(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
2025 年 6 月 5 日 (05.06.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/112801 A1

(51) 国际专利分类号:
G06N 3/098 (2023.01)

(21) 国际申请号: PCT/CN2024/118478

(22) 国际申请日: 2024 年 9 月 12 日 (12.09.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311636365.9 2023年11月30日 (30.11.2023) CN

(71) 申请人: 杭州阿里云飞天信息技术有限公司(**HANGZHOU ALICLOUD APSARA INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国浙江省杭州市余杭区五常街道文一西路969号3幢5层553室 311121 (CN)。

(72) 发明人: 林哲宇(**LIN, Zheyu**); 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。 赵汉宇(**ZHAO, Hanyu**); 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。 肖文聪(**XIAO, Wencong**); 中国北京市朝阳区广善路18号院-阿里巴巴北京朝阳科技园C区 100020 (CN)。 李永(**LI, Yong**); 中国北京市朝阳区广善路18号院-阿里巴巴北京朝阳科技园C区 100020 (CN)。

(74) 代理人: 北京同钧律师事务所(**BEIJING TONGJUN LAW FIRM**); 中国北京市海淀区西直门北大街32号院1号楼7层808室 100082 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI,

(54) **Title:** DEEP LEARNING MODEL TRAINING METHOD AND DEEP LEARNING MODEL TRAINING SYSTEM

(54) 发明名称: 深度学习模型训练方法和深度学习模型训练系统

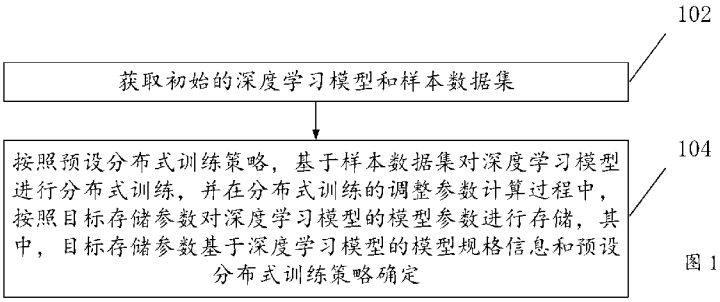

图 1

102 Acquire an initial deep learning model and a sample data set

104 Perform distributed training on the deep learning model on the basis of the sample data set according to a preset distributed training policy, and during the adjustment parameter calculation for distributed training, store model parameters of the deep learning model on the basis of target storage parameters, wherein the target storage parameters are determined on the basis of model specification information of the deep learning model and the preset distributed training policy

(57) **Abstract:** Embodiments of the present disclosure provide a deep learning model training method and a deep learning model training system. The deep learning model training method comprises: acquiring an initial deep learning model and a sample data set; and performing distributed training on the deep learning model on the basis of the sample data set according to a preset distributed training policy, and during the adjustment parameter calculation for distributed training, storing model parameters of the deep learning model on the basis of target storage parameters, wherein the target storage parameters are determined on the basis of model specification information of the deep learning model and the preset distributed training policy. The target storage parameters are determined on the basis of the model specification information of the deep learning model and the preset distributed training policy, which fully takes into account the iteration patterns of distributed training, and during the adjustment parameter calculation, the model parameters of the deep

WO 2025/112801 A1

[见续页]

GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。

learning model are stored on the basis of the target storage parameters, so that high efficiency is achieved while enabling the training of the deep learning model to have high fault tolerance.

(57) 摘要: 本公开实施例提供深度学习模型训练方法和深度学习模型训练系统，其中所述深度学习模型训练方法包括：获取初始的深度学习模型和样本数据集；按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在调整参数计算过程中，按照目标存储参数，完成了对深度学习模型的模型参数的存储，在使得深度学习模型的训练具备高容错能力的同时，具有高效率。

说 明 书

深度学习模型训练方法和深度学习模型训练系统

本公开要求于2023年11月30日提交中国专利局、申请号为2023116363659、申请名称为“深度学习模型训练方法和深度学习模型训练系统”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开实施例涉及深度学习技术领域，特别涉及一种深度学习模型训练方法和深度学习模型训练系统。

背景技术

随着深度学习技术的发展，以大语言模型为代表的大规模深度学习模型在不同场景任务中得到广泛应用。

目前，通过分布式训练，实现对深度学习模型的高效训练，已经成为模型训练的主流方式。然而，在分布式训练过程中，不可避免会出现训练异常，这些训练异常往往导致训练中断，已经更新的模型参数没有得到存储，造成较大的训练进度损失，深度学习模型训练的稳定性不足，而在分布式训练过程中每次迭代后对更新的模型参数进行存储，增加了性能开销，降低了训练效率。

发明内容

有鉴于此，本公开实施例提供了一种深度学习模型训练方法。本公开一个或者多个实施例同时涉及另一种深度学习模型训练方法，一种深度学习模型训练系统，一种深度学习模型训练装置，另一种深度学习模型训练装置，一种计算设备，一种计算机可读存储介质以及一种计算机程序，以解决现有技术中存在的技术缺陷。

根据本公开实施例的第一方面，提供了一种深度学习模型训练方法，包括：

获取初始的深度学习模型和样本数据集；

按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

根据本公开实施例的第二方面，提供了另一种深度学习模型训练方法，应用于云侧设备，云侧设备包括多个分布式节点和存储介质；该方法包括：

获取初始的深度学习模型和样本数据集；

按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定；

在识别到深度学习模型训练异常的情况下，触发多个分布式节点停止对深度学习模型的分布式训练，并确定存储介质当前已存储的目标模型参数；

在接收到恢复训练请求的情况下，从存储介质中获取目标模型参数；

调用多个分布式节点，基于目标模型参数恢复对深度学习模型进行分布式训练。

根据本公开实施例的第三方面，提供了一种深度学习模型训练系统，该系统包括管控单元和多个分布式节点，多个分布式节点包括第一分布式节点，第一分布式节点为多个分布式节点中任一个；

管控单元，用于获取初始的深度学习模型和样本数据集，按照预设分布式训练策略，基于深度学习模型和样本数据集，构建多个分布式数据，将多个分布式数据分发至各分布式节点；

第一分布式节点，用于基于样本数据集对深度学习模型进行分布式训练；并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储其中，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

根据本公开实施例的第四方面，提供了一种深度学习模型训练装置，包括：

第一获取模块，被配置为获取初始的深度学习模型和样本数据集；

第一训练模块，被配置为按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

根据本公开实施例的第五方面，提供了另一种深度学习模型训练装置，应用于云侧设备，云侧设备包括多个分布式节点和存储介质；该装置包括：

第二获取模块，被配置为获取初始的深度学习模型和样本数据集；

第二训练模块，被配置为按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定；

停止模块，被配置为在识别到深度学习模型训练异常的情况下，触发多个分布式节点停止对深度学习模型的分布式训练，并确定存储介质当前已存储的目标模型参数；

参数获取模块，被配置为在接收到恢复训练请求的情况下，从存储介质中获取目标模型参数；

恢复模块，被配置为调用多个分布式节点，基于目标模型参数恢复对深度学习模型进行分布式训练。

根据本公开实施例的第六方面，提供了一种计算设备，包括：

存储器和处理器；

所述存储器用于存储计算机可执行指令，所述处理器用于执行所述计算机可执行指令，该计算机可执行指令被处理器执行时实现上述方法的步骤。

根据本公开实施例的第七方面，提供了一种计算机可读存储介质，其存储有计算机可执行指令，该指令被处理器执行时实现上述方法的步骤。

根据本公开实施例的第八方面，提供了一种计算机程序，其中，当所述计算机程序在

计算机中执行时，令计算机执行上述方法的步骤。

本公开一个实施例中，获取初始的深度学习模型和样本数据集；按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在分布式训练的调整参数计算过程中，按照目标存储参数，完成了对深度学习模型的模型参数的存储，将模型参数的存储过程和分布式训练中的调整参数计算过程重叠，充分节省了性能开销，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率。

附图说明

图 1 是本公开一个实施例提供的一种深度学习模型训练方法的流程图；

图 2 是本公开一个实施例提供的一种深度学习模型训练方法中的系统架构图；

图 3 是本公开一个实施例提供的一种深度学习模型训练方法的流程示意图；

图 4 是本公开一个实施例提供的一种深度学习模型训练方法的原理图；

图 5 是本公开一个实施例提供的另一种深度学习模型训练方法的流程图；

图 6 是本公开一个实施例提供的一种深度学习模型训练系统的结构示意图；

图 7 是本公开一个实施例提供的一种深度学习模型训练装置的结构示意图；

图 8 是本公开一个实施例提供的另一种深度学习模型训练装置的结构示意图；

图 9 是本公开一个实施例提供的一种计算设备的结构框图。

具体实施方式

在下面的描述中阐述了很多具体细节以便于充分理解本公开。但是本公开能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本公开内涵的情况下做类似推广，因此本公开不受下面公开的具体实施的限制。

在本公开一个或多个实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本公开一个或多个实施例。在本公开一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本公开一个或多个实施例中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

应当理解，尽管在本公开一个或多个实施例中可能采用术语第一、第二等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本公开一个或多个实施例范围的情况下，第一也可以被称为第二，类似地，第二也可以被称为第一。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

此外，需要说明的是，本公开一个或多个实施例所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收

集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。

本公开一个或多个实施例中，大模型是指具有大规模模型参数的深度学习模型，通常包含上亿、上百亿、上千亿、上万亿甚至十万亿以上的模型参数。大模型又可以称为基石模型/基础模型（Foundation Model），通过大规模无标注的语料进行大模型的预训练，产出亿级以上参数的预训练模型，这种模型能适应广泛的下游任务，模型具有较好的泛化能力，例如大规模语言模型（Large Language Model，简称 LLM）、多模态预训练模型（multi-modal pre-training model）等。

大模型在实际应用时，仅需少量样本对预训练模型进行微调即可应用于不同的任务中，大模型可以广泛应用于自然语言处理（Natural Language Processing，简称 NLP）、计算机视觉等领域，具体可以应用于如视觉问答（Visual Question Answering，简称 VQA）、图像描述（Image Caption，简称 IC）、图像生成等计算机视觉领域任务，以及基于文本的情感分类、文本摘要生成、机器翻译等自然语言处理领域任务，大模型主要的应用场景包括数字助理、智能机器人、搜索、在线教育、办公软件、电子商务、智能设计等。

首先，对本公开一个或多个实施例涉及的名词术语进行解释。

自然语言处理（Natural Language Processing，简称 NLP）是计算机科学领域与人工智能领域中的一个重要方向，它的目的是使计算机能够理解和使用人类语言，从而执行有用的任务。自然语言处理主要应用于机器翻译、语音识别、文本分析、文本问答等领域。

超参数：是在模型训练过程中计算得到的固定参数，可以理解为一种对模型参数进行更新的策略，用于控制模型参数的更新过程。

梯度权重：是指模型中的梯度与模型参数的关系。梯度描述了损失函数随模型参数变化的方向及其变化率。权重则是指模型参数对损失函数的影响程度。它决定了梯度下降算法更新参数的速度，更大的梯度权重会导致更快的参数更新。

网格搜索调参：是一种常用的超参数调整策略，它通过遍历所有可能的超参数组合，找到目标超参数组合的过程。具体来说，我们首先需要定义一个超参数空间，其中包含了每个超参数的可能取值范围。然后，我们将这些超参数的空间划分为一系列子空间，每个子空间对应一组超参数组合。接着，我们将每一个子空间中的超参数组合都应用于模型，并测量模型的表现。最终，我们会找到目标超参数组合，作为模型参数更新的依据。

贝叶斯优化：是一种常用的超参数调整策略，一种以贝叶斯定理为基础的超参数调整策略。贝叶斯定理是概率论的一种理论，用于估计某一事件的概率，它可以用来推断超参数的目标优化值。在贝叶斯调参中，我们首先定义一个先验分布，它代表了我们对于超参数的知识。然后，我们将先验分布与观测到的实验结果相结合，得到后验分布。通过后验分布，我们可以估算出超参数的目标优化值。

批次（mini-batch）：在深度学习场景中，通过将整个数据集分为若干小的数据集，每一次对小的数据集进行训练，避免了数据集中全部数据一次性参与训练，带来的计算量巨大的问题，另外，在执行梯度更新时，梯度方向和整个数据集不存在较大差别，保证了训练效果。

图形处理单元（Graphics Processing Unit，简称 GPU）：一种用作图形相关计算的微处

理器，随着深度学习技术的发展，因为其并行结构，可以实现高效地矩阵运算，被广泛用作模型训练的计算硬件。

张量处理单元（Tensor Processing Unit，简称 TPU）：一种专门针对深度学习任务的定制化芯片，它可以提供比 GPU 更高的效率和性能。

可编程逻辑门阵列（Field-Programmable Gate Array，简称 FPGA）：一种可编程逻辑门阵列，可以被编程来实现多种功能，包括深度学习中的卷积运算。

专用集成电路（Application Specific Integrated Circuit，简称 ASIC）：一种专门为特定应用而设计的集成电路，可以提供非常高的性能，但成本较高。

中心处理器 （Central Processing Unit，简称 CPU）也可以用于模型训练，但它通常不如 GPU 和 TPU 高效。

深度自注意力模型（Transformer 模型）：一种基于注意力机制（Attention）的深度学习架构，用于处理序列数据，如自然语言。

双向编码表征的深度自注意力模型（Bidirectional Encoder Representations from Transformers，简称 BERT）：一种特殊的 Transformer 模型，使用双向 Transformer 编码器和大规模无标注文本数据进行训练的。BERT 的出色表现使其成为许多自然语言处理任务的标准基线。

大语言模型（Large Language Model，简称 LLM）：在大规模语料库上训练的深度学习模型，用于自然语言处理任务。这些模型一般包含多层神经网络，其输入是一个文本序列，进行文本生成，输出是对该文本序列执行特定自然语言处理任务，生成的任务结果文本。预训练意味着在特定任务之前，模型已经被训练并预先学会处理大量的语言数据。通过预先训练模型，它们可以捕捉到更加复杂的语言和语义规则，从而在各种自然语言处理任务中表现出色，并减少对特定任务的大规模数据需求。

分布式训练：一种利用多个分布式节点来训练深度学习模型的方法，它可以大大提高训练速度并减少计算时间。

模型并行（Model Parallel，简称 MP）：一种分布式训练策略，指将一个模型拆分成多个部分，并将每个部分部署到不同的分布式节点上进行训练。这种方法可以有效地解决模型中参数过多的问题。

数据并行（Data Parallel，简称 DP）：一种分布式训练策略，通过将整个样本集拆分成多个样本子集，并将每个样本子集分配给不同的分布式节点进行训练。这种技术可以提高训练效率，并降低内存占用。

流水式并行（Pipeline Parallel，简称 PP）：一种分布式训练策略，是一种介于模型并行以及数据并行之间的分布式训练策略，核心思想是通过将大模型分解成多个层，并将它们组成一个流水线的方式进行前向传播计算和反向传播计算，从而减少单卡的显存占用，也降低了通信开销。

远距离直接存储读取技术（Remote Direct Memory Access，简称 RDMA）：一种用于远程计算机之间直接读写内存的技术，能够极大地提高数据在网络中的传输速度。

网卡：一种硬件设备，主要用于实现计算机与网络之间的物理连接，并通过发送和接

收数据包来完成数据的传输。

外接元件互连总线拓扑（Peripheral Component Interconnect Express，简称 PCIe）：一组节点之间的物理连接结构，通常由总线构成，可以用于支持多个设备之间的数据传输。

图形处理单元间的高速通道：一种用于连接两个或多个图形处理单元之间的快速通信通道，可以为图形处理单元之间的数据传输提供更高的带宽和更低的延迟。

通信信道：用于传输数据信号的通道，它可以是物理信道或逻辑信道。物理信道是由传输介质和相关通信设备组成的，用于传输实际的数据信号；逻辑信道则是指在物理信道的基础上，通过中间节点实现的逻辑通路，即发送方和接收方之间形成的逻辑通路。

目前，深度学习模型的分布式训练，主要是基于样本集或者深度学习模型的模型参数，构建多个分布式数据，再将多个分布式数据分发至不同的分布式节点，在任一分布式节点上，进行迭代训练，任一次迭代训练过程中，按照传播计算和参数更新的方式，完成对语言模型的训练。

然而，一旦发生训练异常，例如，硬件异常、系统异常、网络异常、或者其他未知异常，在没有对更新的模型参数进行存储的情况下，需要重新执行分布式训练，这对于性能开销较大的深度学习模型是难以承受的，尽管可以在特定检查点（例如，每次迭代过程的参数更新后）对更新的模型参数进行存储，在深度学习模型参数较大的情况下，需要等待模型参数存储完成后，才能继续进行训练。针对模型规格达到百亿迁移级别的深度学习模型，这样的时间开销往往达到几分钟甚至十几分钟，决定了无法频繁进行模型参数存储。在这种情况下，一旦发生训练异常，恢复分布式训练，时间开销可能达到数个小时。因此，迫切需要一种通过较低性能开销，实现在发生训练异常时，预先存储有更新的模型参数，在恢复分布式训练后，无需重新计算，同时具有高稳定性和高效率的深度学习模型训练方法。

针对上述问题，本公开中提供了一种深度学习模型训练方法，本公开同时涉及另一种深度学习模型训练方法，一种深度学习模型训练系统，一种深度学习模型训练装置，另一种深度学习模型训练装置，一种计算设备，一种计算机可读存储介质以及一种计算机程序，在下面的实施例中逐一进行详细说明。

参见图 1，图 1 示出了本公开一个实施例提供的一种深度学习模型训练方法的流程图，包括如下具体步骤：

步骤 102：获取初始的深度学习模型和样本数据集。

本公开实施例应用于具有分布式训练功能的服务端，该服务端上部署有多个分布式节点和存储介质，任一分布式节点包括模型训练的计算硬件，例如，GPU、NPU、FPGA、ASIC 或者 CPU。

深度学习模型是指一类基于深度神经网络结构的机器学习模型，在视觉、语音、自然语言处理等多个领域中广泛应用。深度学习模型具有大规模的模型参数，深度学习模型包括但不限于：语言处理模型、图像处理模型、语音处理模型、代码处理模型等。以语言处理模型为例进行说明，语言处理模型可以执行一种或多种自然语言处理任务，包括但不限于：机器翻译任务、语音识别任务、文本分析任务或者文本问答任务。从功能而言，语言处理模型可以被视作：翻译模型、语音识别模型、文本分析模型和文本问答模型等，在此

不作限定。从模型结构而言，语言处理模型可以为 Transformer 模型、BERT 模型和大语言模型等，在此不作限定。

样本数据集为用于训练的深度学习模型的样本数据的集合，样本数据集包括大规模的样本数据。样本数据集可以为标签样本数据集，也可以为无标签样本数据集。根据训练需求的不同，样本数据可以为不同模态的数据，例如，需要将深度学习模型训练成为具有文本处理功能的模型，文本模态的样本文本，又例如，需要将深度学习模型训练成为具有音频处理功能的模型，音频模态的样本音频，还例如，需要将深度学习模型训练成为具有图像处理功能的模型，图像模态的样本图像，还例如，需要将深度学习模型训练成为具有数值处理功能的模型，数值模态的样本数值。样本数据集可以为从样本数据库中获取得到的，例如，开源样本数据库，也可以为人为构建得到的，例如，利用生成模型生成得到的，还可以为从历史数据库中获取得到，例如，从历史数据库中获取到历史查询文本和历史答案文本，构建得到样本文本集，在此不作限定。

示例性地，在某大语言模型的网站上，目前需要增加虚拟角色对话功能，通过对初始的大语言模型进行训练，使得训练得到的目标大语言模型具有文本问答功能，可以执行文本问答任务。从模型库中获取初始的大语言模型，从开源样本数据库中获取样本文本集，样本文本集包括 10000000 个样本问答文本对。

获取初始的深度学习模型和样本数据集。为后续进行分布式训练奠定了模型基础和样本数据基础。

步骤 104：按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

分布式训练为利用多个分布式节点来训练深度学习模型的训练方法。模型训练为迭代训练的，一次迭代过程包括调整参数计算和模型参数更新。调整参数为对模型参数进行调整的超参数，调整参数计算为对模型参数调整的超参数进行计算，包括但不限于：传播计算、概率计算（通过贝叶斯优化进行模型参数更新）、网格计算（通过网格搜索进行模型参数更新）和扩散推导过程（利用扩散模型进行前向扩散过程和反向推导过程）。

分布式训练策略为采用分布式计算的框架，将深度学习模型的大规模训练任务拆分成多个小规模训练任务，分发到多个分布式节点上执行的策略，包括但不限于：数据并行策略、模型并行策略和流水式并行策略。分布式训练策略是根据深度学习模型的模型属性、样本数据集和/或训练任务来预先设定的，例如，深度学习模型的各模型层之间具有先后执行限定，难以拆分训练，采用数据并行策略，又例如，样本数据集的数据规模过大，采用数据并行策略，还例如，在大规模文本分类的训练任务中，采用数据并行策略，可能造成网络瓶颈，采用模型并行策略，则可能复杂度过高，采用流水式并行策略。不同的分布式训练策略对应模型训练不同的迭代特性，例如，采用数据并行策略，每个分布式节点上的样本数据量小，而模型参数规格大，一次调整参数计算的时间开销高，而次数少，而采用模型并行策略，每个分布式节点上的样本数据量大，而模型参数规格小，一次调整参数计算的时间开销低，而次数多。不同的迭代特征，决定了调整参数计算的时间开销不同，进行存储的时间也不同，需要精细化地确定目标存储参数，将模型参数的存储过程和分布式

训练中的调整参数计算过程重叠。

模型规格信息为深度学习模型的模型参数规格的信息，包括但不限于：模型参数量、基准模型等。模型参数量为模型参数的参数规格量，例如，深度学习模型的模型参数量为百亿规格。基准模型为特定的深度学习模型架构，还例如，在自然语言处理中，深度学习模型的基准模型为 Transformer 模型、BERT 模型和大语言模型等。

本公开实施例中，存储为与调整参数计算过程同步执行的模型参数存储，即存储过程与调整参数计算过程具有偏差不大的时间开销。例如，调整参数计算过程的时间开销为 T，存储过程的时间开销为 T'，如果 T 小于 T'，甚至在模型参数更新过程中，还在执行存储，一旦发生训练异常，难以恢复模型训练，具体参见下述说明。以传播计算为例，存储过程可以在前向传播计算过程中，也可以在反向传播计算过程中，还可以在前向传播计算过程和反向传播计算过程中，在此不做限定。

目标存储参数为对模型参数进行存储的配置参数，目标存储参数是基于深度学习模型的模型规格信息和预设分布式训练策略确定的，包括但不限于：目标存储模型参数规格、目标存储频率、目标存储读写速度、目标存储带宽等。分布式训练策略在确定并执行后，调整参数计算过程的时间开销是难以变更的，为了将模型参数的存储过程和分布式训练中的调整参数计算过程重叠，需要通过目标存储参数，来控制对模型参数进行模型参数存储的时间开销。

需要说明的是，如果在模型参数更新过程中对模型参数进行存储，导致更新的模型参数和未更新的模型参数被混合后存储，一旦发生训练异常，无法通过存储的混合更新的模型参数和未更新的模型参数，恢复模型训练。例如，在完成第 i-1 次迭代训练，当前的模型参数为θ_{i-1}，执行第 i 次迭代训练，在调整参数计算过程中，对当前的模型参数θ_{i-1}进行存储，在完成存储后，如果发生训练异常，可以获取模型参数θ_{i-1}重新执行第 i 次迭代训练。如果在模型参数更新过程中，对当前的模型参数θ_{i-1}进行存储，就会引入部分更新后的模型参数θ_i，进而无法在发生训练异常的情况下，重新执行第 i 次迭代训练。

按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，具体方式为：按照预设分布式训练策略，构建多个分布式数据，将多个分布式数据分发至分布式节点上，执行迭代训练，其中，迭代训练包括调整参数计算和模型参数更新。

按照目标存储参数对深度学习模型的模型参数进行存储，具体方式为：按照目标存储参数，将深度学习模型的模型参数存储至存储介质。其中，存储介质为存储模型参数的硬件设备，包括但不限于：计算硬件缓存（GPU 缓存、NPU 缓存、FPGA 缓存、ASIC 缓存和 CPU 缓存）、内存、硬盘和分布式持久化存储阵列。

基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，具体方式为：基于深度学习模型的模型规格信息和预设分布式训练策略，执行时间开销分析，获得目标存储参数。其中，时间开销分析为分析得到迭代训练中调整参数计算过程和模型参数存储过程的时间开销的操作，进而确定可以实现存储的目标存储参数。

示例性地，预设分布式训练策略为数据并行策略。预先基于大语言模型的模型参数量为10^{13}级别和数据并行策略，执行时间开销分析，确定一次迭代的前向传播计算过程的时间开销 t1 和后向传播计算过程的时间开销 t2，基于该时间开销和模型参数量，获得目标存

储模型参数规格为 P1 和 P2。按照数据并行策略，对样本文本集中的 10000000 个样本问答文本对进行划分。构建 64 个分布式数据，将 64 个分布式数据分发至 64 个分布式节点上，每个分布式节点上部署有 GPU，任一分布式节点上，将任一分布式数据划分为 16 个小批次，利用 GPU 对执行 16 个小批次的迭代训练。在每个迭代训练的前向传播计算过程和反向传播计算过程中，按照目标存储模型参数规格 P1 和 P2，将大语言模型的模型参数θ依次存储至分布式节点的 GPU 缓存、内存和硬盘中。按照上述策略，完成分布式训练后，得到具有文本问答功能的目标大语言模型，将该目标大语言模型部署在该大语言模型的网站的云侧设备上，为用户提供虚拟角色对话功能。

本公开实施例中，获取初始的深度学习模型和样本数据集；按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在分布式训练的调整参数计算过程中，按照目标存储参数，完成了对深度学习模型的模型参数的存储，将模型参数的存储过程和分布式训练中的调整参数计算过程重叠，充分节省了性能开销，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率。

在本公开一种可选实施例中，步骤 104 中按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，包括如下具体步骤：

按照预设分布式训练策略，基于深度学习模型和样本数据集，构建多个分布式数据，其中，预设分布式训练策略包括模型并行训练策略或者数据并行训练策略；

将多个分布式数据分发至各分布式节点；

在第一分布式节点上，基于分布式数据进行传播计算，获得梯度权重，其中，第一分布式节点为多个分布式节点中任一个；

基于各分布式节点上的梯度权重，更新深度学习模型的模型参数，在达到预设训练结束条件的情况下，获得训练完成的深度学习模型。

分布式数据为分发至分布式节点上执行的部分数据，如果将整个模型训练理解为一个大规模训练任务，分布式数据为将该大规模训练任务拆分成得到的小规模训练任务的任务数据。采用不同的分布式训练策略，构建的分布式数据不同，例如，采用模型并行策略，将深度学习模型的模型参数拆分成多个，构建分布式数据，又例如，采用数据并行策略，将样本数据集拆分成多个，构建分布式数据。

传播计算为基于样本数据利用深度学习模型，确定梯度权重的过程，包括前向传播计算过程和反向传播计算过程。前向传播计算为将样本数据输入深度学习模型，输出预测数据的过程。反向传播计算为通过预测数据确定损失值后，反向输入深度学习模型，确定各模型层的梯度权重的过程。模型参数更新为基于梯度权重对模型各模型层的模型参数进行调整的过程。例如，在前向传播计算过程中，将样本数据 X 输入大语言模型，输出预测数据 Z，在反向传播计算过程中，基于预测数据 Z 确定损失值，反向输入大语言模型，确定大语言模型的 n 个模型层的梯度权重$\frac{dL}{dX}$为：$(\frac{dL}{dX})_1$，$(\frac{dL}{dX})_2$，……，$(\frac{dL}{dX})_n$，模型参数更新过程

中，基于梯度权重$\frac{dL}{dX}$，对大语言模型中 n 个模型层的模型参数进行调整。

梯度权重为在反向计算传播过程中，针对每个模型层的模型参数更新分配不同的梯度权重，从而控制每个模型层的参数的更新幅度和模型的参数更新线速度。对每个模型层的参数更新值越大，参数更新得越快，反之亦然。

预设训练结束条件为预先设定的训练停止的判断条件，包括但不限于：预设迭代次数、预设损失值阈值、预设训练时间和预设模型收敛条件。

按照预设分布式训练策略，基于深度学习模型和样本数据集，构建多个分布式数据，具体方式为：按照预设分布式训练策略，对深度学习模型的模型参数和/或样本数据集进行划分，获得多个分布式数据。

基于分布式数据进行传播计算，获得梯度权重，具体方式为：将分布式数据中的样本数据输入深度学习模型，执行前向传播计算，获得预测数据，基于样本数据和预测数据，确定损失值，将损失值反向输入深度学习模型，执行反向传播计算，获得梯度权重。

基于各分布式节点上的梯度权重，更新深度学习模型的模型参数，具体方式为：基于各分布式节点上的梯度权重，通过梯度更新法，更新深度学习模型的模型参数。

示例性地，按照数据并行策略，对样本文本集中的 10000000 个样本问答文本对进行划分，获得 64 个分布式数据。将 64 个分布式数据分发至 64 个分布式节点上，每个分布式节点上部署有 GPU，任一分布式节点上，将任一分布式数据划分为 16 个小批次。任一迭代训练过程中，将小批次的分布式数据中的样本问题文本 X 输入大语言模型，执行前向传播计算，获得预测答案文本Z′，基于样本答案文本 Z 和预测答案文本，确定损失值 Loss，将损失值反向输入大语言模型，执行反向传播计算，获得梯度权重$\frac{dL}{dX}$，基于各分布式节点上的梯度权重，通过梯度更新法，更新大语言模型的模型参数θ，在完成所有小批次的分布式数据训练的情况下，获得训练完成的目标大语言模型，目标大语言模型具有文本问答功能。

按照预设分布式训练策略，基于深度学习模型和样本数据集，构建多个分布式数据，其中，预设分布式训练策略包括模型并行训练策略或者数据并行训练策略；将多个分布式数据分发至各分布式节点；在第一分布式节点上，基于分布式数据进行传播计算，获得梯度权重，其中，第一分布式节点为多个分布式节点中任一个；基于各分布式节点上的梯度权重，更新深度学习模型的模型参数，在达到预设训练结束条件的情况下，获得训练完成的深度学习模型。按照预设分布式训练策略，构建多个分布式数据，在多个分布式节点上执行了分布式训练，得到梯度权重，完成了对深度学习模型的模型参数进行更新，提升了模型训练的效率。

在本公开一种可选实施例中，分布式数据包括多个批次的分布式数据；

在第一分布式节点上，基于分布式数据进行传播计算，获得梯度权重，包括如下具体步骤：

在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重；

对应地，在基于各分布式节点上的梯度权重，更新深度学习模型的模型参数之后，还包括如下具体步骤：

更新当前批次的分布式数据，返回执行在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重的步骤。

当前批次的分布式数据为当前迭代训练过程中，对深度学习模型进行训练的批次的分布式数据，例如，在任一分布式节点上，将分布式数据划分为 16 个批次，需要执行 16 次迭代训练，一次迭代训练过程包括前向传播计算过程、反向传播计算过程和参数更新过程，可选地，在参数更新过程之前，还包括通信过程（整合梯度权重的过程），当前为第 i 次迭代训练过程，当前批次的分布式数据为第 i 批次的分布式数据，当前的模型参数为第 i-1 次更新后的模型参数θ_{i-1}。对应地，更新当前批次的分布式数据，即将第 i 批次的分布式数据更新为第 i+1 批次的分布式数据。

基于当前批次的分布式数据进行传播计算，获得梯度权重，具体方式为：将当前批次的分布式数据中的样本数据输入深度学习模型，执行前向传播计算，获得预测数据，基于样本数据和预测数据，确定损失值，将损失值反向输入深度学习模型，执行反向传播计算，获得梯度权重。

示例性地，将第 i 批次的分布式数据中的样本问题文本X_i输入大语言模型（当前模型参数为第 i-1 次更新后的模型参数θ_{i-1}），执行前向传播计算，获得预测数据预测答案文本Z_i'，基于样本答案文本Z_i和预测答案文本，确定损失值 Loss，将损失值反向输入大语言模型，执行反向传播计算，获得梯度权重$\frac{dL}{dX}$，基于各分布式节点上的梯度权重，通过梯度更新法，将大语言模型的模型参数从θ_{i-1}更新为θ_i，将当前批次的分布式数据从第 i 批次更新为第 i+1 批次，继续执行将第 i+1 批次的分布式数据中的样本问题文本X_{i+1}输入大语言模型的步骤……，在完成所有小批次的分布式数据训练的情况下，获得训练完成的目标大语言模型，目标大语言模型具有文本问答功能。

在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重：更新当前批次的分布式数据，返回执行在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重的步骤。通过将分布式数据划分为多个批次，利用多个批次的分布式数据对深度学习模型进行迭代训练，避免了全部分布式数据一次性参与分布式训练，带来的计算量巨大的问题，造成训练瓶颈的问题，保证了训练效果。

在本公开一种可选实施例中，在第一分布式节点上，对当前批次的分布式数据执行传播计算，获得梯度权重之前，还包括如下具体步骤：

针对各个批次的分布式数据，基于深度学习模型的模型规格信息和预设分布式训练策略，预测传播计算的次数和时间开销；

基于传播计算的次数和时间开销，确定传播计算过程对应的目标存储参数。

本公开实施例中，由于在不同分布式训练策略下，对深度学习模型进行训练过程中，每个批次的分布式数据所需要的前向计算和后向计算的次数和对应的时间开销也是不同的。例如，采用数据并行策略，每个分布式节点上的样本数据量小，而模型参数规格大，一次传播计算的时间开销高，而次数少，而采用模型并行策略，每个分布式节点上的样本数据量大，而模型参数规格小，一次传播计算的时间开销低，而次数多。需要基于传播计算的次数和时间开销，精细化地确定目标存储参数，将每次迭代训练过程中模型参数的存储过程和分布式训练中的传播计算过程重叠。

针对各个批次的分布式数据，基于深度学习模型的模型规格信息和预设分布式训练策略，预测传播计算的次数和时间开销，具体为通过预测算法预测得到，例如，torch.distributed 模块、torch.profiler 工具和 tf.data API 接口等。

基于传播计算的次数和时间开销，确定传播计算过程对应的目标存储参数，具体方式为：以模型参数存储过程的时间开销不超过传播计算过程的时间开销为目标，基于传播计算的次数和时间开销，确定传播计算过程对应的目标存储参数。

示例性地，通过 torch.distributed 模块，针对各个批次的分布式数据，基于大语言模型的模型参数量为10^{13}级别和数据并行策略，预测传播计算的次数为 16 次、前向传播计算过程的时间开销 t1 和后向传播计算过程的时间开销 t2，以模型参数存储过程的时间开销 T' 不超过传播计算过程的时间开销 T=16*（t1+t2）为目标，确定传播计算过程对应的目标存储模型参数规格为 P1 和 P2。

针对各个批次的分布式数据，基于深度学习模型的模型规格信息和预设分布式训练策略，预测传播计算的次数和时间开销；基于传播计算的次数和时间开销，确定传播计算过程对应的目标存储参数。保证了后续对模型参数进行存储的可行性，更为准确地将模型参数的存储过程和分布式训练中的传播计算过程重叠。

在本公开一种可选实施例中，在基于各分布式节点上的梯度权重，更新深度学习模型的模型参数之前，还包括如下具体步骤：

通过各分布式节点之间的通信通道，整合各分布式节点上的梯度权重。

各分布式节点之间的通信通道为各分布式节点之间的数据传输的信道连接，可以是物理连接，例如，网卡、PCIe 拓扑或光纤等，也可以是虚拟连接，例如，GPU 之间的高速通道、各分布式节点之间的高速通道等。通信通道可以通过 RDMA 技术实现，提升了传输速度。

示例性地，通过 64 个分布式节点上的网卡构建的高速通道，利用 RDMA 技术，整合各分布式节点上的梯度权重$\frac{dL}{dX}$，基于各分布式节点上的梯度权重，通过梯度更新法，更新大语言模型的模型参数θ。

通过各分布式节点之间的通信通道，整合各分布式节点上的梯度权重。通过通信通道，集中完成了迭代训练的通信过程，提升了模型训练的效率和稳定性。

在本公开一种可选实施例中，任一分布式节点包括连接存储介质的第一通信通道和与其他分布式节点连接的第二通信通道；

步骤 104 中按照目标存储参数对深度学习模型的模型参数进行存储，包括如下具体步骤：

按照目标存储参数，通过第一通信通道，将深度学习模型的模型参数存储至存储介质；

对应地，通过各分布式节点之间的通信通道，整合各分布式节点上的梯度权重，包括如下具体步骤：

通过第二通信通道，整合各分布式节点上的梯度权重。

在本公开实施例中，存储介质独立于分布式节点，避免任一分布式节点上出现异常，造成数据损失，保证了整个分布式系统的可靠性，所以分布式节点需要和存储介质之间建

立通信通道，进行数据存储，然而，分布式节点上的资源性能有限，在模型参数存储过程中需要用到通信通道，在通信过程中，即整合各个分布式节点上的梯度权重，也需要用到通信通道，这对于大规模的深度学习模型是难以实现的，因而，需要将两个过程的通信通道区分，将两个数据传输过程进行隔离，避免引入额外的集合通信，确保不对分布式训练带来干扰。

连接存储介质的第一通信通道为用于模型参数存储过程的通信通道，包括分布式节点和存储介质之间的物理连接和虚拟连接，例如，网卡、PCIe 拓扑、光纤、分布式节点和存储介质之间的读写通道（数据总线）等。

与其他分布式节点连接的第二通信通道为用于通信过程的通信通道，包括分布式节点和存储介质之间的物理连接和虚拟连接，例如，网卡、PCIe 拓扑、光纤、分布式节点、GPU 之间的高速通道、各分布式节点之间的高速通道等。第二通信通道可以通过 RDMA 技术实现，提升了传输速度。

示例性地，第一通信通道为 64 个分布式节点上的网卡分别与分布式持久化存储阵列之间的物理连接，按照目标存储模型参数规格 P1 和 P2，通过第一通信通道，将大语言模型的模型参数θ存储至分布式持久化存储阵列。第二通信通道为 64 个分布式节点上的网卡之间的物理连接，以及各分布式节点上插有 GPU 的 PCIe 拓扑，配合有 GPU 之间的高速通道和各分布式节点之间的高速通道的虚拟连接。通过第二通信通道，利用 RDMA 技术，整合各分布式节点上的梯度权重$\frac{dL}{dX}$，基于各分布式节点上的梯度权重，通过梯度更新法，更新大语言模型的模型参数θ。

按照目标存储参数，通过第一通信通道，将深度学习模型的模型参数存储至存储介质；通过第二通信通道，整合各分布式节点上的梯度权重。将模型参数存储过程和通信过程这两个数据传输过程进行隔离，避免引入额外的集合通信，确保不对分布式训练带来干扰，提升了分布式训练的可靠性，提升了模型训练的效果。

在本公开一种可选实施例中，传播计算包括前向传播计算和反向传播计算；

在步骤 104 中分布式训练的传播计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储之前，还包括如下具体步骤：

基于深度学习模型的模型规格信息和预设分布式训练策略，确定前向传播计算过程对应的第一目标存储参数和反向传播计算过程对应的第二目标存储参数；

对应地，步骤 104 中在分布式训练的传播计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，包括如下具体步骤：

在前向传播计算过程中，按照第一目标存储参数对深度学习模型的模型参数进行存储；

在反向传播计算过程中，按照第二目标存储参数对深度学习模型的模型参数进行存储。

传播计算过程包括前向传播计算过程和反向传播计算过程，两者需要的时间开销是不同的，因而，需要细化区分，确定在前向传播计算过程中的目标存储参数和反向传播计算过程中的目标存储参数。

第一目标存储参数为对前向传播计算过程中、模型参数进行存储的配置参数，第一目标存储参数是基于深度学习模型的模型规格信息和预设分布式训练策略确定的，包括但不

限于：第一目标存储模型参数规格、第一目标存储频率、第一目标存储读写速度、第一目标存储带宽等。

第二目标存储参数为对反向传播计算过程中、模型参数进行存储的配置参数，第二目标存储参数是基于深度学习模型的模型规格信息和预设分布式训练策略确定的，包括但不限于：第二目标存储模型参数规格、第二目标存储频率、第二目标存储读写速度、第二目标存储带宽等。

基于深度学习模型的模型规格信息和预设分布式训练策略，确定前向传播计算过程对应的第一目标存储参数和反向传播计算过程对应的第二目标存储参数，具体方式为：针对各个批次的分布式数据，基于深度学习模型的模型规格信息和预设分布式训练策略，预测前向传播计算和反向传播计算的次数和时间开销，基于前向传播计算和反向传播计算的次数和时间开销，确定前向传播计算对应的第一目标存储参数和反向传播计算过程对应的第二目标存储参数。

在前向传播计算过程中，按照第一目标存储参数对深度学习模型的模型参数进行存储，具体方式为：在前向传播计算过程中，按照第一目标存储参数，将深度学习模型的模型参数存储至存储介质。

在反向传播计算过程中，按照第二目标存储参数对深度学习模型的模型参数进行存储，具体方式为：在反向传播计算过程中，按照第二目标存储参数，将深度学习模型的模型参数存储至存储介质。

示例性地，通过多进程并行性通信模块，针对各个批次的分布式数据，基于大语言模型的模型参数量为10^{13}级别和数据并行策略，预测前向传播计算的次数为 16 次和时间开销 t1，反向传播计算的次数为 16 次和时间开销 t2，以模型参数存储过程的时间开销 T’不超过传播计算过程的时间开销 T=16*（t1+t2）为目标，确定前向传播计算过程对应的第一目标存储模型参数规格为 P1 和反向传播计算过程对应的第二目标存储模型参数规格为 P2。在前向传播计算过程中，按照第一目标存储模型参数规格，将大语言模型的模型参数θ存储至分布式节点的 GPU 缓存、内存和分布式持久化存储阵列中。在反向传播计算过程中，按照第二目标存储模型参数规格，将大语言模型的模型参数θ存储至分布式节点的 GPU 缓存、内存和分布式持久化存储阵列中。

基于深度学习模型的模型规格信息和预设分布式训练策略，确定前向传播计算过程对应的第一目标存储参数和反向传播计算过程对应的第二目标存储参数；在前向传播计算过程中，按照第一目标存储参数对深度学习模型的模型参数进行存储；在反向传播计算过程中，按照第二目标存储参数对深度学习模型的模型参数进行存储。更为精细地划分了前向传播和反向传播对应的目标存储参数，将模型参数的存储过程和分布式训练中的前向传播计算过程、反向传播计算过程重叠，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率。

在本公开一种可选实施例中，步骤 104 中按照目标存储参数对深度学习模型的模型参数进行存储，包括如下具体步骤：

按照目标存储参数和多个存储介质的存储性能优先级，将深度学习模型的模型参数存储至多个存储介质中。

不同的存储介质具有不同的存储性能，包括但不限于：读写速度和持久性。例如，内存的读写速度高于硬盘的读写速度，但是内存的持久性低于硬盘的持久性，相比于硬盘，内存的异常容忍能力更低。

按照目标存储参数和多个存储介质的存储性能优先级，将深度学习模型的模型参数存储至多个存储介质中，具体方式为：根据目标存储参数和多个存储介质的存储性能优先级，确定存储策略，按照存储策略，将深度学习模型的模型参数存储至多个存储介质中。本公开实施例提供一种存储策略：建立存储介质层级：从上到下分别为 GPU 缓存、内存和硬盘，1、尽可能使用更上层，读写速度更快的存储介质，来最大化保存模型参数的参数规格量和异常恢复能力；2、即使上层的存储介质不可用，依然可以依靠下层的、更高持久性的存储介质保证持久性地存储有当前的模型参数；3、通过上下层之间的异步执行，节省开销。

可选地，按照目标存储参数，将深度学习模型的模型参数存储至多层次存储介质，包括如下具体步骤：按照目标存储参数，将深度学习模型的模型参数分别存储至第一存储介质和第二存储介质，其中，第一存储介质的读写速度快于第二存储介质。

示例性地，按照目标存储模型参数规格 P1 和 P2，将大语言模型的模型参数θ分别存储至分布式节点的 GPU 缓存、内存和硬盘中。

可选地，按照目标存储参数，将深度学习模型的模型参数存储至多层次存储介质，包括如下具体步骤：按照目标存储参数，将深度学习模型的模型参数存储至第一存储介质，以使将模型参数从第一存储介质转存至第二存储介质，其中，第一存储介质的读写速度快于第二存储介质。

示例性地，按照目标存储模型参数规格 P1 和 P2，将大语言模型的模型参数θ存储至 GPU 缓存，以使将模型参数从 GPU 缓存转存至内存，以使将模型参数从内存转存至硬盘中。

按照目标存储参数和多个存储介质的存储性能优先级，将深度学习模型的模型参数存储至多个存储介质中。充分利用不同存储介质的存储性能，实现了高频率的模型参数存储，且存储有当前的模型参数，节省了开销。

在本公开一种可选实施例中，该方法还包括如下具体步骤：

在接收到恢复训练请求的情况下，获取已存储的目标模型参数，其中，恢复训练请求为确定分布式训练的训练异常恢复后生成的，目标模型参数为发生训练异常前存储的模型参数；

基于目标模型参数，恢复对深度学习模型进行分布式训练。

训练异常为在模型训练过程中出现的异常情况，异常情况可能导致模型训练无法正常运行，或者影响训练得到模型的性能和准确率，包括但不限于：硬件异常、系统异常、网络异常、或者其他未知异常。训练异常恢复为发生训练异常的情况下，通过采取对应措施来恢复模型训练。恢复训练请求为用于将恢复模型训练的指令请求。

目标模型参数为发生训练异常前存储的模型参数，例如，在完成第 i-1 次迭代训练，目标模型参数为θ_{i-1}，执行第 i 次迭代训练，在传播计算过程中，对目标模型参数θ_{i-1}进行存储，在完成存储后，如果发生训练异常，可以获取模型参数θ_{i-1}重新执行第 i 次迭代训

练。

获取已存储的目标模型参数，具体方式为：从存储介质中，获取已存储的目标模型参数。

示例性地，在接收到恢复训练请求的情况下，从硬盘中获取已存储的目标模型参数θ。基于目标模型参数θ，恢复对大语言模型进行分布式训练，获得训练完成的目标大语言模型，目标大语言模型具有文本问答功能。

在接收到恢复训练请求的情况下，获取已存储的目标模型参数，其中，恢复训练请求为确定分布式训练的训练异常恢复后生成的，目标模型参数为发生训练异常前存储的模型参数； 基于目标模型参数，恢复对深度学习模型进行分布式训练。增加了深度学习模型训练的容错性，避免了重新进行深度学习模型训练，具有稳定性的同时保证了训练效率，降低了训练成本。

图 2 示出了本公开一个实施例提供的一种深度学习模型训练方法中的系统架构图，如图 2 所示：

系统架构包括多层级存储介质，从上到下依次为：图形处理单元的缓存、内存和硬盘。多层级存储介质从上到下，持久性越来越高，多层级存储介质从下到上，读写速度越来越快。在分布式训练过程中，将深度学习模型的模型参数从图形处理单元存储到图形处理单元的缓存中，进行非持久性的高速读写。将深度学习模型的模型参数从图形处理单元存储到内存中，进行非持久性的高速读写。通过内存将深度学习模型的模型参数转存至硬盘中，进行持久性存储，在需要恢复训练的情况下，从多层级存储介质中获取深度学习模型的模型参数，重新进行模型训练。

图 3 示出了本公开一个实施例提供的一种深度学习模型训练方法的流程示意图，如图 3 所示：

目前，模型的迭代训练中，每次迭代包括传播计算、参数更新和参数存储。在第 i-1 次迭代中，完成传播计算和参数更新，对更新的模型参数进行参数存储，开启第 i 次迭代，同样完成传播计算和参数更新，对更新的模型参数进行参数存储。

本公开图 1 实施例中，每次迭代包括传播计算、通信整合和参数更新过程。在第 i-1 次迭代的传播计算过程中，对模型参数进行参数存储，接着执行通信整合和参数更新过程，开启第 i 次迭代，在第 i 次迭代的传播计算过程中，对模型参数进行参数存储，接着执行通信整合和参数更新过程。

两者相比，节省了时间开销，提升了模型训练的效率。

图 4 示出了本公开一个实施例提供的一种深度学习模型训练方法的原理图，如图 4 所示：

系统包括多个分布式节点（图中为两个分布式节点）和存储介质（图中为分布式持久化存储阵列），任一分布式节点包括内存、多个图形处理单元、第一网卡和第二网卡。分布式节点中的各图形处理单元之间构建有高速通信通道，分布式节点上的第一网卡和分布式存储节点之间建立了第一通信通道，各分布式节点的第二网卡之间建立了第二通信通道。

参见图 5，图 5 示出了本公开一个实施例提供的另一种深度学习模型训练方法的流程图，应用于云侧设备，云侧设备包括多个分布式节点和存储介质；该方法包括如下具体

步骤：

步骤 502：获取初始的深度学习模型和样本数据集。

步骤 504：按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

步骤 506：在识别到深度学习模型训练异常的情况下，触发多个分布式节点停止对深度学习模型的分布式训练，并确定存储介质当前已存储的目标模型参数。

步骤 508：在接收到恢复训练请求的情况下，从存储介质中获取目标模型参数。

步骤 510：调用多个分布式节点，基于目标模型参数恢复对深度学习模型进行分布式训练。

本公开实施例应用于具有分布式训练功能的云侧设备，该云侧设备为一种网络云设备，为一种虚拟设备，是由多个分布式节点和存储介质构成的，任一分布式节点包括模型训练的计算硬件，例如，GPU、NPU、FPGA、ASIC 或者 CPU。

本公开实施例与上述图 1 说明书实施例出于同一发明构思，步骤 502、504、508 和 510 的具体方式已在上述图 1 说明书实施例中详细说明，在此不再赘述。

示例性地，在某大语言模型的网站上，目前需要增加虚拟角色对话功能，通过对初始的大语言模型进行训练，使得训练得到的目标大语言模型具有文本问答功能，可以执行文本问答任务。从模型库中获取初始的大语言模型，从开源样本数据库中获取样本文本集，样本文本集包括 10000000 个样本问答文本对，预设分布式训练策略为数据并行策略。预先基于大语言模型的模型参数量为10^{13}级别和数据并行策略，执行时间开销分析，确定一次迭代的前向传播计算过程的时间开销为 t1 和反向传播计算过程的时间开销为 t2，基于该时间开销和模型参数量，获得目标存储模型参数规格为 P1 和 P2。按照数据并行策略，对样本文本集中的 10000000 个样本问答文本对进行划分。构建 64 个分布式数据，将 64 个分布式数据分发至 64 个分布式节点上，每个分布式节点上部署有 GPU，任一分布式节点上，将任一分布式数据划分为 16 个小批次，利用 GPU 对执行 16 个小批次的迭代训练。在每个迭代训练的前向传播计算过程和反向传播计算过程中，按照目标存储模型参数规格为 P1 和 P2，将大语言模型的模型参数θ依次存储至分布式节点的 GPU 缓存、内存和硬盘中，在识别到大语言模型训练异常的情况下，触发 64 个分布式节点停止对大语言模型的分布式训练，并确定硬盘当前已存储的目标模型参数θ，在接收到恢复训练请求的情况下，从硬盘中获取已存储的目标模型参数θ。基于目标模型参数θ，恢复对大语言模型进行分布式训练，获得训练完成的目标大语言模型，目标大语言模型具有文本问答功能，将该目标大语言模型部署在该大语言模型的网站的云侧设备上，为用户提供虚拟角色对话功能。

本公开实施例中，获取初始的深度学习模型和样本数据集；按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定；在识别到深度学习模型训练异常的情况下，触发多个分布式节点停止对深度学习模型的分布式

训练，并确定存储介质当前已存储的目标模型参数；在接收到恢复训练请求的情况下，从存储介质中获取目标模型参数；调用多个分布式节点，基于目标模型参数恢复对深度学习模型进行分布式训练。基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在分布式训练的调整参数计算过程中，按照目标存储参数，将深度学习模型的模型参数的存储至存储介质，实现了模型参数的存储过程和分布式训练中的调整参数计算过程重叠，充分节省了性能开销，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率，在识别到深度学习模型训练异常的情况下，从存储介质中获取存储的目标模型参数，恢复对深度学习模型的分布式训练，增加了深度学习模型训练的容错性，避免了重新进行深度学习模型训练，具有稳定性的同时保证了训练效率，降低了训练成本。

在本公开一种可选实施例中，存储介质包括多个存储性能不同的存储介质；

步骤 504 中在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，包括如下具体步骤：

按照目标存储参数和多个存储介质的存储性能优先级，将深度学习模型的模型参数存储至多个存储介质中；

对应地，步骤 508 中从存储介质中获取目标模型参数，包括如下具体步骤：

从第一存储介质中获取目标模型参数；

若未获取到，则从第二存储介质中获取目标模型参数，其中，第一存储介质的存储性能优先级高于第二存储介质。

按照目标存储参数和多个存储介质的存储性能优先级，将深度学习模型的模型参数存储至多个存储介质中的步骤已在上述图 1 说明书实施例中详细说明，在此不再赘述。

考虑到第一存储介质的存储性能优先级高于第二存储介质，以图 2 为例，第一存储介质为内存，第二存储介质为硬盘，内存的读写速度高于硬盘，如果可以获取到，训练效率高于硬盘，但是在内存上是非持久性存储，因而，可能无法获取到目标模型参数，需要从硬盘中进行获取。

示例性地，从内存中获取已存储的目标模型参数θ，若未获取到，则从硬盘中获取目标模型参数θ。

从第一存储介质中获取目标模型参数；若未获取到，则从第二存储介质中获取目标模型参数，其中，第一存储介质的存储性能优先级高于第二存储介质。充分利用了存储介质的存储性能优先级差异，在发生训练异常的情况下，优先从高存储性能优先级的存储介质获取目标模型参数，同时保证了高存储性能优先级的存储介质中持久性存储有目标模型参数，提升了模型训练效率的同时，保证了模型训练的可靠性。

在本公开一种可选实施例中，云侧设备还包括与各分布式节点连接的第一通信通道、与存储介质连接的第二通信通道；

步骤 504 中按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，包括如下具体步骤：

按照预设分布式训练策略，通过第一通信通道调用多个分布式节点，基于样本数据集

对深度学习模型进行分布式训练；

对应地，步骤 504 中按照目标存储参数将深度学习模型的模型参数存储至存储介质，包括如下具体步骤：

按照目标存储参数，通过第二通信通道将深度学习模型的模型参数存储至存储介质。

连接存储介质的第二通信信道为用于模型参数存储过程的通信信道，包括分布式节点和存储介质之间的物理信道和逻辑信道，例如，网卡、PCIe 拓扑、光纤、分布式节点和存储介质之间的读写通道（数据总线）等。

与其他分布式节点连接的第一通信信道为用于通信过程的通信信道，包括分布式节点和存储介质之间的物理信道和逻辑信道，例如，网卡、PCIe 拓扑、光纤、分布式节点、GPU 之间的高速通道、各分布式节点之间的高速通道等。第一通信信道可以通过 RDMA 技术实现，提升了传输速度。

示例性地，第二通信信道为 64 个分布式节点上的网卡分别与分布式持久化存储阵列之间的物理信道，按照目标存储模型参数规格 P，通过第二通信信道，将大语言模型的模型参数θ存储至分布式持久化存储阵列。第一通信信道为 64 个分布式节点上的网卡之间的物理信道，以及各分布式节点上插有 GPU 的 PCIe 拓扑，配合有 GPU 之间的高速通道和各分布式节点之间的高速通道的逻辑信道。通过第一通信信道，利用 RDMA 技术，整合各分布式节点上的梯度权重$\frac{dL}{dX}$，基于各分布式节点上的梯度权重，通过梯度更新法，更新大语言模型的模型参数θ。

按照预设分布式训练策略，通过第一通信通道调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练；按照目标存储参数，通过第二通信通道将深度学习模型的模型参数存储至存储介质。将模型参数存储过程和通信过程这两个数据传输过程进行隔离，避免引入额外的集合通信，确保不对分布式训练带来干扰，提升了分布式训练的可靠性，提升了模型训练的效果。

与上述方法实施例相对应，本公开还提供了深度学习模型训练系统实施例，图 6 示出了本公开一个实施例提供的一种深度学习模型训练系统的结构示意图。如图 6 所示，该系统包括管控单元 602 和多个分布式节点，多个分布式节点包括第一分布式节点 604，第一分布式节点 604 为多个分布式节点中任一个；

管控单元 602，用于获取初始的深度学习模型和样本数据集，按照预设分布式训练策略，基于深度学习模型和样本数据集，构建多个分布式数据，将多个分布式数据分发至各分布式节点；

第一分布式节点 604，用于基于样本数据集对深度学习模型进行分布式训练；并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储其中，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

可选地，系统还包括持久性存储介质，第一分布式节点 604 包括非持久性存储介质；

对应地，第一分布式节点 604，还用于在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至非持久性存储介质，以使将深度学习模型的模型参数从非持久性存储介质转存至持久性存储介质。

可选地，调整参数为梯度权重，调整参数计算为传播计算；第一分布式节点 604 还包括与各分布式节点连接的第一通信通道、与存储介质连接的第二通信通道；

对应地，第一分布式节点 604，还用于通过第一通信通道，整合各分布式节点上的梯度权重；

对应地，第一分布式节点 604，还用于通过第二通信通道，将深度学习模型的模型参数从非持久性存储介质发送至持久性存储介质进行存储。

本公开实施例中，基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在分布式训练的调整参数计算过程中，按照目标存储参数，完成了对深度学习模型的模型参数的存储，将模型参数的存储过程和分布式训练中的调整参数计算过程重叠，充分节省了性能开销，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率。

上述为本实施例的一种深度学习模型训练系统的示意性方案。需要说明的是，该深度学习模型训练系统的技术方案与上述的深度学习模型训练方法的技术方案属于同一构思，深度学习模型训练系统的技术方案未详细描述的细节内容，均可以参见上述深度学习模型训练方法的技术方案的描述。

与上述方法实施例相对应，本公开还提供了深度学习模型训练装置实施例，图 7 示出了本公开一个实施例提供的一种深度学习模型训练装置的结构示意图。如图 7 所示，该装置包括：

第一获取模块 702，被配置为获取初始的深度学习模型和样本数据集；

第一训练模块 704，被配置为按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定。

可选地，第一训练模块 704 被进一步配置为：

按照预设分布式训练策略，基于深度学习模型和样本数据集，构建多个分布式数据，其中，预设分布式训练策略包括模型并行训练策略或者数据并行训练策略；将多个分布式数据分发至各分布式节点；在第一分布式节点上，基于分布式数据进行传播计算，获得梯度权重，其中，第一分布式节点为多个分布式节点中任一个；基于各分布式节点上的梯度权重，更新深度学习模型的模型参数，在达到预设训练结束条件的情况下，获得训练完成的深度学习模型。

可选地，分布式数据包括多个批次的分布式数据；

第一训练模块 704 被进一步配置为：

在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重；

对应地，该装置还包括：

迭代模块，被配置为更新当前批次的分布式数据，返回执行在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重的步骤。

可选地，该装置还包括：

存储参数确定模块，被配置为针对各个批次的分布式数据，基于深度学习模型的模型规格信息和预设分布式训练策略，预测传播计算的次数和时间开销；基于传播计算的次数和时间开销，确定传播计算过程对应的目标存储参数。

可选地，该装置还包括：

梯度权重整合模块，被配置为通过各分布式节点之间的通信通道，整合各分布式节点上的梯度权重。

可选地，任一分布式节点包括连接存储介质的第一通信通道和与其他分布式节点连接的第二通信通道；

第一训练模块 704 被进一步配置为：

按照目标存储参数，通过第一通信通道，将深度学习模型的模型参数存储至存储介质；

对应地，梯度权重整合模块被进一步配置为：

通过第二通信通道，整合各分布式节点上的梯度权重。

可选地，传播计算包括前向传播计算和反向传播计算；

该装置还包括：

前反向存储参数确定模块，被配置为基于深度学习模型的模型规格信息和预设分布式训练策略，确定前向传播计算过程对应的第一目标存储参数和反向传播计算过程对应的第二目标存储参数；

对应地，第一训练模块 704 被进一步配置为：

在前向传播计算过程中，按照第一目标存储参数对深度学习模型的模型参数进行存储；在反向传播计算过程中，按照第二目标存储参数对深度学习模型的模型参数进行存储。

可选地，该装置还包括：

恢复训练模块，被配置为在接收到恢复训练请求的情况下，获取已存储的目标模型参数，其中，恢复训练请求为确定分布式训练的训练异常恢复后生成的，目标模型参数为发生训练异常前存储的模型参数；基于目标模型参数，恢复对深度学习模型进行分布式训练。

本公开实施例中，基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在分布式训练的传播计算过程中，按照目标存储参数，完成了对深度学习模型的模型参数的存储，将模型参数的存储过程和分布式训练中的传播计算过程重叠，充分节省了性能开销，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率。

上述为本实施例的一种深度学习模型训练装置的示意性方案。需要说明的是，该深度学习模型训练装置的技术方案与上述的深度学习模型训练方法的技术方案属于同一构思，深度学习模型训练装置的技术方案未详细描述的细节内容，均可以参见上述深度学习模型训练方法的技术方案的描述。

与上述方法实施例相对应，本公开还提供了深度学习模型训练装置实施例，图 8 示出了本公开一个实施例提供的另一种深度学习模型训练装置的结构示意图。如图 8 所示，应

用于云侧设备，云侧设备包括多个分布式节点和存储介质；该装置包括：

第二获取模块 802，被配置为获取初始的深度学习模型和样本数据集；

第二训练模块 804，被配置为按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定；

停止模块 806，被配置为在识别到深度学习模型训练异常的情况下，触发多个分布式节点停止对深度学习模型的分布式训练，并确定存储介质当前已存储的目标模型参数；

参数获取模块 808，被配置为在接收到恢复训练请求的情况下，从存储介质中获取目标模型参数；

恢复模块 810，被配置为调用多个分布式节点，基于目标模型参数恢复对深度学习模型进行分布式训练。

可选地，存储介质包括多个存储性能不同的存储介质；

第二训练模块 804 被进一步配置为：

按照目标存储参数和多个存储介质的存储性能优先级，将深度学习模型的模型参数存储至多个存储介质中；

对应地，参数获取模块 808 被进一步配置为：

从第一存储介质中获取目标模型参数；若未获取到，则从第二存储介质中获取目标模型参数，其中，第一存储介质的存储性能优先级高于第二存储介质。

可选地，云侧设备还包括与各分布式节点连接的第一通信通道、与存储介质连接的第二通信通道；

第二训练模块 804 被进一步配置为：

按照预设分布式训练策略，通过第一通信通道调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练；按照目标存储参数，通过第二通信通道将深度学习模型的模型参数存储至存储介质。

本公开实施例中，基于深度学习模型的模型规格信息和预设分布式训练策略确定目标存储参数，充分考虑了分布式训练的迭代规律，在分布式训练的调整参数计算过程中，按照目标存储参数，将深度学习模型的模型参数的存储至存储介质，实现了模型参数的存储过程和分布式训练中的调整参数计算过程重叠，充分节省了性能开销，以接近零性能开销的方式，完成了对深度学习模型参数的实时存储，使得深度学习模型训练具备高容错能力的同时，具有高效率，在识别到深度学习模型训练异常的情况下，从存储介质中获取存储的目标模型参数，恢复对深度学习模型的分布式训练，增加了深度学习模型训练的容错性，避免了重新进行深度学习模型训练，具有稳定性的同时保证了训练效率，降低了训练成本。

上述为本实施例的一种深度学习模型训练装置的示意性方案。需要说明的是，该深度学习模型训练装置的技术方案与上述的深度学习模型训练方法的技术方案属于同一构思，深度学习模型训练装置的技术方案未详细描述的细节内容，均可以参见上述深度学习模型训练方法的技术方案的描述。

图 9 示出了本公开一个实施例提供的一种计算设备的结构框图。该计算设备 900 的部件包括但不限于存储器 910 和处理器 920。处理器 920 与存储器 910 通过总线 930 相连接，数据库 950 用于保存数据。

计算设备 900 还包括接入设备 940，接入设备 940 使得计算设备 900 能够经由一个或多个网络 960 通信。这些网络的示例包括公用交换电话网（Public Switched Telephone Network，简称 PSTN）、局域网（LAN，Local Area Network）、广域网（Wide Area Network，简称 WAN）、个域网（Personal Area Network，简称 PAN）或诸如因特网的通信网络的组合。接入设备 940 可以包括有线或无线的任何类型的网络接口（例如，网络接口卡（network interface controller，简称 NIC））中的一个或多个，诸如 IEEE802.11 无线局域网（Wireless Local Area Network，简称 WLAN）无线接口、全球微波互联接入（Worldwide Interoperability for Microwave Access，简称 Wi-MAX）接口、以太网接口、通用串行总线（Universal Serial Bus，简称 USB）接口、蜂窝网络接口、蓝牙接口、近场通信（Near Field Communication，简称 NFC）。

在本公开的一个实施例中，计算设备 900 的上述部件以及图 9 中未示出的其他部件也可以彼此相连接，例如通过总线。应当理解，图 9 所示的计算设备结构框图仅仅是出于示例的目的，而不是对本公开范围的限制。本领域技术人员可以根据需要，增添或替换其他部件。

计算设备 900 可以是任何类型的静止或移动计算设备，包括移动计算机或移动计算设备（例如，平板计算机、个人数字助理、膝上型计算机、笔记本计算机、上网本等）、移动电话（例如，智能手机）、可佩戴的计算设备（例如，智能手表、智能眼镜等）或其他类型的移动设备，或者诸如台式计算机或个人计算机（Personal Computer，简称 PC）的静止计算设备。计算设备 900 还可以是移动式或静止式的服务器。

其中，处理器 920 用于执行如下计算机可执行指令，该计算机可执行指令被处理器执行时实现上述深度学习模型训练方法的步骤。

上述为本实施例的一种计算设备的示意性方案。需要说明的是，该计算设备的技术方案与上述的深度学习模型训练方法的技术方案属于同一构思，计算设备的技术方案未详细描述的细节内容，均可以参见上述深度学习模型训练方法的技术方案的描述。

本公开一实施例还提供一种计算机可读存储介质，其存储有计算机可执行指令，该计算机可执行指令被处理器执行时实现上述深度学习模型训练方法的步骤。

上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是，该存储介质的技术方案与上述的深度学习模型训练方法的技术方案属于同一构思，存储介质的技术方案未详细描述的细节内容，均可以参见上述深度学习模型训练方法的技术方案的描述。

本公开一实施例还提供一种计算机程序，其中，当所述计算机程序在计算机中执行时，令计算机执行上述深度学习模型训练方法的步骤。

上述为本实施例的一种计算机程序的示意性方案。需要说明的是，该计算机程序的技术方案与上述的深度学习模型训练方法的技术方案属于同一构思，计算机程序的技术方案未详细描述的细节内容，均可以参见上述深度学习模型训练方法的技术方案的描述。

上述对本公开特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一

些情况下，在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

所述计算机指令包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（Read-Only Memory，简称ROM）、随机存取存储器（Random Access Memory，简称RAM）、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读介质包含的内容可以根据专利实践的要求进行适当的增减，例如在某些地区，根据专利实践，计算机可读介质不包括电载波信号和电信信号。

需要说明的是，对于前述的各方法实施例，为了简便描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本公开实施例并不受所描述的动作顺序的限制，因为依据本公开实施例，某些步骤可以采用其它顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定都是本公开实施例所必须的。

在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其它实施例的相关描述。

以上公开的本公开优选实施例只是用于帮助阐述本公开。可选实施例并没有详尽叙述所有的细节，也不限制该发明仅为所述的具体实施方式。显然，根据本公开实施例的内容，可作很多的修改和变化。本公开选取并具体描述这些实施例，是为了更好地解释本公开实施例的原理和实际应用，从而使所属技术领域技术人员能很好地理解和利用本公开。本公开仅受权利要求书及其全部范围和等效物的限制。

权 利 要 求 书

1、一种深度学习模型训练方法，包括：

获取初始的深度学习模型和样本数据集；

按照预设分布式训练策略，基于所述样本数据集对所述深度学习模型进行分布式训练，并在所述分布式训练的调整参数计算过程中，按照目标存储参数对所述深度学习模型的模型参数进行存储，其中，所述目标存储参数基于所述深度学习模型的模型规格信息和所述预设分布式训练策略确定。

2、根据权利要求 1 所述的方法，所述调整参数为梯度权重，调整参数计算为传播计算；

所述按照预设分布式训练策略，基于所述样本数据集对所述深度学习模型进行分布式训练，包括：

按照预设分布式训练策略，基于所述深度学习模型和样本数据集，构建多个分布式数据，其中，所述预设分布式训练策略包括模型并行训练策略或者数据并行训练策略；

将所述多个分布式数据分发至各分布式节点；

在第一分布式节点上，基于分布式数据进行传播计算，获得梯度权重，其中，所述第一分布式节点为多个分布式节点中任一个；

基于各分布式节点上的梯度权重，更新所述深度学习模型的模型参数，在达到预设训练结束条件的情况下，获得训练完成的所述深度学习模型。

3、根据权利要求 2 所述的方法，所述分布式数据包括多个批次的分布式数据；

所述在第一分布式节点上，基于分布式数据进行传播计算，获得梯度权重，包括：

在第一分布式节点上，基于当前批次的分布式数据进行传播计算，获得梯度权重；

在所述基于各分布式节点上的梯度权重，更新所述深度学习模型的模型参数之后，还包括：

更新所述当前批次的分布式数据，返回执行所述在第一分布式节点上，基于当前批次的分布式数据进行计算，获得梯度权重的步骤。

4、根据权利要求 3 所述的方法，在所述在第一分布式节点上，对当前批次的分布式数据执行计算，获得梯度权重之前，还包括：

针对各个批次的分布式数据，基于所述深度学习模型的模型规格信息和所述预设分布式训练策略，预测计算的次数和时间开销；

基于所述计算的次数和时间开销，确定所述计算过程对应的目标存储参数。

5、根据权利要求 2 所述的方法，在所述基于各分布式节点上的梯度权重，更新所述深度学习模型的模型参数之前，还包括：

通过各分布式节点之间的通信通道，整合所述各分布式节点上的梯度权重。

6、根据权利要求 5 所述的方法，所述任一分布式节点包括连接存储介质的第一通信通道和与其他分布式节点连接的第二通信通道；

所述按照目标存储参数对所述深度学习模型的模型参数进行存储，包括：

按照目标存储参数，通过所述第一通信通道，将所述深度学习模型的模型参数存储至所述存储介质；

所述通过各分布式节点之间的通信通道，整合所述各分布式节点上的梯度权重，包括：

通过所述第二通信通道，整合所述各分布式节点上的梯度权重。

7、根据权利要求 1 至 6 中任一项所述的方法，所述传播计算包括前向传播计算和反向传播计算；

在所述分布式训练的计算过程中，按照目标存储参数对所述深度学习模型的模型参数进行存储之前，还包括：

基于所述深度学习模型的模型规格信息和所述预设分布式训练策略，确定前向传播计算过程对应的第一目标存储参数和反向传播计算过程对应的第二目标存储参数；

所述在所述分布式训练的计算过程中，按照目标存储参数对所述深度学习模型的模型参数进行存储，包括：

在所述前向传播计算过程中，按照所述第一目标存储参数对所述深度学习模型的模型参数进行存储；

在所述反向传播计算过程中，按照所述第二目标存储参数对所述深度学习模型的模型参数进行存储。

8、根据权利要求 1 至 7 中任一项所述的方法，还包括：

在接收到恢复训练请求的情况下，获取已存储的目标模型参数，其中，所述恢复训练请求为确定分布式训练的训练异常恢复后生成的，所述目标模型参数为发生所述训练异常前存储的模型参数；

基于所述目标模型参数，恢复对所述深度学习模型进行分布式训练。

9、一种深度学习模型训练方法，应用于云侧设备，所述云侧设备包括多个分布式节点和存储介质；所述方法包括：

获取初始的深度学习模型和样本数据集；

按照预设分布式训练策略，调用所述多个分布式节点，基于所述样本数据集对所述深度学习模型进行分布式训练，并在所述分布式训练的调整参数计算过程中，按照目标存储参数将所述深度学习模型的模型参数存储至所述存储介质，其中，所述目标存储参数基于所述深度学习模型的模型规格信息和所述预设分布式训练策略确定；

在识别到所述深度学习模型训练异常的情况下，触发所述多个分布式节点停止对所述深度学习模型的分布式训练，并确定所述存储介质当前已存储的目标模型参数；

在接收到恢复训练请求的情况下，从所述存储介质中获取所述目标模型参数；

调用所述多个分布式节点，基于所述目标模型参数恢复对所述深度学习模型进行分布式训练。

10、根据权利要求 9 所述的方法，所述存储介质包括多个存储性能不同的存储介质；

所述在所述分布式训练的计算过程中，按照目标存储参数将所述深度学习模型的模型参数存储至所述存储介质，包括：

按照目标存储参数和多个存储介质的存储性能优先级，将所述深度学习模型的模型参数存储至所述多个存储介质中；

所述从所述存储介质中获取所述目标模型参数，包括：

从第一存储介质中获取所述目标模型参数；

若未获取到，则从第二存储介质中获取所述目标模型参数，其中，所述第一存储介质的存储性能优先级高于所述第二存储介质。

11、根据权利要求 9 所述的方法，所述云侧设备还包括与各分布式节点连接的第一通信通道、与所述存储介质连接的第二通信通道；

按照预设分布式训练策略，调用所述多个分布式节点，基于所述样本数据集对所述深度学习模型进行分布式训练，包括：

按照预设分布式训练策略，通过所述第一通信通道调用所述多个分布式节点，基于所述样本数据集对所述深度学习模型进行分布式训练；

所述按照目标存储参数将所述深度学习模型的模型参数存储至所述存储介质，包括：

按照目标存储参数，通过所述第二通信通道将所述深度学习模型的模型参数存储至所述存储介质。

12、一种深度学习模型训练系统，所述系统包括管控单元和多个分布式节点，所述多个分布式节点包括第一分布式节点，所述第一分布式节点为多个分布式节点中任一个；

所述管控单元，用于获取初始的深度学习模型和样本数据集，按照预设分布式训练策略，基于所述深度学习模型和样本数据集，构建多个分布式数据，将所述多个分布式数据分发至各分布式节点；

所述第一分布式节点，用于基于所述样本数据集对所述深度学习模型进行分布式训练；并在所述分布式训练的调整参数计算过程中，按照目标存储参数对所述深度学习模型的模型参数进行存储其中，其中，所述目标存储参数基于所述深度学习模型的模型规格信息和所述预设分布式训练策略确定。

13、根据权利要求 12 所述的系统，所述系统还包括持久性存储介质，所述第一分布式节点包括非持久性存储介质；

所述第一分布式节点，还用于在所述分布式训练的调整参数计算过程中，按照目标存储参数将所述深度学习模型的模型参数存储至所述非持久性存储介质，以使将所述深度学习模型的模型参数从所述非持久性存储介质转存至所述持久性存储介质。

14、根据权利要求 13 所述的系统，所述调整参数为梯度权重，调整参数计算为传播计算；

所述第一分布式节点还包括与各分布式节点连接的第一通信通道、与所述存储介质连接的第二通信通道；

所述第一分布式节点，还用于通过所述第一通信通道，整合所述各分布式节点上的梯度权重；

所述第一分布式节点，还用于通过所述第二通信通道，将所述深度学习模型的模型参数从所述非持久性存储介质发送至所述持久性存储介质进行存储。

15、一种计算设备，包括：

存储器和处理器；

所述存储器用于存储计算机可执行指令，所述处理器用于执行所述计算机可执行指令，该计算机可执行指令被处理器执行时实现权利要求 1 至 11 任意一项所述方法的步骤。

16、一种计算机可读存储介质，其存储有计算机可执行指令，该计算机可执行指令被处理器执行时实现权利要求 1 至 11 任意一项所述方法的步骤。

17、一种计算机程序，所述计算机程序被处理器执行时实现如权利要求 1 至 11 中任一项所述的方法。

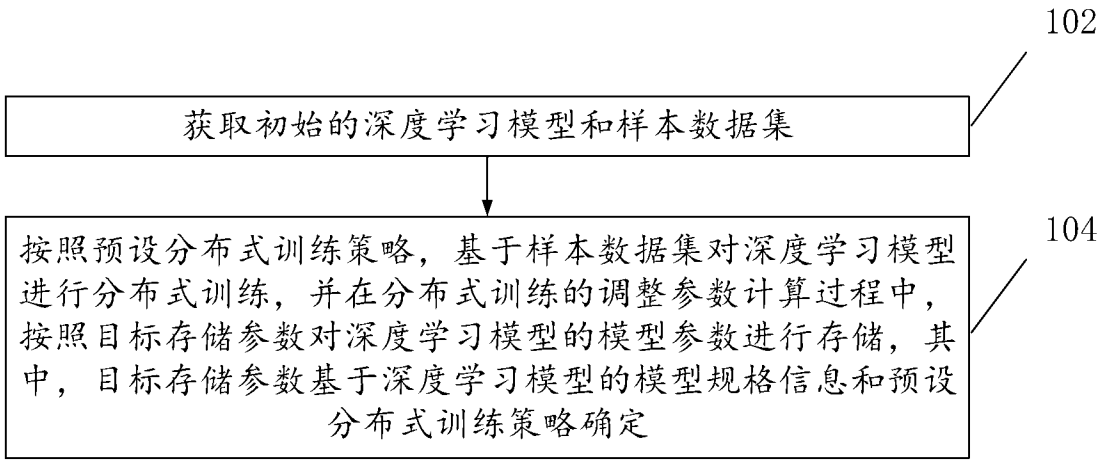
102
获取初始的深度学习模型和样本数据集
104
按照预设分布式训练策略，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数对深度学习模型的模型参数进行存储，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定

图 1

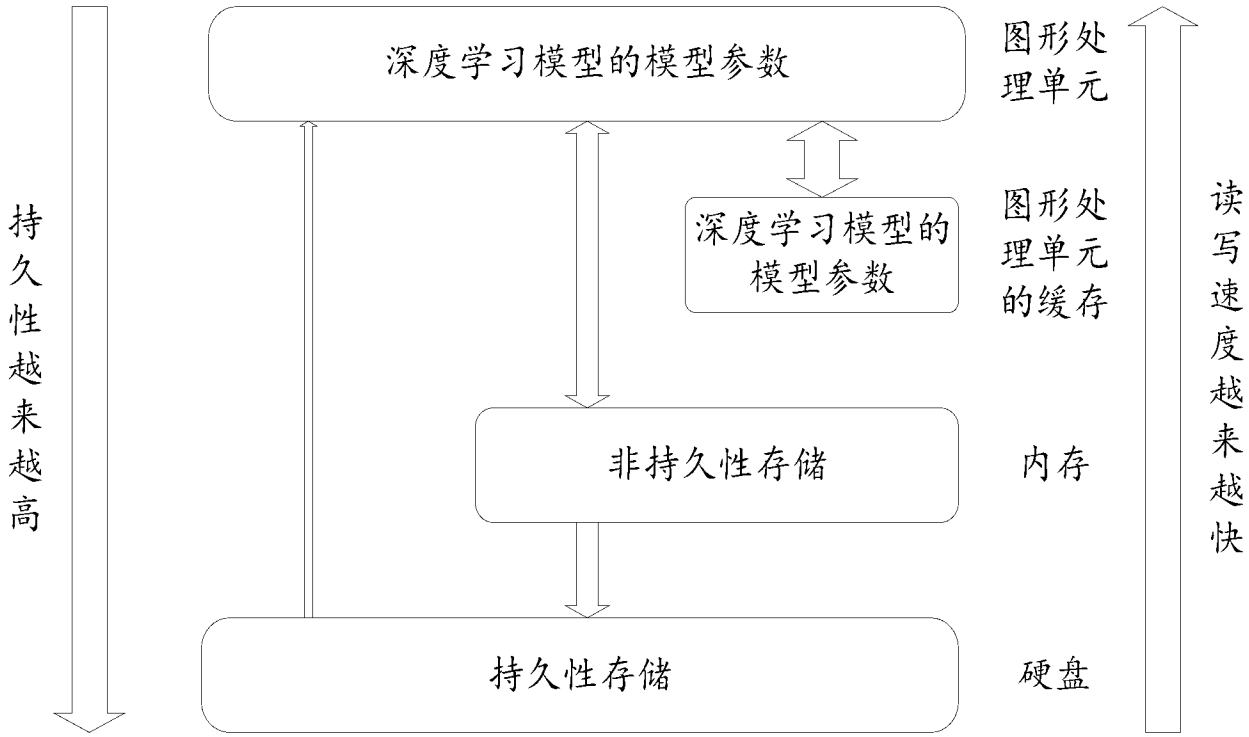
深度学习模型的模型参数
图形处理单元
深度学习模型的模型参数
图形处理单元的缓存
非持久性存储
内存
持久性存储
硬盘
持久性越来越高
读写速度越来越快

图 2

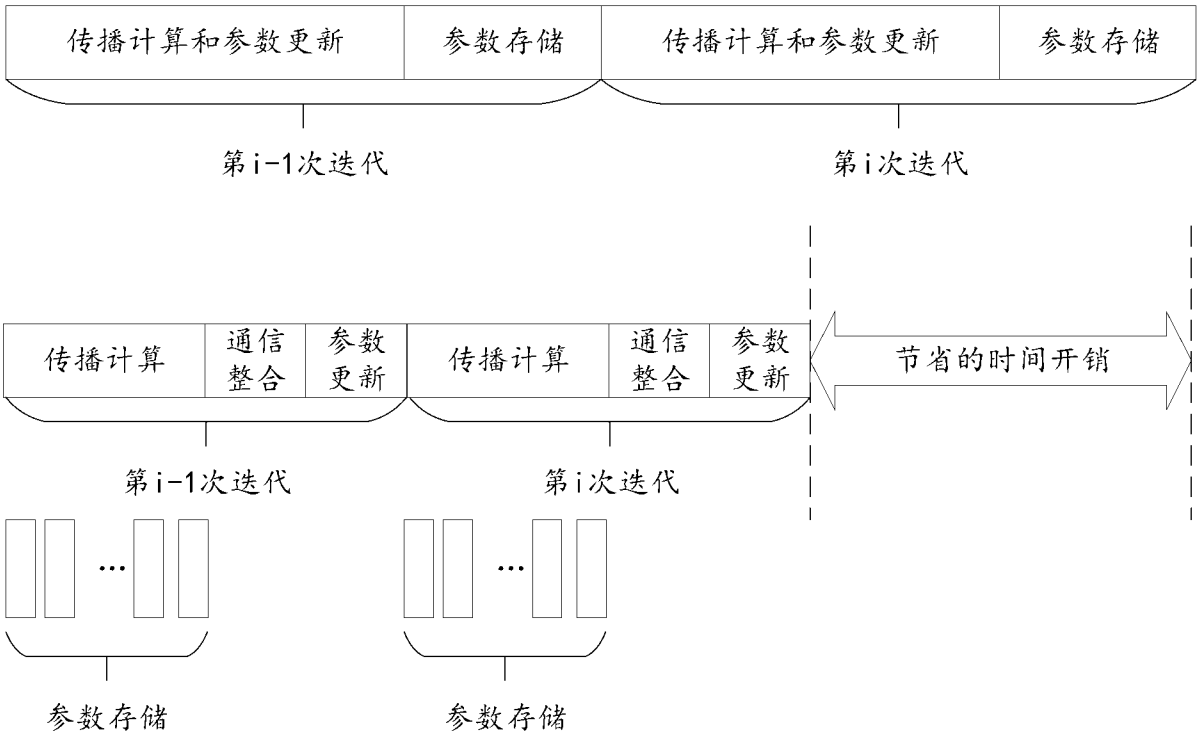
传播计算和参数更新
参数存储
传播计算和参数更新
参数存储
第i-1次迭代
第i次迭代
传播计算
通信整合
参数更新
传播计算
通信整合
参数更新
节省的时间开销
第i-1次迭代
第i次迭代
参数存储
参数存储

图 3

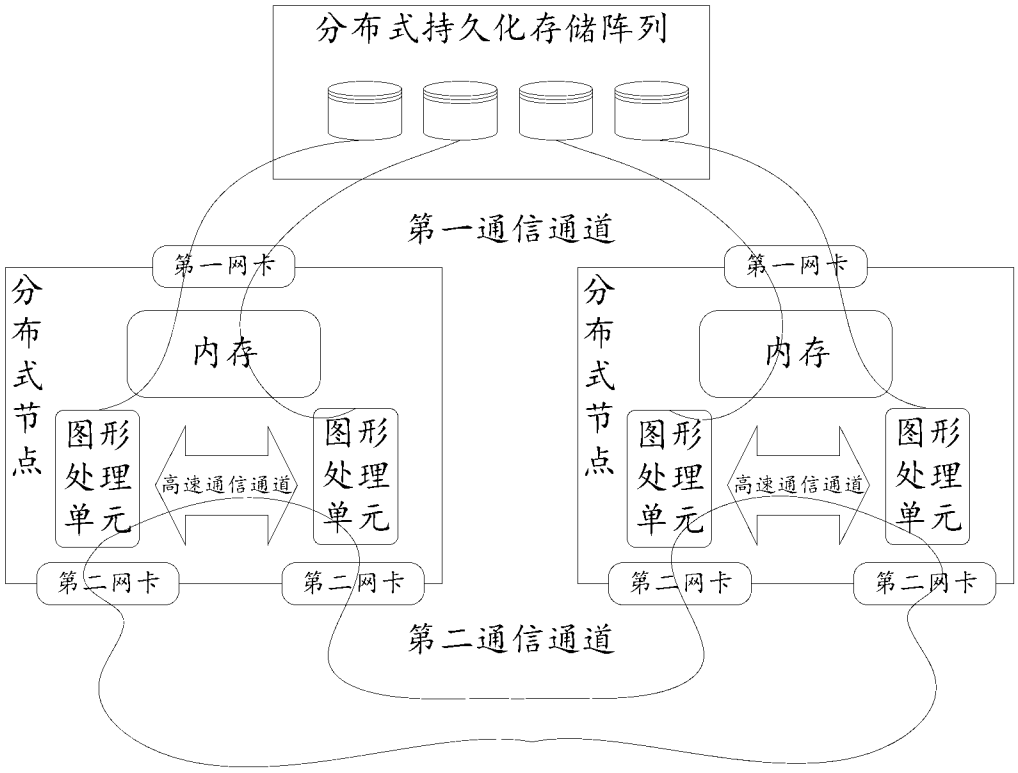
分布式持久化存储阵列
第一通信通道
第一网卡
分布式节点
内存
图形处理单元
高速通信通道
图形处理单元
第二网卡
第二网卡
第一网卡
分布式节点
内存
图形处理单元
高速通信通道
图形处理单元
第二网卡
第二网卡
第二通信通道

图 4

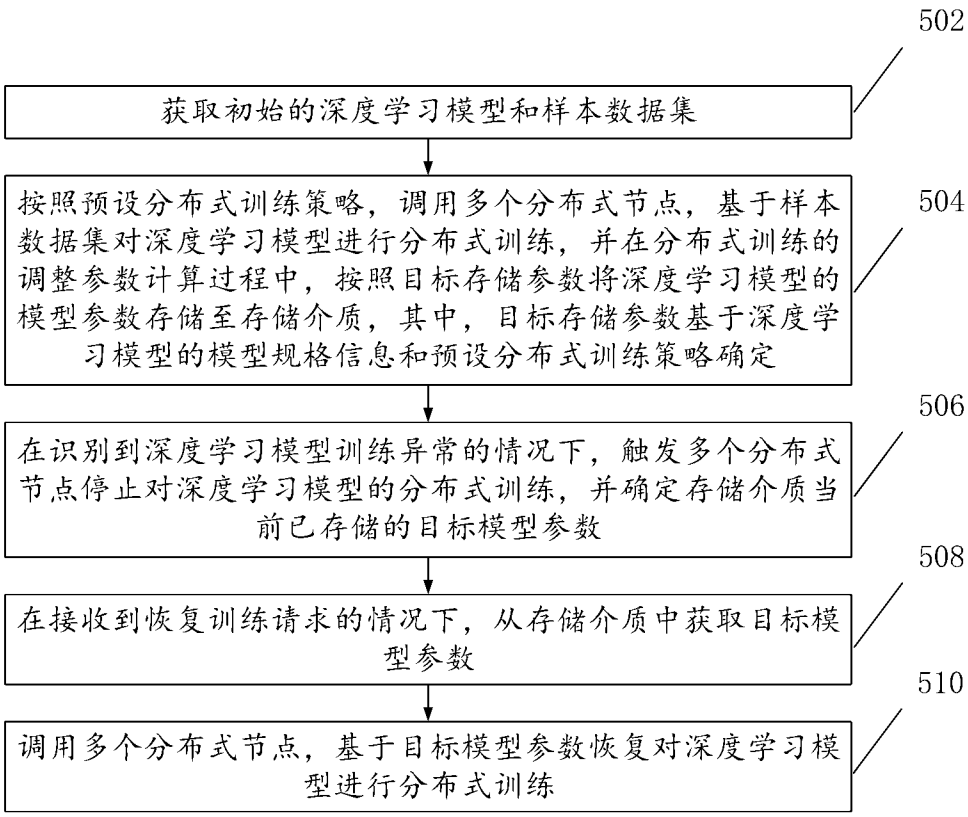
502
获取初始的深度学习模型和样本数据集
504
按照预设分布式训练策略，调用多个分布式节点，基于样本数据集对深度学习模型进行分布式训练，并在分布式训练的调整参数计算过程中，按照目标存储参数将深度学习模型的模型参数存储至存储介质，其中，目标存储参数基于深度学习模型的模型规格信息和预设分布式训练策略确定
506
在识别到深度学习模型训练异常的情况下，触发多个分布式节点停止对深度学习模型的分布式训练，并确定存储介质当前已存储的目标模型参数
508
在接收到恢复训练请求的情况下，从存储介质中获取目标模型参数
510
调用多个分布式节点，基于目标模型参数恢复对深度学习模型进行分布式训练

图 5

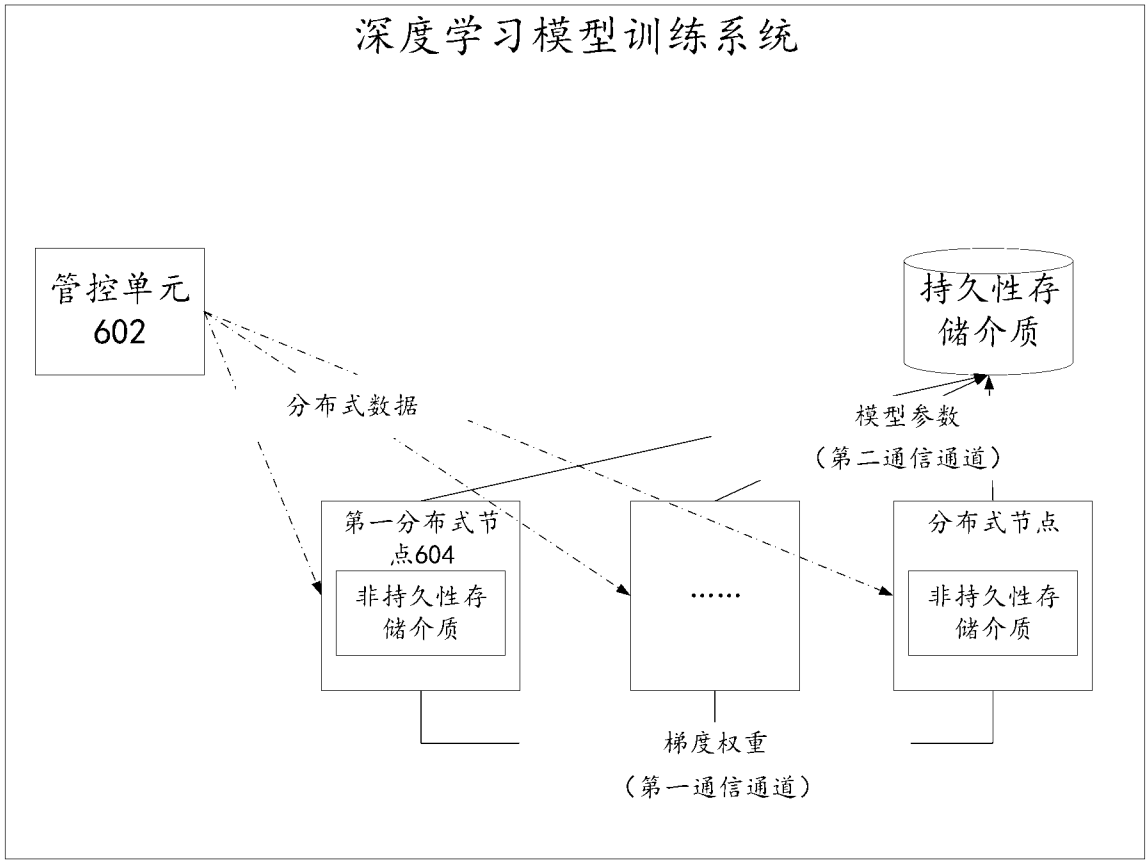
深度学习模型训练系统
管控单元
602
持久性存储介质
分布式数据
模型参数
（第二通信通道）
第一分布式节点604
非持久性存储介质
……
分布式节点
非持久性存储介质
梯度权重
（第一通信通道）

图 6

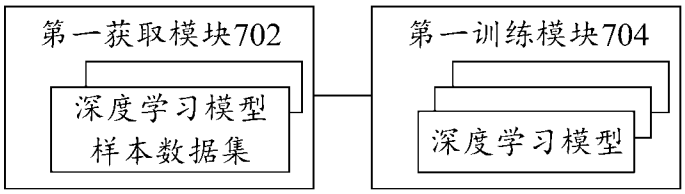
第一获取模块702
深度学习模型样本数据集
第一训练模块704
深度学习模型

图 7

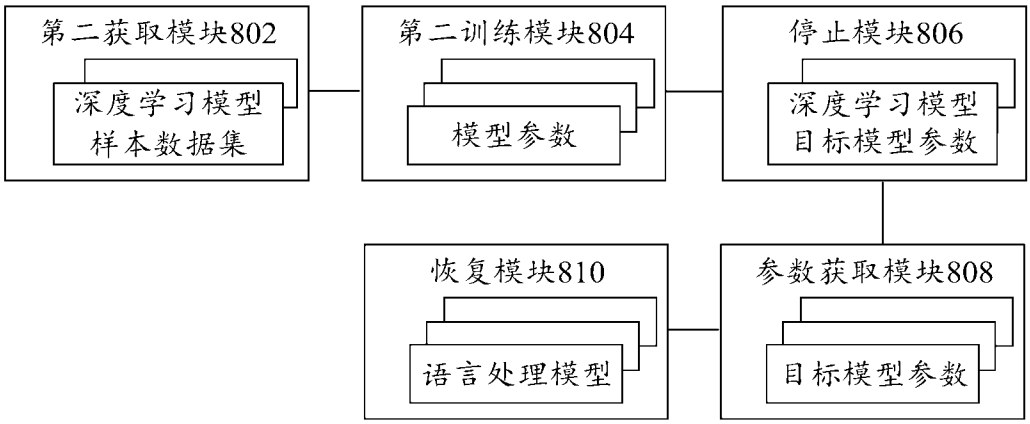
第二获取模块802
深度学习模型样本数据集
第二训练模块804
模型参数
停止模块806
深度学习模型目标模型参数
恢复模块810
语言处理模型
参数获取模块808
目标模型参数

图 8

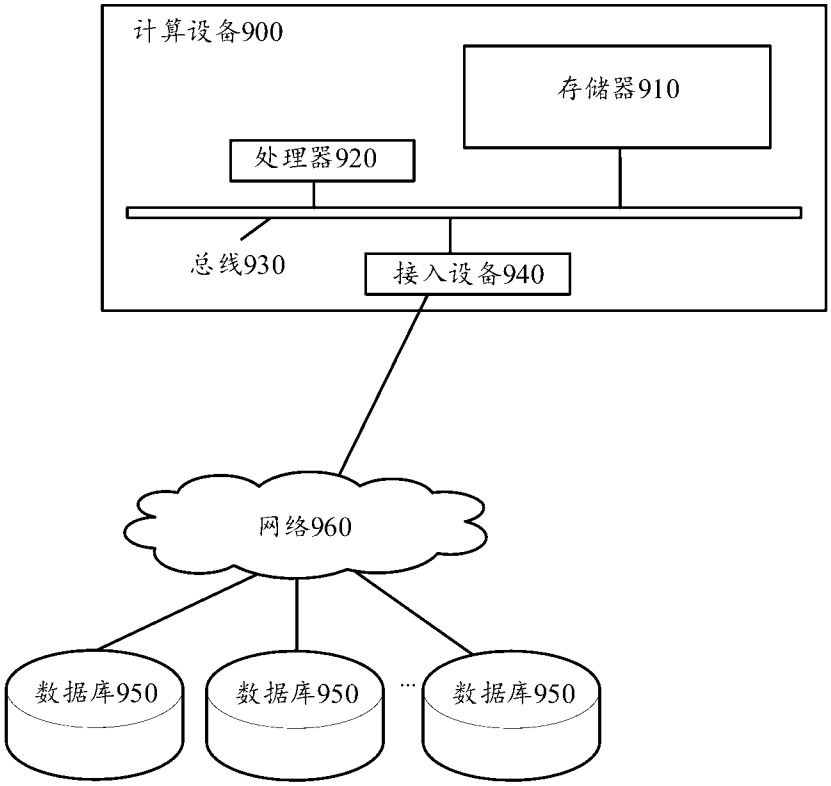
计算设备900
存储器910
处理器920
总线930
接入设备940
网络960
数据库950
数据库950
…
数据库950

图 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/118478

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/098(2023.01)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)

IPC: G06N 3/-

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)

CNTXT, ENTXTC, CNKI, 万方, WANGFANG: 模型, 网络, 分布式, 训练, 优化, 参数, 权重, 计算, 调整, 更新, 前向, 后向, 过程中, 步骤中, 时间内, 处理中, 同步, 同时, 实时, 重叠, 存储, 储存, 保存, 迭代, 处理, 次数, 时间; ENTXT, VEN, Web of Science, IEEE: model, network, distributed, train, optimization, parameter, weight, calculation, adjustment, update, forward, backward, during, within, time, step, process synchronized, real-time, overlap, storage, store, save, preservation, iteration

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117669700 A (HANGZHOU ALICLOUD FEITIAN INFORMATION TECHNOLOGY CO., LTD.) 08 March 2024 (2024-03-08) entire document	1-17
A	CN 113515370 A (ZHEJIANG LAB et al.) 19 October 2021 (2021-10-19) description, paragraphs [0047]-[0080]	1-17
A	CN 109754060 A (ALIBABA GROUP HOLDING LIMITED) 14 May 2019 (2019-05-14) entire document	1-17
A	CN 113705801 A (HUAWEI TECHNOLOGIES CO., LTD.) 26 November 2021 (2021-11-26) entire document	1-17
A	CN 111788585 A (HUAWEI TECHNOLOGIES CO., LTD.) 16 October 2020 (2020-10-16) entire document	1-17
A	CN 117093871 A (ZHEJIANG LAB) 21 November 2023 (2023-11-21) entire document	1-17

☐ 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
"D" document cited by the applicant in the international application
"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	Date of mailing of the international search report
19 November 2024	**26 November 2024**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/118478

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 117669700 A	08 March 2024	None	
CN 113515370 A	19 October 2021	CN 113515370 B	12 March 2024
CN 109754060 A	14 May 2019	CN 109754060 B	25 August 2023
CN 113705801 A	26 November 2021	WO 2021232907 A1	25 November 2021
		EP 4148624 A1	15 March 2023
		US 2023087642 A1	23 March 2023
		EP 4148624 A4	18 October 2023
CN 111788585 A	16 October 2020	WO 2020147142 A1	23 July 2020
		EP 3889846 A1	06 October 2021
		EP 3889846 A4	01 June 2022
		US 2021342696 A1	04 November 2021
		CN 111788585 B	12 April 2024
CN 117093871 A	21 November 2023	CN 117093871 B	13 February 2024

国际检索报告

国际申请号 PCT/CN2024/118478

A. 主题的分类

G06N 3/098(2023.01)i

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)

IPC: G06N 3/-

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））

CNTXT,ENTXTC,CNKI,万方:模型,网络,分布式,训练,优化,参数,权重,计算,调整,更新,前向,后向,过程中,步骤中,时间内,处理中,同步,同时,实时,重叠,存储,储存,保存,迭代,处理,次数,时间; ENTXT,VEN,Web of Science,IEEE:model,network,distributed,train,optimization,parameter,weight,calculation,adjustment,update,forward,backward,during,within,time,step,process synchronized,real-time,overlap,storage,store,save,preservation,iteration.

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 117669700 A (杭州阿里云飞天信息技术有限公司) 2024年3月8日 (2024 - 03 - 08) 全文	1-17
A	CN 113515370 A (之江实验室 等) 2021年10月19日 (2021 - 10 - 19) 说明书第[0047]-[0080]段	1-17
A	CN 109754060 A (阿里巴巴集团控股有限公司) 2019年5月14日 (2019 - 05 - 14) 全文	1-17
A	CN 113705801 A (华为技术有限公司) 2021年11月26日 (2021 - 11 - 26) 全文	1-17
A	CN 111788585 A (华为技术有限公司) 2020年10月16日 (2020 - 10 - 16) 全文	1-17
A	CN 117093871 A (之江实验室) 2023年11月21日 (2023 - 11 - 21) 全文	1-17

☐ 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

"D" 申请人在国际申请中引证的文件

"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	国际检索报告邮寄日期
2024年11月19日	2024年11月26日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	受权官员 曾贞 电话号码 (+86) 028-62969610

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/118478

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	117669700	A	2024年3月8日	无			
CN	113515370	A	2021年10月19日	CN	113515370	B	2024年3月12日
CN	109754060	A	2019年5月14日	CN	109754060	B	2023年8月25日
CN	113705801	A	2021年11月26日	WO	2021232907	A1	2021年11月25日
				EP	4148624	A1	2023年3月15日
				US	2023087642	A1	2023年3月23日
				EP	4148624	A4	2023年10月18日
CN	111788585	A	2020年10月16日	WO	2020147142	A1	2020年7月23日
				EP	3889846	A1	2021年10月6日
				EP	3889846	A4	2022年6月1日
				US	2021342696	A1	2021年11月4日
				CN	111788585	B	2024年4月12日
CN	117093871	A	2023年11月21日	CN	117093871	B	2024年2月13日