(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 117114148 B
(45)授权公告日 2024.04.09

(21)申请号 202311046071.0
(22)申请日 2023.08.18
(65)同一申请的已公布的文献号
申请公布号 CN 117114148 A
(43)申请公布日 2023.11.24
(73)专利权人 湖南工商大学
地址 410205 湖南省长沙市岳麓区岳麓大道569号
(72)发明人 梁伟 石家帅 黄素珍 周晓康
(74)专利代理机构 长沙轩荣专利代理有限公司 43235
专利代理师 李崇章
(51)Int.Cl.
G06N 20/20(2019.01)
G06N 3/0464(2023.01)
G06N 3/098(2023.01)
G06N 3/082(2023.01)

(56)对比文件
CN 113505210 A,2021.10.15
CN 109886397 A,2019.06.14
CN 109389043 A,2019.02.26
CN 113205863 A,2021.08.03
CN 113705712 A,2021.11.26
CN 114154643 A,2022.03.08
CN 114547315 A,2022.05.27
CN 114663791 A,2022.06.24
CN 114882582 A,2022.08.09
CN 115018039 A,2022.09.06
CN 115272738 A,2022.11.01
CN 115358419 A,2022.11.18
CN 115511108 A,2022.12.23

审查员 张钰柔

权利要求书2页 说明书9页 附图2页

(54)发明名称

一种轻量级联邦学习训练方法

(57)摘要

本发明提供了一种轻量级联邦学习训练方法,包括:中央服务器将深度卷积神经网络模型的参数进行初始化,得到初始化深度卷积神经网络模型并传输至多个客户端;客户端将初始化深度卷积神经网络模型的模型反向蒸馏至本地深度卷积神经网络模型;并将本地图像数据输入本地深度卷积神经网络模型对本地深度卷积神经网络的参数进行更新,得到训练后的本地深度卷积神经网络模型;通过剪枝算法对训练后的本地深度卷积神经网络模型进行剪枝,得到轻量化深度卷积神经网络模型并正向蒸馏至本地局部模型;将本地局部模型输入中央服务器进行聚合,得到全局模型;与现有技术相比,本发明能够在提高通信和聚合效率的同时提升模型的精确性。

CN 117114148 B

1.一种轻量级联邦学习训练方法,其特征在于,包括:

步骤1,中央服务器将深度卷积神经网络模型的参数进行初始化,得到初始化深度卷积神经网络模型,并将所述初始化深度卷积神经网络模型传输至多个客户端;

步骤2,所述客户端通过设定蒸馏温度,将所述初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型;

步骤3,所述客户端将采集的本地图像数据输入所述本地深度卷积神经网络模型进行图像识别,得到识别结果并计算损失函数,通过所述损失函数对所述本地深度卷积神经网络的参数进行更新,得到训练后的本地深度卷积神经网络模型;

步骤4,所述客户端通过剪枝算法对训练后的本地深度卷积神经网络模型中的编码器和分类器分别进行剪枝,得到剪枝后的编码器和剪枝后的分类器,并将剪枝后的编码器和剪枝后的分类器进行拼接,得到轻量化深度卷积神经网络模型;

根据所述训练后的本地深度卷积神经网络模型的网络结构,将所述训练后的本地深度卷积神经网络模型分为编码器和分类器;

利用结构化剪枝的方式,评估所述编码器中每个卷积层中每个过滤器的影响系数,并所述影响系数低于预设值的过滤器进行修剪,得到剪枝后的编码器;

利用非结构化剪枝的方式,将所述分类器的权重绝对值小于预设阈值的权重进行修剪,得到剪枝后的分类器;

将所述剪枝后的编码器和所述剪枝后的分类器进行拼接,得到轻量化深度卷积神经网络模型;

步骤5,所述客户端通过设定蒸馏温度,将所述轻量化深度卷积神经网络模型的参数正向知识蒸馏至本地局部模型,并将所述本地局部模型输入所述中央服务器;

步骤6,所述中央服务器将多个所述客户端上传的本地局部模型进行聚合,得到全局模型,并判断所述全局模型是否满足预设训练条件;若是,则训练结束,将待识别的图像数据输入所述全局模型进行图像识别,得到识别结果;否则,将所述全局模型作为所述步骤1中的初始化深度卷积神经网络模型传输至多个客户端,并返回执行步骤2。

2.根据权利要求1所述的轻量级联邦学习训练方法,其特征在于,在所述客户端将采集的本地图像数据输入所述本地深度卷积神经网络模型进行图像识别之前,还包括:

对采集的本地图像数据进行数据标签规范化处理和异常数据删除处理,得到处理后的本地图像数据;

所述客户端将采集的本地图像数据输入所述本地深度卷积神经网络模型进行图像识别。

3.根据权利要求2所述的轻量级联邦学习训练方法,其特征在于,

根据所述训练后的本地深度卷积神经网络模型的网络特性,通过对所述训练后的本地深度卷积神经网络模型进行正则化,得到编码器和分类器,正则化的表达式为:

$$R(W) = R_{Enc}(W_E) + R_{Cls}(W_C)$$

$$R_{Enc}(W_E) = \sum_{l=1}^{W_E}\left(\sum_{f_l=1}^{F_l}\left\|W^{(l)}_{E\,f_{l,:,:,:}}\right\|_g + \sum_{ch_l=1}^{Ch_l}\left\|W^{(l)}_{E\,ch_{l,:,:,:}}\right\|_g\right)$$

$$R_{Cls}(W_C)=\sum_{l=1}^{W_C}\left(\sum_{row_l=1}^{Row_l}\left\|W^{(l)}_{C\,row_{l,:}}\right\|_g+\sum_{col_l=1}^{Col_l}\left\|W^{(l)}_{C\,col_{l,:}}\right\|_g\right)$$

其中,R(W)表示本地深度卷积神经网络模型的剪枝权重,R_{Enc}表示编码器的剪枝权重,R_{Cls}表示分类器的剪枝权重,W_E表示编码器的权重,W_C表示分类器的权重,$||\cdot||_g$是group Lasso算法,F_l是第l个卷积层中滤波器的数量,Ch_l是第l个卷积层中通道的个数,Row_l代表分类器中第l层的行数,Col_l代表分类器中第l层的列数。

4.根据权利要求3所述的轻量级联邦学习训练方法,其特征在于,所述轻量化深度卷积神经网络模型的损失函数为:

$F(W)=F_D(W)+\lambda R(W)$

其中,$F_D(W)$是轻量化深度卷积神经网络模型的损失函数,λ是结构化稀疏正则化的系数。

5.根据权利要求4所述的轻量级联邦学习训练方法,其特征在于,所述所述本地局部模型的损失函数为:

$$L_m=\beta L_{C_m}+(1-\beta)D_{KL}(p_l\|p_m)$$

其中,β表示控制来自数据或其他模型知识比例的超参数,L_{C_m}表示本地局部模型的交叉熵损失函数,D_{KL}表示KL散度,p_l表示本地深度卷积神经网络模型的预测值,p_m表示本地局部模型的预测值。

6.根据权利要求5所述的轻量级联邦学习训练方法,其特征在于,所述全局模型训练终止的条件为:

直至所述全局模型的精度达到预设训练精度或迭代次数达到预设上限时,终止训练。

一种轻量级联邦学习训练方法

技术领域

[0001] 本发明涉及信息技术领域,特别涉及一种轻量级联邦学习训练方法。

背景技术

[0002] 随着移动设备的功能越来越强大,越来越多的基于神经网络的智能应用已被开发用于移动设备,例如图像识别、视频分析、目标检测等。为了使智能应用能够达到预计效果,通常会通过大量的数据训练智能应用的神经网络模型,然而,单个移动设备的数据量是有限的,不太可能帮助神经网络达到理想的精度。同时,考虑到隐私保护和通信量过大等原因,将数据从许多移动设备传输到一个中央服务器并进行集中训练将不再可行。在联邦学习中的中央服务器的编排下,以分散的方式训练共享全局模型,实现在保护用户数据隐私的同时,最大化提升模型的训练效率和模型的整体精度。

[0003] 目前,由于联邦学习在解决隐私保护和数据孤岛等问题方面的优势,已经逐步成为流行的机器学习范式。此类方法通常分为四个步骤:首先,在每轮通信中,每个参与设备从中央服务器下载当前模型;其次,通过本地数据训练局部模型;第三,通过中央服务器聚合所有局部模型;第四,将聚合后的全局模型发送回设备。然而,由于移动设备通信成本高且通信传输不稳定,联邦学习通信负载较大等问题,常规的联邦学习方法难以在一定设备尤其是告诉移动设备中使用。因此,目前的面向移动设备的联邦学习方法却存在以下不可忽略的技术问题:

[0004] 传统的联邦学习训练方法主要考虑的是稳定通信的设备或者是慢速的移动设备,从而忽略了联邦学习算法应用在高速移动设备上的挑战。在高速移动场景下,例如高速车联网中,车辆的高速移动性带来了信号质量的下降,导致车载网络无法实现最佳带宽和通信速度,这意味着参与训练的设备将会消耗大量的时间和资源在模型的传输过程中。同时,由于不同设备的网络时延不同,中央服务器的聚合过程将会导致更长的等待时间,这将导致联邦学习的效率进一步降低,这些问题严重影响了传统联邦学习在移动场景下的应用效果。

发明内容

[0005] 本发明提供了一种轻量级联邦学习训练方法,其目的是为了节约模型传输过程中的传输时间和减少模型聚合过程中的等待时间。

[0006] 为了达到上述目的,本发明提供了一种轻量级联邦学习训练方法,包括:

[0007] 步骤1,中央服务器将深度卷积神经网络模型的参数进行初始化,得到初始化深度卷积神经网络模型,并将初始化深度卷积神经网络模型传输至多个客户端;

[0008] 步骤2,客户端通过设定蒸馏温度,将初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型;

[0009] 步骤3,客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别,得到识别结果并计算损失函数,通过损失函数对本地深度卷积神经网络的参数进行

更新,得到训练后的本地深度卷积神经网络模型;

[0010] 步骤4,客户端通过剪枝算法对训练后的本地深度卷积神经网络模型中的编码器和分类器分别进行剪枝,得到剪枝后的编码器和剪枝后的分类器,并将剪枝后的编码器和剪枝后的分类器进行拼接,得到轻量化深度卷积神经网络模型;

[0011] 步骤5,客户端通过设定蒸馏温度,将轻量化深度卷积神经网络模型的参数正向知识蒸馏至本地局部模型,并将本地局部模型输入中央服务器;

[0012] 步骤6,中央服务器将多个客户端上传的本地局部模型进行聚合,得到全局模型,并判断全局模型是否满足预设训练条件;若是,则训练结束,将待识别的图像数据输入全局模型进行图像识别,得到识别结果;否则,将全局模型作为步骤1中的初始化深度卷积神经网络模型传输至多个客户端,并返回执行步骤2。

[0013] 进一步来说,在客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别之前,还包括:

[0014] 对采集的本地图像数据进行数据标签规范化处理和异常数据删除处理,得到处理后的本地图像数据;

[0015] 客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别。

[0016] 进一步来说,步骤4包括:

[0017] 根据训练后的本地深度卷积神经网络模型的网络特性,将训练后的本地深度卷积神经网络模型分为编码器和分类器;

[0018] 利用结构化剪枝的方式,将编码器的权重绝对值小于预设阈值的权重进行修剪,得到剪枝后的编码器;

[0019] 利用非结构化剪枝的方式,评估分类器中每个卷积层中每个过滤器的影响系数,并影响系数低于预设值的过滤器进行修剪,得到剪枝后的分类器;

[0020] 将剪枝后的编码器和剪枝后的分类器进行拼接,得到轻量化深度卷积神经网络模型。

[0021] 进一步来说,根据训练后的本地深度卷积神经网络模型的网络特性,通过对训练后的本地深度卷积神经网络模型进行正则化,得到编码器和分类器,正则化的表达式为:

[0022] $R(W)=R_{Enc}(W_E)+R_{Cls}(W_C)$

[0023] $$R_{Enc}(W_E)=\sum_{l=1}^{W_E}\left(\sum_{f_l=1}^{F_l}\left\|W^{(l)}_{E\,f_{l,:,:,:}}\right\|_g+\sum_{ch_l=1}^{Ch_l}\left\|W^{(l)}_{E\,ch_{l,:,:,:}}\right\|_g\right)$$

[0024] $$R_{Cls}(W_C)=\sum_{l=1}^{W_C}\left(\sum_{row_l=1}^{Row_l}\left\|W^{(l)}_{C\,row_{l,:}}\right\|_g+\sum_{col_l=1}^{Col_l}\left\|W^{(l)}_{C\,col_{l,:}}\right\|_g\right)$$

[0025] 其中,$R(W)$表示本地深度卷积神经网络模型的剪枝权重,R_{Enc}表示编码器的剪枝权重,R_{Cls}表示分类器的剪枝权重,W_E表示编码器的权重,W_C表示分类器的权重,$||\cdot||_g$是group Lasso算法,F_l是第l个卷积层中滤波器的数量,Ch_l是第l个卷积层中通道的个数,Row_l代表分类器中第l层的行数,Col_l代表分类器中第l层的列数。

[0026] 进一步来说,轻量化深度卷积神经网络模型的损失函数为:

[0027] $F(W) = F_D(W) + \lambda R(W)$

[0028] 其中，$F_D(W)$是轻量化深度卷积神经网络模型的损失函数，λ是结构化稀疏正则化的系数。

[0029] 进一步来说，本地局部模型的损失函数为：

[0030] $$L_m = \beta L_{C_m} + (1-\beta) D_{KL}(p_l \| p_m)$$

[0031] 其中，β表示控制来自数据或其他模型知识比例的超参数，L_{C_m}表示本地局部模型的交叉熵损失函数，D_{KL}表示KL散度，p_l表示本地深度卷积神经网络模型的预测值，p_m表示本地局部模型的预测值。

[0032] 进一步来说，训练终止的条件为：

[0033] 直至全局模型的精度达到预设训练精度或迭代次数达到预设上限时，终止训练。

[0034] 本发明的上述方案有如下的有益效果：

[0035] 本发明通过中央服务器将深度卷积神经网络模型的参数进行初始化，得到初始化深度卷积神经网络模型，并将初始化深度卷积神经网络模型传输至多个客户端；客户端通过设定蒸馏温度，将初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型；客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别，得到识别结果并计算损失函数，通过损失函数对本地深度卷积神经网络的参数进行更新，得到训练后的本地深度卷积神经网络模型；客户端通过剪枝算法对训练后的本地深度卷积神经网络模型中的编码器和分类器分别进行剪枝，得到剪枝后的编码器和剪枝后的分类器，并将剪枝后的编码器和剪枝后的分类器进行拼接，得到轻量化深度卷积神经网络模型；客户端通过设定蒸馏温度，将轻量化深度卷积神经网络模型的参数正向知识蒸馏至本地局部模型，并将本地局部模型输入中央服务器；中央服务器将多个客户端上传的本地局部模型进行聚合，得到全局模型，并判断全局模型是否满足预设训练条件；若是，则训练结束，将待识别的图像数据输入全局模型进行图像识别，得到识别结果；否则，将全局模型作为步骤1中的初始化深度卷积神经网络模型传输至多个客户端，并返回执行步骤2；与现有技术相比，本发明采用双向蒸馏的方式压缩模型的参数，极大程度上提高了通信效率并减少了聚合时的等待时间，通过剪枝算法对模型做进一步的压缩，有效的去除了局部模型中多余的参数以减少模型参数量，从而能够在提高通信和聚合效率的同时提升模型的精确性。

[0036] 本发明的其它有益效果将在随后的具体实施方式部分予以详细说明。

附图说明

[0037] 图1为本发明实施例的流程示意图；

[0038] 图2为本发明实施例中轻量级联邦学习训练框架示意图。

具体实施方式

[0039] 为使本发明要解决的技术问题、技术方案和优点更加清楚，下面将结合附图及具体实施例进行详细描述。显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的

所有其他实施例,都属于本发明保护的范围。

[0040] 在本发明的描述中,需要说明的是,术语“中心”、“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本发明和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。此外,术语“第一”、“第二”、“第三”仅用于描述目的,而不能理解为指示或暗示相对重要性。

[0041] 在本发明的描述中,需要说明的是,除非另有明确的规定和限定,术语“安装”、“相连”、“连接”应做广义理解,例如,可以是锁定连接,也可以是可拆卸连接,或一体地连接;可以是机械连接,也可以是电连接;可以是直接相连,也可以通过中间媒介间接相连,可以是两个元件内部的连通。对于本领域的普通技术人员而言,可以具体情况理解上述术语在本发明中的具体含义。

[0042] 此外,下面所描述的本发明不同实施方式中所涉及的技术特征只要彼此之间未构成冲突就可以相互结合。

[0043] 本发明针对现有的问题,提供了一种轻量级联邦学习训练方法。

[0044] 如图1所示,本发明的实施例提供了一种轻量级联邦学习训练方法,包括:

[0045] 步骤1,中央服务器将深度卷积神经网络模型的参数进行初始化,得到初始化深度卷积神经网络模型,并将初始化深度卷积神经网络模型传输至多个客户端;

[0046] 步骤2,客户端通过设定蒸馏温度,将初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型;

[0047] 步骤3,客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别,得到识别结果并计算损失函数,通过损失函数对本地深度卷积神经网络的参数进行更新,得到训练后的本地深度卷积神经网络模型;

[0048] 步骤4,客户端通过剪枝算法对训练后的本地深度卷积神经网络模型中的编码器和分类器分别进行剪枝,得到剪枝后的编码器和剪枝后的分类器,并将剪枝后的编码器和剪枝后的分类器进行拼接,得到轻量化深度卷积神经网络模型;

[0049] 步骤5,客户端通过设定蒸馏温度,将轻量化深度卷积神经网络模型的参数正向知识蒸馏至本地局部模型,并将本地局部模型输入中央服务器;

[0050] 步骤6,中央服务器将多个客户端上传的本地局部模型进行聚合,得到全局模型,并判断全局模型是否满足预设训练条件;若是,则训练结束,将待识别的图像数据输入全局模型进行图像识别,得到识别结果;否则,将全局模型作为步骤1中的初始化深度卷积神经网络模型传输至多个客户端,并返回执行步骤2。

[0051] 具体来说,基于移动设备的数据质量、处理器性能、通信质量等因素,选择多个高质量的客户端与中央服务器建立联系并加入联邦学习的训练过程;中央服务器对深度卷积神经网络模型的参数进行初始化,并通过无线网络将初始化深度卷积神经网络模型分别传输至相应的客户端,初始化深度卷积神经网络模型由19个卷积层、5个池化层、3个全连接层和softmax层组成。

[0052] 需要说明的是,本发明实施例中所提到的客户端搭载于具备摄像功能的网联汽车上,网联汽车通过采集模块采集道路图像数据存储至客户端。

[0053] 具体来说,在客户端设定合适的蒸馏温度,将初始化深度卷积神经网络模型参数

反向蒸馏至本地的深度卷积神经网络模型中;联邦学习初始化阶段通过反向蒸馏的方式间接的将本地深度卷积神经网络模型初始化,进而使得整个联邦学习过程加快。

[0054] 具体来说,客户端将采集的本地道路图像数据输入本地深度卷积神经网络模型进行图像识别,得到识别结果和每张图像对应的标签值结果,然后反向传播通过导数链式法则计算损失函数对各参数的梯度,并根据梯度进行参数的更新,得到训练后的本地深度卷积神经网络模型。

[0055] 具体来说,在客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别之前,还包括:

[0056] 对采集的本地图像数据进行数据标签规范化处理和异常数据删除处理,得到处理后的本地图像数据;

[0057] 客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别。

[0058] 本发明实施例以客户端i为例,通过客户端i采集本地图像数据;对本地图像数据进行数据标签规范化处理和异常数据删除处理,得到处理后的本地图像数据;将处理后的本地图像数据输入本地深度卷积神经网络模型进行图像识别,得到识别结果并计算损失函数,通过损失函数对本地深度卷积神经网络的参数进行更新,得到训练后的本地深度卷积神经网络模型$local_i$。

[0059] 具体来说,步骤4包括:

[0060] 根据训练后的本地深度卷积神经网络模型$local_i$的网络结构,将训练后的本地深度卷积神经网络模型$local_i$分为编码器Encoder和分类器Classifier,如图2所示;其中编码器Encoder由卷积神经网络CNN组成,分类器Classifier由全连接神经网络组成;将模型分为编码器和分类器再分别进行剪枝是根据编码器和分类器在网络中的角色和性质不同。编码器中的滤波器负责提取局部特征,对于图像的不同部分有不同的响应。而分类器中的全连接层负责整合卷积层提取的特征,对于整体任务的影响较大,全连接层剪枝时需要保留对任务性能影响较大的神经元。通过分别考虑滤波器和全连接层的剪枝,以求得最大限度地对模型进行压缩,减少计算复杂度,同时保持模型的性能。

[0061] 利用结构化剪枝的方式,评估编码器中每个卷积层中每个过滤器的影响系数,并影响系数低于预设值的过滤器进行修剪,得到剪枝后的编码器Encoder;

[0062] 利用非结构化剪枝的方式,将分类器的权重绝对值小于预设阈值的权重进行修剪,得到剪枝后的分类器Classifier;

[0063] 将剪枝后的编码器Encoder和剪枝后的分类器Classifier进行拼接,得到轻量化深度卷积神经网络模型,轻量化深度卷积神经网络模型包括13个卷积层、5个池化层、3个全连接层以及softmax层组成。

[0064] 本发明实施例提出的剪枝算法用于减小模型大小和通信开销,包括基于结构化剪枝的编码器Encoder修剪方法和基于非结构化剪枝的分类器Classifier修剪方法。

[0065] 非结构化剪枝通常适用于全连接神经网络,根据设定的阈值,将权重绝对值小于阈值的参数定义为不重要的参数直接设为零,具有很高的灵活性;结构化剪枝通常用于卷积神经网络CNN中,通过一些方法评估CNN中每个卷积层过滤器的影响系数,然后将其中影响系数较低的卷积层过滤器移除,该方法虽然灵活性较低,但是能更大程度上压缩模型,基于上述讨论,对于本地深度卷积神经网络模型$local_i$的正则化可以表示为:

[0066] $R(W)=R_{Enc}(W_E)+R_{Cls}(W_C)$

[0067] $$R_{Enc}(W_E)=\sum_{l=1}^{W_E}\left(\sum_{f_l=1}^{F_l}\left\|W^{(l)}_{E\,f_{l,:,:,:}}\right\|_g+\sum_{ch_l=1}^{Ch_l}\left\|W^{(l)}_{E\,ch_{l,:,:,:}}\right\|_g\right)$$

[0068] $$R_{Cls}(W_C)=\sum_{l=1}^{W_C}\left(\sum_{row_l=1}^{Row_l}\left\|W^{(l)}_{C\,row_{l,:}}\right\|_g+\sum_{col_l=1}^{Col_l}\left\|W^{(l)}_{C\,col_{l,:}}\right\|_g\right)$$

[0069] 其中,$R(W)$表示本地深度卷积神经网络模型的剪枝权重,R_{Enc}表示编码器的剪枝权重,R_{Cls}表示分类器的剪枝权重,W_E表示编码器Encoder的权重,W_C表示分类器Classifier的权重,$||\cdot||_g$表示group Lasso分组最小角回归算法,F_1表示第1个卷积层中过滤器的数量,Ch_1表示第1个卷积层中通道的个数,Row_1表示分类器中第1层的行数,Col_1代表分类器中第1层的列数。

[0070] $$\left\|W^l_{Mod}\right\|_g=\sqrt{\sum_{i=1}^{|W^l_{Mod}|}\left(W^l_{Mod_i}\right)^2}$$

[0071] 其中Modules={C:Classifer,E:Encoder},$\left|W^l_{Mod}\right|$代表W^l_{Mod}的参数量。

[0072] 应用上述的正则化方法后,轻量化深度卷积神经网络模型的训练损失函数为:

[0073] $F(W)=F_D(W)+\lambda R(W)$

[0074] 其中,$F_D(W)$是轻量化深度卷积神经网络模型的损失函数,λ是结构化剪枝正则化的系数。

[0075] 通过使用客户端收集的本地图像数据优化轻量化深度卷积神经网络模型中的损失函数,可以识别轻量化深度卷积神经网络模型中的零值和非零值参数。

[0076] 具体来说,终止训练的条件为:直至全局模型的精度达到预设训练精度或迭代次数达到预设上限时,终止训练。

[0077] 本发明实施例通过剪枝算法得到轻量化深度卷积神经网络模型,再使用双向知识蒸馏算法进一步压缩轻量化深度卷积神经网络模型的模型参数,以便于联邦学习过程中知识的上传和下载,首先通过正向知识蒸馏将轻量化深度卷积神经网络模型提取到更紧凑、轻量的本地局部模型中,然后将本地局部模型输入中央服务器进行模型聚合,得到全局模型并传输至各个客户端,用于更新客户端中的初始化深度卷积神经网络模型;客户端将全局模型替换为初始化深度卷积神经网络模型,最后通过反向知识蒸馏将初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型。

[0078] 具体来说,本地局部模型的损失函数为:

[0079] $$L_m=\beta L_{C_m}+(1-\beta)D_{KL}(p_l\|p_m)$$

[0080] 其中,β表示控制来自数据或其他模型知识比例的超参数,L_{C_m}表示本地局部模型的交叉熵损失函数,D_{KL}表示KL散度,p_1表示本地深度卷积神经网络模型的预测值,p_m表示本地局部模型的预测值。

[0081] 具体来说,中央服务器基于FedAug算法对本地局部模型进行加权聚合,得到本轮全局模型,FedAug算法如下所示:

[0082] $$W_{t+1}=\sum_{i=1}^{K}\frac{1}{K}W_{t+1}^{i}$$

[0083] 其中,W_{t+1}^{i}是第t+1轮中第i个客户端的本地局部模型的参数,W_{t+1}是第t+1轮联邦学习中全局模型的参数。

[0084] 具体来说,中央服务器将全局模型传输至每个客户端中,客户端利用接收到的全局模型代替初始化深度卷积神经网络模型作为下一轮的初始化深度卷积神经网络模型,即$W_{t+2}^{i}=W_{t+1},\ \mathrm{i}=\{1,2,\ldots,\mathrm{k}\}$。

[0085] 初始化深度卷积神经网络模型的损失函数为:

[0086] $$L_{l}=\alpha L_{C_{l}}+(1-\alpha)D_{KL}(p_{m}\|p_{l})$$

[0087] 其中,$L_{C_{l}}$是初始化深度卷积神经网络模型的交叉熵损失函数,α是控制来自数据或其他模型知识比例的超参数。

[0088] 本发明实施例通过网联汽车上的采集模块采集道路图像数据,并将道路图像数据输入全局模型进行图像识别,得到识别结果,识别结果包括:道路上存在行人、道路上存在非静止的障碍物、道路上存在静止的障碍物。

[0089] 下面结合具体的实例对本发明实施例所提出的训练方法进行验证,具体如下:

[0090] 本发明实施例利用CIFAR10和MNIST数据集进行测试。CIFAR10由60000张32*32彩色图像组成,图像有10个类,每个类有6000个图像,它分别包含有50000个训练图像和10000个测试图像;MNIST由70000张28*28像素的灰度手写数字图像,图像有10个类,每个类有7000个样本,它分别包含有60000个训练图像和10000个测试图像;具体如表1所示:

[0091] 表1

[0092]

	图像尺寸	图像通道数	图像类数	训练集数量	测试集数量
CIFAR10	32*32	3	10	50000	10000
MNIST	28*28	1	10	60000	10000

[0093] 由于在高速移动场景下,每个客户端中间的数据集常常不满足独立同分布不假设,因此本发明实施例额外采用Dirichlet分布来为每个客户端划分数据集,是每个客户端上样本标签分布不同。

[0094] 表2

[0095]

Method	IID	CIFAR-10				MNIST			
		Acc	Params(M)	CR	TCC	Acc	Params(M)	CR	TCC
FL	0	55.7	133.05	174	185205.6	94.32	133.05	72	76636.8
FL+KD	0	58.2	133.05	152	161788.8	95.93	133.05	67	71314.8
FL	1	60.9	133.05	152	161788.8	95.78	133.05	51	54284.4
FL+KD	1	61.5	133.05	122	129856.8	95.93	133.05	47	50026.8

[0096] 为了评估和验证本发明实施例所训练出的全局模型的性能,本发明实施例首先测算了FL(Federated Learining)、FL+KD(Federated Learining+Knowledge Distillation)分别在IID和Non-IID情况下的通讯开销,采用CR(communication rounds)、TCC(total communication cost)作为通信开销的主要评价指标,根据表2可以得出,本发明实施例所提供的方法在评价指标中都取得了较好的数值测算结果,对于模型的性能表现,本发明实施例使用目前流行的Basic(Centralized Machine Learning,集中式机器学习集中式机器学习)、联邦平均算法(FederatedAveragingAlgorithm,FedAVG)、联邦学习框架FedProx作为基准测试模型,并采用Acc、Precision、Recall、F1作为模型的主要评价指标,结果如表3所示:

[0097] 表3

[0098]

Method	IID	CIFAR-10				MNIST			
		Acc	Precision	Recall	F1	Acc	Precision	Recall	F1
Basic	0	71.2	70.3	72.3	72.1	99.32	99.53	99.39	99.41
FedAVG	0	47.55	48.45	48.19	48.31	94.78	95.17	95.03	95.10
FedProx	0	37.37	38.81	36.73	37.74	99.13	99.27	98.85	99.06
Our	0	65.4	66.5	66.1	66.3	90.3	90.79	90.13	90.46
Basic	1	71.2	70.3	72.3	71.3	99.32	99.53	99.39	99.41
FedAVG	1	53.7	54.7	53.4	54.0	98.01	99.2	98.3	98.7
FedProx	1	48.98	48.11	47.99	48.05	98.96	99.11	99.13	99.12
Our	1	69.7	68.60	70.46	69.52	99.06	99.17	99.13	99.15

[0099] 从上表3可看出,本发明所述方法在评价指标中都取得了较高的性能表现,并超过基准测试(FedAVG、FedProx)模型。

[0100] 本发明实施例中央服务器将深度卷积神经网络模型的参数进行初始化,得到初始化深度卷积神经网络模型,并将初始化深度卷积神经网络模型传输至多个客户端;客户端通过设定蒸馏温度,将初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型;客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别,得到识别结果并计算损失函数,通过损失函数对本地深度卷积神经网络的参数进行更新,得到训练后的本地深度卷积神经网络模型;客户端通过剪枝算法对训练后的本地深度卷积神经网络模型中的编码器和分类器分别进行剪枝,得到剪枝后的编码器和剪枝后的分

类器,并将剪枝后的编码器和剪枝后的分类器进行拼接,得到轻量化深度卷积神经网络模型;客户端通过设定蒸馏温度,将轻量化深度卷积神经网络模型的参数正向知识蒸馏至本地局部模型,并将本地局部模型输入中央服务器;中央服务器将多个客户端上传的本地局部模型进行聚合,得到全局模型,并判断全局模型是否满足预设训练条件;若是,则训练结束,将待识别的图像数据输入全局模型进行图像识别,得到识别结果;否则,将全局模型作为步骤1中的初始化深度卷积神经网络模型传输至多个客户端,并返回执行步骤2;与现有技术相比,本发明采用双向蒸馏的方式压缩模型的参数,极大程度上提高了通信效率并减少了聚合时的等待时间,通过剪枝算法对模型做进一步的压缩,有效的去除了局部模型中多余的参数以减少模型参数量,从而能够在提高通信和聚合效率的同时提升模型的精确性。

[0101] 以上所述是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明所述原理的前提下,还可以作出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

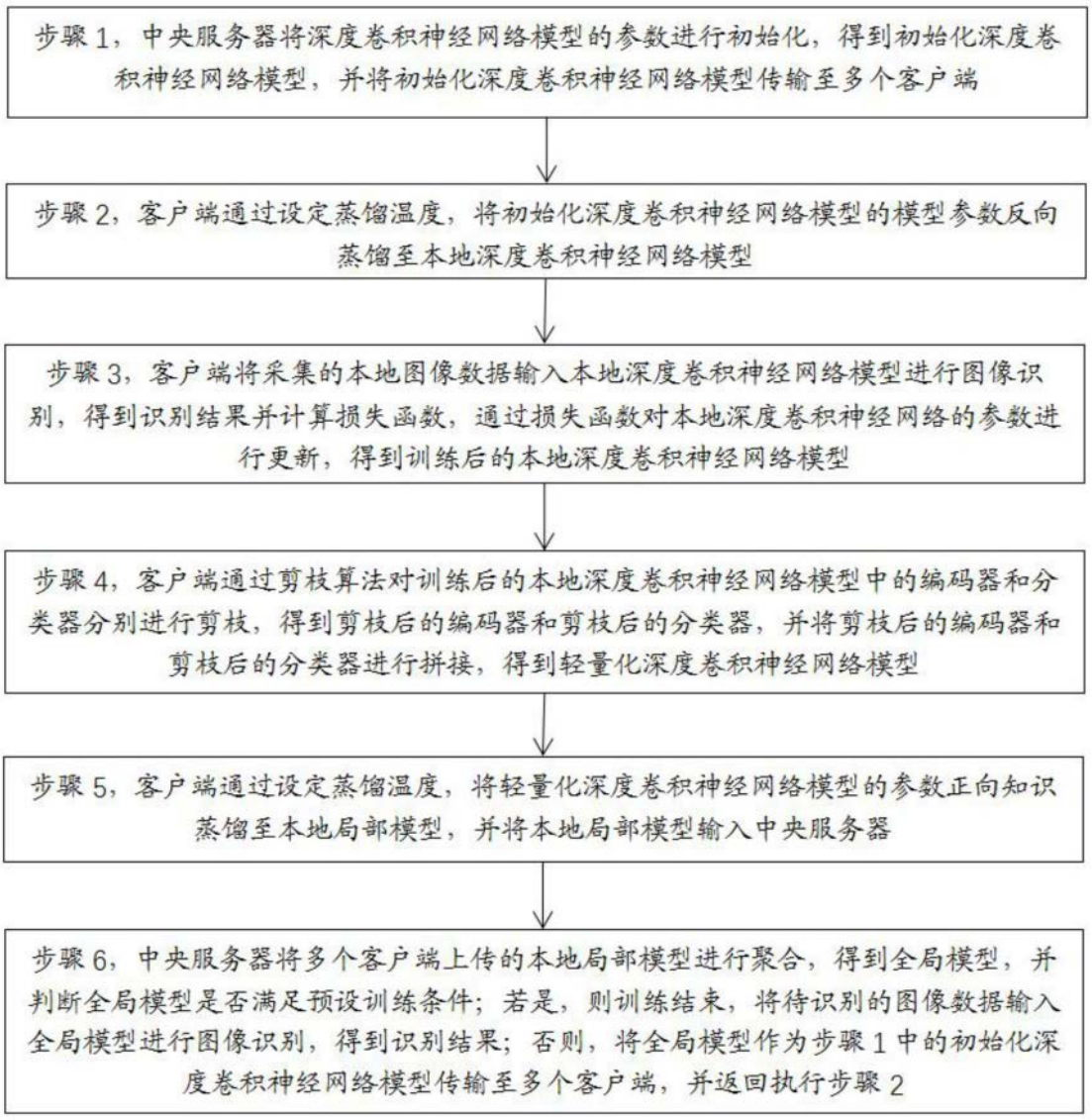
步骤1，中央服务器将深度卷积神经网络模型的参数进行初始化，得到初始化深度卷积神经网络模型，并将初始化深度卷积神经网络模型传输至多个客户端
步骤2，客户端通过设定蒸馏温度，将初始化深度卷积神经网络模型的模型参数反向蒸馏至本地深度卷积神经网络模型
步骤3，客户端将采集的本地图像数据输入本地深度卷积神经网络模型进行图像识别，得到识别结果并计算损失函数，通过损失函数对本地深度卷积神经网络的参数进行更新，得到训练后的本地深度卷积神经网络模型
步骤4，客户端通过剪枝算法对训练后的本地深度卷积神经网络模型中的编码器和分类器分别进行剪枝，得到剪枝后的编码器和剪枝后的分类器，并将剪枝后的编码器和剪枝后的分类器进行拼接，得到轻量化深度卷积神经网络模型
步骤5，客户端通过设定蒸馏温度，将轻量化深度卷积神经网络模型的参数正向知识蒸馏至本地局部模型，并将本地局部模型输入中央服务器
步骤6，中央服务器将多个客户端上传的本地局部模型进行聚合，得到全局模型，并判断全局模型是否满足预设训练条件；若是，则训练结束，将待识别的图像数据输入全局模型进行图像识别，得到识别结果；否则，将全局模型作为步骤1中的初始化深度卷积神经网络模型传输至多个客户端，并返回执行步骤2

图1

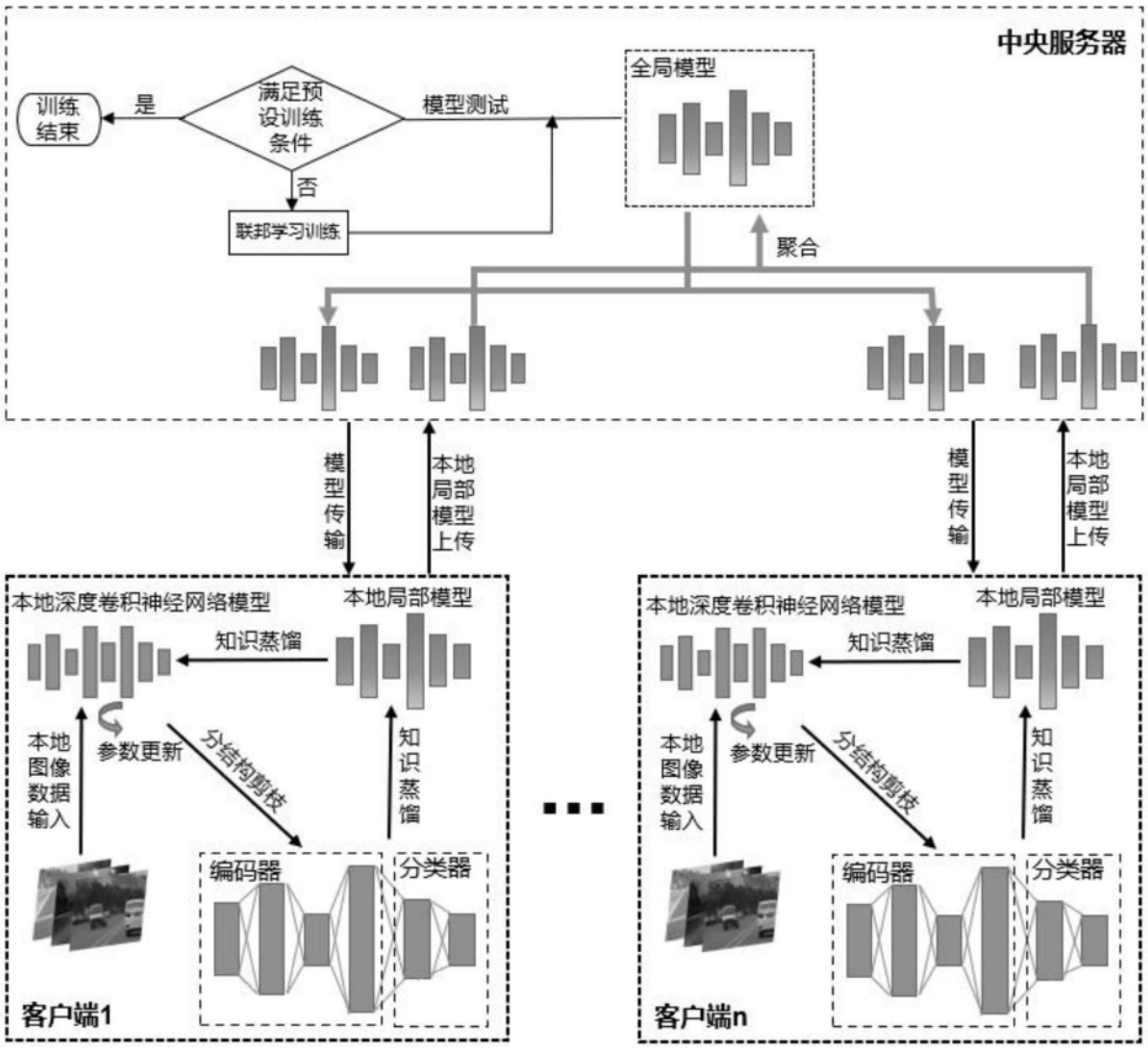
中央服务器
全局模型
训练结束
是
满足预设训练条件
否
模型测试
联邦学习训练
聚合
模型传输
本地局部模型上传
模型传输
本地局部模型上传
本地深度卷积神经网络模型
本地局部模型
知识蒸馏
本地图像数据输入
参数更新
分结构剪枝
知识蒸馏
编码器
分类器
客户端1
本地深度卷积神经网络模型
本地局部模型
知识蒸馏
本地图像数据输入
参数更新
分结构剪枝
知识蒸馏
编码器
分类器
客户端n

图2