



(12)发明专利申请

(10)申请公布号 CN 108256555 A

(43)申请公布日 2018.07.06

(21)申请号 201711394566.7

(22)申请日 2017.12.21

(71)申请人 北京达佳互联信息技术有限公司
地址 100084 北京市海淀区中关村东路1号
院8号楼20层B2201

(72)发明人 张志伟 杨帆

(74)专利代理机构 北京润泽恒知识产权代理有
限公司 11319
代理人 莎日娜

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

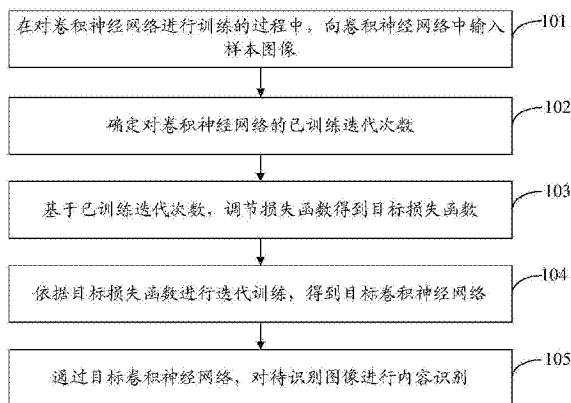
权利要求书2页 说明书11页 附图4页

(54)发明名称

图像内容识别方法、装置及终端

(57)摘要

本发明实施例提供了一种图像内容识别方法、装置及终端,其中,所述方法包括:在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像,其中,所述样本图像用于对所述卷积神经网络进行迭代训练;确定对所述卷积神经网络的已训练迭代次数;基于所述已训练迭代次数,调节损失函数得到目标损失函数;依据所述目标损失函数进行迭代训练,得到目标卷积神经网络;通过所述目标卷积神经网络,对待识别图像进行内容识别。通过本发明实施例提供的卷积神经网络训练方案,能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数,从而在保证卷积神经网络识别结果准确率的情况下,增加样本的召回率。



1. 一种图像内容识别方法,其特征在于,所述方法包括:

在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像,其中,所述样本图像用于对所述卷积神经网络进行迭代训练;

确定对所述卷积神经网络的已训练迭代次数;

基于所述已训练迭代次数,调节损失函数得到目标损失函数;

依据所述目标损失函数进行迭代训练,得到目标卷积神经网络;

通过所述目标卷积神经网络,对待识别图像进行内容识别。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述已训练迭代次数,调节预设损失函数得到目标损失函数的步骤,包括:

提取预设损失函数,判断所述已训练迭代次数是否大于第一预设次数;若否,将所述预设损失函数中的超参数调节为0,得到目标损失函数;

若是,将所述预设损失函数中的超参数调节为预设值,得到目标损失函数。

3. 根据权利要求1所述的方法,其特征在于,所述预设损失函数如下:

$$\text{sinFocalLoss} = -(1-p_t)^{\gamma \sin(2\pi * \text{clip}(s-i, 0, i/2) / i)} \log(p_t)$$

其中, p_t 为概率值, γ 为超参数, i 为迭代次数上限值, s 为已训练迭代次数;

$$\text{clip}(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases}.$$

4. 根据权利要求3所述的方法,其特征在于,所述基于所述已训练迭代次数,调节预设损失函数得到目标损失函数的步骤,包括:

确定迭代次数上限值;

将所述迭代次数上限值和所述已训练迭代次数,代入所述预设损失函数中,得到目标损失函数。

5. 根据权利要求1-4任一项所述的方法,其特征在于,依据所述目标损失函数进行一次迭代训练的步骤,包括:

通过所述卷积神经网络确定所述样本图像对应的特征图;

将所述特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量;其中,所述第一特征向量中包括多个点,每个点对应一个所述卷积神经网络中的分类标签以及一个概率值;

基于所述目标损失函数计算所述卷积神经网络的平均损失值;

计算所述目标损失函数在所述特征向量中各点的偏导数得到梯度值,依据所述梯度值对所述卷积神经网络对应的模型参数进行更新。

6. 一种图像内容识别装置,其特征在于,所述装置包括:

输入模块,被配置为在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像,其中,所述样本图像用于对所述卷积神经网络进行迭代训练;

确定模块,被配置为确定对所述卷积神经网络的已训练迭代次数;

损失函数调节模块,被配置为基于所述已训练迭代次数,调节损失函数得到目标损失函数;

训练模块,被配置为依据所述目标损失函数进行本次迭代训练,得到目标卷积神经网络;

预测模块,被配置为通过所述目标卷积神经网络,对待识别图像进行内容识别。

7. 根据权利要求6所述的装置,其特征在于,所述损失函数调节模块包括:

提取子模块,被配置为提取预设损失函数,判断所述已训练迭代次数是否大于第一预设次数;

第一调节子模块,被配置为若否,将所述预设损失函数中的超参数调节为0,得到目标损失函数;

第二调节子模块,被配置为若是,将所述预设损失函数中的超参数调节为预设值,得到目标损失函数。

8. 根据权利要求6所述的装置,其特征在于,所述预设损失函数如下:

$$\text{sinFocalLoss} = -(1-p_t)^{\gamma \sin(2\pi * \text{clip}(s-i, 0, i/2) / i)} \log(p_t)$$

其中, p_t 为概率值, γ 为超参数, i 为迭代次数上限值, s 为已训练迭代次数;

$$\text{clip}(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases} .$$

9. 根据权利要求8所述的装置,其特征在于,所述损失函数调节模块包括:

上限值确定子模块,被配置为确定迭代次数上限值;

代入子模块,被配置为将所述迭代次数上限值和所述已训练迭代次数,代入所述预设损失函数中,得到目标损失函数。

10. 根据权利要求6-9任一项所述的装置,其特征在于,所述训练模块包括:

特征图确定子模块,被配置为通过所述卷积神经网络确定所述样本图像对应的特征图;

处理子模块,被配置为将所述特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量;其中,所述第一特征向量中包括多个点,每个点对应一个所述卷积神经网络中的分类标签以及一个概率值;

计算子模块,被配置为基于所述目标损失函数计算所述卷积神经网络的平均损失值;

更新子模块,被配置为计算所述目标损失函数在所述特征向量中各点的偏导数得到梯度值,依据所述梯度值对所述卷积神经网络对应的模型参数进行更新。

11. 一种终端,其特征在于,包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的卷积神经网络训练程序,所述图像内容识别程序被所述处理器执行时实现如权利要求1至5中任一项所述的图像内容识别方法的步骤。

12. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有图像内容识别程序,所述图像内容识别程序被处理器执行时实现如权利要求1至5中任一项所述的图像内容识别方法的步骤。

图像内容识别方法、装置及终端

技术领域

[0001] 本发明涉及图像识别技术领域，特别是涉及一种图像内容识别方法、装置及终端。

背景技术

[0002] 深度学习在视频图像、语音识别、自然语言处理等相关领域得到了广泛应用。卷积神经网络作为深度学习的一个重要分支，由于其超强的拟合能力以及端到端的全局优化能力，使得其在目标检测、分类等计算机视觉任务中所得预测结果的精度大幅提升。

[0003] 但是在实际应用中，一般不会直接使用卷积神经网络产生的结果。以一个二分类任务为例，对于一个输入数据卷积神经网络会给出它在某一个类别上的概率。概率阈值会根据具体的应用场景设置，通常情况下会设置一个较高的阈值来获得较高的准确率，但是图像样本的召回率就会相应下降，显然识别结果的准确率与图像样本的召回率成反比。可见，目前迫切需要本领域技术人员迫切解决的技术问题为：如何在保证卷积神经网络识别结果准确率的情况下，增加样本的召回率。

发明内容

[0004] 本发明实施例提供一种图像内容识别方法、装置及终端，以解决现有技术中存在的卷积神经网络识别结果准确率与样本的召回率无法兼顾的问题。

[0005] 依据本发明的一个方面，提供了一种图像内容识别方法，所述方法包括：在对卷积神经网络进行训练的过程中，向卷积神经网络中输入样本图像，其中，所述样本图像用于对所述卷积神经网络进行迭代训练；确定对所述卷积神经网络的已训练迭代次数；基于所述已训练迭代次数，调节损失函数得到目标损失函数；依据所述目标损失函数进行迭代训练，得到目标卷积神经网络；通过所述目标卷积神经网络，对待识别图像进行内容识别。

[0006] 可选地，所述基于所述已训练迭代次数，调节预设损失函数得到目标损失函数的步骤，包括：提取预设损失函数，判断所述已训练迭代次数是否大于第一预设次数；若否，将所述预设损失函数中的超参数调节为0，得到目标损失函数；若是，将所述预设损失函数中的超参数调节为预设值，得到目标损失函数。

[0007] 可选地，所述预设损失函数如下：

[0008] $\text{sinFocalLoss} = -(1-p_t)^{\gamma \sin(2\pi * \text{clip}(s-i, 0, i/2) / i)} \log(p_t)$

[0009] 其中， p_t 为概率值， γ 为超参数， i 为迭代次数上限值， s 为已训练迭代次数；

[0010]
$$\text{clip}(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases} .$$

[0011] 可选地，所述基于所述已训练迭代次数，调节预设损失函数得到目标损失函数的步骤，包括：确定迭代次数上限值；将所述迭代次数上限值和所述已训练迭代次数，代入所述预设损失函数中，得到目标损失函数。

[0012] 可选地,所述依据所述目标损失函数进行本次迭代训练的步骤,包括:通过所述卷积神经网络确定所述样本图像对应的特征图;将所述特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量;其中,所述特征向量中包括多个点,每个点对应一个所述卷积神经网络中的分类标签以及一个概率值;基于所述目标损失函数计算所述卷积神经网络的平均损失值;计算所述目标损失函数在所述特征向量中各点的偏导数得到梯度值,依据所述梯度值对所述卷积神经网络对应的模型参数进行更新。

[0013] 依据本发明的另一个方面,提供了一种图像内容识别装置,所述装置包括:输入模块,被配置为在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像,其中,所述样本图像用于对所述卷积神经网络进行迭代训练;确定模块,被配置为确定对所述卷积神经网络的已训练迭代次数;损失函数调节模块,被配置为基于所述已训练迭代次数,调节损失函数得到目标损失函数;训练模块,被配置为依据所述目标损失函数进行本次迭代训练,得到目标卷积神经网络;预测模块,被配置为通过所述目标卷积神经网络,对待识别图像进行内容识别。

[0014] 可选地,所述损失函数调节模块包括:提取子模块,被配置为提取预设损失函数,判断所述已训练迭代次数是否大于第一预设次数;第一调节子模块,被配置为若否,将所述预设损失函数中的超参数调节为0,得到目标损失函数;第二调节子模块,被配置为若是,将所述预设损失函数中的超参数调节为预设值,得到目标损失函数。

[0015] 可选地,所述预设损失函数如下:

$$[0016] \quad \text{sinFocalLoss} = -(1-p_t)^{\gamma \sin(2\pi * \text{clip}(s-i, 0, i/2) / i)} \log(p_t)$$

[0017] 其中, p_t 为概率值, γ 为超参数, i 为迭代次数上限值, s 为已训练迭代次数;

$$[0018] \quad \text{clip}(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases} .$$

[0019] 可选地,所述损失函数调节模块包括:上限值确定子模块,被配置为确定迭代次数上限值;代入子模块,被配置为将所述迭代次数上限值和所述已训练迭代次数,代入所述预设损失函数中,得到目标损失函数。

[0020] 可选地,所述训练模块包括:特征图确定子模块,被配置为通过所述卷积神经网络确定所述样本图像对应的特征图;处理子模块,被配置为将所述特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量;其中,所述特征向量中包括多个点,每个点对应一个所述卷积神经网络中的分类标签以及一个概率值;计算子模块,被配置为基于所述目标损失函数计算所述卷积神经网络的平均损失值;更新子模块,被配置为计算所述目标损失函数在所述特征向量中各点的偏导数得到梯度值,依据所述梯度值对所述卷积神经网络对应的模型参数进行更新。

[0021] 根据本发明的再一方面,提供了一种终端,包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的图像内容识别程序,所述图像内容识别程序被所述处理器执行时实现本发明中所述的任意一种图像内容识别方法的步骤。

[0022] 根据本发明的又一方面,提供了一种计算机可读存储介质,所述计算机可读存储介质上存储有图像内容识别程序,所述图像内容识别程序被处理器执行时实现本发明中

所述的任意一种图像内容识别方法的步骤。

[0023] 与现有技术相比,本发明具有以下优点:

[0024] 本发明实施例提供的图像内容识别方案,在依据输入的样本图像对卷积神经网络进行迭代训练时,基于迭代次数动态调整每次用于迭代训练的目标损失函数,能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数,从而在保证卷积神经网络识别结果准确率的情况下,增加样本的召回率。

[0025] 上述说明仅是本发明技术方案的概述,为了能够更清楚了解本发明的技术手段,而可依照说明书的内容予以实施,并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂,以下特举本发明的具体实施方式。

附图说明

[0026] 通过阅读下文优选实施方式的详细描述,各种优点和益处对于本领域普通技术人员将变得清楚了。附图仅用于示出优选实施方式,而并不认为是对本发明的限制。而且在整个附图中,用相同的参考符号表示相同的部件。在附图中:

[0027] 图1是根据本发明实施例一的一种图像内容识别方法的步骤流程图;

[0028] 图2是根据本发明实施例二的一种图像内容识别方法的步骤流程图;

[0029] 图3是根据本发明实施例三的一种图像内容识别装置的结构框图;

[0030] 图4是根据本发明实施例四的一种终端的结构框图。

具体实施方式

[0031] 下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例,然而应当理解,可以以各种形式实现本公开而不 应被这里阐述的实施例所限制。相反,提供这些实施例是为了能够更透彻地理解本公开,并且能够将本公开的范围完整的传达给本领域的技术人员。

[0032] 实施例一

[0033] 参照图1,示出了本发明实施例一的一种图像内容识别方法的步骤流程图。

[0034] 本发明实施例的图像内容识别方法可以包括以下步骤:

[0035] 步骤101:在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像。

[0036] 其中,样本图像用于对卷积神经网络进行迭代训练。

[0037] 本发明实施例的卷积神经网络可以为多分类内容识别模型,可以识别图像所属的类别;也可以为二分类内容识别模型,可以识别出图像是否属于某一类别。卷积神经网络建模完成后,需要采用大量的样本图像对其进行多次迭代训练,以确保卷积神经网络的收敛性确保预测结果的准确性。各次通过样本图像对卷积神经网络进行训练的具体流程相同,本发明实施例中以输入一个样本图像对卷积神经网络进行一次迭代训练为例进行说明。

[0038] 步骤102:确定对卷积神经网络的已训练迭代次数。

[0039] 由于对卷积神经网络进行训练时,需要对其进行多次迭代训练,每次输入一个样本图像则需要进行一次迭代训练。在训练时,系统累加记录对卷积神经网络的训练迭代次

数,所记录的训练迭代次数用于下一次迭代训练时调整损失函数。

[0040] 例如:在执行步骤101前已对卷积神经网络进行了100次迭代训练,步骤101中输入样本图像则进行第101次迭代训练,故确定已训练迭代次数为100次。

[0041] 步骤103:基于已训练迭代次数,调节损失函数得到目标损失函数。

[0042] 损失函数中设置有可调节变量,可调节变量随迭代次数的变化而变化。从而实现目标损失函数随迭代次数的变化而动态调节。

[0043] 基于迭代次数动态调整每次用于迭代训练的目标损失函数,能够降低简单样本在参数训练中的梯度,能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数。

[0044] 步骤104:依据目标损失函数进行迭代训练,得到目标卷积神经网络。

[0045] 对于基于目标损失函数对卷积神经网络进行迭代训练的具体过程,参数相关技术即可,本发明实施例中对此不做具体限制。通过目标损失函数可以检测卷积神经网络的收敛程度,并且还可以借助目标损失函数计算得到的梯度值对卷积神经网络对应的模型参数进行更新。

[0046] 重复执行多次卷积神经网络迭代训练,待卷积神经网络收敛到预设程度时,生成目标卷积神经网络,后续可通过目标卷积神经网络进行图像内容识别。

[0047] 步骤105:通过目标卷积神经网络,对待识别图像进行内容识别。

[0048] 若目标卷积神经网络为二分类识别模型,则对待识别图像进行内容识别后,可输出指示待识别图像是否为特定类别图像的结果。

[0049] 其中,目标卷积神经网络可以被训练为识别任意特定类别的图像,例如:A类别的图像,相应地目标卷积神经网络可识别出待识别图像是否属于A类别的图像。

[0050] 本发明实施例提供的图像内容识别方法,在依据输入的样本图像对卷积神经网络进行迭代训练时,基于迭代次数动态调整每次用于迭代训练的目标损失函数,能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数,从而在保证卷积神经网络识别结果准确率的情况下,增加样本的召回率。

[0051] 实施例二

[0052] 参照图2,示出了本发明实施例二的图像内容识别方法的步骤流程图。

[0053] 本发明实施例的图像内容识别方法具体可以包括以下步骤:

[0054] 步骤201:在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像。

[0055] 其中,样本图像用于对卷积神经网络进行迭代训练。卷积神经网络建模完成后,需要采用大量的样本图像对其进行多次迭代训练,以确保卷积神经网络的收敛性确保预测结果的准确性。各次通过样本图像对卷积神经网络进行训练的具体流程相同,本发明实施例中以输入一个样本图像对卷积神经网络进行一次迭代训练为例进行说明。

[0056] 步骤202:确定对卷积神经网络的已训练迭代次数。

[0057] 系统在对卷积神经网络进行迭代训练的过程中,累加各次迭代训练次数。例如:前一次迭代训练为第50次,则本次迭代训练为第51次,相应地下一次迭代训练为第52次。

[0058] 步骤203:基于已训练迭代次数,调节损失函数得到目标损失函数。

[0059] 在具体实现过程中,本领域技术人员可以根据实际需求设置不同的损失函数。所

设置的损失函数不同,则基于已训练迭代次数对损失函数的具体调节方式也不同。无论如何设置损失函数、如何调节损失函数能够保证通过目标损失函数对卷积神经网络进行迭代训练时,降低简单样本在参数训练中的提取,使模型更好地拟合到复杂样本分布。

[0060] 可选地,可将损失函数预设为: $FocalLoss = -(1-p_t)^\gamma \log(p_t)$;其中, p_t 为概率值, γ 为超参数。 p_t 大的样本图像为简单样本图像,反之 p_t 小的样本图像为复杂样本图像。

[0061] 针对上述预设的损失函数,基于已训练迭代次数,调节预设损失函数得到目标损失函数时:提取预设损失函数,判断已训练迭代次数是否大于第一预设次数;若否,将所述预设损失函数中的超参数调节为0,得到目标损失函数;若是,将预设损失函数中的超参数调节为预设值,得到目标损失函数。

[0062] 上述调节预设损失函数的方式,模型训练初期设置 $\gamma = 0$ 训练模型一段时间,待模型几乎收敛时,为了让模型更好地学习困难样本的分布,此时将 γ 调整为预设值,预设值为非0值。该种调节预设损失函数的方式,能够拟合到复杂样本的分布,减小中间概率值分布的样本图像个数。不足之处在于超参 γ 会在训练的过程中出现0到预设值的突变,这样会对模型参数瞬间造成极大地影响。

[0063] 本发明实施例中还提供了一种优选地预设函数,优选地预设损失函数如下:

[0064] $\sinFocalLoss = -(1-p_t)^{\gamma \sin(2\pi * clip(s-i, 0, i/2) / i)} \log(p_t)$

[0065] 其中, p_t 为概率值, γ 为超参数, i 为迭代次数上限值, s 为已训练迭代次数;

[0066] $clip(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases}$ 。

[0067] 针对该优选地预设损失函数,基于已训练迭代次数,调节预设损失函数得到目标损失函数的方式如下:首先,确定迭代次数上限值;其次,将迭代次数上限值和已训练迭代次数,代入预设损失函数中,得到目标损失函数。

[0068] 迭代次数上限值可以由本领域技术人员根据实际需求进行设置,本发明实施例中对于该值不做具体限制。该种优选地预设损失函数、以及调节预设损失函数的方式,不仅能够拟合到复杂样本的分布,减小中间概率值分布的样本图像个数。由于 $\gamma \sin(2\pi * clip(s-i, 0, i/2) / i)$ 是渐变的,故损失函数变化不会对模型参数瞬间造成极大影响。

[0069] 步骤204:通过卷积神经网络确定样本图像对应的特征图。

[0070] 本发明实施例中样本图像可以为视频中的单帧图像,也可以仅为一个多媒体图像。一张图像输入到卷积神经网络中,经过卷积层或者池化层之后会得到特征图。对于将样本图像输入卷积神经网络中,得到特征图的具体处理方式,参照现有相关技术即可,本发明实施例中对此不作具体限制。

[0071] 步骤205:将特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量。

[0072] 其中,特征向量中包括多个点,每个点对应一个卷积神经网络中的分类标签以及一个概率值。该卷积神经网络中包含多个标签,概率值为样本图像与分类标签匹配度。需要说明的是,若卷积神经网络二分类内容识别模型,卷积神经网络中包含两个分类标签,分别为用于指示为该分类的标签和用于指示不为该分类的标签;若卷积神经网络为多分

类内容识别模型,卷积神经网络中将包含各分类对应的分类标签。

[0073] 步骤206:基于目标损失函数计算卷积神经网络的平均损失值。

[0074] 计算平均损失值时,首先通过目标损失函数计算特征向量中各点对应的损失值;然后计算各点对应的损失值的均值,得到平均损失值。

[0075] 通过平均损失值可以判断是否可以结束卷积神经网络的迭代训练。具体地,判断平均损失值是否小于预设损失值;若是,则结束对卷积神经网络的迭代训练无需再返回步骤201向卷积神经网络中输入样本图像;若否,则返回执行步骤201继续向卷积神经网络中输入样本图像,对卷积神经网络进行迭代训练,直至平均损失值小于预设损失值后对卷积神经网络的迭代训练。

[0076] 平均损失值小于预设损失值则可确定卷积神经网络收敛到预设标准。预设损失值可以由本领域技术人员根据实际需求进行设置,本发明实施例中对此不作具体限制。预设损失值越小,则训练完成后的卷积神经网络的收敛性越好;预设损失值越大,则卷积神经网络的迭代训练越容易。

[0077] 步骤207:计算目标损失函数在特征向量中各点的偏导数得到梯度值,依据梯度值对卷积神经网络对应的模型参数进行更新,得到目标卷积神经网络。

[0078] 具体地为计算目标损失函数,在特征向量中各点的偏导数得到梯度值。对卷积神经网络的迭代训练实质上是对模型参数的不断更新,直至卷积神经网络收敛到预设标准后即可进行图像内容预测。

[0079] 需要说明的是,在具体实现过程中步骤207并不局限于在步骤206之后执行,还可以在步骤206之前执行。

[0080] 步骤208:通过目标卷积神经网络,对待识别图像进行内容识别。

[0081] 通过步骤201至步骤207中训练得到的目标卷积神经网络中能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数。因此,将待识别图像输入目标卷积神经网络中进行内容识别时,能够得到精准的识别结果。

[0082] 本发明实施例提供的图像内容识别方法,在依据输入的样本图像对卷积神经网络进行迭代训练时,基于迭代次数动态调整每次用于迭代训练的目标损失函数,能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数,从而在保证卷积神经网络识别结果准确率的情况下,增加样本的召回率。

[0083] 实施例三

[0084] 参照图3,示出了本发明实施例三的一种图像内容识别装置的结构框图。

[0085] 本发明实施例的图像内容识别装置可以包括:输入模块301,被配置为在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像,其中,所述样本图像用于对所述卷积神经网络进行迭代训练;确定模块302,被配置为确定对所述卷积神经网络的已训练迭代次数;损失函数调节模块303,被配置为基于所述已训练迭代次数,调节损失函数得到目标损失函数;训练模块304,被配置为依据所述目标损失函数进行本次迭代训练,得到目标卷积神经网络;预测模块305,被配置为通过所述目标卷积神经网络,对待识别图像进行内容识别。

[0086] 优选地,所述损失函数调节模块303可以包括:提取子模块3031,被配置为提取预设损失函数,判断所述已训练迭代次数是否大于第一预设次数;第一调节子模块3032,被

配置为若否,将所述预设损失函数中的超参数调节为0,得到目标损失函数;第二调节子模块3033,被配置为若是,将所述预设损失函数中的超参数调节为预设值,得到目标损失函数。

[0087] 优选地,所述预设损失函数如下:

[0088] $\text{sinFocalLoss} = -(1-p_t)^{\gamma \sin(2\pi * \text{clip}(s-i, 0, i/2)/i)} \log(p_t)$

[0089] 其中, p_t 为概率值, γ 为超参数 i 为迭代次数上限值, s 为已训练迭代次数;

[0090]
$$\text{clip}(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases} .$$

[0091] 优选地,所述损失函数调节模块303可以包括:上限值确定子模块3034,被配置为确定迭代次数上限值;代入子模块3035,被配置为将所述迭代次数上限值和所述已训练迭代次数,代入所述预设损失函数中,得到目标损失函数。

[0092] 优选地,所述训练模块304可以包括:特征图确定子模块3041,被配置为通过所述卷积神经网络确定所述样本图像对应的特征图;处理子模块3042,被配置为将所述特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量;其中,所述第一特征向量中包括多个点,每个点对应一个所述卷积神经网络中的分类标签以及一个概率值;计算子模块3043,被配置为基于所述目标损失函数计算所述卷积神经网络的平均损失值;更新子模块3044,被配置为计算所述目标损失函数在所述特征向量中各点的偏导数得到梯度值,依据所述梯度值对所述卷积神经网络对应的模型参数进行更新。

[0093] 本发明实施例的图像内容识别装置用于实现前述实施例一、实施例二中相应的图标内容识别方法,并具有与方法实施例相应的有益效果,在此不再赘述。

[0094] 实施例四

[0095] 参照图4,示出了本发明实施例四的一种用于图标内容识别的终端的结构框图。

[0096] 本发明实施例的终端可以包括:存储器、处理器及存储在存储器上并可在处理器上运行的图标内容识别程序,图标内容识别程序被处理器执行时实现本发明中所述的任意一种图标内容识别方法的步骤。

[0097] 图4是根据一示例性实施例示出的一种卷积神经网络训练终端600的框图。例如,终端600可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等。

[0098] 参照图4,终端600可以包括以下一个或多个组件:处理组件602,存储器604,电源组件606,多媒体组件608,音频组件610,输入/输出(I/O)的接口612,传感器组件614,以及通信组件616。

[0099] 处理组件602通常控制装置600的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件602可以包括一个或多个处理器620来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件602可以包括一个或多个模块,便于处理组件602和其他组件之间的交互。例如,处理部件602可以包括多媒体模块,以方便多媒体组件608和处理组件602之间的交互。

[0100] 存储器604被配置为存储各种类型的数据以支持在终端600的操作。这些数据的

示例包括用于在终端600上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器604可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0101] 电源组件606为终端600的各种组件提供电力。电源组件606可以包括电源管理系统,一个或多个电源,及其他与为终端600生成、管理和分配电力相关联的组件。

[0102] 多媒体组件608包括在所述终端600和用户之间的提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件608包括一个前置摄像头和/或后置摄像头。当终端600处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0103] 音频组件610被配置为输出和/或输入音频信号。例如,音频组件610包括一个麦克风(MIC),当终端600处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器604或经由通信组件616发送。在一些实施例中,音频组件610还包括一个扬声器,用于输出音频信号。

[0104] I/O接口612为处理组件602和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0105] 传感器组件614包括一个或多个传感器,用于为终端600提供各个方面的状态评估。例如,传感器组件614可以检测到终端600的打开/关闭状态,组件的相对定位,例如所述组件为终端600的显示器和小键盘,传感器组件614还可以检测终端600或终端600一个组件的位置改变,用户与终端600接触的存在或不存在,装置600方位或加速/减速和终端600的温度变化。传感器组件614可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件614还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件614还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0106] 通信组件616被配置为便于终端600和其他设备之间有线或无线方式的通信。终端600可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信部件616经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信部件616还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0107] 在示例性实施例中,终端600可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行卷积神经网络训练方

法,具体地图像内容识别方法包括:在对卷积神经网络进行训练的过程中,向卷积神经网络中输入样本图像,其中,所述样本图像用于对所述卷积神经网络进行迭代训练;确定对所述卷积神经网络的已训练迭代次数;基于所述已训练迭代次数,调节损失函数得到目标损失函数;依据所述目标损失函数进行迭代训练,得到目标卷积神经网络;通过所述目标卷积神经网络,对待识别图像进行内容识别。

[0108] 优选地,所述基于所述已训练迭代次数,调节预设损失函数得到目标损失函数的步骤,包括:提取预设损失函数,判断所述已训练迭代次数是否大于第一预设次数;若否,将所述预设损失函数中的超参数调节为0,得到目标损失函数;若是,将所述预设损失函数中的超参数调节为预设值,得到目标损失函数。

[0109] 优选地,所述预设损失函数如下:

$$[0110] \quad \text{sinFocalLoss} = -(1-p_t)^{\gamma \sin(2\pi * \text{clip}(s-i, 0, i/2) / i)} \log(p_t)$$

[0111] 其中, p_t 为概率值, γ 为超参数, i 为迭代次数上限值, s 为已训练迭代次数;

$$[0112] \quad \text{clip}(s-i, 0, i/2) = \begin{cases} 0 & \text{if } s-i < 0 \\ i/2 & \text{if } s-i > i/2 \\ s-i & \text{if } 0 \leq s-i \leq i/2 \end{cases} .$$

[0113] 优选地,所述基于所述已训练迭代次数,调节预设损失函数得到目标损失函数的步骤,包括:确定迭代次数上限值;将所述迭代次数上限值和所述已训练迭代次数,代入所述预设损失函数中,得到目标损失函数。

[0114] 优选地,所述依据所述目标损失函数进行一次迭代训练的步骤,包括:通过所述卷积神经网络确定所述样本图像对应的特征图;将所述特征图进行平均池化,对平均池化后的特征图进行降维处理,得到特征向量;其中,所述第一特征向量中包括多个点,每个点对应一个所述卷积神经网络中的分类标签以及一个概率值;基于所述目标损失函数计算所述卷积神经网络的平均损失值;计算所述目标损失函数在所述特征向量中各点的偏导数得到梯度值,依据所述梯度值对所述卷积神经网络对应的模型参数进行更新。

[0115] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器604,上述指令可由终端600的处理器620执行以完成上述卷积神经网络训练方法。例如,所述非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。当存储介质中的指令由终端的处理器执行时,使得终端能够执行本发明中所述的任意一种卷积神经网络训练方法的步骤。

[0116] 本发明实施例提供的终端在依据输入的样本图像对卷积神经网络进行迭代训练时,基于迭代次数动态调整每次用于迭代训练的目标损失函数,能够更好地拟合到复杂图像样本的分布,减小中间概率值分布的样本图像个数,从而在保证卷积神经网络识别结果准确率的情况下,增加样本的召回率。

[0117] 对于装置实施例而言,由于其与方法实施例基本相似,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0118] 在此提供的图像内容识别方案不与任何特定计算机、虚拟系统或者其他设备固有相关。各种通用系统也可以与基于在此的示教一起使用。根据上面的描述,构造具有本发明方案的系统所要求的结构是显而易见的。此外,本发明也不针对任何特定编程语言。

应当明白,可以利用各种编程语言实现在此描述的本发明的内容,并且上面对特定语言所做的描述是为了披露本发明的最佳实施方式。

[0119] 在此处所提供的说明书中,说明了大量具体细节。然而,能够理解,本发明的实施例可以在没有这些具体细节的情况下实践。在一些实施例中,并未详细示出公知的方法、结构和技术,以便不模糊对本说明书的理解。

[0120] 类似地,应当理解,为了精简本公开并帮助理解各个发明方面中的一个或多个,在上面对本发明的示例性实施例的描述中,本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而,并不应将该公开的方法解释成反映如下意图:即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说,如权利要求书所反映的那样,发明方面在于少于前面公开的单个实施例的所有特征。因此,遵循具体实施方式的权利要求书由此明确地并入该具体实施方式,其中每个权利要求本身都作为本发明的单独实施例。

[0121] 本领域那些技术人员可以理解,可以对实施例中的设备中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个设备中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件,以及此外可以把它们分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外,可以采用任何组合对本说明书(包括伴随的权利要求、摘要和附图)中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述,本说明书(包括伴随的权利要求、摘要和附图)中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

[0122] 此外,本领域的技术人员能够理解,尽管在此所述的一些实施例包括其它实施例中包括的某些特征而不是其它特征,但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如,在权利要求书中,所要求保护的实施例的任意之一都可以以任意的组合方式使用。

[0123] 本发明的各个部件实施例可以以硬件实现,或者以在一个或者多个处理器上运行的软件模块实现,或者以它们的组合实现。本领域的技术人员应当理解,可以在实践中使用微处理器或者数字信号处理器(DSP)来实现根据本发明实施例的图像内容识别方案中的一些或者全部部件的一些或者全部功能。本发明还可以实现为用于执行这里所描述的方法的一部分或者全部的设备或者装置程序(例如,计算机程序和计算机程序产品)。这样的实现本发明的程序可以存储在计算机可读介质上,或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下载得到,或者在载体信号上提供,或者以任何其他形式提供。

[0124] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制,并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中,这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解

释为名称。

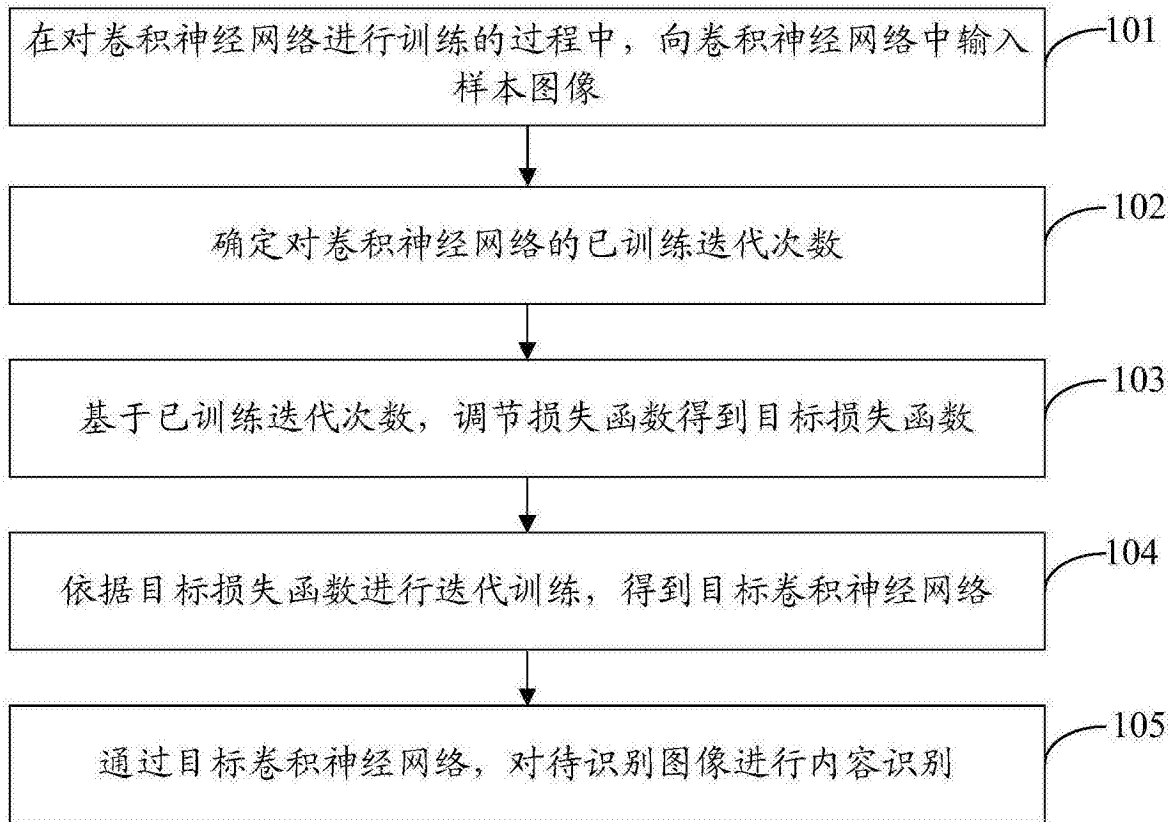


图1

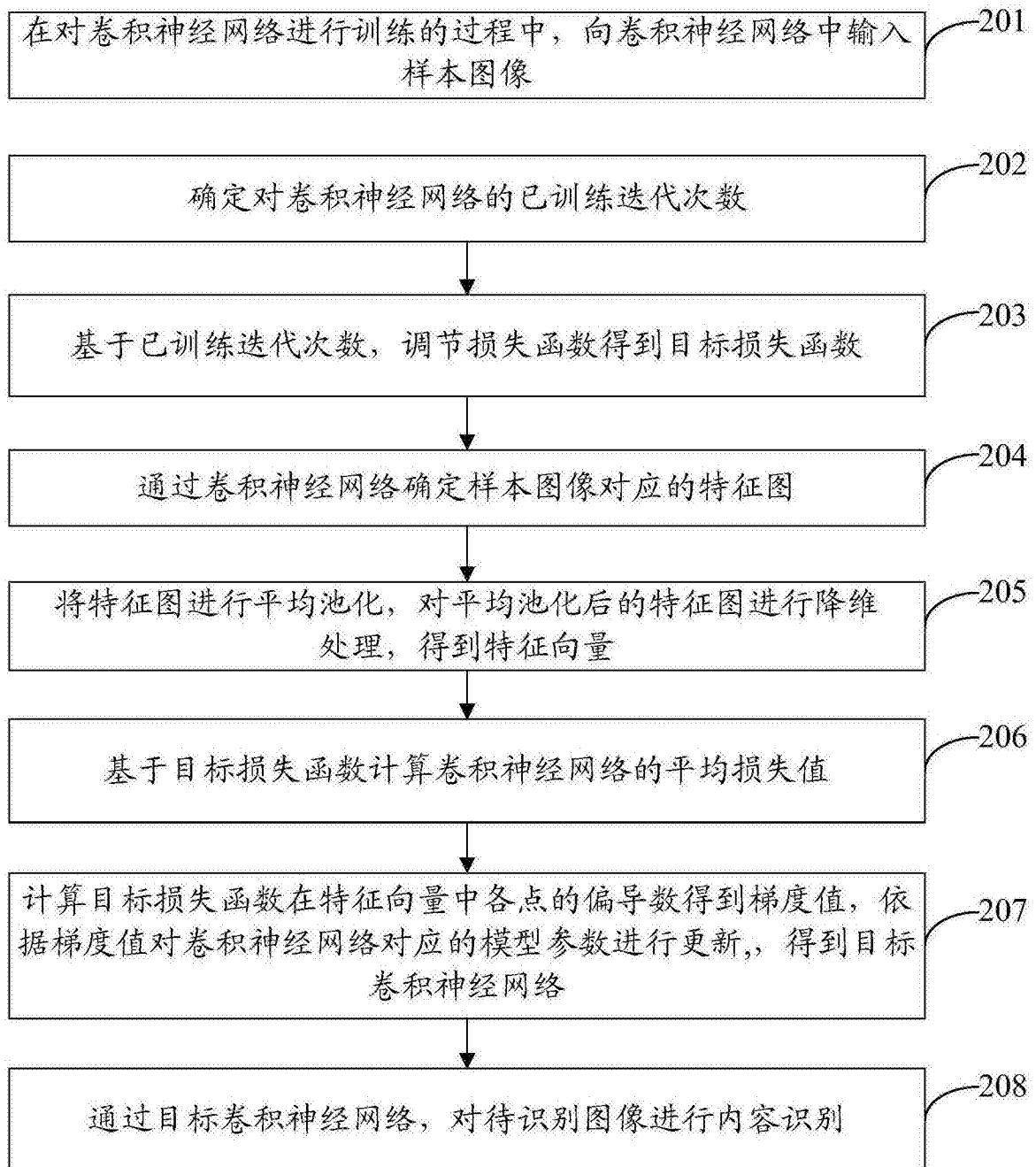


图2

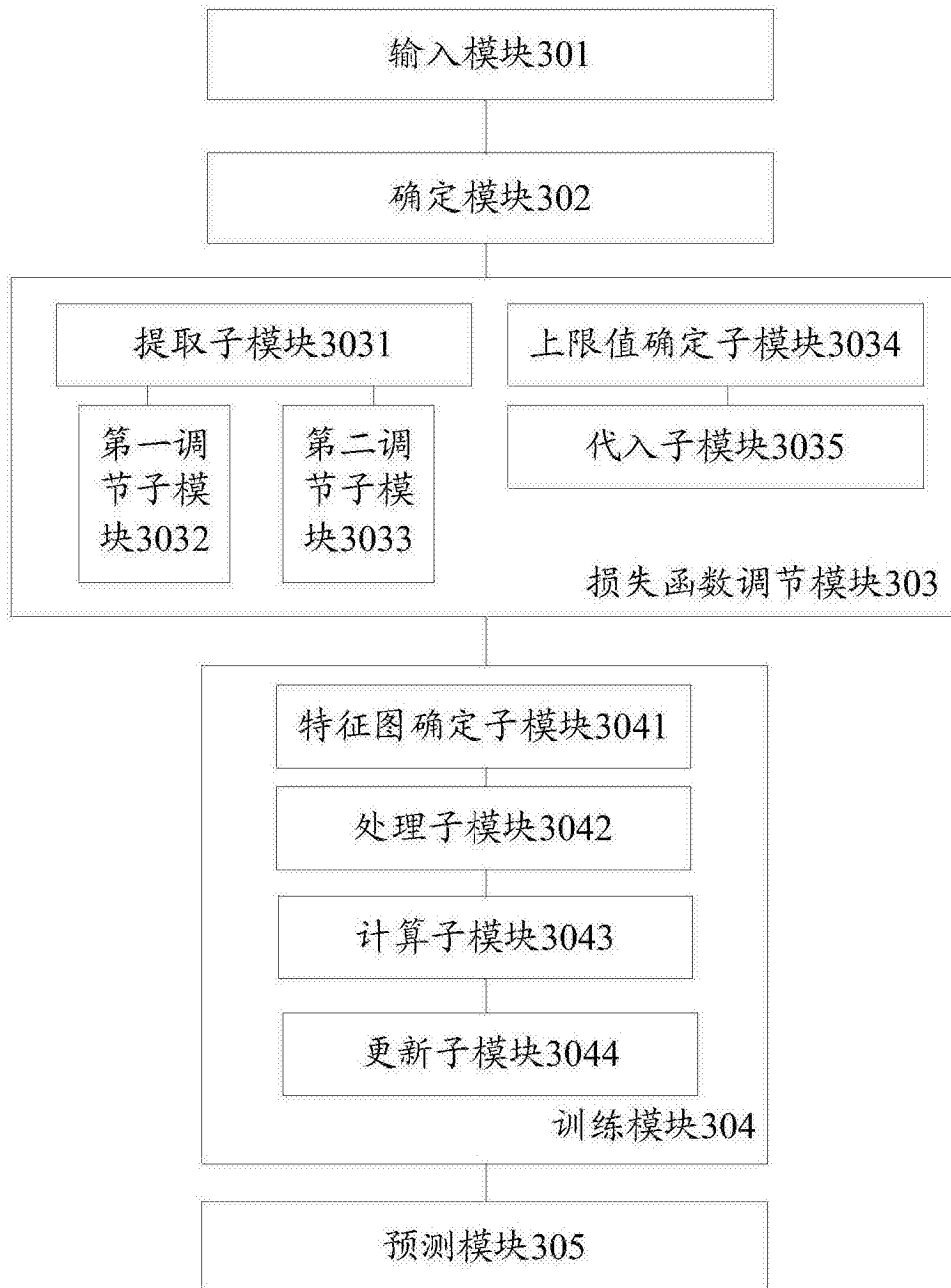


图3

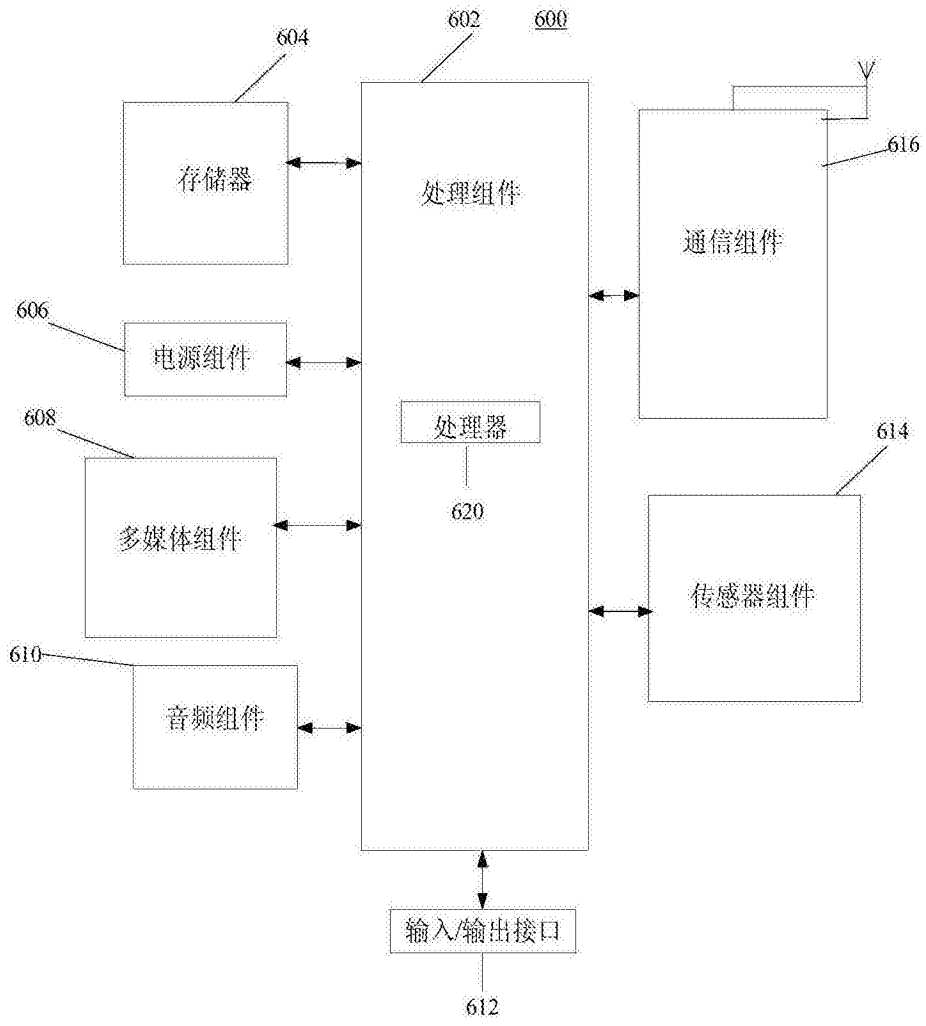


图4