

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 10 月 15 日 (15.10.2020)



(10) 国际公布号
WO 2020/207174 A1

(51) 国际专利分类号:
G06N 3/063 (2006.01) *G06N 3/08* (2006.01)

(21) 国际申请号: PCT/CN2020/078586

(22) 国际申请日: 2020 年 3 月 10 日 (10.03.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910288941.2 2019年4月11日 (11.04.2019) CN

(71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街30号院3号楼2层B-0035房间, Beijing 100041 (CN)。

(72) 发明人: 刘阳 (LIU, Yang); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。

(74) 代理人: 北京天达共和律师事务所 (EAST & CONCORD PARTNERS); 中国北京市朝阳区东三环北路8号亮马河大厦写字楼2座19层关刚, Beijing 100004 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,

(54) Title: METHOD AND APPARATUS FOR GENERATING QUANTIZED NEURAL NETWORK

(54) 发明名称: 用于生成量化神经网络的方法和装置

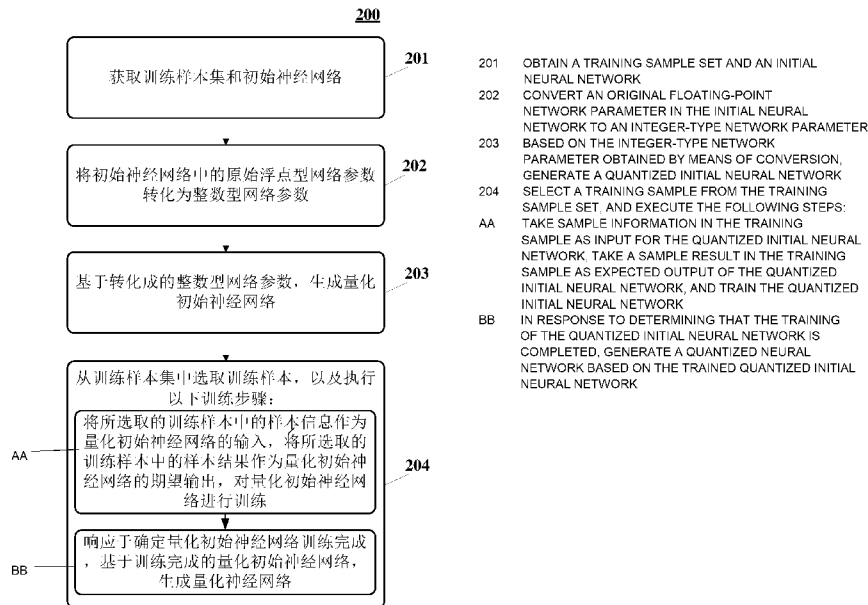


图 2

(57) Abstract: Disclosed are a method and apparatus for generating a quantized neural network. A particular embodiment of the method comprises: obtaining a training sample set and an initial neural network; converting an original floating-point network parameter in the initial neural network to an integer-type network parameter; based on the integer-type network parameter obtained by means of conversion, generating a quantized initial neural network; selecting a training sample from the training sample set, and executing the following steps: taking sample information in the training sample as input for the quantized initial neural network, taking a sample

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

result in the training sample as expected output of the quantized initial neural network, and training the quantized initial neural network; and in response to determining that the training of the quantized initial neural network is completed, generating a quantized neural network based on the trained quantized initial neural network. The embodiment facilitates reduction of a storage space occupied by the neural network, and reduction of CPU consumption when the neural network is used for information processing, such that information processing efficiency is improved.

(57) 摘要: 本公开的实施例公开了用于生成量化神经网络的方法和装置。该方法的一具体实施方式包括: 获取训练样本集和初始神经网络; 将初始神经网络中的原始浮点型网络参数转化为整数型网络参数; 基于转化成的整数型网络参数, 生成量化初始神经网络; 从训练样本集中选取训练样本, 执行以下训练步骤: 将训练样本中的样本信息作为量化初始神经网络的输入, 将训练样本中的样本结果作为量化初始神经网络的期望输出, 对量化初始神经网络进行训练; 响应于确定量化初始神经网络训练完成, 基于训练完成的量化初始神经网络, 生成量化神经网络。该实施方式有助于减小神经网络所占用的存储空间, 以及利用神经网络进行信息处理时对CPU的消耗, 提高信息处理的效率。

用于生成量化神经网络的方法和装置

5 相关申请的交叉引用

本申请基于申请号为 201910288941.2、申请日为 2019 年 04 月 11 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此引入本申请作为参考。

10 技术领域

本公开的实施例涉及计算机技术领域，尤其涉及用于生成量化神经网络的方法和装置。

背景技术

15 目前，为了加速神经网络的训练，通常会在神经网络包括的卷积层后面连接一个 BN（Batch Normalization，批量归一化）层。BN 层用于在将卷积层的输出传递给其他层之前，对卷积层的输出进行归一化处理，以此，提高神经网络的收敛速度。

BN 层包括用于与卷积层的输出相乘，以对卷积层的输出进行归一化处理的归一化参数。卷积层的输出由卷积层的权重与卷积层的输入进行卷积获得。实践中，权重和归一化参数的数据类型通常为浮点型。

发明内容

25 本公开的实施例提出了用于生成量化神经网络的方法和装置。

第一方面，本公开的实施例提供了一种用于生成量化神经网络的方法，该方法包括：获取训练样本集和初始神经网络，其中，训练样本包括样本信息和针对样本信息预先确定的样本结果，初始神经网络包括原始浮点型网络参数，原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化

30

参数的积；将初始神经网络中的原始浮点型网络参数转化为整数型网络参数；基于转化成的整数型网络参数，生成量化初始神经网络；从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

在一些实施例中，基于转化成的整数型网络参数，生成量化初始神经网络，包括：将转化成的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的初始神经网络确定为量化初始神经网络。

在一些实施例中，将初始神经网络中的原始浮点型网络参数转化为整数型网络参数，包括：将原始浮点型网络参数所对应的浮点型权重转化为整数型权重，以及将原始浮点型网络参数所对应的浮点型归一化参数转化为整数型归一化参数；对转化成的整数型权重和整数型归一化参数进行求积，获得整数型网络参数。

在一些实施例中，该方法还包括：响应于确定量化初始神经网络未训练完成，执行以下步骤：从训练样本集包括的未被选取的训练样本中选取训练样本；调整量化初始神经网络的参数，获得新的浮点型网络参数；将新的浮点型网络参数转化为新的整数型网络参数，以及基于新的整数型网络参数，生成新的量化初始神经网络；使用最近一次选取的训练样本和最新生成的量化初始神经网络，继续执行训练步骤。

在一些实施例中，基于新的整数型网络参数，生成新的量化初始神经网络，包括：将新的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的量化初始神经网络确定为新的量化初始神经网络。

在一些实施例中，该方法还包括：将量化神经网络发送给用户终端，以使用户终端对接收的量化神经网络进行存储。

第二方面，本公开的实施例提供了一种用于处理信息的方法，该

方法包括：获取待处理信息和目标量化神经网络，其中，目标量化神经网络为采用如上述第一方面中任一实施例的方法生成的；将待处理信息输入目标量化神经网络，获得处理结果及输出。

第三方面，本公开的实施例提供了一种用于生成量化神经网络的装置，该装置包括：第一获取单元，被配置成获取训练样本集和初始神经网络，其中，训练样本包括样本信息和针对样本信息预先确定的样本结果，初始神经网络包括原始浮点型网络参数，原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积；转化单元，被配置成将初始神经网络中的原始浮点型网络参数转化为整数型网络参数；生成单元，被配置成基于转化成的整数型网络参数，生成量化初始神经网络；第一执行单元，被配置成从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

在一些实施例中，生成单元进一步被配置成：将转化成的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的初始神经网络确定为量化初始神经网络。

在一些实施例中，转化单元包括：转化模块，被配置成将原始浮点型网络参数所对应的浮点型权重转化为整数型权重，以及将原始浮点型网络参数所对应的浮点型归一化参数转化为整数型归一化参数；求积模块，被配置成对转化成的整数型权重和整数型归一化参数进行求积，获得整数型网络参数。

在一些实施例中，该装置还包括：第二执行单元，被配置成响应于确定量化初始神经网络未训练完成，执行以下步骤：从训练样本集包括的未被选取的训练样本中选取训练样本；调整量化初始神经网络的参数，获得新的浮点型网络参数；将新的浮点型网络参数转化为新的整数型网络参数，以及基于新的整数型网络参数，生成新的量化初始神经网络；使用最近一次选取的训练样本和最新生成的量化初始神

神经网络，继续执行训练步骤。

在一些实施例中，第二执行单元进一步被配置成：将新的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的量化初始神经网络确定为新的量化初始神经网络。

5 在一些实施例中，该装置还包括：发送单元，被配置成将量化神经网络发送给用户终端，以使用户终端对接收的量化神经网络进行存储。

第四方面，本公开的实施例提供了一种用于处理信息的装置，该装置包括：第二获取单元，被配置成获取待处理信息和目标量化神经网络，其中，目标量化神经网络为采用如上述第一方面中任一实施例的方法生成的；输入单元，被配置成将待处理信息输入目标量化神经网络，获得处理结果及输出。

第五方面，本公开的实施例提供了一种电子设备，包括：一个或多个处理器；存储装置，其上存储有一个或多个程序，当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现上述第一方面或第二方面中任一实施例的方法。

第六方面，本公开的实施例提供了一种计算机可读介质，其上存储有计算机程序，该程序被处理器执行时实现上述第一方面或第二方面中任一实施例的方法。

20 本公开的实施例提供的用于生成量化神经网络的方法和装置，通过获取训练样本集和初始神经网络，其中，初始神经网络包括原始浮点型网络参数，而后将初始神经网络中的原始浮点型网络参数转化为整数型网络参数，以及基于转化成的整数型网络参数，生成量化初始神经网络，最后从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络，从而在神经网络的训练过程中，将神经网络中的浮点型网络参数转化为
25 整数型网络参数，以此，为神经网络的网络参数添加了量化约束，有
30

助于减小神经网络所占用的存储空间，以及利用神经网络进行信息处理时对 CPU 的消耗，提高信息处理的效率；并且，与现有技术中的直接对训练完成的神经网络中的网络参数进行量化，生成量化神经网络相比，本公开的方案可以减小对网络参数量化所引起的精度损失，提高量化神经网络的准确度，进而，利用本公开的量化神经网络进行信息处理的电子设备，相较于现有技术中的利用量化神经网络进行信息处理的电子设备，可以具有更准确的信息处理功能。

附图说明

通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本公开的其它特征、目的和优点将会变得更明显：

图 1 是本公开的一个实施例可以应用于其中的示例性系统架构图；

图 2 是根据本公开的用于生成量化神经网络的方法的一个实施例的流程图；

图 3 是根据本公开的实施例的用于生成量化神经网络的方法的一个应用场景的示意图；

图 4 是根据本公开的用于生成量化神经网络的方法的又一个实施例的流程图；

图 5 是根据本公开的用于生成量化神经网络的装置的一个实施例的结构示意图；

图 6 是适于用来实现本公开的实施例的电子设备的计算机系统的结构示意图。

具体实施方式

下面结合附图和实施例对本公开作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释相关发明，而非对该发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关发明相关的部分。

需要说明的是，在不冲突的情况下，本公开中的实施例及实施例

中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本公开。

图 1 示出了可以应用本公开的用于生成量化神经网络的方法、用于生成量化神经网络的装置、用于处理信息的方法或用于处理信息的装置的实施例的示例性系统架构 100。

如图 1 所示，系统架构 100 可以包括终端设备 101、102、103，网络 104 和服务器 105。网络 104 用以在终端设备 101、102、103 和服务器 105 之间提供通信链路的介质。网络 104 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

10 用户可以使用终端设备 101、102、103 通过网络 104 与服务器 105 交互，以接收或发送消息等。终端设备 101、102、103 上可以安装有各种通讯客户端应用，例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

15 终端设备 101、102、103 可以是硬件，也可以是软件。当终端设备 101、102、103 为硬件时，可以是各种电子设备，包括但不限于智能手机、平板电脑、电子书阅读器、MP3 播放器（Moving Picture Experts Group Audio Layer III，动态影像专家压缩标准音频层面 3）、MP4（Moving Picture Experts Group Audio Layer IV，动态影像专家压缩标准音频层面 4）播放器、膝上型便携计算机和台式计算机等等。当终端设备 101、102、103 为软件时，可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块（例如用来提供分布式服务的多个软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

25 服务器 105 可以是提供各种服务的服务器，例如对终端设备 101、102、103 发送的初始神经网络进行处理的模型处理服务器。模型处理服务器可以对接收到的初始神经网络等数据进行分析等处理，并将处理结果（例如量化神经网络）反馈给终端设备。

需要说明的是，本公开的实施例所提供的用于生成量化神经网络的方法一般由服务器 105 执行，相应地，用于生成量化神经网络的装置一般设置于服务器 105 中；另外，本公开的实施例所提供的用于处

理信息的的方法一般由终端设备 101、102、103 执行，相应地，用于处理信息的装置一般设置于 101、102、103 中。

需要说明的是，服务器可以是硬件，也可以是软件。当服务器为硬件时，可以实现成多个服务器组成的分布式服务器集群，也可以实现成单个服务器。当服务器为软件时，可以实现成多个软件或软件模块（例如用来提供分布式服务的多个软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

应该理解，图 1 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务

继续参考图 2，示出了根据本公开的用于生成量化神经网络的方法的一个实施例的流程 200。该用于生成量化神经网络的方法，包括以下步骤：

步骤 201，获取训练样本集和初始神经网络。

在本实施例中，用于生成量化神经网络的方法的执行主体（例如图 1 所示的服务器）可以通过有线连接方式或者无线连接方式从远程或本地获取训练样本集和初始神经网络。其中，训练样本集中的训练样本包括样本信息和针对样本信息预先确定的样本结果。样本信息为初始神经网络所能处理的信息，可以包括但不限于以下至少一项：文本、图像、音频、视频。例如，初始神经网络可以为用于进行人脸识别的神经网络，则样本信息可以为样本人脸图像。样本结果为利用初始神经网络对样本信息进行处理所能获得的期望结果（例如用于表征样本人脸图像所对应的人物的性别的性别信息）。

初识神经网络可以为未经训练的神经网络，也可以为经过训练的神经网络。初始神经网络的功能或者说输入、输出可以是预先确定的。进而，上述执行主体可以获取到用于训练初始神经网络的训练样本集。

在本实施例中，初始神经网络包括原始浮点型网络参数。原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积。

具体的，初始神经网络包括卷积层和批量归一化层。卷积层包括

浮点型权重。浮点型权重可以用于与卷积层的输入进行卷积运算，获得卷积层的输出。批量归一化层可以与卷积层连接，用于对卷积层的输出进行归一化处理。具体的，批量归一化层包括用于与卷积层的输出相乘，以对卷积层的输出进行归一化处理的浮点型归一化参数。在这里，可以将卷积层和批量归一化层作为一个网络结构，该网络结构的输出由卷积层的输出与浮点型归一化参数的乘积确定，由于卷积层的输出由浮点型权重与卷积层的输入的卷积确定，因此，上述网络结构的输出可以通过对浮点型归一化参数和浮点型权重进行求积，再与卷积层的输入进行卷积确定。可以理解，卷积层的输入即为上述网络结果的输入变量，浮点型归一化参数与浮点型权重的积即为上述网络结构的参数。进而，在本实施例中，可以将浮点型归一化参数与浮点型权重的积确定为浮点型网络参数。而原始浮点型网络参数即为初始神经网络包括的、待对其进行量化的网络参数。

实践中，对浮点型数据的量化指的是将浮点型数据转化为某个取值范围内的整数型数据。这里，取值范围由整数型数据的比特位数限定。例如欲转化成的整数型数据为 8 比特位（即 8bit），则取值范围为（0，255）。需要说明的是，在本实施例中，在对原始浮点型网络参数进行量化时，欲量化成的整数型网络参数的比特位数可以由技术人员预先确定。

可以理解，对于位数相同的浮点型数据和整数型数据，由于浮点型数据可以记录小数点之后的数据信息，因而具有更高的精度。而整数型数据由于不记录小数点之后的数据信息，因此，可以占用更少的存储空间，且利用整数型数据进行计算时，计算速度更快。

需要说明的是，为了获得更高的精度，现有技术中的神经网络中的权重和归一化参数通常存储为浮点型。

步骤 202，将初始神经网络中的原始浮点型网络参数转化为整数型网络参数。

在本实施例中，基于步骤 201 中得到的初始神经网络，上述执行主体可以将初始神经网络中的原始浮点型网络参数转化为整数型网络参数。

具体的，上述执行主体可以首先确定所要转化的整数型网络参数的比特位数，然后，采用现有的各种方法将初始神经网络中的原始浮点型网络参数转化为整数型网络参数。可以理解，将初始神经网络中的原始浮点型网络参数转化为整数型网络参数，相当于对初始神经网络添加量化约束。

作为示例，原始浮点型网络参数包括数值“21.323”，预先确定出欲转化成的整数型权重的比特位数为八位，即可以确定整数型网络参数的取值范围为(0, 255)，进而可以直接采用四舍五入的方式将原始浮点型网络参数中的数值“21.323”转化为整数型网络参数“21”。

在本实施例的一些可选的实现方式中，上述执行主体可以通过以下步骤将初始神经网络中的原始浮点型网络参数转化为整数型网络参数：首先，上述执行主体可以将原始浮点型网络参数所对应的浮点型权重转化为整数型权重，以及将原始浮点型网络参数所对应的浮点型归一化参数转化为整数型归一化参数。然后，上述执行主体可以对转化成的整数型权重和整数型归一化参数进行求积，获得整数型网络参数。本实现方式首先对浮点型权重和浮点型归一化参数添加了量化约束，而后利用添加了量化约束的整数型权重和整数型归一化参数获得整数型网络参数，以此，可以减少量化后的整数型网络参数的精度损失，有助于提高量化后的初始神经网络的准确度。

步骤 203，基于转化成的整数型网络参数，生成量化初始神经网络。

在本实施例中，基于步骤 202 中转化成的整数型网络参数，上述执行主体可以生成量化初始神经网络。

具体的，上述执行主体可以直接将包括转化成的整数型网络参数的初始神经网络确定为量化神经网络；或者，上述执行主体也可以对包括转化成的整数型网络参数的初始神经网络进行处理，并将处理后的初始神经网络确定为量化初始神经网络。

在本实施例的一些可选的实现方式中，上述执行主体可以通过以下步骤生成量化初始神经网络：上述执行主体可以将转化成的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参

数的初始神经网络确定为量化初始神经网络。

这里，将整数型网络参数转化为浮点型网络参数为上述将原始浮点型网络参数转化为整数型网络参数的逆过程，可以参考将原始浮点型网络参数化为整数型网络参数的步骤对转化成的整数型网络参数进行转化，获得浮点型网络参数。

继续上述示例，转化成的整数型网络参数为“21”，由原始浮点型网络参数中的“21.323”可知，浮点型网络参数精确到小数点后三位。所以，这里可以将整数型网络参数“21”转化为浮点型网络参数“21.000”。

需要说明的是，浮点型数据相较于整数型数据可以具有更高的精度。所以为初始神经网络添加量化约束后，再将整数型网络参数转化为浮点型网络参数，有助于在后续对初始神经网络的训练过程中，提高训练精度，获得更为准确的训练结果。

步骤 204，从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

在本实施例中，基于步骤 201 中获得的训练样本集，上述执行主体可以从训练样本集中选取训练样本，以及执行以下训练步骤：

步骤 2041，将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练。

在这里，上述执行主体可以利用机器学习方法，对量化初始神经网络进行训练。具体的，上述执行主体将样本信息输入量化初始神经网络，获得实际结果，然后利用预设的损失函数计算所得到的实际结果与训练样本中的样本结果之间的差异，例如，可以采用 L2 范数作为损失函数计算所得到的实际结果与训练样本中的样本结果之间的差异。

步骤 2042，响应于确定量化初始神经网络训练完成，基于训练完

成的量化初始神经网络，生成量化神经网络。

具体的，上述执行主体可以确定当前对量化初始神经网络的训练是否满足预先设置的完成条件，若满足，则可以确定量化初始神经网络训练完成。其中，完成条件可以包括但不限于以下至少一项：训练
5 时间超过预设时长；训练次数超过预设次数；计算所得的差异小于预设差异阈值。

在本实施例中，上述执行主体可以响应于确定训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。其中，量化神经网络为训练完成的、所包括的网络参数为整数型网络参数的神经网络。

10 具体的，响应于确定训练完成的量化初始神经网络中的网络参数为整数型权重，则上述执行主体可以直接将训练完成的量化初始神经网络确定为量化神经网络；响应于确定训练完成的量化初始神经网络中的网络参数为浮点型网络参数，则上述执行主体可以将训练完成的
15 量化初始神经网络中的浮点型网络参数转化为整数型网络参数，进而将包括转化成的整数型网络参数的、训练完成的量化初始神经网络确定为量化神经网络。

在本实施例中，上述执行主体还可以响应于确定量化初始神经网络未训练完成，执行以下步骤：从训练样本集包括的未被选取的训练样本中选取训练样本；调整量化初始神经网络的参数，获得新的浮点
20 型网络参数；将新的浮点型网络参数转化为新的整数型网络参数，以及基于新的整数型网络参数，生成新的量化初始神经网络；使用最近一次选取的训练样本和最新生成的量化初始神经网络，继续执行上述训练步骤（步骤 2041-2042）。

这里，可以采用各种实现方式基于计算得到的实际结果与训练样本中的样本结果之间的差异调整量化初始神经网络的参数。例如，可以
25 采用 BP（Back Propagation，反向传播）算法及 SGD（Stochastic Gradient Descent，随机梯度下降）算法来调整量化初始神经网络的参数。需要说明的是，调整参数时，为了不影响训练的收敛情况，保证训练能够重复执行，通常会将参数调整为浮点型。因此，调整参数后
30 量化初始神经网络会获得新的浮点型网络参数，进而，上述执行主体

可以将新的浮点型网络参数转化为新的整数型网络参数，以重新为包括新的浮点型网络参数的量化初始神经网络添加量化约束，生成新的量化初始神经网络。

具体的，上述执行主体可以基于新的整数型网络参数，采用各种方式生成新的量化初始神经网络。例如，可以直接将包括新的整数型网络参数的量化初始神经网络确定为新的量化初始神经网络。或者，上述执行主体也可以对包括新的整数型网络参数的量化初始神经网络进行处理，并将处理后的量化初始神经网络确定为新的量化初始神经网络。

在本实施例的一些可选的实现方式中，上述执行主体可以通过以下步骤生成新的量化初始神经网络：上述执行主体可以将新的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的量化初始神经网络确定为新的量化初始神经网络。

在本实施例的一些可选的实现方式中，上述执行主体可以将量化神经网络发送给用户终端，以便用户终端对接收的量化神经网络进行存储。在这里，添加了量化约束的量化神经网络可以占用更少的存储空间，通过本实现方式，可以节省用户终端的存储资源。

继续参见图 3，图 3 是根据本实施例的用于生成神经网络的方法的应用场景的一个示意图。在图 3 的应用场景中，服务器 301 首先可以获取训练样本集 302 和初始神经网络 303，其中，训练样本集 302 中的训练样本包括样本信息和针对样本信息预先确定的样本结果。初始神经网络 303 包括原始浮点型网络参数 304（例如“2.134”）。原始浮点型网络参数 304 为初始神经网络 303 中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积。然后，服务器 301 将初始神经网络 303 中的原始浮点型网络参数 304 转化为整数型网络参数 305（例如“2”）。接着，服务器 301 基于转化成的整数型网络参数 305，生成量化初始神经网络 306。最后，服务器 301 可以从训练样本集 302 中选取训练样本 3021，以及执行以下训练步骤：将所选取的训练样本 3021 中的样本信息 30211 作为量化初始神经网络 306

的输入,将所选取的训练样本 3021 中的样本结果 30212 作为量化初始神经网络 306 的期望输出,对量化初始神经网络 306 进行训练;响应于确定量化初始神经网络 306 训练完成,基于训练完成的量化初始神经网络 306,生成量化神经网络 307。

5 本公开的上述实施例提供的方法在神经网络的训练过程中,将神经网络中的浮点型网络参数转化为整数型网络参数,以此,为神经网络的网络参数添加了量化约束,有助于减小神经网络所占用的存储空间,以及利用神经网络进行信息处理时对 CPU 的消耗,提高信息处理的效率;并且,与现有技术中的直接对训练完成的神经网络中的网络
10 参数进行量化,生成量化神经网络相比,本公开的方案可以减小对网络参数量化所引起的精度损失,提高量化神经网络的准确度,进而,利用本公开的量化神经网络进行信息处理的电子设备,相较于现有技术中的利用量化神经网络进行信息处理的电子设备,可以具有更准确的信息处理功能。

15

进一步参考图 4,其示出了用于处理信息的方法的一个实施例的流程 400。该用于处理信息的方法的流程 400,包括以下步骤:

步骤 401,获取待处理信息和目标量化神经网络。

在本实施例中,用于处理信息的方法的执行主体(例如图 1 所示
20 的终端设备)可以通过有线连接方式或者无线连接方式从远程或本地获取待处理信息和目标量化神经网络。其中,目标量化神经网络为采用图 2 对应的实施例中的任一实施例的方法生成的。目标量化神经网络为待利用其进行信息处理的量化神经网络。待处理信息可以为目标量化神经网络所能够处理的信息。可以包括但不限于以下至少一项:
25 文本、图像、音频、视频。作为示例,目标量化神经网络为用于进行人脸识别的模型,则待处理信息可以为人脸图像。待处理信息可以预先存储在上述执行主体上,也可以由其他电子设备发送给上述执行主体。处理结果可以为目标量化神经网络的输出结果。

步骤 402,将待处理信息输入目标量化神经网络,获得处理结果
30 及输出。

具体的,上述执行主体可以将待处理信息输入目标量化神经网络,获得目标量化神经网络输出的处理结果。

在这里,获得处理结果后,上述执行主体可以输出处理结果。具体的,上述执行主体可以将处理结果输出给通信连接的其他电子设备,也可以对处理结果进行输出显示。

实践中,由于用户终端的存储空间较小,而一般的神经网络需要消耗大量的存储资源,所以通常神经网络在用户终端上难以适用。本公开的实施例提供的方法采用图 2 对应的任一实施例中生成的量化神经网络,可以使得量化神经网络适用于用户终端,同时有助于减小对用户终端的存储资源的消耗;并且,当用户终端在利用量化神经网络进行信息处理时,由于量化神经网络的复杂程度低,所以可以提高用户终端进行信息处理的效率,减小对用户终端的 CPU 的消耗;此外,由于发送给用户终端的量化神经网络是通过在训练过程中添加量化约束而获得的神经网络,与现有技术中的、通过为训练完成的神经网络添加量化约束而生成的量化神经网络相比,本公开的量化神经网络的精度损失更小,进而,用户终端利用本公开的量化神经网络可以实现更为准确的信息处理及输出。

进一步参考图 5,作为对上述各图所示方法的实现,本公开提供了一种用于生成量化神经网络的装置的一个实施例,该装置实施例与图 2 所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

如图 5 所示,本实施例的用于生成量化神经网络的装置 500 包括:第一获取单元 501、转化单元 502、生成单元 503 和第一执行单元 504。其中,第一获取单元 501 被配置成获取训练样本集和初始神经网络,其中,训练样本包括样本信息和针对样本信息预先确定的样本结果,初始神经网络包括原始浮点型网络参数,原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积;转化单元 502 被配置成将初始神经网络中的原始浮点型网络参数转化为整数型网络参数;生成单元 503 被配置成

基于转化成的整数型网络参数，生成量化初始神经网络；第一执行单元 504 被配置成从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

在本实施例中，用于生成量化神经网络的装置 500 的第一获取单元 501 可以通过有线连接方式或者无线连接方式从远程或本地获取训练样本集和初始神经网络。其中，训练样本集中的训练样本包括样本信息和针对样本信息预先确定的样本结果。样本信息为初始神经网络所能处理的信息，可以包括但不限于以下至少一项：文本、图像、音频、视频。初识神经网络可以为未经训练的神经网络，也可以为经过训练的神经网络。

在本实施例中，初始神经网络包括原始浮点型网络参数。原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积。

在本实施例中，基于第一获取单元 501 得到的初始神经网络，转化单元 502 可以将初始神经网络中的原始浮点型网络参数转化为整数型网络参数。

在本实施例中，基于转化单元 502 转化成的整数型网络参数，生成单元 503 可以生成量化初始神经网络。

在本实施例中，基于第一获取单元 501 获得的训练样本集，第一执行单元 504 可以从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

在本实施例的一些可选的实现方式中，生成单元 503 可以进一步被配置成：将转化成的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的初始神经网络确定为量化初始神经网络。

网络。

在本实施例的一些可选的实现方式中，转化单元 502 可以包括：转化模块（图中未示出），被配置成将原始浮点型网络参数所对应的浮点型权重转化为整数型权重，以及将原始浮点型网络参数所对应的浮点型归一化参数转化为整数型归一化参数；求积模块（图中未示出），
5 被配置成对转化成的整数型权重和整数型归一化参数进行求积，获得整数型网络参数。

在本实施例的一些可选的实现方式中，装置 500 还可以包括：第二执行单元（图中未示出），被配置成响应于确定量化初始神经网络未
10 训练完成，执行以下步骤：从训练样本集包括的未被选取的训练样本中选取训练样本；调整量化初始神经网络的参数，获得新的浮点型网络参数；将新的浮点型网络参数转化为新的整数型网络参数，以及基于新的整数型网络参数，生成新的量化初始神经网络；使用最近一次选取的训练样本和最新生成的量化初始神经网络，继续执行训练步骤。

在本实施例的一些可选的实现方式中，第二执行单元可以进一步
15 被配置成：将新的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的量化初始神经网络确定为新的量化初始神经网络。

在本实施例的一些可选的实现方式中，装置 500 还可以包括：发
20 送单元（图中未示出），被配置成将量化神经网络发送给用户终端，以便用户终端对接收的量化神经网络进行存储。

可以理解的是，该装置 500 中记载的诸单元与参考图 2 描述的方法中的各个步骤相对应。由此，上文针对方法描述的操作、特征以及产生的有益效果同样适用于装置 500 及其中包含的单元，在此不再赘
25 述。

本公开的上述实施例提供的装置 500 在神经网络的训练过程中，将神经网络中的浮点型权重转化为整数型权重，以此，为神经网络的权重添加了量化约束，有助于减小神经网络所占用的存储空间，以及利用神经网络进行信息处理时对 CPU 的消耗，提高信息处理的效率；
30 并且，与现有技术中的直接对训练完成的神经网络中的权重进行量化，

生成量化神经网络相比，本公开的方案可以减小对权重量化所引起的精度损失，提高量化神经网络的准确度，进而，利用本公开的量化神经网络进行信息处理的电子设备，相较于现有技术中的利用量化神经网络进行信息处理的电子设备，可以具有更准确的信息处理功能。

5

进一步参考图 6，作为对上述各图所示方法的实现，本公开提供了一种用于处理信息的装置的一个实施例，该装置实施例与图 4 所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图 6 所示，本实施例的用于处理信息的装置 600 包括：第二获取单元 601 和输入单元 602。其中，第二获取单元 601 被配置成获取待处理信息和目标量化神经网络，其中，目标量化神经网络为采用如图 2 对应的实施例中的任一实施例的方法生成的；输入单元 602 被配置成将待处理信息输入目标量化神经网络，获得处理结果及输出。

在本实施例中，用于处理信息的装置 600 的第二获取单元 601 可以通过有线连接方式或者无线连接方式从远程或本地获取待处理信息和目标量化神经网络。其中，目标量化神经网络为采用图 2 对应的实施例中的任一实施例的方法生成的。目标量化神经网络为待利用其进行信息处理的量化神经网络。待处理信息可以为目标量化神经网络所能够处理的信息。可以包括但不限于以下至少一项：文本、图像、音频、视频。

在本实施例中，输入单元 602 可以将待处理信息输入目标量化神经网络，获得目标量化神经网络输出的处理结果及输出。

可以理解的是，该装置 600 中记载的诸单元与参考图 4 描述的方法中的各个步骤相对应。由此，上文针对方法描述的操作、特征以及产生的有益效果同样适用于装置 600 及其中包含的单元，在此不再赘述。

本公开的上述实施例提供的装置 600 采用图 2 对应的任一实施例中生成的量化神经网络，可以使得量化神经网络适用于用户终端，同时有助于减小对用户终端的存储资源的消耗；并且，当用户终端在利用量化神经网络进行信息处理时，由于量化神经网络的复杂程度低，

所以可以提高用户终端进行信息处理的效率，减小对用户终端的 CPU 的消耗；此外，由于发送给用户终端的量化神经网络是通过在训练过程中添加量化约束而获得的神经网络，与现有技术中的、通过为训练完成的神经网络添加量化约束而生成的量化神经网络相比，本公开的
5 量化神经网络的精度损失更小，进而，用户终端利用本公开的量化神经网络可以实现更为准确的信息处理及输出。

下面参考图 7，其示出了适于用来实现本公开实施例的电子设备（例如图 1 中的终端设备或服务器）700 的结构示意图。本公开实施
10 例中的终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA（个人数字助理）、PAD（平板电脑）、PMP（便携式多媒体播放器）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字 TV、台式计算机等等的固定终端。图 7 示出的电子设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

如图 7 所示，电子设备 700 可以包括处理装置（例如中央处理器、图形处理器等）701，其可以根据存储在只读存储器（ROM）702 中的程序或者从存储装置 708 加载到随机访问存储器（RAM）703 中的程序而执行各种适当的动作和处理。在 RAM 703 中，还存储有电子设备 700 操作所需的各种程序和数据。处理装置 701、ROM 702 以及 RAM
15 703 通过总线 704 彼此相连。输入/输出（I/O）接口 705 也连接至总线 704。

通常，以下装置可以连接至 I/O 接口 705：包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 606；包括例如液晶显示器（LCD）、扬声器、振动器等等的输出装置 707；
25 包括例如磁带、硬盘等的存储装置 708；以及通信装置 709。通信装置 709 可以允许电子设备 700 与其他设备进行无线或有线通信以交换数据。虽然图 7 示出了具有各种装置的电子设备 700，但是应理解的是，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以
30

被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 709 从网络上被下载和安装，或者从存储装置 708 被安装，或者从 ROM 702 被安装。在该计算机程序被处理装置 701 执行时，执行本公开实施例的方法中限定的上述功能。

需要说明的是，本公开所述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、RF（射频）等等，或者上述的任意合适的组合。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备：获取训练样本集和初始神经网络，其中，训练样本包括样本信息和针对样本信息预先确定的样本结果，初始神经网络包括

原始浮点型网络参数，原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积；将初始神经网络中的原始浮点型网络参数转化为整数型网络参数，以及基于转化成的整数型网络参数，生成量化初始神经网络；从训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

此外，当上述一个或者多个程序被该电子设备执行时，还可以使得该电子设备：获取待处理信息和目标量化神经网络，其中，目标量化神经网络为采用图 2 对应的实施例中的任一实施例的方法生成的；将待处理信息输入目标量化神经网络，获得处理结果及输出。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时

也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

- 5 描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元的名称在某种情况下并不构成对该单元本身的限定，例如，第一获取单元还可以被描述为“获取训练样本集和初始神经网络的单元”。

- 10 以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

15

权 利 要 求 书

1、一种用于生成量化神经网络的方法，包括：

获取训练样本集和初始神经网络，其中，训练样本包括样本信息
5 和针对样本信息预先确定的样本结果，初始神经网络包括原始浮点型网络参数，原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积；

将初始神经网络中的原始浮点型网络参数转化为整数型网络参数；

10 基于转化成的整数型网络参数，生成量化初始神经网络；

从所述训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，
15 基于训练完成的量化初始神经网络，生成量化神经网络。

2、根据权利要求 1 所述的方法，其中，所述基于转化成的整数型网络参数，生成量化初始神经网络，包括：

将转化成的整数型网络参数转化为浮点型网络参数，以及将包括
20 转化成的浮点型网络参数的初始神经网络确定为量化初始神经网络。

3、根据权利要求 1 所述的方法，其中，所述将初始神经网络中的原始浮点型网络参数转化为整数型网络参数，包括：

将原始浮点型网络参数所对应的浮点型权重转化为整数型权重，
25 以及将原始浮点型网络参数所对应的浮点型归一化参数转化为整数型归一化参数；

对转化成的整数型权重和整数型归一化参数进行求积，获得整数型网络参数。

30 4、根据权利要求 1 所述的方法，其中，所述方法还包括：

响应于确定量化初始神经网络未训练完成，执行以下步骤：从所述训练样本集包括的未被选取的训练样本中选取训练样本；调整量化初始神经网络的参数，获得新的浮点型网络参数；将新的浮点型网络参数转化为新的整数型网络参数，以及基于新的整数型网络参数，生成新的量化初始神经网络；使用最近一次选取的训练样本和最新生成的量化初始神经网络，继续执行所述训练步骤。

5、根据权利要求4所述的方法，其中，所述基于新的整数型网络参数，生成新的量化初始神经网络，包括：

10 将新的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的量化初始神经网络确定为新的量化初始神经网络。

6、根据权利要求1-5之一所述的方法，其中，所述方法还包括：

15 将量化神经网络发送给用户终端，以便用户终端对接收的量化神经网络进行存储。

7、一种用于处理信息的方法，包括：

20 获取待处理信息和目标量化神经网络，其中，所述目标量化神经网络为采用如权利要求1-6中任一所述的方法生成的；

将所述待处理信息输入所述目标量化神经网络，获得处理结果及输出。

8、一种用于生成量化神经网络的装置，包括：

25 第一获取单元，被配置成获取训练样本集和初始神经网络，其中，训练样本包括样本信息和针对样本信息预先确定的样本结果，初始神经网络包括原始浮点型网络参数，原始浮点型网络参数为初始神经网络中的卷积层的浮点型权重和与卷积层连接的批量归一化层的浮点型归一化参数的积；

30 转化单元，被配置成将初始神经网络中的原始浮点型网络参数转

化为整数型网络参数；

生成单元，被配置成基于转化成的整数型网络参数，生成量化初始神经网络；

5 第一执行单元，被配置成从所述训练样本集中选取训练样本，以及执行以下训练步骤：将所选取的训练样本中的样本信息作为量化初始神经网络的输入，将所选取的训练样本中的样本结果作为量化初始神经网络的期望输出，对量化初始神经网络进行训练；响应于确定量化初始神经网络训练完成，基于训练完成的量化初始神经网络，生成量化神经网络。

10

9、根据权利要求8所述的装置，其中，所述生成单元进一步被配置成：

将转化成的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的初始神经网络确定为量化初始神经网络。

15

10、根据权利要求8所述的装置，其中，所述转化单元包括：

转化模块，被配置成将原始浮点型网络参数所对应的浮点型权重转化为整数型权重，以及将原始浮点型网络参数所对应的浮点型归一化参数转化为整数型归一化参数；

20 求积模块，被配置成对转化成的整数型权重和整数型归一化参数进行求积，获得整数型网络参数。

11、根据权利要求8所述的装置，其中，所述装置还包括：

25 第二执行单元，被配置成响应于确定量化初始神经网络未训练完成，执行以下步骤：从所述训练样本集包括的未被选取的训练样本中选取训练样本；调整量化初始神经网络的参数，获得新的浮点型网络参数；将新的浮点型网络参数转化为新的整数型网络参数，以及基于新的整数型网络参数，生成新的量化初始神经网络；使用最近一次选取的训练样本和最新生成的量化初始神经网络，继续执行所述训练步骤。

30

12、根据权利要求 11 所述的装置，其中，所述第二执行单元进一步被配置成：

5 将新的整数型网络参数转化为浮点型网络参数，以及将包括转化成的浮点型网络参数的量化初始神经网络确定为新的量化初始神经网络。

13、根据权利要求 8-12 之一所述的装置，其中，所述装置还包括：

10 发送单元，被配置成将量化神经网络发送给用户终端，以便用户终端对接收的量化神经网络进行存储。

14、一种用于处理信息的装置，包括：

15 第二获取单元，被配置成获取待处理信息和目标量化神经网络，其中，所述目标量化神经网络为采用如权利要求 1-6 中任一所述的方法生成的；

输入单元，被配置成将所述待处理信息输入所述目标量化神经网络，获得处理结果及输出。

15、一种电子设备，包括：

20 一个或多个处理器；

存储装置，其上存储有一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-7 中任一所述的方法。

25 16、一种计算机可读介质，其上存储有计算机程序，其中，该程序被处理器执行时实现如权利要求 1-7 中任一所述的方法。

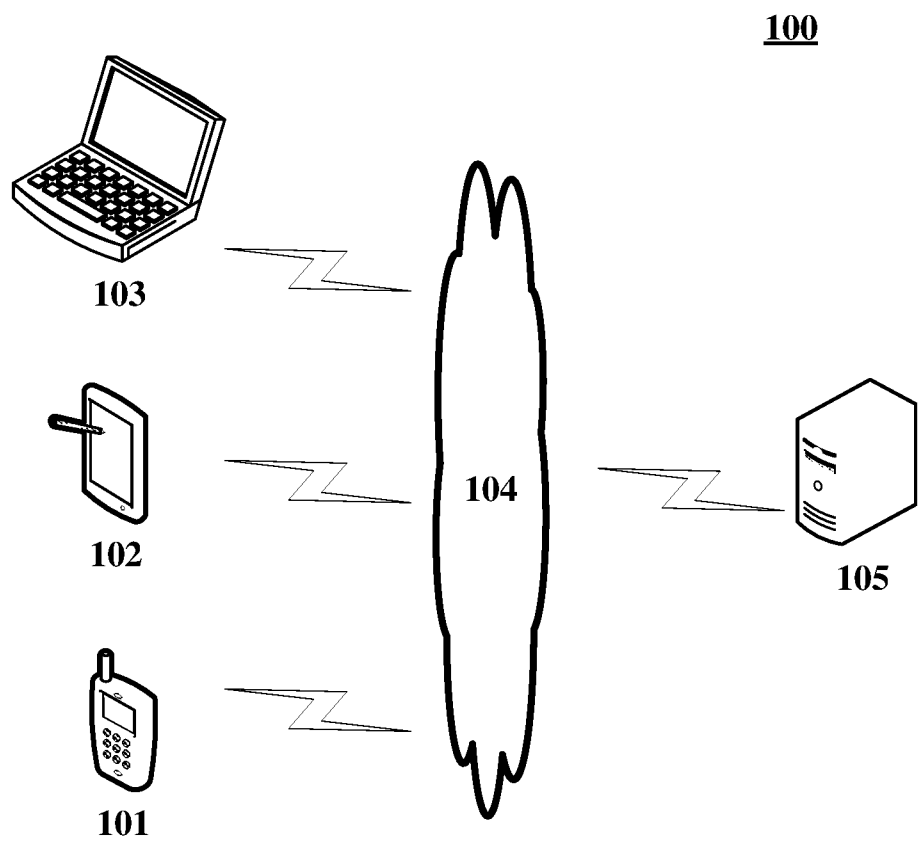


图 1

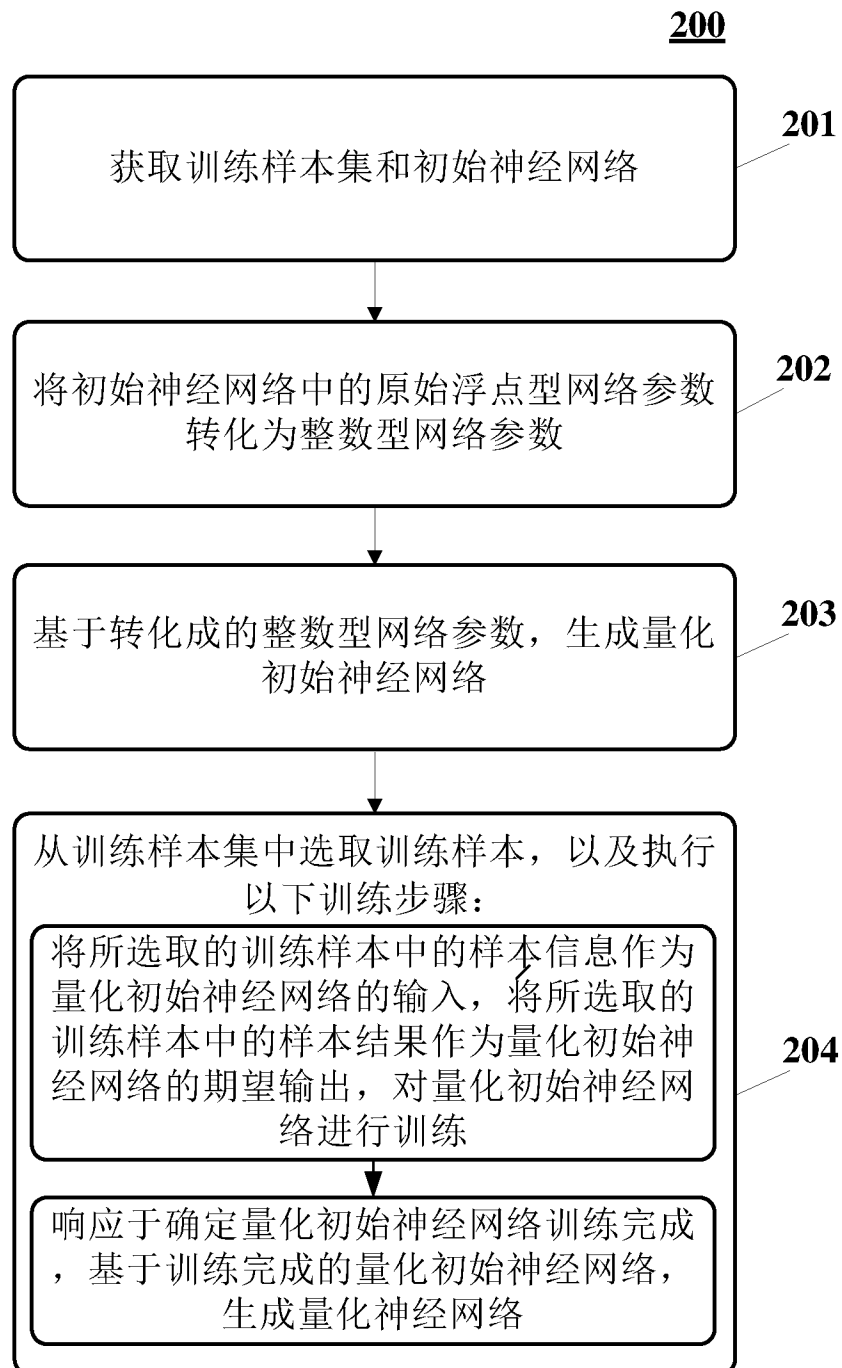


图 2

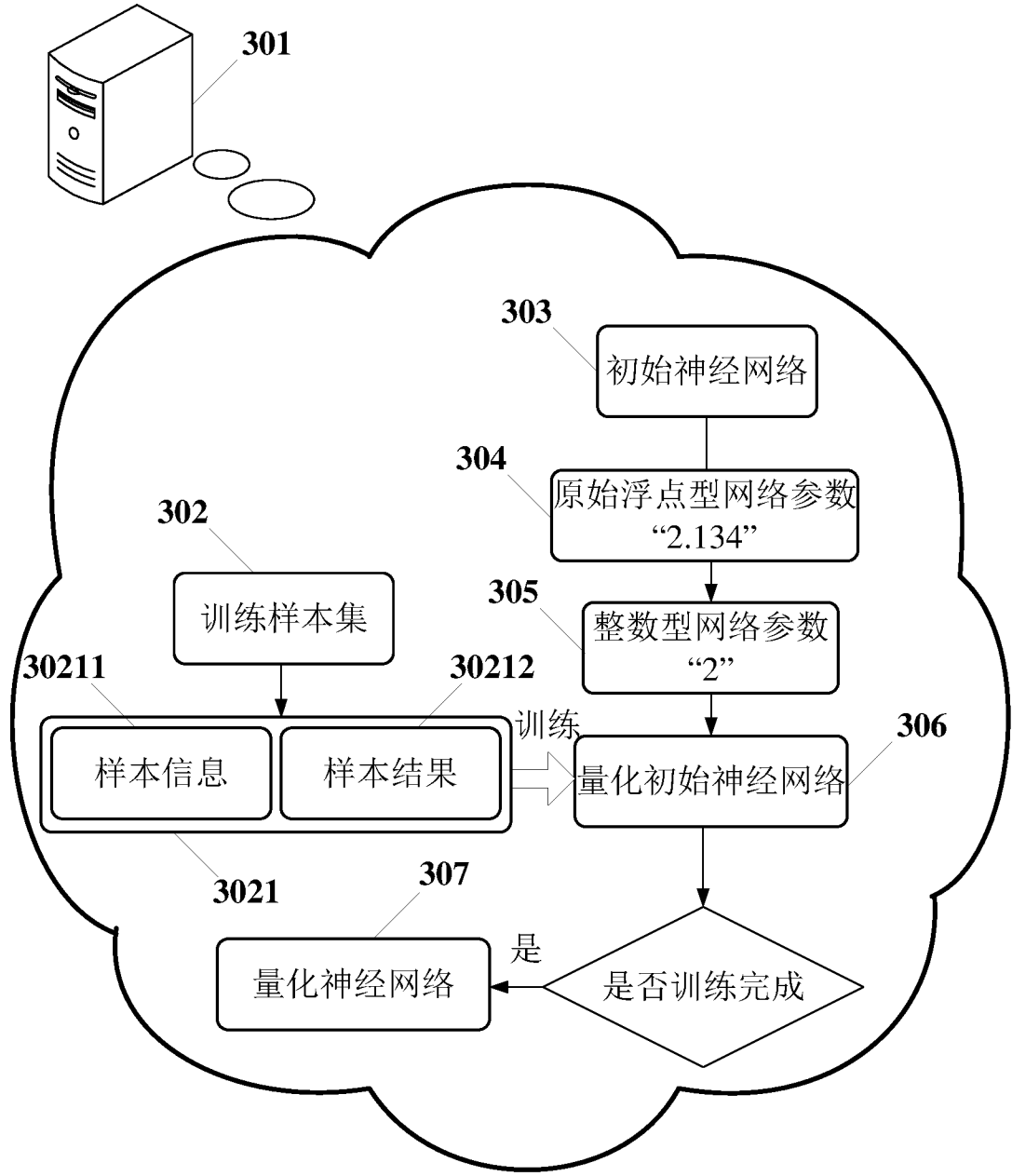


图 3

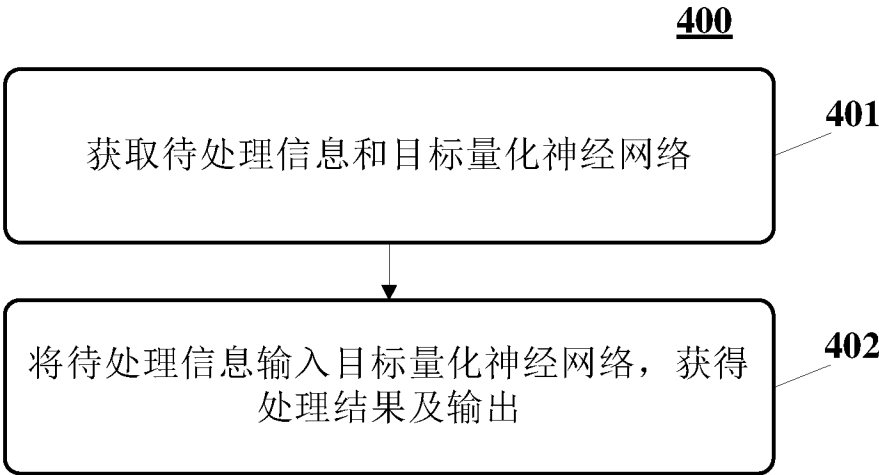


图 4

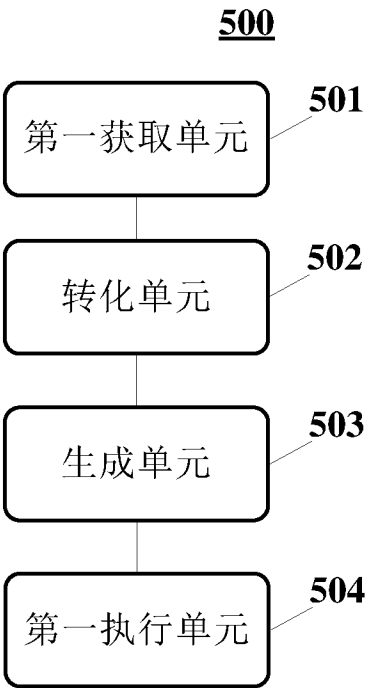


图 5

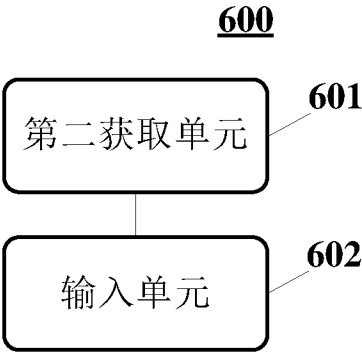


图 6

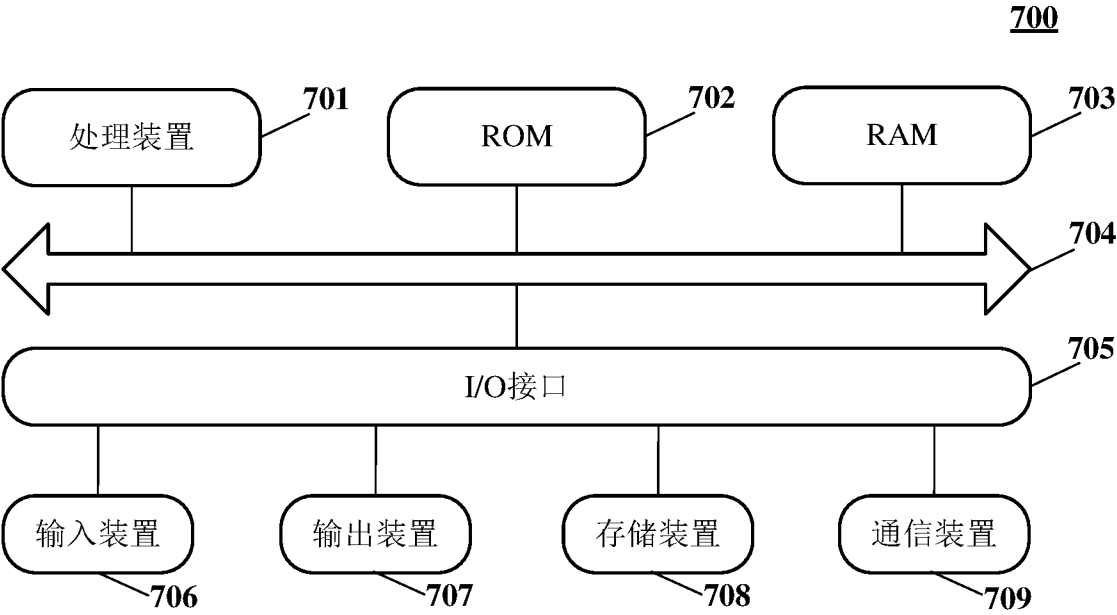


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/078586

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/063(2006.01)i; G06N 3/08(2006.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; CNKI; SIPOABS; DWPI; USTXT; WOTXT; EPTXT: 神经网络, 机器学习, 训练, 压缩, 卷积, 参数, 权重, 归一化, 全连接, 浮点, 整型, 定点, 优化, neural, network, training, compress, normalization, weight, learning, floating, fixed, optimize

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109961141 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 02 July 2019 (2019-07-02) claims 1-16, description, paragraphs [0030]-[0113], figures 1-7	1-16
X	CN 109165736 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 08 January 2019 (2019-01-08) description, paragraphs [0026]-[0088], figures 1-6	1-16
A	CN 107590460 A (BEIJING DAJIA INTERCONNECTION INFORMATION TECHNOLOGY CO., LTD.) 16 January 2018 (2018-01-16) entire document	1-16
A	CN 108509179 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 07 September 2018 (2018-09-07) entire document	1-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

29 March 2020

Date of mailing of the international search report

03 June 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/078586

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109961141	A	02 July 2019	None			
CN	109165736	A	08 January 2019	None			
CN	107590460	A	16 January 2018	CN	107590460	B	03 May 2019
CN	108509179	A	07 September 2018	None			

国际检索报告

国际申请号

PCT/CN2020/078586

A. 主题的分类

G06N 3/063 (2006.01)i; G06N 3/08 (2006.01)n

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS;CNTXT;CNKI;SIPOABS;DWPI;USTXT;WOTXT;EPTXT; 神经网络, 机器学习, 训练, 压缩, 卷积, 参数, 权重, 归一化, 全连接, 浮点, 整型, 定点, 优化, neural, network, training, compress, normalization, weight, learning, floating, fixed, optimize

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 109961141 A (北京字节跳动网络技术有限公司) 2019年 7月 2日 (2019 - 07 - 02) 权利要求第1-16项, 说明书第[0030]-[0113]段, 附图1-7	1-16
X	CN 109165736 A (北京字节跳动网络技术有限公司) 2019年 1月 8日 (2019 - 01 - 08) 说明书第[0026]-[0088]段, 附图1-6	1-16
A	CN 107590460 A (北京达佳互联信息技术有限公司) 2018年 1月 16日 (2018 - 01 - 16) 全文	1-16
A	CN 108509179 A (百度在线网络技术北京有限公司) 2018年 9月 7日 (2018 - 09 - 07) 全文	1-16

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2020年 3月 29日

国际检索报告邮寄日期

2020年 6月 3日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

任洪潮

电话号码 (86-512) 88995842

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/078586

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109961141	A	2019年 7月 2日	无	
CN	109165736	A	2019年 1月 8日	无	
CN	107590460	A	2018年 1月 16日	CN 107590460 B	2019年 5月 3日
CN	108509179	A	2018年 9月 7日	无	