

(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
2016年8月4日 (04.08.2016)



(10) 国际公布号
WO 2016/119429 A1

- (51) 国际专利分类号:
G06N 3/08 (2006.01)
- (21) 国际申请号: PCT/CN2015/086011
- (22) 国际申请日: 2015年8月4日 (04.08.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510036813.0 2015年1月26日 (26.01.2015) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 陈嘉 (CHEN, Jia); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。曾嘉 (ZENG, Jia); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京龙双利达知识产权代理有限公司 (LONGSUN LEAD IP LTD.); 中国北京市海淀区丹棱街16号海兴大厦C座1108, Beijing 100080 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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)。

根据细则 4.17 的声明:

- 关于申请人有权申请并被授予专利(细则 4.17(ii))

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

(54) Title: SYSTEM AND METHOD FOR TRAINING PARAMETER SET IN NEURAL NETWORK

(54) 发明名称: 用于神经网络中训练参数集的系统和方法

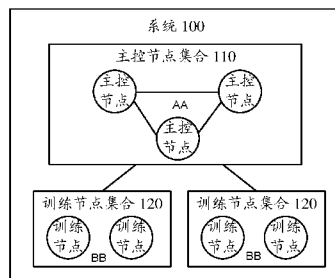
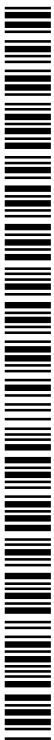


图 1

(57) Abstract: A system and method for a training parameter set in a neural network. The system comprises a master control set (110) and N training node sets (120). The master control set is used for controlling a training process and storing a data set and a parameter set used in training. The master control set comprises M master control nodes, every two of the M master control nodes are in communication connection, and at least one of the M master control nodes is used for backing up the parameter set. The training node sets comprise multiple training nodes. The training nodes are used for performing training according to the data set and the parameter set delivered by the master control set, and sending training results to the corresponding master control nodes. The system and the method can avoid a case in which the whole training fails due to the failure of a master control node, thereby improving the reliability of the training process; training is performed in parallel by configuring multiple training node sets, thereby improving the training efficiency.

(57) 摘要:

[见续页]



WO 2016/119429 A1

一种用于神经网络中训练参数集的系统和方法，该系统包括：主控节点集合（110），用于控制训练过程，并用于存储训练所使用的数据集和参数集，主控节点集合包括 M 个主控节点，M 个主控节点两两之间通信连接，M 个主控节点中的至少一个主控节点用于备份参数集；以及 N 个训练节点集合（120），训练节点集合包括多个训练节点，训练节点用于根据主控节点集合下发的数据集和参数集进行训练，并将训练结果发送给相应的主控节点。该系统和方法可以避免当某一个主控节点失效时导致的整个训练失败情况，提高训练过程的可靠性，通过配置多个训练节点集合并行地进行训练，提高训练效率。

用于神经网络中训练参数集的系统和方法

技术领域

本发明涉及数据处理领域，尤其涉及数据处理领域中的用于神经网络中
5 训练参数集的系统和方法。

背景技术

神经网络是一种模拟大脑神经突触结构来进行信息处理的数学模型，是
对人脑的抽象、简化和模拟，可以反映人脑的基本特性。神经网络由大量的
10 节点（也称为神经元）和相互之间的加权连接构成。每个节点代表一种特定的
输出函数，称为激励函数；而每两个节点间的连接都代表一个对于通过该
连接信号的加权值，称为权重。神经网络用数学函数可以表示为 $Y = f(X, W)$ ，
其中， X 代表网络的输入， Y 代表网络的输出， W 代表网络的参数集。

下面以监督学习为例来简单描述神经网络的训练问题。神经网络的训练
15 即是要寻找上述函数中的参数集 W 。神经网络的训练过程为：给定训练的数
据集 $D = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$ ，对每一个训练数据 (X_i, Y_i) ，定义其价
值函数为 $E_i = \frac{1}{2}(Y_i - f(X_i, W))^2$ ，确定 W ，使得 $\frac{1}{N} \sum_{i=1}^N E_i$ 的值最小。

深度学习是针对神经网络的训练方法之一。目前，深度学习已经可以很
好的用于解决语音识别、图象识别及文本处理等实际应用问题。神经网络通
20 常需要使用大规模的训练数据进行训练，以保证神经网络的运算结果达到一
定的准确度。相应地，训练数据规模越大就会使得计算量越大，训练所需的
时间也越长。为了加快神经网络的训练速度，图形处理单元（Graphic
Processing Unit, GPU）等协处理器被广泛应用于深度学习训练计算中。但是
这些协处理器的内存相对较小，无法容纳大型神经网络的加权参数集。

25 并且，现有技术通过主控节点将神经网络各副本发送给运算节点，
并指示运算节点进行训练。每个运算节点至少配备一个 GPU 进行运算处理。
主控节点在运算节点进行训练时定时查询运算节点状态，并在运算节点达到
停止状态后更新主控节点以及运算节点上副本神经网络加权参数。现有技术
中，使用众多计算节点协同训练一个大型神经网络，采用传统的同步更新的

方式，系统中所有的计算节点只能同时基于相同的参数集 W 进行训练，系统的整体性能会被最慢的节点以及系统网络带宽所限制。当某个或某些节点失效时，会对整个训练过程带来严重影响。

因此，现有的神经网络的训练系统可靠性较差，仅支持一个主控节点，
5 当主控节点失效时会导致整个训练的失败。并且，现有的训练系统的运算节点只能同时基于相同的参数集 W 进行训练，系统的规模和整体性能受限于主控节点以及运算节点的内存大小。

发明内容

10 本发明实施例提供了一种用于神经网络中训练参数集的系统和方法，能够提高神经网络训练过程的可靠性和训练效率。

第一方面，提供了一种用于神经网络中训练参数集的系统，所述系统包括：

15 主控节点集合，所述主控节点集合包括 M 个主控节点，所述主控节点集合用于控制所述神经网络中训练参数集的过程，并用于存储所述训练参数集的过程所使用的数据集和参数集，所述数据集包括多个数据子集，所述参数集包括多个参数子集，所述多个参数子集分别存储于不同的主控节点上，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集，所述 M 个主控节点两两之间通信连接，所述 M 个主控节点中的至少一个主
20 控节点用于备份所述参数集，其中， M 为大于1的正整数；以及

N 个训练节点集合，所述 N 个训练节点集合中的每一个训练节点集合与所述主控节点集合通信连接，所述训练节点集合包括多个训练节点，所述训练节点用于接收所述主控节点集合下发的数据子集和所述参数集，根据接收的所述数据子集和所述参数集，对自身负责的参数子集进行训练，并将训练
25 结果发送给存储所述参数子集的主控节点，其中， N 为大于1的正整数，所述 N 个训练节点集合中的任意两个训练节点集合训练所使用的数据子集不同，所述每个训练节点集合中的所有训练节点所训练的参数子集的合集为所述参数集。

结合第一方面，在第一方面的第一种可能的实现方式中，所述训练结果
30 为训练节点根据接收的所述数据子集和所述参数集，对自身负责的参数子集进行训练得到的自身负责的参数子集的参数变化量，所述主控节点集合中的

主控节点还用于:

接收所述训练节点发送的所述参数变化量;

根据所述参数变化量, 对所述主控节点中存储的参数子集进行更新。

结合第一方面或第一方面的第一种可能的实现方式, 在第一方面的第二
5 种可能的实现方式中, 所述主控节点集合具体用于:

将所述参数集划分为多个参数子集;

将所述多个参数子集分别存储于不同的主控节点上, 其中, 所述主控节点
集合中的所有主控节点存储的参数子集的合集为所述参数集;

根据多个所述参数子集的大小确定所述 N 个训练节点集合中的每个训
10 练节点。

结合第一方面和第一方面的第一种至第二种可能的实现方式中的任一
种可能的实现方式, 在第一方面的第三种可能的实现方式中, 所述主控节点
具体用于:

在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量
15 对所述主控节点中存储的参数子集进行更新;

在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量
对所述主控节点中存储的参数子集进行更新。

结合第一方面和第一方面的第一种至第三种可能的实现方式中的任一
种可能的实现方式, 在第一方面的第四种可能的实现方式中, 所述主控节点
20 集合具体用于:

根据训练结果的准确性, 确定是否停止所述训练参数集的过程。

结合第一方面和第一方面的第一种至第四种可能的实现方式中的任一
种可能的实现方式, 在第一方面的第五种可能的实现方式中, 所述训练节点
还用于:

25 接收所述主控节点集合发送的指令, 停止所述训练参数集的过程。

结合第一方面和第一方面的第一种至第五种可能的实现方式中的任一
种可能的实现方式, 在第一方面的第六种可能的实现方式中, 同一所述训练
节点集合中的训练节点两两之间通信连接。

第二方面, 提供了一种用于神经网络中训练参数集的方法, 所述方法执
30 行于权利要求 1 至 7 中任一项所述的用于神经网络中训练参数集的系统中的
主控节点集合, 所述系统还包括 N 个训练节点集合, 其中, 所述主控节点集

合包括 M 个主控节点，所述 M 个主控节点两两之间通信连接，其中， M 为大于 1 的正整数， N 为大于 1 的正整数，所述方法包括：

所述主控节点集合存储训练所使用的数据集和参数集，所述数据集包括多个数据子集，所述参数集包括多个参数子集，所述多个参数子集分别存储于不同的主控节点上，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集，所述 M 个主控节点中的至少一个主控节点用于备份所述参数集；

所述主控节点集合中的主控节点向负责训练自身存储的参数子集的训练节点下发数据子集和所述参数子集；

10 所述主控节点集合中的主控节点接收所述训练节点发送的训练结果，其中所述训练节点属于训练节点集合，所述训练节点集合与所述主控节点集合通信连接，所述训练节点集合包括多个训练节点，所述训练结果是根据接收的所述主控节点集合下发的数据子集和参数集进行训练得到的。

结合第二方面，在第二方面的第一种可能的实现方式中，所述训练结果为训练节点根据接收的所述主控节点集合下发的所述数据子集和所述参数集，对自身负责的参数子集进行训练得到的参数子集的参数变化量，所述方法还包括：

所述主控节点集合中的主控节点接收所述训练节点发送的所述参数变化量；

20 所述主控节点集合中的主控节点根据所述参数变化量，对所述主控节点中存储的参数子集进行更新。

结合第二方面或第二方面的第一种可能的实现方式，在第二方面的第二种可能的实现方式中，所述主控节点集合存储训练所使用的数据集和参数集，包括：

25 所述主控节点集合将所述参数集划分为多个参数子集；

将所述多个参数子集分别存储于不同的主控节点上，其中，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集；

所述方法还包括：

30 所述主控节点集合根据多个所述参数子集的大小确定所述 N 个训练节点集合中的每个训练节点。

结合第二方面和第二方面的第一种至第二种可能的实现方式中的任一

种可能的实现方式，在第二方面的第三种可能的实现方式中，所述主控节点集合中的主控节点根据所述参数变化量，对所述主控节点中存储的参数子集进行更新，包括：

5 所述主控节点集合中的主控节点在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量对所述主控节点中存储的参数子集进行更新；

所述主控节点集合中的主控节点在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量对所述主控节点中存储的参数子集进行更新。

10 结合第二方面和第二方面的第一种至第三种可能的实现方式中的任何一种可能的实现方式，在第二方面的第四种可能的实现方式中，所述方法还包括：

所述主控节点集合根据训练结果的准确性，确定是否停止所述训练参数集的过程。

15 结合第二方面和第二方面的第一种至第四种可能的实现方式中的任何一种可能的实现方式，在第二方面的第五种可能的实现方式中，一个所述参数子集由至少一个主控节点存储并负责，并对应的由至少两个训练节点负责，所述至少两个训练节点属于不同的训练节点集合，所述多个训练节点集合中的任意两个训练节点集合训练所使用的数据子集不同，所述每个训练节点集
20 合中的所有训练节点所训练的参数子集的合集为所述参数集。

结合第二方面和第二方面的第一种至第五种可能的实现方式中的任何一种可能的实现方式，在第二方面的第六种可能的实现方式中，同一所述训练节点集合中的训练节点两两之间通信连接。

25 基于上述技术方案，本发明实施例提供的用于神经网络中训练参数集的系统和方法，通过由多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程，可以避免当某一个主控节点失效时导致的整个训练失败情况，能够提高训练过程的可靠性，通过配置多个训练节点集合并行地对参数集进行训练，可以提高训练效率。

30 附图说明

为了更清楚地说明本发明实施例的技术方案，下面将对本发明实施例中

所需要使用的附图作简单地介绍，显而易见地，下面所描述的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

5 图 1 是根据本发明实施例的用于神经网络中训练参数集的系统 100 的示意性框图。

图 2 是根据本发明实施例的计算设备的示意性框图。

图 3 是根据本发明实施例的用于神经网络中训练参数集的系统工作流程的示意图。

图 4 是根据本发明实施例的训练过程的示意性流程图。

10 图 5 是根据本发明实施例的用于神经网络中训练参数集的方法的示意性流程图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明的一部分实施例，而不是全部实施例。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的前提下所获得的所有其他实施例，都应属于本发明保护的范围。

图 1 示出了根据本发明实施例的用于神经网络中训练参数集的系统 100 的示意性框图。如图 1 所示，系统 100 包括：

20 主控节点集合 110，该主控节点集合 110 包括 M 个主控节点，该主控节点集合 110 用于控制该神经网络中训练参数集的过程，并用于存储该训练参数集的过程所使用的数据集和参数集，该数据集包括多个数据子集，该参数集包括多个参数子集，该多个参数子集分别存储于不同的主控节点上，该主控节点集合 110 中的所有主控节点存储的参数子集的合集为该参数集，该 M 个主控节点两两之间通信连接，该 M 个主控节点中的至少一个主控节点用于备份该参数集，其中， M 为大于 1 的正整数；以及

30 N 个训练节点集合 120，该 N 个训练节点集合 120 中的每一个训练节点集合与该主控节点集合 110 通信连接，该训练节点集合 120 包括多个训练节点，该训练节点用于接收该主控节点集合 110 下发的数据子集和该参数集，根据接收的该数据子集和该参数集，对自身负责的参数子集进行训练，并将训练结果发送给存储该参数子集的主控节点，其中， N 为大于 1 的正整数，

该N个训练节点集合120中的任意两个训练节点集合120训练所使用的数据子集不同，该每个训练节点集合120中的所有训练节点所训练的参数子集的合集为该参数集。

因此，本发明实施例提供的用于神经网络中训练参数集的系统，通过由多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程，可以避免当某一个主控节点失效时导致的整个训练失败情况，能够提高训练过程的可靠性，通过配置多个训练节点集合并行地对参数集进行训练，可以提高训练效率。

具体而言，训练参数集的系统100包括一个主控节点集合110和至少两个训练节点集合120。主控节点集合110包括至少两个主控节点，主控节点两两之间通信连接，至少一个主控节点用于备份参数集，可以提高训练过程的可靠性。训练节点集合120可以是主控节点集合110根据数据处理的规模和用于形成训练节点集合120的训练节点的性能（如内存大小等）划分的。

本发明实施例的训练参数集的系统100可以应用于神经网络的训练过程。神经网络的训练过程的输入为神经网络函数 $Y=f(X,W)$ 、初始的参数集和训练的数据集 D ，输出为训练后的神经网络的参数集 W 。主控节点集合110用于控制训练过程，例如控制训练过程的开始或结束，控制各训练节点集合使用的数据子集，以及确定训练节点集合中的每个训练节点等。主控节点集合110还用于存储训练过程的所使用的数据集 D 和参数集 W 。参数集 W 包括多个参数子集，多个参数子集分别存储于不同的主控节点上，该主控节点集合110中的所有主控节点存储的参数子集的合集为该参数集 W 。

训练节点集合120中的训练节点用于接收该主控节点集合110下发的数据子集和当前的参数集 W ，根据接收的数据子集和当前的参数集 W 对自身负责的参数子集进行训练，并将根据该数据子集和当前的参数集 W 训练可以得到的用于更新的参数变化量 ΔW 发送给主控节点。训练过程中，该N个训练节点集合120中的任意两个训练节点集合120训练所使用的数据子集不同，该每个训练节点集合120中的所有训练节点所训练的参数子集的合集为该参数集。即，多个训练节点集合120并行处理不同的数据子集，对同一参数子集而言，同一时刻有多个训练节点对其进行训练，可以提高训练过程的效率。

因此，本发明实施例提供的用于神经网络中训练参数集的系统，通过由多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程，可

以避免当某一个主控节点失效时导致的整个训练失败情况，能够提高训练过程的可靠性，通过配置多个训练节点集合并行地对参数集进行训练，可以提高训练效率。

在本发明实施例中，数据集包括多个数据子集，参数集包括多个参数子集，该 N 个训练节点集合 120 中的任意两个训练节点集合 120 训练所使用的数据子集不同，至少存在两个训练节点训练同一个参数子集，该两个训练节点属于不同的训练节点集合 120。

具体而言，训练参数集的系统 100 包括多于一个训练节点集合 120。此时，主控节点集合 110 存储的数据集包括多个数据子集，训练时主控节点集合 110 将不同的数据子集下发给不同的训练节点集合 120。主控节点集合 110 存储的参数集包括多个参数子集，主控节点集合 110 中的主控节点分别存储和负责维护不同的参数子集。训练节点集合 120 中负责某一参数子集的训练节点从相应的主控节点接收其存储和负责维护的该参数子集，从多个主控节点接收的参数子集的合集为参数集。根据数据子集和参数集训练自身负责的参数子集。其中，至少存在两个训练节点训练同一个参数子集，这两个训练节点属于不同的训练节点集合 120。即，当有多个训练节点集合 120 时，多个训练节点集合 120 并行处理不同的数据子集，对同一参数子集而言，同一时刻有多个训练节点对其进行训练，可以提高训练过程的效率。

应理解，图 1 中示出系统 100 中的主控节点集合 110 中主控节点的个数、训练节点集合 120 的个数以及训练节点集合 120 中训练节点的个数均为示意性的。主控节点集合 110 中包括多于 1 个主控节点，系统 100 中包括至少两个训练节点集合 120，训练节点集合 120 中包括多于 1 个训练节点。

因此，本发明实施例提供的用于神经网络中训练参数集的系统，通过由多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程，可以避免当某一个主控节点失效时导致的整个训练失败情况，能够提高训练过程的可靠性。并且，通过配置多个训练节点集合并行地对参数集进行训练，可以提高训练的效率。

主控节点集合 110 中的主控节点和训练节点集合 120 中的训练节点均为计算设备。图 2 示出了根据本发明实施例的计算设备的示意性框图。如图 2 所示，计算设备可以包含处理模块、存储模块、用于计算的协处理模块（例如，图形处理器（Graphic Processing Unit, GPU）、英特尔超多核心（Intel Many

Integrated Core, Intel MIC) 处理器、现场可编程门阵列 (Field-Programmable Gate Array, FPGA) 等) 和用于在训练节点和与主控节点进行通信或者在主控节点集合 110 内部通信的通信模块。

5 可选地, 作为一个实施例, 在同一时刻, 该 N 个训练节点集合 120 中的至少一个训练节点集合 120 训练所使用的参数集与当前该主控节点集合 110 中存储的参数集不同。

或者, 可选地, 作为一个实施例, 主控节点具体用于:

在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量对该主控节点中存储的参数子集进行更新;

10 在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量对还主控节点中存储的参数子集进行更新。

具体而言, 系统 100 中的每个训练节点集合 120 均独立并行地运作, 互不影响。任何一个训练节点集合 120 的失效, 不影响整个系统 100 继续进行训练。在训练过程中的某一时刻, N 个训练节点集合 120 中的至少一个训练节点集合 120 计算所使用的参数集与当前该主控节点集合 110 中存储的该参数集不同。或者说, 在训练过程中的某一时刻, N 个训练节点集合 120 中的至少一个训练节点集合 120 训练所使用的参数集与其它的训练节点集合 120 训练所使用的参数集不同。即, 主控节点集合 110 对参数集 W 的更新是异步的, 主控节点在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量对该主控节点中存储的参数子集进行更新; 在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量对该主控节点中存储的参数子集进行更新。在某一时刻, 主控节点集合 110 当前的参数集 W 可能已经和训练节点集合 120 正在训练所使用的参数集 W 不同了。

可选地, 作为一个实施例, 主控节点集合 110 具体可以用于:

25 将该参数集划分为多个参数子集;

将该多个参数子集分别存储于不同的主控节点上, 其中, 该主控节点集合 110 中的所有主控节点存储的参数子集的合集为该参数集;

根据多个该参数子集的大小确定该 N 个训练节点集合 120 中的每个训练节点。

30 具体而言, 主控节点集合 110 在训练的最开始进行初始化工作, 例如, 划分训练节点集合 120、配置训练的数据集和参数集、初始化原始模型等等。

其中配置训练的参数集 W 具体为，将参数集 W 划分为多个参数子集 W_1, W_2, \dots, W_k 。每个主控节点负责维护一个或多个参数子集。如果参数子集 W_i 由主控节点 M_j 负责存储、更新和维护，则称 M_j 是 W_i 的宿主节点。

5 根据参数集 W 的大小以及每个用于形成训练节点集合 120 的训练节点的内存（或者训练节点的协处理器的内存）大小，主控节点集合 110 对所有的用于形成训练节点集合 120 的训练节点进行划分。通常而言，参数子集的大小越大，则需要为其分配的训练节点的能力应越强。假设共有 P 个训练节点集合 120，记为 C^1, C^2, \dots, C^P 。每个训练节点负责至少一个参数子集，每个训练节点集合 120 协同存储和处理参数集 W 的一个完整副本。

10 可选地，作为一个实施例，主控节点集合 110 采用磁盘阵列 RAID0/1/5/6 或者纠删码对参数集进行备份。

具体而言，为了保证系统 100 的可靠性，主控节点集合 110 可以采用 RAID0/1/5/6 或者纠删码（Erasure Coding）的编码方法对参数集进行备份。这样，在某些主控节点失效的情况下，系统 100 可以通过相应的解码运算来恢复失效的参数子集而维持正常运作。应理解，还可以采用其它的编码方法来保证系统 100 的可靠性，本发明实施例对此不作限定。

可选地，作为一个实施例，训练节点具体可以用于：

接收该主控节点集合 110 发送的指令，停止该训练参数集的过程。

20 具体而言，对于训练节点集合 C^k 的训练节点需要访问其负责的参数子集的宿主节点，并下载最新的参数子集的副本。训练节点集合 C^k 的所有的训练节点通过通信网络获取的所有最新的参数子集的合集即为最新参数集，记作 W^k 。不同的训练节点集合可能会在不同的时刻从主控节点集合 110 获取最新的参数集 W ，而参数集 W 是不断变化的。因此，在同一时刻，不同的训练节点集合计算所使用的参数集 W 的副本可能是不同的。

25 要进行训练，训练节点集合 C^k 的训练节点还需从该主控节点集合 110 获取数据集的一部分数据，即数据子集，其中，同一训练节点集合中的训练节点所获取的数据子集相同。进而，训练节点根据该参数集 W^k 和该数据子集，进行训练，以获得自身负责的参数子集 W_i 对应的参数变化量 ΔW_i^k 。训练节点将训练得到的参数变化量 ΔW_i^k 发送给负责对应的参数子集 W_i 的主控节点，即宿主节点。训练节点集合 C^k 的所有的训练节点计算得到的参数

变化量 ΔW_i^k 合集记为 ΔW^k 。对于训练节点从主控节点集合 110 获取参数子集和数据的方式，本发明实施例不作限定。

5 在训练过程中，训练节点不断地重复接收参数集、接收数据子集的进行训练，直至从主控节点集合 110 接收到主控节点集合 110 发送的停止训练的指令，训练节点停止训练参数集的过程。

可选地，作为一个实施例，如果在训练的过程中，训练节点集合中的训练节点之间的参数是相互关联的，则训练节点之间需要进行必要的数据交换，此时，同一训练节点集合中的训练节点两两之间可以通信连接。

10 可选地，作为一个实施例，训练结果为训练节点根据接收的该数据子集和该参数集，对训练节点自身负责的参数子集进行训练得到的自身负责的参数子集的参数变化量，该主控节点集合 110 中的主控节点还用于：

接收该训练节点发送的该参数变化量；

根据该参数变化量，对该主控节点中存储的参数子集进行更新。

15 具体而言，主控节点集合 110 中的主控节点从某个训练节点集合 C^k 的训练节点接收该训练节点根据该数据集和该参数集训练得到的用于更新的参数变化量 ΔW_i^k ，从而对主控节点集合的该主控节点负责的参数子集 W_i 进行更新。亦即，主控节点集合从某个训练节点集合 C^k 接收到完整的参数集变化量 ΔW^k 后，对神经网络的参数集 W 进行更新。主控节点集合对参数集 W 的更新是异步的，也就是说，在同一时刻，主控节点集合当前的参数集 W 可能
20 已经和训练节点集合 C^k 在训练过程中使用的参数集 W^k 不同。这种异步的更新方式可以充分的利用所有训练节点集合的训练能力。此外，本发明实施例对主控节点集合对参数集 W 的具体更新方法不作限定。

可选地，作为一个实施例，主控节点集合具体用于：

根据训练结果的准确性，确定是否停止训练参数集的过程。

25 具体而言，主控节点集合 110 确根据训练结果是否准确，确定是否应停止当前的训练。例如，主控节点集合 110 可以在当参数集 W 的变化 ΔW^k 小于一定的阈值时，确定停止训练过程；或者，当更新的参数集 W 使得根据参数集 W 和神经网络的数学函数 $Y = f(X, W)$ 计算得到的结果 Y 的变化值小于一定的阈值时，确定停止训练过程，本发明实施例对此不作限定。

30 下面将结合具体的例子对本发明实施例提供的系统 100 的工作流程进行详细说明。

本发明实施例提供的系统 100 应用于基于深层卷积神经网络的图像分类系统，并使用基于小批量随机梯度下降 (Mini-batch Stochastic Gradient Descent) 的优化算法进行训练。该深层卷积神经网络的输入 X 为图像输出 Y 为图像类别，训练过程的数据集 $D = \{(X_i, Y_i)\}$ 。卷积神经网络的参数集为 W ，系统训练的参数集包括的参数为小批量的大小 m 和学习率 α 。图 3 是根据本发明实施例的数据处理的系统工作流程的示意图。深层卷积神经网络的参数集为 W 被分成两个参数子集 W_1 和 W_2 。主控节点集合包括三个主控节点 M_1 ， M_2 ，和 M_3 。主控节点 M_1 是参数子集 W_1 的宿主节点，主控节点 M_2 是参数子集 W_2 的宿主节点，主控节点 M_3 保存 $W_3 = W_1 \oplus W_2$ 。本发明实施例中 \oplus 表示异或训练。每个训练节点集合 C^k 包括两个训练节点 C_1^k 和 C_2^k ，分别用来负责参数子集 W_1 和 W_2 的训练。

图 4 是根据本发明实施例的训练过程 200 的示意性流程图。训练过程 200 包括：

210，系统 100 包括 P 个训练节点集合 C^k ($k=1, 2, \dots, P$)。训练节点 C_1^k 和 C_2^k 分别从主控节点 M_1 和 M_2 下载最新的参数子集 W_1 和 W_2 ，记作 W_1^k 和 W_2^k 。如果主控节点 M_1 失效，训练节点 C_1^k 可以从主控节点 M_2 下载参数子集 W_2 ，从主控节点 M_3 下载 $W_3 = W_1 \oplus W_2$ ，然后通过训练 $W_2 \oplus W_3$ 得到参数子集 W_1^k 。如果主控节点 M_2 失效，训练节点 C_2^k 可以从主控节点 M_1 下载参数子集 W_1 ，从主控节点 M_3 下载 $W_3 = W_1 \oplus W_2$ ，然后通过训练 $W_1 \oplus W_3$ 得到参数子集 W_2^k 。

220，训练节点 C_1^k 和 C_2^k 都从主控节点集合接收同一批训练数据 $\{D_i^k = (X_i^k, Y_i^k) | i=1, 2, \dots, m\}$ ，分别基于参数子集 W_1^k 和 W_2^k 进行正向传递训练。训练过程中训练节点 C_1^k 和 C_2^k 可以相互通信，以进行必要的交换。

230，对于每一个训练数据 $D_i^k = (X_i^k, Y_i^k)$ ， $i=1, 2, \dots, m$ ，训练节点 C_1^k 和 C_2^k 分别训练出其对应的误差 $E_i^k = \frac{1}{2}(Y_i^k - f(X_i^k, W^k))^2$ ， $i=1, 2, \dots, m$ ，然后通过误差反向传播 (Error Back Propagation, BP) 算法进行反向传递训练，分别训练出 $\Delta W_{i,1}^k = \frac{\partial E_i^k}{\partial W_1^k}$ ， $\Delta W_{i,2}^k = \frac{\partial E_i^k}{\partial W_2^k}$ 。训练过程中，训练节点 C_1^k 和 C_2^k 可以相互通信，以进行必要的交换。

240, 训练节点 C_1^k 和 C_2^k 分别求得参数变化量 $\Delta W_1^k = \frac{1}{m} \sum_{i=1}^m \Delta W_{i,1}^k$,

$$\Delta W_2^k = \frac{1}{m} \sum_{i=1}^m \Delta W_{i,2}^k。$$

250, 训练节点 C_1^k 和 C_2^k 分别把 ΔW_1^k 和 ΔW_2^k 上传到主控节点 M_1 和 M_2 。
训练节点 C_1^k 和 C_2^k 重复步骤 210 至 250, 直至从主控节点集合接收到终止训
5 练的指令。

260, 主控节点集合包括主控节点 M_1 和 M_2 , 步骤 260 与步骤 210 至 250
并行进行。主控节点 M_1 和 M_2 分别从训练节点集合的训练节点 C_1^k 和 C_2^k 接收
参数变化量 ΔW_1^k 和 ΔW_2^k 。根据参数变化量 ΔW_1^k 和 ΔW_2^k , 主控节点 M_1 和 M_2
根据以下公式分别对参数子集 W_1 和 W_2 进行更新: $W_1 = W_1 - \alpha \Delta W_1^k$,
10 $W_2 = W_2 - \alpha \Delta W_2^k$ 。主控节点 M_1 和 M_2 将更新过的参数子集 W_1 和 W_2 传输给主
控节点 M_3 。主控节点 M_3 根据以下公式对 W_3 进行更新: $W_3 = W_1 \oplus W_2$ 。

270, 主控节点集合根据训练结果的准确性, 确定是否停止训练过程。
如果不满足训练停止条件, 重复步骤 210 至 270; 如果满足训练停止条件,
执行步骤 280。

15 280, 主控节点集合向训练节点集合发出终止训练的指令。

因此, 本发明实施例提供的用于神经网络中训练参数集的系统, 通过由
多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程, 可
以避免当某一个主控节点失效时导致的整个训练失败情况, 能够提高训练过
程的可靠性, 并且, 通过配置多个训练节点集合并行地对参数集进行训练,
20 可以提高训练的效率。

下面对对应于本发明实施例的用于神经网络中训练参数集的方法 200 进行
详细的说明。

图 5 示出了根据本发明实施例的用于神经网络中训练参数集的方法
300, 该方法 300 执行于上述用于神经网络中训练参数集的系统中的主控节
25 点集合, 该系统还包括 N 个训练节点集合, 其中, 该主控节点集合包括 M

个主控节点，该 M 个主控节点两两之间通信连接，其中， M 为大于 1 的正整数， N 为大于 1 的正整数，该方法 300 包括：

S310，该主控节点集合存储训练所使用的数据集和参数集，该数据集包括多个数据子集，该参数集包括多个参数子集，该多个参数子集分别存储于不同的主控节点上，该主控节点集合中的所有主控节点存储的参数子集的合集为该参数集，该 M 个主控节点中的至少一个主控节点用于备份该参数集；

S320，该主控节点集合中的主控节点向负责训练自身存储的参数子集的训练节点下发数据子集和该参数子集；

S330，该主控节点集合中的主控节点接收该训练节点发送的训练结果，其中该训练节点属于训练节点集合，该训练节点集合与该主控节点集合通信连接，该训练节点集合包括多个训练节点，该训练结果是根据接收的该主控节点集合下发的数据子集和参数集进行训练得到的。

因此，本发明实施例提供的用于神经网络中训练参数集的方法，通过由多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程，可以避免当某一个主控节点失效时导致的整个训练失败情况，能够提高训练过程的可靠性，通过多个训练节点集合并行地对参数集进行训练，可以提高训练效率。

可选地，作为一个实施例，该训练结果为训练节点根据接收的该主控节点集合下发的该数据子集和该参数集，对自身负责的参数子集进行训练得到的参数子集的参数变化量，该方法 300 还包括：

该主控节点集合中的主控节点接收该训练节点发送的该参数变化量；

该主控节点集合中的主控节点根据该参数变化量，对该主控节点中存储的参数子集进行更新。

可选地，作为一个实施例，该主控节点集合存储训练所使用的数据集和参数集，包括：

该主控节点集合将该参数集划分为多个参数子集；

将该多个参数子集分别存储于不同的主控节点上，其中，该主控节点集合中的所有主控节点存储的参数子集的合集为该参数集；

该方法 300 还包括：

该主控节点集合根据多个该参数子集的大小确定该 N 个训练节点集合中的每个训练节点。

可选地，作为一个实施例，该主控节点集合中的主控节点根据该参数变化量，对该主控节点中存储的参数子集进行更新，包括：

该主控节点集合中的主控节点在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量对该主控节点中存储的参数子集进行更新；

- 5 该主控节点集合中的主控节点在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量对该主控节点中存储的参数子集进行更新。

可选地，作为一个实施例，该方法 300 还包括：

该主控节点集合根据训练结果的准确性，确定是否停止该训练参数集的过程。

- 10 可选地，作为一个实施例，一个该参数子集由至少一个主控节点存储并负责，并对应的由至少两个训练节点负责，该至少两个训练节点属于不同的训练节点集合，该多个训练节点集合中的任意两个训练节点集合训练所使用的数据子集不同，该每个训练节点集合中的所有训练节点所训练的参数子集的合集为该参数集。

- 15 可选地，作为一个实施例，同一该训练节点集合中的训练节点两两之间通信连接。

- 因此，本发明实施例提供的用于神经网络中训练参数集的方法，通过由多个两两之间通信连接的多个主控节点形成主控节点集合控制训练过程，可以避免当某一个主控节点失效时导致的整个训练失败情况，能够提高训练过程的可靠性，并且，通过配置多个训练节点集合并行进行训练，可以提高训练的效率。
- 20

应理解，在本发明实施例中，“与 X 相应的 Y”表示 Y 与 X 相关联，根据 X 可以确定 Y。但还应理解，根据 X 确定 Y 并不意味着仅仅根据 X 确定 Y，还可以根据 X 和/或其它信息确定 Y。

- 25 本领域普通技术人员可以意识到，结合本文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、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。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发明的范围。
- 30

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，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。

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，所述单元的划分，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。另外，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，也可以是电的，机械的或其它的形式连接。

所述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发明实施例方案的目的。

另外，在本发明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，也可以是两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元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，本发明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，服务器，或者网络设备等等）执行本发明各个实施例所述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（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，仅为本发明的具体实施方式，但本发明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发明揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以权利要求的保护范围为准。

权利要求

1. 一种用于神经网络中训练参数集的系统，其特征在于，所述系统包括：

5 主控节点集合，所述主控节点集合包括 M 个主控节点，所述主控节点集合用于控制所述神经网络中训练参数集的过程，并用于存储所述训练参数集的过程所使用的数据集和参数集，所述数据集包括多个数据子集，所述参数集包括多个参数子集，所述多个参数子集分别存储于不同的主控节点上，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集，所述 M 个主控节点两两之间通信连接，所述 M 个主控节点中的至少一个主控节点用于备份所述参数集，其中， M 为大于 1 的正整数；以及

10 N 个训练节点集合，所述 N 个训练节点集合中的每一个训练节点集合与所述主控节点集合通信连接，所述训练节点集合包括多个训练节点，所述训练节点用于接收所述主控节点集合下发的数据子集和所述参数集，根据接收的所述数据子集和所述参数集，对自身负责的参数子集进行训练，并将训练结果发送给存储所述参数子集的主控节点，其中， N 为大于 1 的正整数，所述 N 个训练节点集合中的任意两个训练节点集合训练所使用的数据子集不同，所述每个训练节点集合中的所有训练节点所训练的参数子集的合集为所述参数集。

2. 根据权利要求 1 所述的系统，其特征在于，所述训练结果为训练节点根据接收的所述数据子集和所述参数集，对自身负责的参数子集进行训练得到的自身负责的参数子集的参数变化量，所述主控节点集合中的主控节点还用于：

接收所述训练节点发送的所述参数变化量；

根据所述参数变化量，对所述主控节点中存储的参数子集进行更新。

25 3. 根据权利要求 1 或 2 所述的系统，其特征在于，所述主控节点集合具体用于：

将所述参数集划分为多个参数子集；

将所述多个参数子集分别存储于不同的主控节点上，其中，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集；

30 根据多个所述参数子集的大小确定所述 N 个训练节点集合中的每个训练节点。

4. 根据权利要求 1 至 3 中任一项所述的系统，其特征在于，所述主控节点具体用于：

在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量对所述主控节点中存储的参数子集进行更新；

5 在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量对所述主控节点中存储的参数子集进行更新。

5. 根据权利要求 1 至 4 中任一项所述的系统，其特征在于，所述主控节点集合具体用于：

根据训练结果的准确性，确定是否停止所述训练参数集的过程。

10 6. 根据权利要求 1 至 5 中任一项所述的系统，其特征在于，所述训练节点还用于：

接收所述主控节点集合发送的指令，停止所述训练参数集的过程。

7. 根据权利要求 1 至 6 中任一项所述的系统，其特征在于，同一所述训练节点集合中的训练节点两两之间通信连接。

15 8. 一种用于神经网络中训练参数集的方法，其特征在于，所述方法执行于权利要求 1 至 7 中任一项所述的用于神经网络中训练参数集的系统中的主控节点集合，所述系统还包括 N 个训练节点集合，其中，所述主控节点集合包括 M 个主控节点，所述 M 个主控节点两两之间通信连接，其中，M 为大于 1 的正整数，N 为大于 1 的正整数，所述方法包括：

20 所述主控节点集合存储训练所使用的数据集和参数集，所述数据集包括多个数据子集，所述参数集包括多个参数子集，所述多个参数子集分别存储于不同的主控节点上，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集，所述 M 个主控节点中的至少一个主控节点用于备份所述参数集；

25 所述主控节点集合中的主控节点向负责训练自身存储的参数子集的训练节点下发数据子集和所述参数子集；

所述主控节点集合中的主控节点接收所述训练节点发送的训练结果，其中所述训练节点属于训练节点集合，所述训练节点集合与所述主控节点集合通信连接，所述训练节点集合包括多个训练节点，所述训练结果是根据接收的所述主控节点集合下发的数据子集和参数集进行训练得到的。

30 9. 根据权利要求 8 所述的方法，其特征在于，所述训练结果为训练节

点根据接收的所述主控节点集合下发的所述数据子集和所述参数集，对自身负责的参数子集进行训练得到的参数子集的参数变化量，所述方法还包括：

所述主控节点集合中的主控节点接收所述训练节点发送的所述参数变化量；

5 所述主控节点集合中的主控节点根据所述参数变化量，对所述主控节点中存储的参数子集进行更新。

10. 根据权利要求 8 或 9 所述的方法，其特征在于，所述主控节点集合存储训练所使用的数据集和参数集，包括：

所述主控节点集合将所述参数集划分为多个参数子集；

10 将所述多个参数子集分别存储于不同的主控节点上，其中，所述主控节点集合中的所有主控节点存储的参数子集的合集为所述参数集；

所述方法还包括：

所述主控节点集合根据多个所述参数子集的大小确定所述 N 个训练节点集合中的每个训练节点。

15 11. 根据权利要求 8 至 10 中任一项所述的方法，其特征在于，所述主控节点集合中的主控节点根据所述参数变化量，对所述主控节点中存储的参数子集进行更新，包括：

所述主控节点集合中的主控节点在第一时刻根据第一训练节点集合的第一训练节点发送的参数变化量对所述主控节点中存储的参数子集进行更新；

20 所述主控节点集合中的主控节点在第二时刻根据第二训练节点集合的第二训练节点发送的参数变化量对所述主控节点中存储的参数子集进行更新。

25 12. 根据权利要求 8 至 11 中任一项所述的方法，其特征在于，所述方法还包括：

所述主控节点集合根据训练结果的准确性，确定是否停止所述训练参数集的过程。

30 13. 根据权利要求 8 至 12 中任一项所述的方法，其特征在于，一个所述参数子集由至少一个主控节点存储并负责，并对应的由至少两个训练节点负责，所述至少两个训练节点属于不同的训练节点集合，所述多个训练节点集合中的任意两个训练节点集合训练所使用的数据子集不同，所述每个训练

节点集合中的所有训练节点所训练的参数子集的合集为所述参数集。

14. 根据权利要求 8 至 13 中任一项所述的方法，其特征在于，同一所述训练节点集合中的训练节点两两之间通信连接。

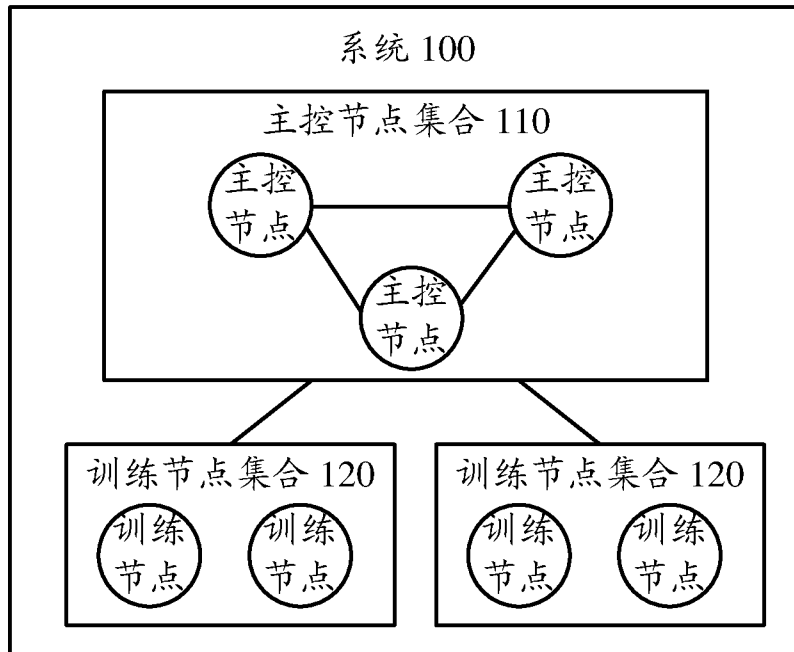


图 1

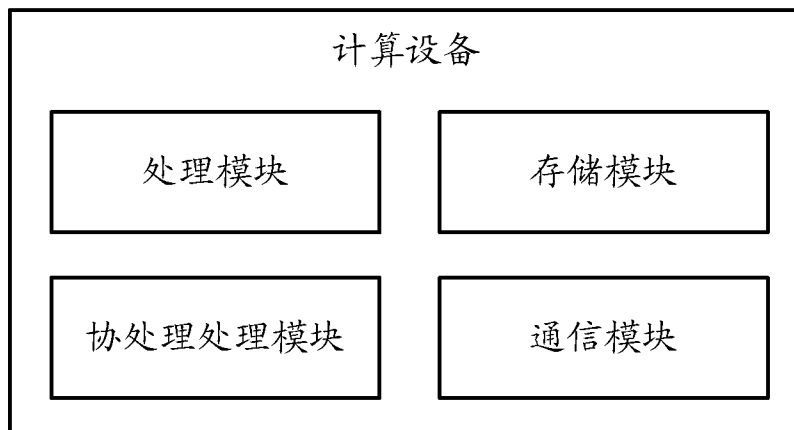


图 2

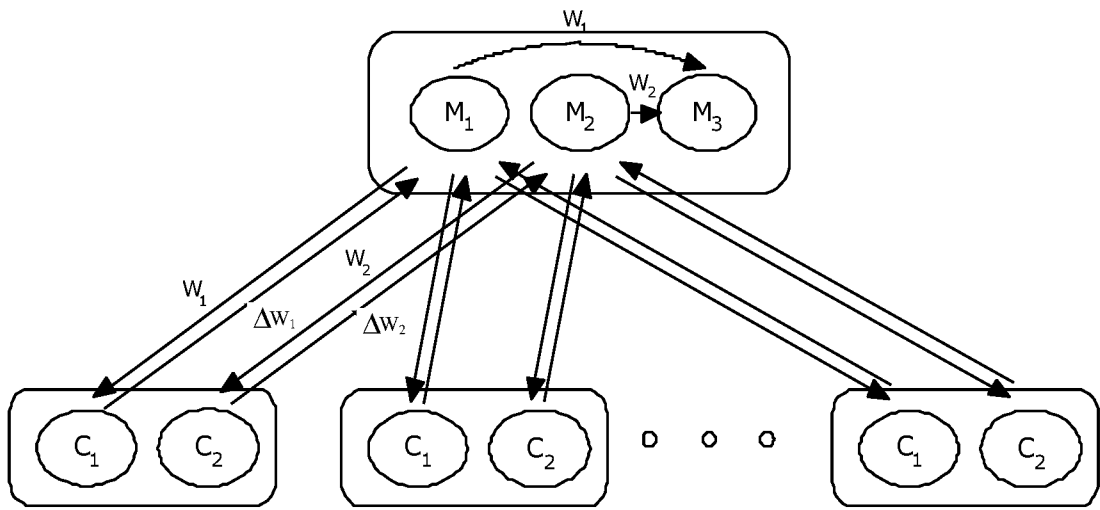


图 3

200

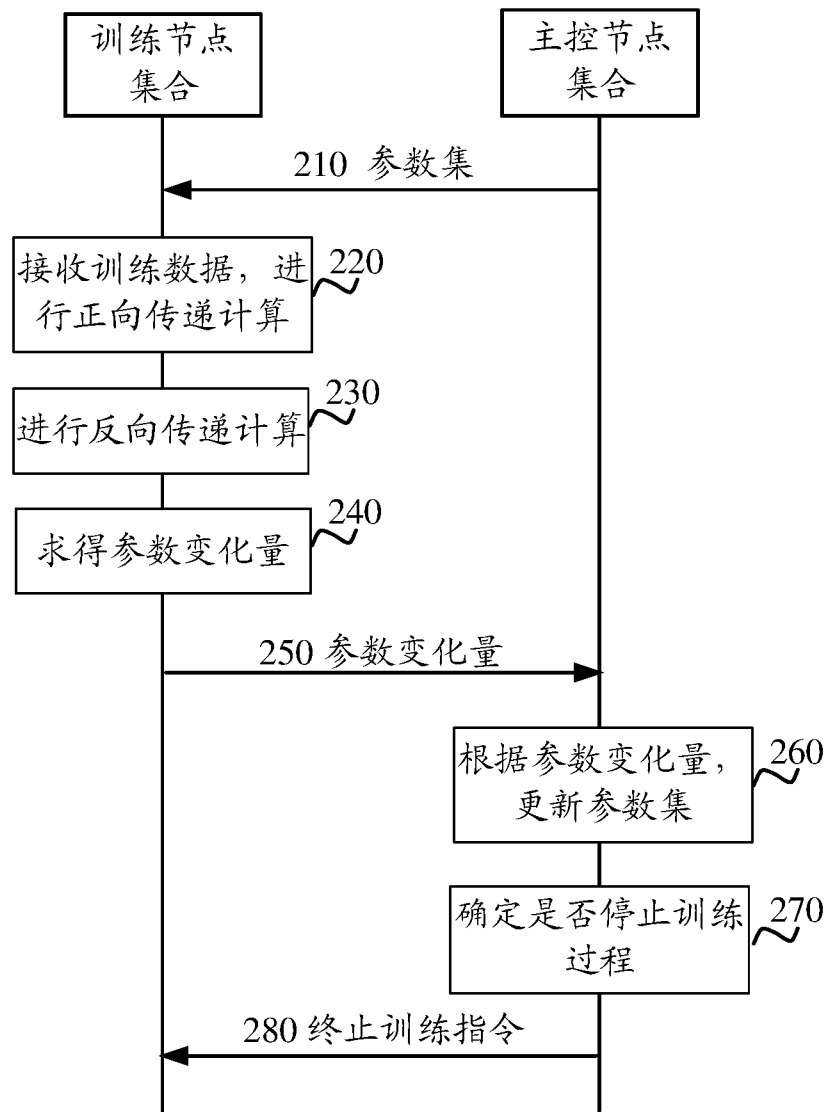


图 4

300

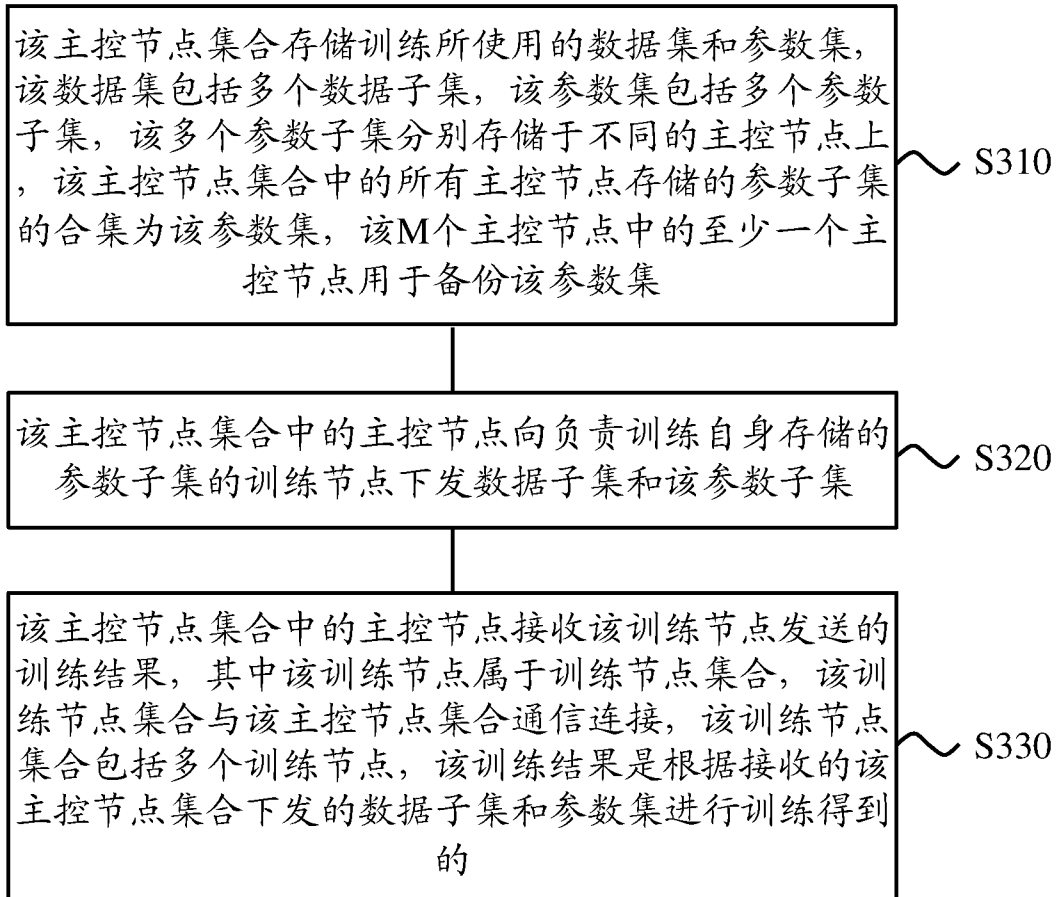


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/086011

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/08 (2006.01) i

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)

CNPAT, CNKI, WPI, EPODOC, GOOGLE: neural network, train, parameter, subset, group, main node, coprocessor, parallel, educate, subclass, sub node

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 104463322 A (INSPUR BEIJING ELECTRIC INFORMATION INDUSTRY CO., LTD.) 25 March 2015 (25.03.2015) description, paragraphs [0030] to [0035]	1-14
PA	CN 104463324 A (CHANGSHA MASHA ELECTRONIC TECHNOLOGY CO., LTD) 25 March 2015 (25.03.2015) the whole document	1-14
Y	CN 104035751 A (TENCENT INC.) 10 September 2014 (10.09.2014) description, paragraphs [0072] to [0095] and figures 4 and 5	1-14
Y	CN 102735747 A (UNIV NANJING AERONAUTICS) 17 October 2012 (17.10.2012) description, paragraphs [0012] to [0019] and figure 2	1-14
A	CN 103077347 A (CHINA ELECTRIC POWER RES INST et al.) 01 May 2013 (01.05.2013) the whole document	1-14

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
26 September 2015

Date of mailing of the international search report
28 October 2015

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10) 62019451

Authorized officer

ZHANG, Xia

Telephone No. (86-10) 61648105

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CN2015/086011

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2004133531 A1 (CHEN, DINGDING et al.) 08 July 2004 (08.07.2004) the whole document	1-14
A	WO 0126026 A2 (UNIVERSITY OF FLORIDA) 12 April 2001 (12.04.2001) the whole document	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/086011

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104463322 A	25 March 2015	None	
CN 104463324 A	25 March 2015	None	
CN 104035751 A	10 September 2014	None	
CN 102735747 A	17 October 2012	None	
CN 103077347 A	01 May 2013	None	
US 2004133531 A1	08 July 2004	CA 2512640 A1	29 July 2004
		GB 2411503 A	31 August 2005
		BR 0317938 A	29 November 2005
		AU 2003299896 A1	10 August 2004
		WO 2004063769 A2	29 July 2004
		MX PA05007294 A	30 September 2005
		WO 0126026 A2	12 April 2001

国际检索报告

国际申请号

PCT/CN2015/086011

<p>A. 主题的分类</p> <p>G06N 3/08(2006.01) i</p> <p>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</p>																																				
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06N</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT, CNKI, WPI, EPODOC, GOOGLE: 神经网络, 训练, 参数, 子集, 分组, 主控节点, 主节点, 协处理器, 并行, neural network, educate, parameter, subset, subclass, group, main node, sub node, coprocessor, parallel</p>																																				
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 104463322 A (浪潮北京电子信息产业有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0030]-[0035]段</td> <td>1-14</td> </tr> <tr> <td>PA</td> <td>CN 104463324 A (长沙马沙电子科技有限公司) 2015年 3月 25日 (2015 - 03 - 25) 全文</td> <td>1-14</td> </tr> <tr> <td>Y</td> <td>CN 104035751 A (深圳市腾讯计算机系统有限公司) 2014年 9月 10日 (2014 - 09 - 10) 说明书第[0072]-[0095]段, 图4-5</td> <td>1-14</td> </tr> <tr> <td>Y</td> <td>CN 102735747 A (南京航空航天大学) 2012年 10月 17日 (2012 - 10 - 17) 说明书第[0012]-[0019]段, 图2</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 103077347 A (中国电力科学研究院 等) 2013年 5月 1日 (2013 - 05 - 01) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>US 2004133531 A1 (CHEN, DINGDING 等) 2004年 7月 8日 (2004 - 07 - 08) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>WO 0126026 A2 (UNIVERSITY OF FLORIDA) 2001年 4月 12日 (2001 - 04 - 12) 全文</td> <td>1-14</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 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</p> <table border="1"> <tr> <td>国际检索实际完成的日期</td> <td>国际检索报告邮寄日期</td> </tr> <tr> <td>2015年 9月 26日</td> <td>2015年 10月 28日</td> </tr> <tr> <td>ISA/CN的名称和邮寄地址</td> <td>受权官员</td> </tr> <tr> <td>中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国</td> <td>张霞</td> </tr> <tr> <td>传真号 (86-10)62019451</td> <td>电话号码 (86-10)61648105</td> </tr> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 104463322 A (浪潮北京电子信息产业有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0030]-[0035]段	1-14	PA	CN 104463324 A (长沙马沙电子科技有限公司) 2015年 3月 25日 (2015 - 03 - 25) 全文	1-14	Y	CN 104035751 A (深圳市腾讯计算机系统有限公司) 2014年 9月 10日 (2014 - 09 - 10) 说明书第[0072]-[0095]段, 图4-5	1-14	Y	CN 102735747 A (南京航空航天大学) 2012年 10月 17日 (2012 - 10 - 17) 说明书第[0012]-[0019]段, 图2	1-14	A	CN 103077347 A (中国电力科学研究院 等) 2013年 5月 1日 (2013 - 05 - 01) 全文	1-14	A	US 2004133531 A1 (CHEN, DINGDING 等) 2004年 7月 8日 (2004 - 07 - 08) 全文	1-14	A	WO 0126026 A2 (UNIVERSITY OF FLORIDA) 2001年 4月 12日 (2001 - 04 - 12) 全文	1-14	国际检索实际完成的日期	国际检索报告邮寄日期	2015年 9月 26日	2015年 10月 28日	ISA/CN的名称和邮寄地址	受权官员	中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国	张霞	传真号 (86-10)62019451	电话号码 (86-10)61648105
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																																		
PX	CN 104463322 A (浪潮北京电子信息产业有限公司) 2015年 3月 25日 (2015 - 03 - 25) 说明书第[0030]-[0035]段	1-14																																		
PA	CN 104463324 A (长沙马沙电子科技有限公司) 2015年 3月 25日 (2015 - 03 - 25) 全文	1-14																																		
Y	CN 104035751 A (深圳市腾讯计算机系统有限公司) 2014年 9月 10日 (2014 - 09 - 10) 说明书第[0072]-[0095]段, 图4-5	1-14																																		
Y	CN 102735747 A (南京航空航天大学) 2012年 10月 17日 (2012 - 10 - 17) 说明书第[0012]-[0019]段, 图2	1-14																																		
A	CN 103077347 A (中国电力科学研究院 等) 2013年 5月 1日 (2013 - 05 - 01) 全文	1-14																																		
A	US 2004133531 A1 (CHEN, DINGDING 等) 2004年 7月 8日 (2004 - 07 - 08) 全文	1-14																																		
A	WO 0126026 A2 (UNIVERSITY OF FLORIDA) 2001年 4月 12日 (2001 - 04 - 12) 全文	1-14																																		
国际检索实际完成的日期	国际检索报告邮寄日期																																			
2015年 9月 26日	2015年 10月 28日																																			
ISA/CN的名称和邮寄地址	受权官员																																			
中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国	张霞																																			
传真号 (86-10)62019451	电话号码 (86-10)61648105																																			

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/086011

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	104463322	A	2015年 3月 25日	无			
CN	104463324	A	2015年 3月 25日	无			
CN	104035751	A	2014年 9月 10日	无			
CN	102735747	A	2012年 10月 17日	无			
CN	103077347	A	2013年 5月 1日	无			
US	2004133531	A1	2004年 7月 8日	CA	2512640	A1	2004年 7月 29日
				GB	2411503	A	2005年 8月 31日
				BR	0317938	A	2005年 11月 29日
				AU	2003299896	A1	2004年 8月 10日
				WO	2004063769	A2	2004年 7月 29日
				MX	PA05007294	A	2005年 9月 30日
WO	0126026	A2	2001年 4月 12日	AU	7749100	A	2001年 5月 10日

表 PCT/ISA/210 (同族专利附件) (2009年7月)