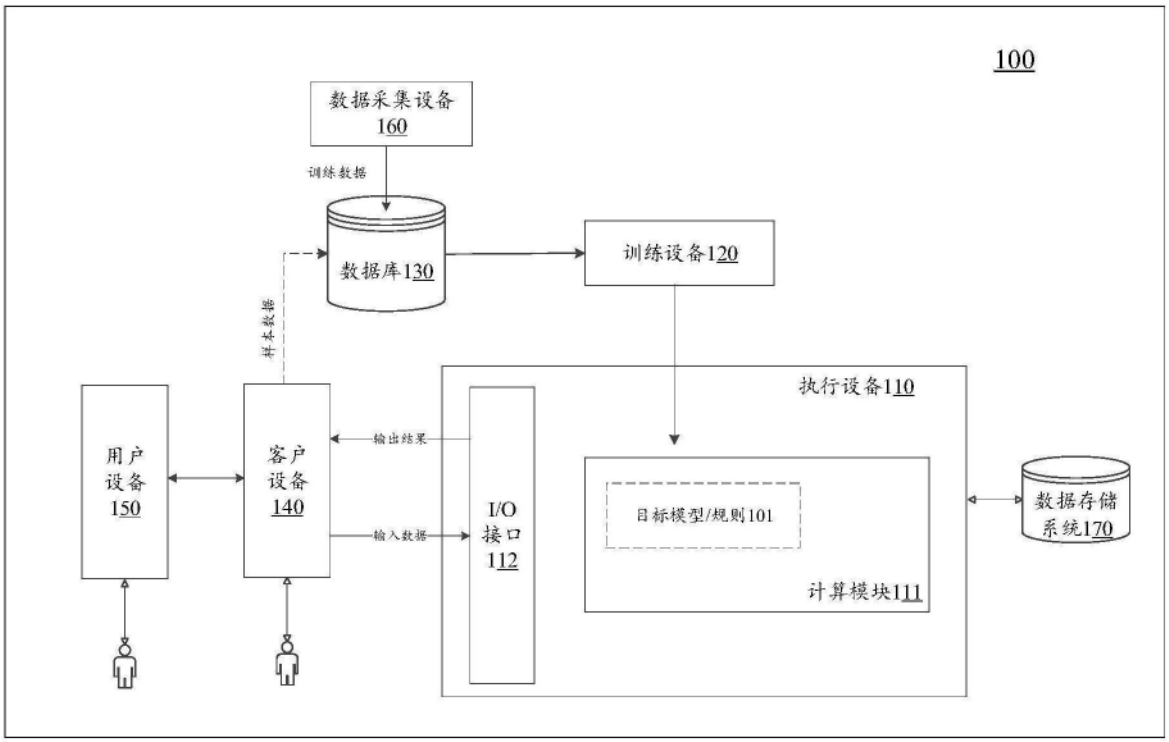

100
数据采集设备
160
训练数据
数据库130
训练设备120
样本数据
执行设备110
用户
设备
150
客户
设备
140
输出结果
输入数据
I/O
接口
112
目标模型/规则101
计算模块111
数据存储
系统170

图3

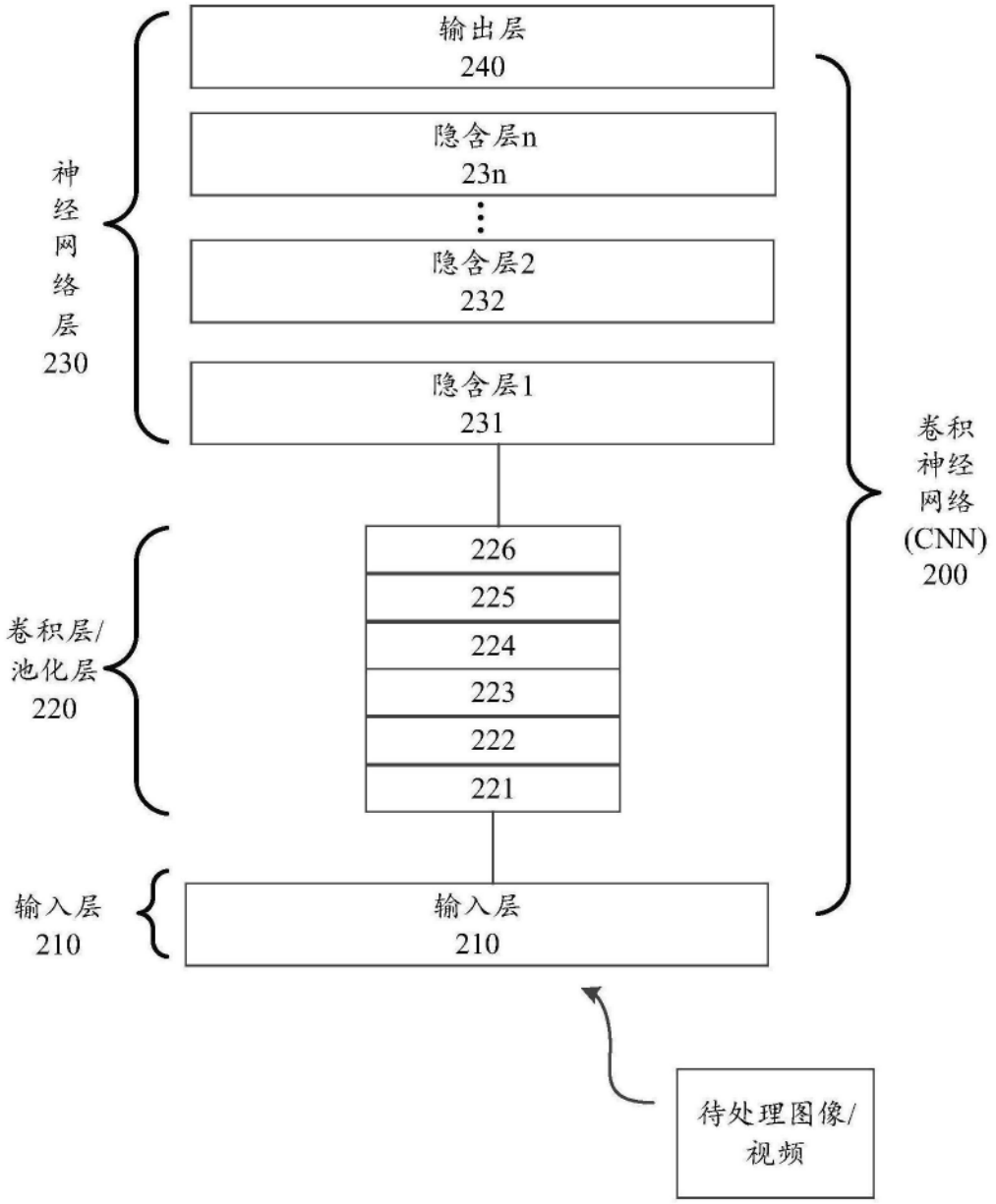

输出层
240
隐含层n
23n
隐含层2
232
隐含层1
231
神经网络层
230
226
225
224
223
222
221
卷积层/池化层
220
输入层
210
输入层
210
卷积神经网络(CNN)
200
待处理图像/视频

图4A

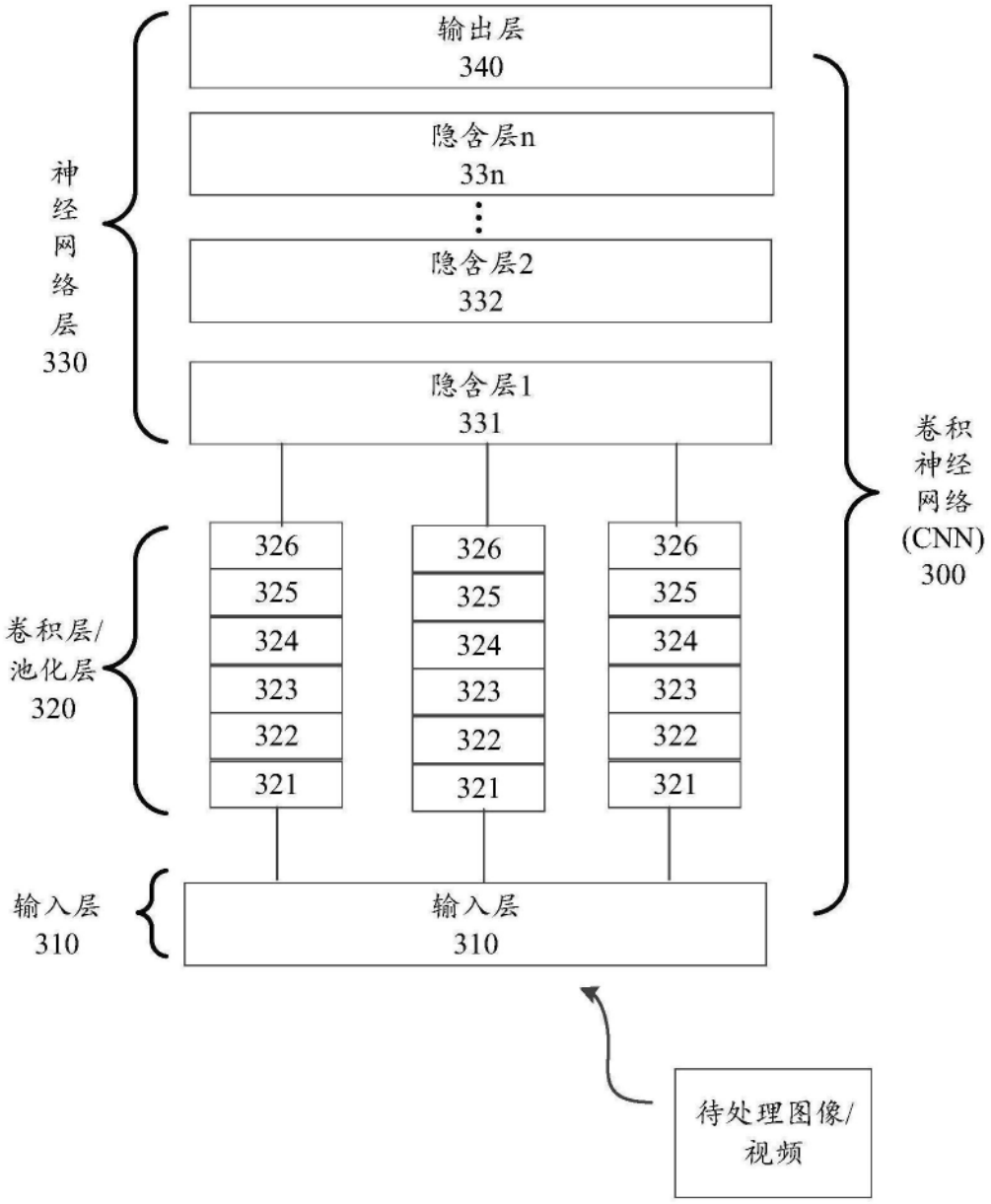

输出层
340
隐含层n
33n
隐含层2
332
隐含层1
331
神经网络层
330
326
325
324
323
322
321
卷积层/池化层
320
输入层
310
卷积神经网络(CNN)
300
待处理图像/视频

图4B

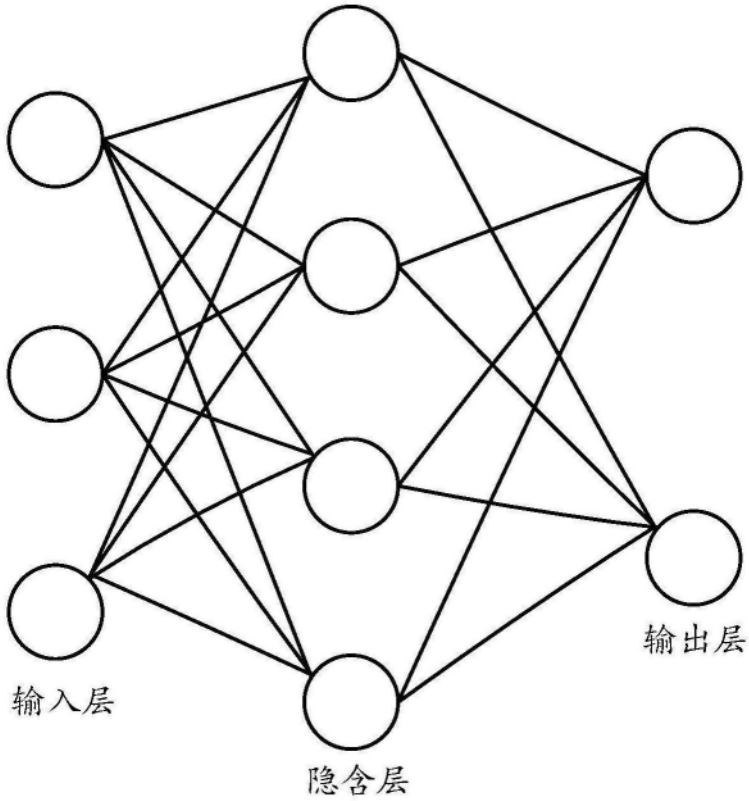

输入层
隐含层
输出层

图4C

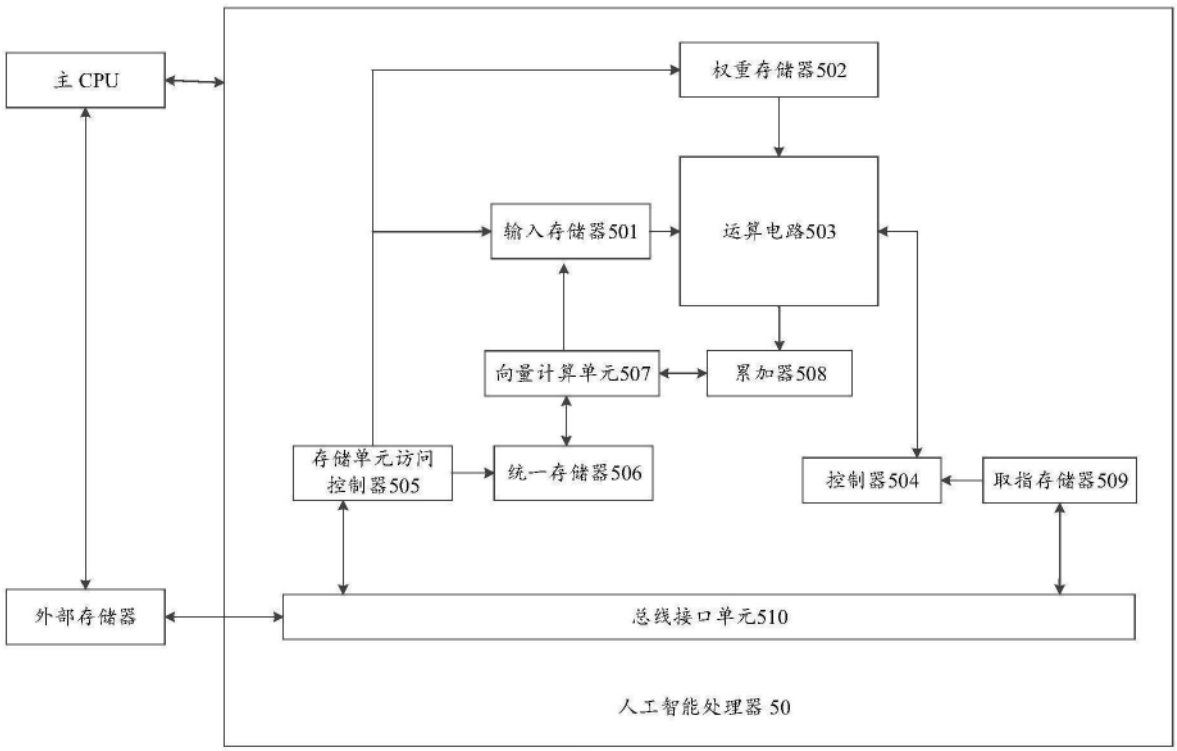

主 CPU
权重存储器502
输入存储器501
运算电路503
向量计算单元507
累加器508
存储单元访问
控制器505
统一存储器506
控制器504
取指存储器509
外部存储器
总线接口单元510
人工智能处理器 50

图5

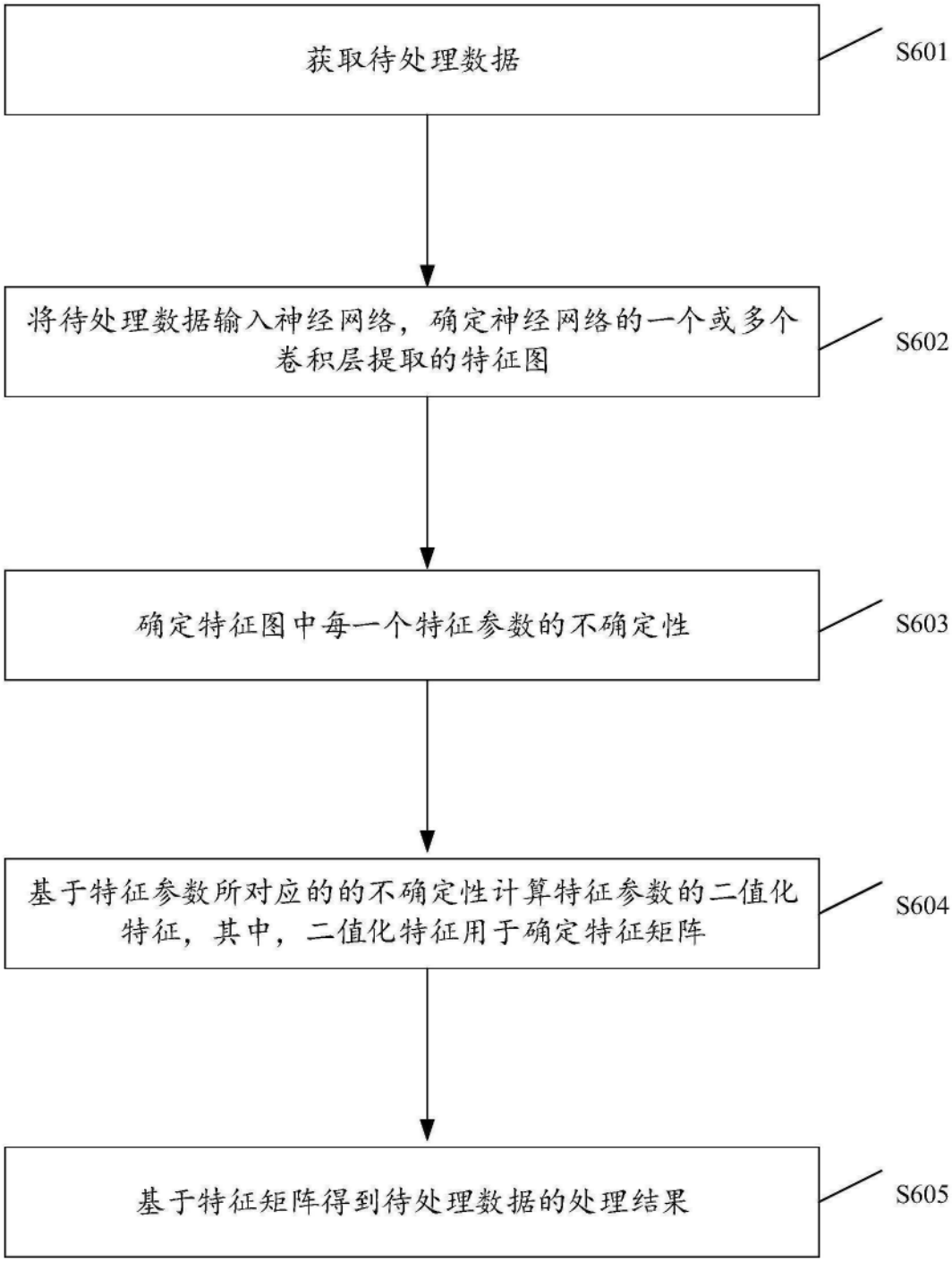

获取待处理数据
S601
将待处理数据输入神经网络，确定神经网络的一个或多个卷积层提取的特征图
S602
确定特征图中每一个特征参数的不确定性
S603
基于特征参数所对应的的不确定性计算特征参数的二值化特征，其中，二值化特征用于确定特征矩阵
S604
基于特征矩阵得到待处理数据的处理结果
S605

图6

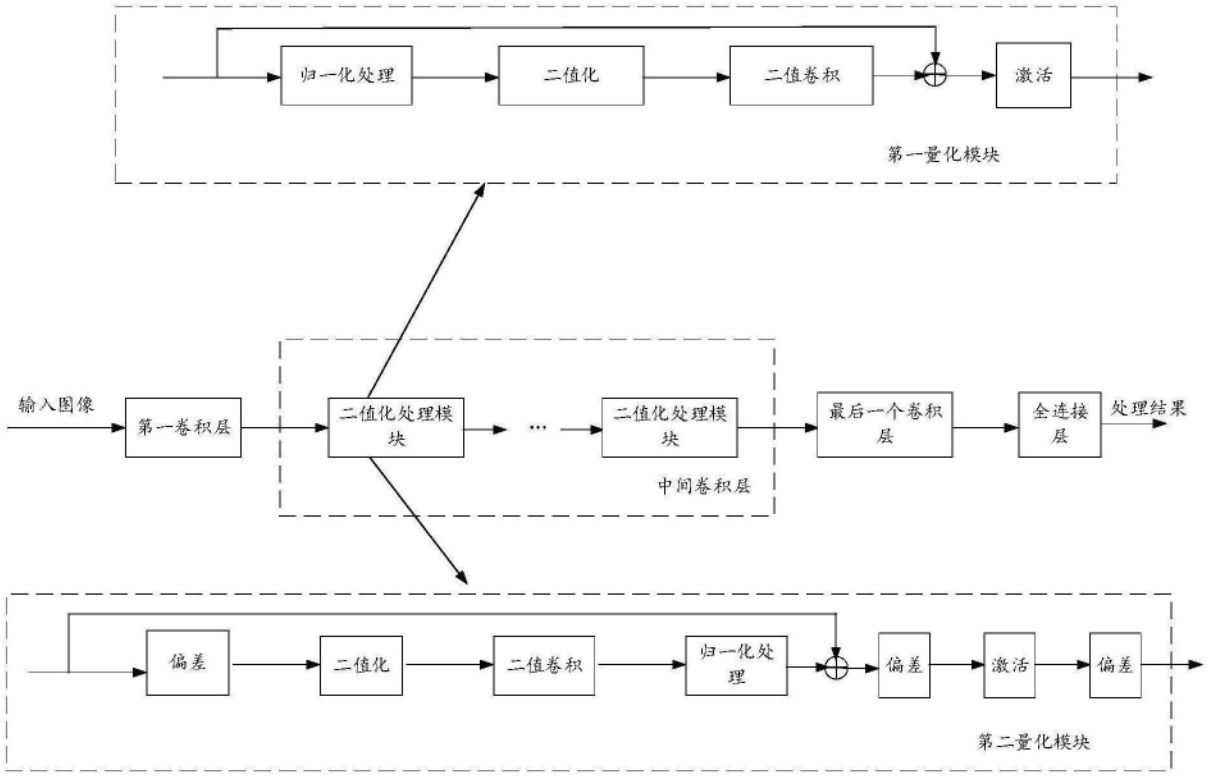

归一化处理
二值化
二值卷积
激活
第一量化模块
输入图像
第一卷积层
二值化处理模块
...
二值化处理模块
中间卷积层
最后一个卷积层
全连接层
处理结果
偏差
二值化
二值卷积
归一化处理
偏差
激活
偏差
第二量化模块

图7

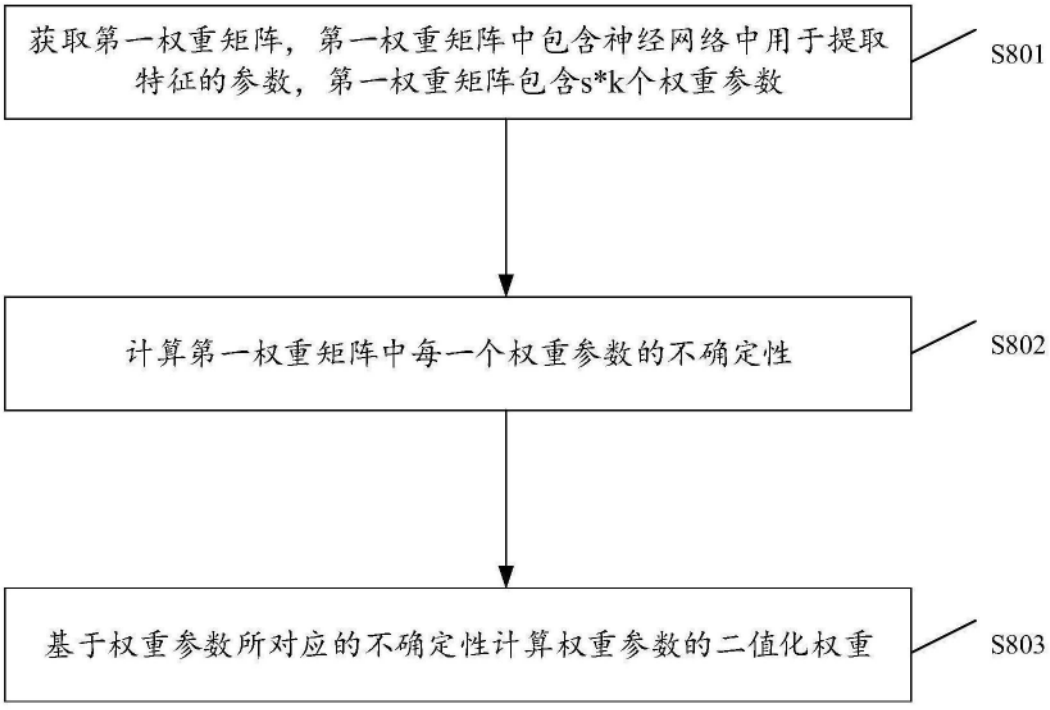

获取第一权重矩阵，第一权重矩阵中包含神经网络中用于提取特征的参数，第一权重矩阵包含s*k个权重参数
S801
计算第一权重矩阵中每一个权重参数的不确定性
S802
基于权重参数所对应的不确定性计算权重参数的二值化权重
S803

图8A

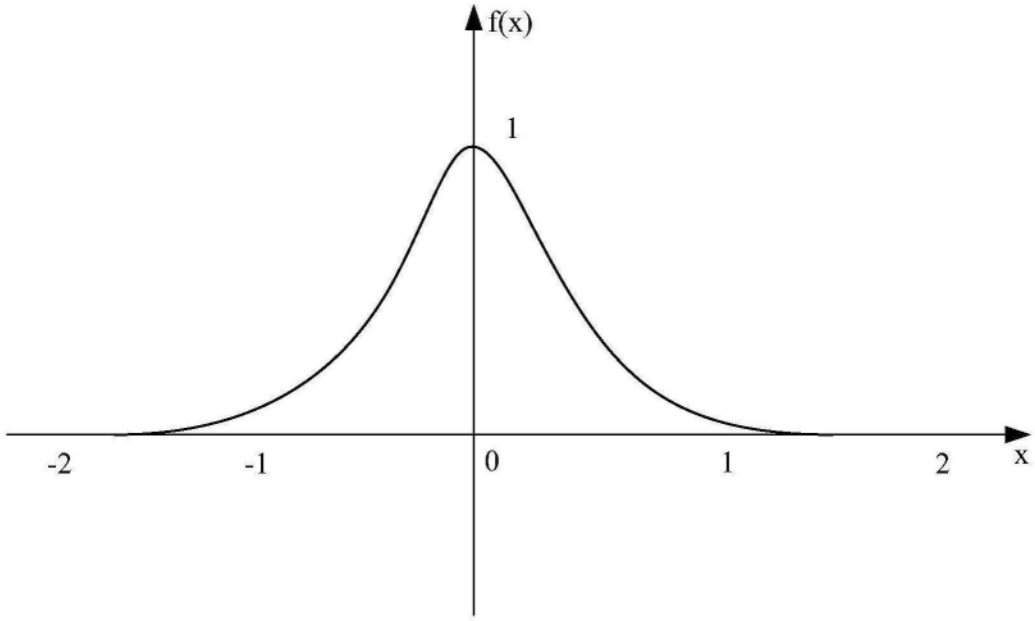

f(x)
1
-2
-1
0
1
2
x

图8B

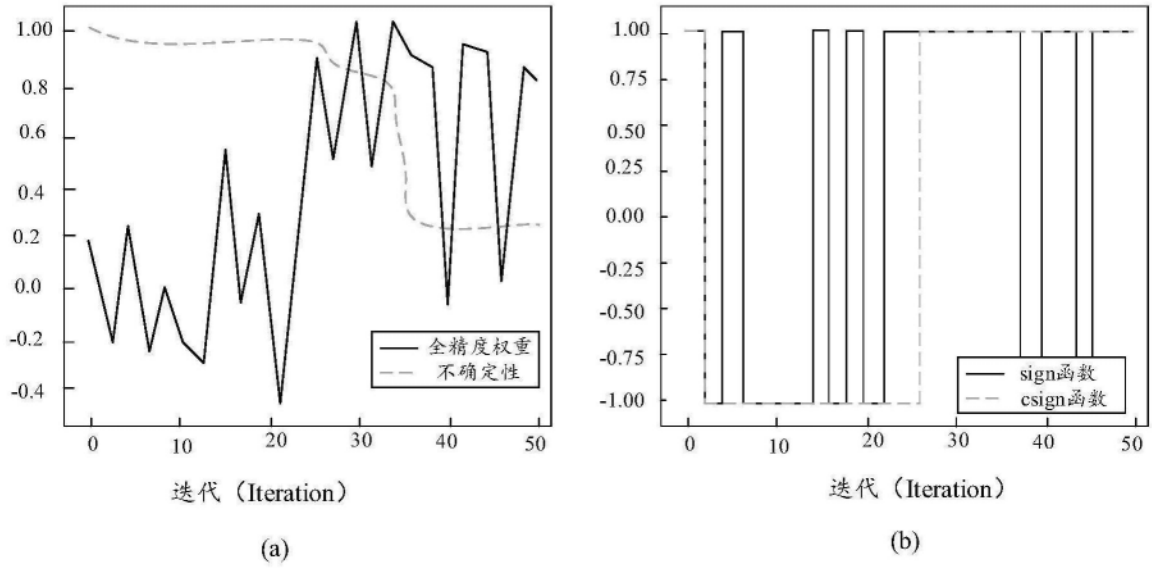

1.00
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
10
20
30
40
50
全精度权重
不确定性
迭代（Iteration）
(a)
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
sign函数
csign函数
迭代（Iteration）
(b)

图8C

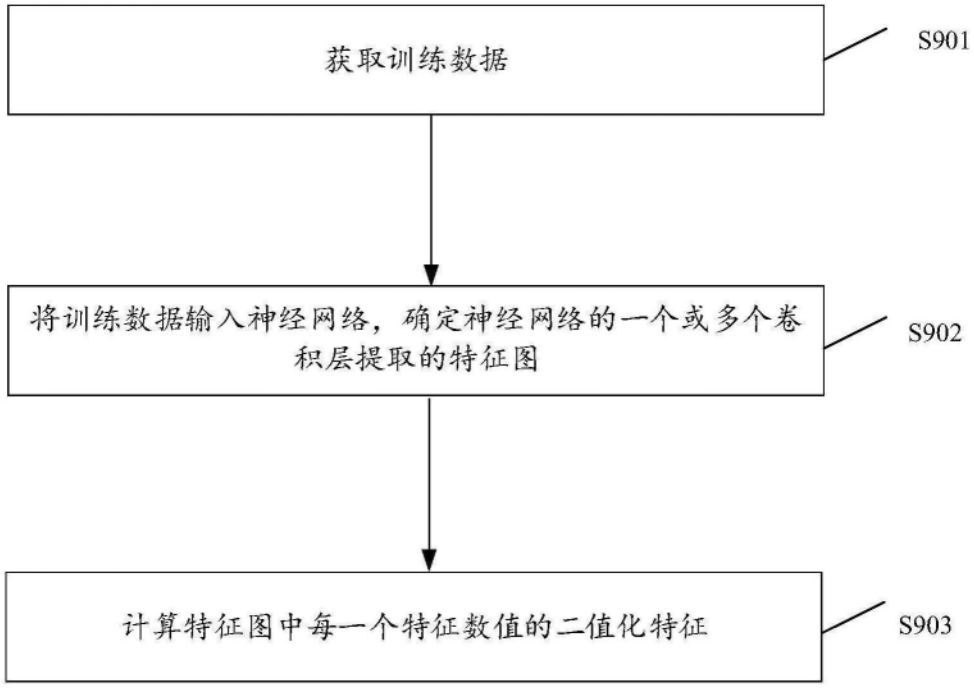

获取训练数据
S901
将训练数据输入神经网络，确定神经网络的一个或多个卷积层提取的特征图
S902
计算特征图中每一个特征数值的二值化特征
S903

图9

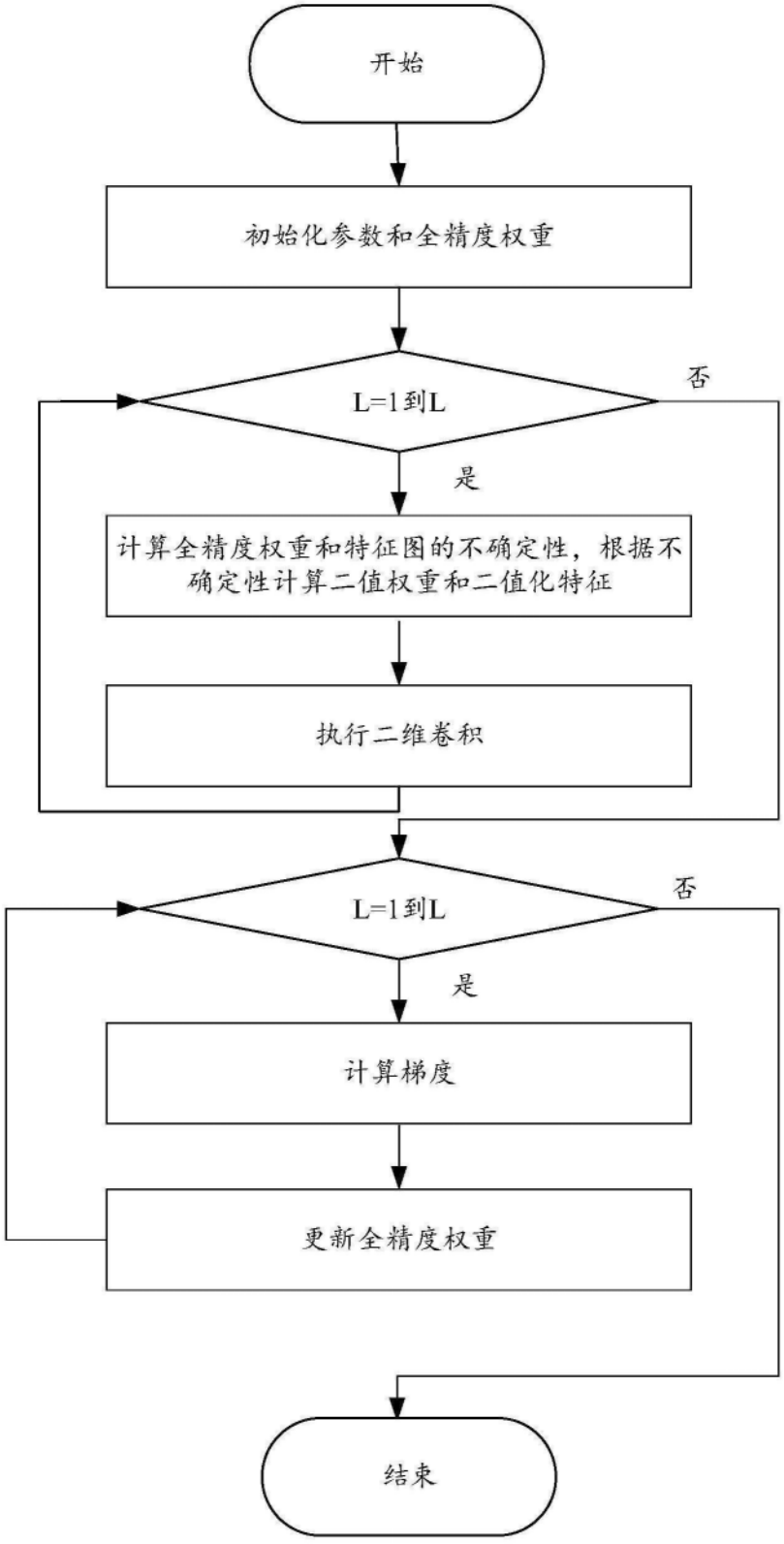

开始
初始化参数和全精度权重
L=1到L
否
是
计算全精度权重和特征图的不确定性，根据不确定性计算二值权重和二值化特征
执行二维卷积
L=1到L
否
是
计算梯度
更新全精度权重
结束

图10

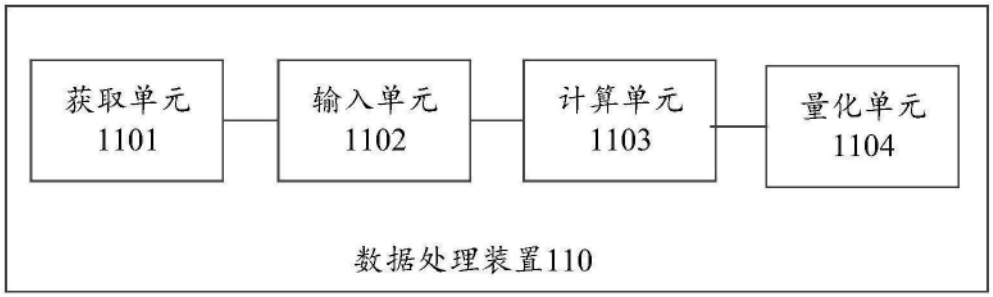

获取单元
1101
输入单元
1102
计算单元
1103
量化单元
1104
数据处理装置110

图11

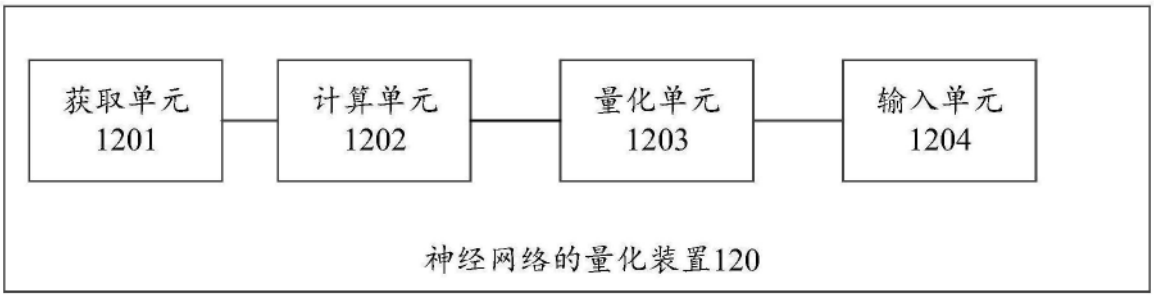

获取单元
1201
计算单元
1202
量化单元
1203
输入单元
1204
神经网络的量化装置120

图12

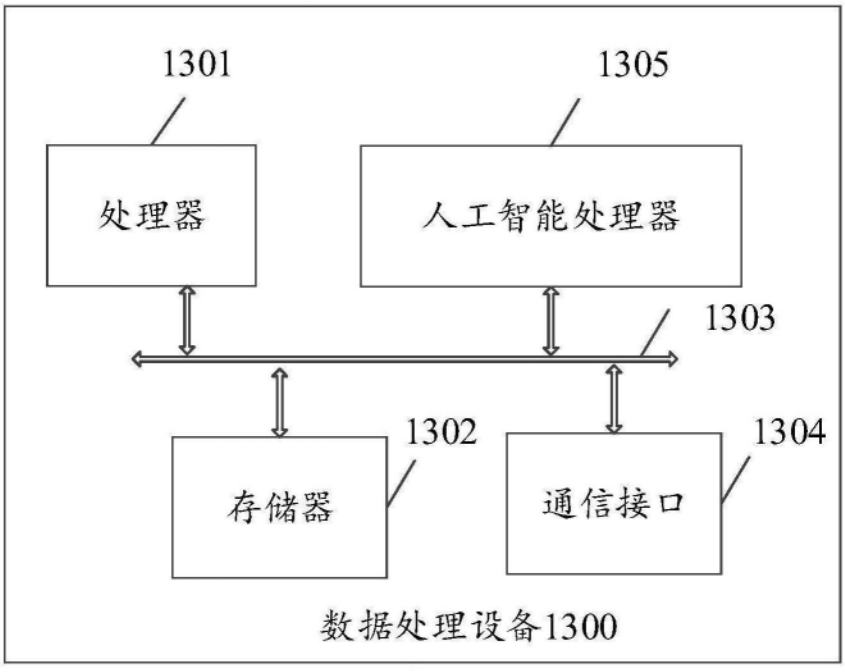

1301
1305
处理器
人工智能处理器
1303
1302
1304
存储器
通信接口
数据处理设备1300

图13

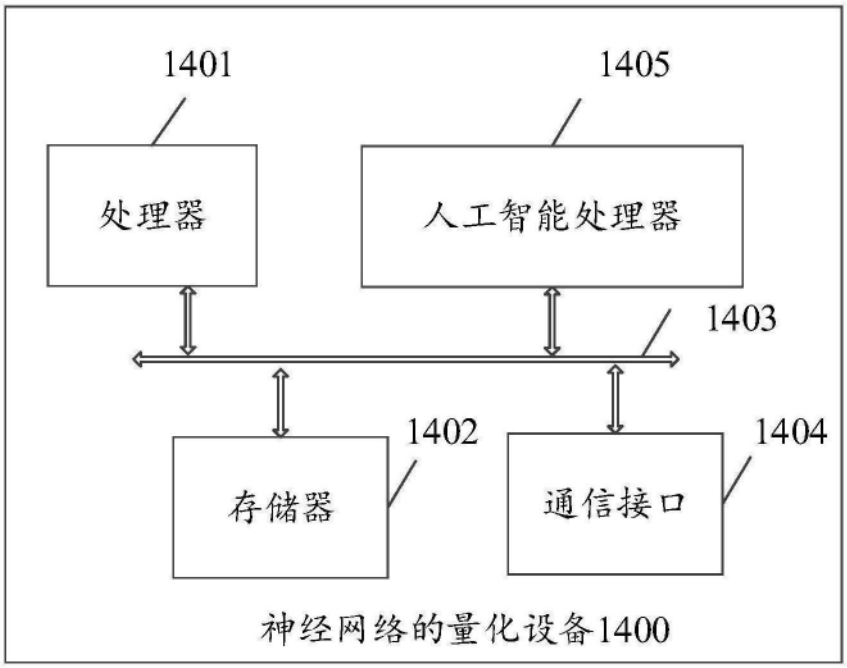

1401
1405
处理器
人工智能处理器
1403
1402
1404
存储器
通信接口
神经网络的量化设备1400

图14