

模型训练配置方法以及相关装置

技术领域

[0001] 本申请涉及通信技术领域，尤其涉及一种模型训练配置方法以及相关装置。

背景技术

[0002] 日益成熟的(artificial intelligence, AI)技术将对未来移动通信网络技术的演进产生重要的推动作用。目前，已将AI技术应用于网络层(如网络优化，移动性管理，资源分配等)和物理层(如信道编译码，信道预测、接收机等)等方面。

[0003] AI技术应用于通信设备中。例如，第一通信设备与第二通信设备是两个AI训练节点。第一通信设备与第二通信设备之间的信道是未知的，第一通信设备可以将数据或参数映射到空口资源上进行传输，即将数据或参数转换为空口上待发送的信号，并向第二通信设备发送待发送的信号。实现第一通信设备与第二通信设备之间的联合训练。第一通信设备与第二通信设备无需采用物理层进行信道编码、符号调制，再生成信号发送。

[0004] 由此可知，如何实现第一通信设备将数据或参数映射到空口资源上进行传输，即如何实现第一通信设备将数据或参数转换为空口上待发送的信号，是值得考虑的问题。

[0005] 发明内容

[0006] 本申请提供一种模型训练配置方法以及相关装置，用于AI任务节点为第一AI训练节点配置至少一个第一附加模块。有利于实现第一AI训练节点通过该至少一个第一附加模块实现与第二AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。

[0007] 本申请第一方面提供一种模型训练配置方法，包括：

[0008] AI任务节点向第一AI训练节点发送第一配置信息，第一配置信息用于配置至少一个第一附加模块，至少一个第一附加模块用于将第一AI训练节点的第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号，和/或，至少一个

第一附加模块用于将第一AI训练节点在空口上接收到的符号转换为用于更新第一子模型的更新参数。从而实现第一AI训练节点通过该至少一个第一附加模块实现与第二AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。

[0009] 基于第一方面，一种可能的实现方式中，方法还包括：AI任务节点向第一AI训练节点发送第一请求，第一请求用于请求第一AI训练节点对第一子模型进行训练。从而实现向第一AI训练节点下发模型训练任务。以便于实现第一AI训练节点与第二AI训练节点进行联合训练。

[0010] 本申请第二方面提供一种模型训练配置方法，包括：

[0011] 第一AI训练节点接收来自AI任务节点的第一配置信息，第一配置信息用于配置至少一个第一附加模块；第一AI训练节点通过至少一个第一附加模块将第一AI训练节点的第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号，并向第二AI训练节点发送所述待输出的符号；和/或，第一AI训练节点通过至少一个第一附加模块将第一AI训练节点在空口上接收到来自第二AI训练节点的符号转换为用于更新第一子模型的更新参数，第一AI训练节点通过更新参数更新第一子模型。由此可知，第一AI训练节点接收来自AI任务节点的第一配置信息，第一配置信息用于配置至少一个第一附加模块。从而实现第一AI训练节点通过该至少一个第一附加模块实现与第二AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。

[0012] 基于第二方面，一种可能的实现方式中，方法还包括：第一AI训练节点接收来自AI任务节点的第一请求，第一请求用于请求第一AI训练节点对第一子模型进行训练。从而实现第一AI训练节点对第一子模型进行训练。以便于实现第一AI训练节点与第二AI训练节点进行联合训练。

[0013] 基于第一方面或第二方面，一种可能的实现方式中，第一配置信息包括以下至少一项：至少一个第一附加模块的索引、类型、结构参数、第一指示信息、第二指示信息、第三指示信息、第四指示信息、或第五指示信息；其中，第一指示信息用于指示至少一个第一附加模块是否参与第一子模型的训练；第二指示信息用于指示至少一个第一附加模块是否作为第一子模型的一部分上报给AI任

务节点；第三指示信息用于指示至少一个第一附加模块中各个第一附加模块的部署方式，部署方式包括串联部署或并联部署；第四指示信息用于指示是否向第一AI训练节点指示至少一个第一附加模块的类型、结构参数；第五指示信息用于指示至少一个第一附加模块中各个第一附加模块插入第一子模型的位置。在该实现方式中，示出了第一配置信息包括的内容，有利于方案的实施。例如，第一配置信息可以包括该至少一个第一附加模块的一些配置。例如，该至少一个第一附加模块的类型、结构参数等。第一AI训练节点通过第三指示信息可以确定各个第一附加模块的部署方式，通过第五指示信息可以确定各个第一附加模块插入第一子模型的位置。从而便于，从而便于第一AI训练节点确定该至少一个第一附加模块。第一AI训练节点通过第一指示信息可以确定是否在第一子模型的训练过程中也训练该至少一个第一附加模块。第一AI训练节点通过第二指示信息可以确定是否在上报经过训练的第一子模型时上报该至少一个第一附加模块。

[0014] 基于第一方面或第二方面，一种可能的实现方式中，第一配置信息还包括第一超网络的索引，第一超网络用于生成至少一个第一附加模块。在该实现方式中，AI任务节点可以向第一AI训练节点指示第一超网络的索引。从而便于第一AI训练节点生成该至少一个第一附加模块。有利于降低AI任务节点配置该至少一个第一附加模块产生的信令开销。

[0015] 基于第一方面或第二方面，一种可能的实现方式中，第五指示信息包括第一比特位图，第一比特位图中的比特与第一子模型中的神经网络层对应，第一比特位图中的比特用于指示在该比特对应的神经网络层是否插入第一附加模块。在该实现方式中，示出了第五指示信息的一种具体的实现形式，通过该第一比特位图指示各个神经网络层是否插入第一附加模块。

[0016] 基于第一方面或第二方面，一种可能的实现方式中，第一比特位图中的一个比特与第一子模型中的一层神经网络对应；当第一比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块；当第一比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一

层神经网络之间不插入第一附加模块；或者，当第一比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块；当第一比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块。在该实现方式中，第一比特位图中的一个比特对应一层神经网络，从而通过该比特的取值指示在该比特对应的神经网络层是否插入第一附加模块。该实现方式产生的比特开销较小，有利于方案的实施。

- [0017] 基于第一方面或第二方面，一种可能的实现方式中，至少一个第一附加模块包括多个第一附加模块；第五指示信息用于指示在第一子模型的第X层神经网络与第X+1层神经网络之间插入多个第一附加模块，X为大于或等于1且小于第一子模型包括的神经网络层数；其中，多个第一附加模块的输入维度之和等于第X层神经网络的输出维度，多个第一附加模块的输出维度之和等于第X+1层神经网络的输入维度；或者，多个第一附加模块的输入维度之和大于第X层神经网络的输出维度，第X层神经网络在输出数据上补零，使得第X层神经网络的输出维度与多个第一附加模块的输入维度之和相等，多个第一附加模块中的一个或多个第一附加模块在输出数据上降维，使得多个第一附加模块的输出维度之和等于第X+1层神经网络的输入维度；或者，
- [0018] 第五指示信息用于指示在第一子模型的最外层神经网络插入多个第一附加模块；其中，多个第一附加模块的输入维度之和等于最外层神经网络的输出维度；或者，多个第一附加模块的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与多个第一附加模块的输入维度之和相等。
- [0019] 在该实现方式中，AI任务节点可以通过第五指示信息指示在一层神经网络插入多个第一附加模块。从而使得该多个第一附加模块的维度与该层神经网络的维度。在该实现方式中，对于第一AI训练节点来说，第一AI训练节点可以存储维度较小的第一附加模块，当一个第一附加模块的维度不足以适配一层神经网络的维度时，AI任务节点可以配置多个第一附加模块插入一层神经网络。因此第

一AI训练节点无需依据神经网络的维度存储各种不同维度的第一附加模块。从而避免第一AI训练节点存储第一附加模块的存储压力问题。需要说明的是，该实现方式也可以作为独立的一方面，即独立于第一方面或第二方面的另一方面提供的一种技术方案。具体该实现方式不依赖于第一方面或第二方面。

[0020] 基于第一方面或第二方面，一种可能的实现方式中，至少一个第一附加模块中各个第一附加模块为神经网络，或者为信号处理模块。示出了至少一个第一附加模块的实现形式，有利于方案的执行。

[0021] 基于第一方面或第二方面，一种可能的实现方式中，第一请求包括以下至少一项：第一子模型的结构、第一超参数、第一损失函数、第一训练停止条件、第一子模型在多个子模型中的编号、或数据集；其中，多个子模型是AI任务节点对AI模型节点请求训练的第一模型拆分得到的，第一训练停止条件是第一AI训练节点停止训练第一子模型的条件，数据集用于作为第一子模型的输入数据。在该实现方式中，提供了第一请求包括的一些可能的内容。从而实现第一AI训练节点结合第一请求包括的内容实现对第一子模型的训练。

[0022] 基于第一方面，一种可能的实现方式中，在AI任务节点向第一AI训练节点发送第一配置信息之前，方法还包括：AI任务节点向第一AI训练节点发送第二请求，第二请求用于请求第一AI训练节点与第二AI训练节点之间的信道状态信息；AI任务节点接收来自第一AI训练节点的信道状态信息；AI任务节点根据任务类型和信道状态信息确定至少一个第一附加模块，任务类型是AI模型节点向AI任务节点请求的任务类型。在该实现方式中，提供了AI任务节点确定至少一个第一附加模块的一种可能的实现方式。该至少一个第一附加模块应当实现对信道干扰的抵抗，且与任务类型相关。因此，AI任务节点根据任务类型和信道状态信息确定至少一个第一附加模块。从而有利于AI任务节点为第一AI训练节点配置合适的第一附加模块。提升第一AI训练节点与第二AI训练节点之间的传输性能。

[0023] 基于第二方面，一种可能的实现方式中，在第一AI训练节点接收来自第一AI任务节点的第一配置信息之前，方法还包括：第一AI训练节点接收来自AI任务节点的第二请求，第二请求用于请求第一AI训练节点与第二AI训练节点之间的信

道状态信息；第一AI训练节点向第二AI训练节点发送信道测量信号；第一AI训练节点接收来自第二AI训练节点的信道状态信息，信道状态信息是第二AI训练节点测量信道测量信号得到的；第一AI训练节点向AI任务节点发送信道状态信息。从而实现AI任务节点获取该信道状态信息，便于AI任务节点为第一AI训练节点配置合适的第一附加模块。提升第一AI训练节点与第二AI训练节点之间的传输性能。

[0024] 基于第一方面，一种可能的实现方式中，方法还包括：AI任务节点向第二AI训练节点发送第二配置信息，第二配置信息用于配置至少一个第二附加模块，至少一个第二附加模块用于将用于更新第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号，和/或，至少一个第二附加模块用于将第二AI训练节点在空口上接收到的符号映射为第二AI训练节点的第二子模型输入的符号。从而实现第二AI训练节点通过该至少一个第二附加模块实现与第一AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。

[0025] 基于第一方面，一种可能的实现方式中，方法还包括：AI任务节点向第二AI训练节点发送第三请求，第三请求用于请求第二AI训练节点对第二子模型进行训练。从而实现向第二AI训练节点下发模型训练任务。以便于实现第一AI训练节点与第二AI训练节点进行联合训练。

[0026] 基于第二方面，一种可能的实现方式中，至少一个第一附加模块包括第一附加模块1；第五指示信息用于指示在第一子模型的第X层神经网络与第X+1层神经网络之间插入第一附加模块1，X为大于或等于1且小于第一子模型包括的神经网络层数，第一附加模块1的输入维度小于第X层神经网络的输出维度；方法还包括：第一AI训练节点确定第一附加模块2，第一AI训练节点在第X层神经网络与第X+1层神经网络之间插入第一附加模块1和第一附加模块2；其中，第一附加模块1的输入维度与第一附加模块2的输入维度之和等于第X层神经网络的输出维度，第一附加模块1的输出维度与第一附加模块2的输出维度之和等于第X+1层神经网络的输入维度；或者，第一附加模块1的输入维度与第一附加模块2的输入维度之和大于第X层神经网络的输出维度，第X层神经网络在输出数据上补

零，使得第X层神经网络的输出维度与第一附加模块1的输入维度与第一附加模块2的输入维度之和相等，第一附加模块1和/或第一附加模块2在输出数据上降维，使得第一附加模块1的输出维度与第一附加模块2的输出维度之和等于第X+1层神经网络的输入维度；或者，

[0027] 至少一个第一附加模块包括第一附加模块1；第五指示信息用于指示在第一子模型的最外层神经网络插入第一附加模块1；方法还包括：第一AI训练节点确定第一附加模块2，第一AI训练节点在最外层神经网络插入第一附加模块1和第一附加模块2；其中，第一附加模块1的输入维度与第一附加模块2的输入维度之和等于最外层神经网络的输出维度；或者，第一附加模块1的输入维度与第一附加模块2的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与第一附加模块1的输入维度与第一附加模块2的输入维度之和相等。

[0028] 在该实现方式中，AI任务节点可以通过第五指示信息指示在一层神经网络插入第一附加模块1。而第一附加模块1的维度远小于该层神经网络的维度。因此第一AI训练节点可以自行确定第二附加模块2，并将第一附加模块1和第一附加模块2插入该层神经网络。从而使得第一附加模块1的维度和第一附加模块2的维度之和与该层神经网络的维度。在该实现方式中，对于第一AI训练节点来说，第一AI训练节点可以存储维度较小的第一附加模块，当一个第一附加模块的维度不足以适配一层神经网络的维度时，第一AI训练节点可以自行确定多个第一附加模块插入一层神经网络。因此第一AI训练节点无需依据神经网络的维度存储各种不同维度的第一附加模块。从而避免第一AI训练节点存储第一附加模块的存储压力问题。需要说明的是，该实现方式也可以作为独立的一方面，即独立于第一方面或第二方面的另一方面提供的一种技术方案。具体该实现方式不依赖于第一方面或第二方面。

[0029] 基于第二方面，一种可能的实现方式中，方法还包括：第一AI训练节点通过第一冗余处理模块对待输出的符号进行冗余处理，再向第二AI训练节点发送经过冗余处理的待输出的符号；和/或，第一AI训练节点通过第一冗余处理模块对第一AI训练节点在空口上接收到来自第二AI训练节点的符号进行去冗余处理，再

通过至少一个第一附加模块将经过去冗余处理的符号转换为更新参数。在该实现方式中，第一AI训练节点可以对待输出的符号进行冗余处理，从而实现冗余发送。有利于提升第一AI训练节点发送的信号功率增益，还有利于提升数据增强的增益。需要说明的是，该实现方式也可以作为独立的一方面，即独立于第一方面或第二方面的另一方面提供的一种技术方案。具体该实现方式不依赖于第一方面或第二方面。

[0030] 基于第二方面，一种可能的实现方式中，第一冗余处理模块的维度等于冗余倍数与至少一个第一附加模块的维度的乘积。从而实现对该至少一个第一附加模块输出的符号进行冗余发送。

[0031] 基于第二方面，一种可能的实现方式中，方法还包括：第一AI训练节点向第二AI训练节点发送信道测量信号；第一AI训练节点接收来自第二AI训练节点的冗余倍数，冗余倍数是第二AI训练节点通过测量信道测量信号确定的。从而实现第一AI训练节点获取该冗余倍数。

[0032] 基于第二方面，一种可能的实现方式中，方法还包括：第一AI训练节点向第二AI训练节点发送信道测量信号；第一AI训练节点接收来自第二AI训练节点的信道状态信息，信道状态信息是第二AI训练节点测量信道测量信号得到的；第一AI训练节点根据信道状态信息确定冗余倍数。从而实现第一AI训练节点获取该冗余倍数。

[0033] 基于第二方面，一种可能的实现方式中，第一冗余处理模块为神经网络，或信号处理模块。

[0034] 本申请第三方面提供一种模型训练配置方法，包括：

[0035] AI任务节点向第二AI训练节点发送第二配置信息，第二配置信息用于配置至少一个第二附加模块，至少一个第二附加模块用于将用于更新第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号，和/或，至少一个第二附加模块用于将第二AI训练节点在空口上接收到的符号映射为第二AI训练节点的第二子模型输入的符号。从而实现第二AI训练节点通过该至少一个第二附加模块实现与第一AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。

- [0036] 基于第三方面，一种可能的实现方式中，方法还包括：AI任务节点向第二AI训练节点发送第三请求，第三请求用于请求第二AI训练节点对第二子模型进行训练。从而实现向第二AI训练节点下发模型训练任务。以便于实现第一AI训练节点与第二AI训练节点进行联合训练。
- [0037] 本申请第四方面提供一种模型训练配置方法，包括：
- [0038] 第二AI训练节点接收来自AI任务节点的第二配置信息，第二配置信息用于配置至少一个第二附加模块；第二AI训练节点通过至少一个第二附加模块将用于更新第一AI训练节点的第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号，并向第一AI训练节点发送所述待输出的符号；和/或，第二AI训练节点通过至少一个第二附加模块将第二AI训练节点在空口上接收到来自第一AI训练节点的符号映射为待输入第二AI训练节点的第二子模型的符号。由此可知，第二AI训练节点接收来自AI任务节点的第二配置信息，第二配置信息用于配置至少一个第二附加模块。从而实现第二AI训练节点通过该至少一个第二附加模块实现与第一AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。
- [0039] 基于第四方面，一种可能的实现方式中，方法还包括：第二AI训练节点接收来自AI任务节点的第三请求，第三请求用于请求第二AI训练节点对第二子模型进行训练。从而实现第二AI训练节点对第二子模型进行训练。以便于实现第一AI训练节点与第二AI训练节点进行联合训练。
- [0040] 基于第三方面或第四方面，一种可能的实现方式中，第二配置信息包括以下至少一项：至少一个第二附加模块的索引、类型、结构参数、第六指示信息、第七指示信息、第八指示信息、第九指示信息、或第十指示信息；其中，第六指示信息用于指示至少一个第二附加模块是否参与第二子模型的训练；第七指示信息用于指示至少一个第二附加模块是否作为第二子模型的一部分上报给AI任务节点；第八指示信息用于指示至少一个第二附加模块中各个第二附加模块的部署方式，部署方式包括串联部署或并联部署；第九指示信息用于指示是否向第二AI训练节点指示至少一个第二附加模块的类型、结构参数；第十指示信息用于指示至少一个第二附加模块中各个第二附加模块插入第二子模型的位置。

置。关于该实现方式的有益效果与前述第一方面或第二方面中第一配置信息包括的内容的实现方式类似，具体可以参阅前述的相关介绍。

[0041] 基于第三方面或第四方面，一种可能的实现方式中，第二配置信息还包括第二超网络的索引，第二超网络用于生成至少一个第二附加模块。该实现方式与前述第一方面或第二方面中第一配置信息还包括第一超网络的索引的实现方式类似，具体可以参阅前述的相关介绍。

[0042] 基于第三方面或第四方面，一种可能的实现方式中，第十指示信息包括第二比特位图，第二比特位图中的比特与第二子模型中的神经网络层对应，第二比特位图中的比特用于指示在该比特对应的神经网络层是否插入第二附加模块。该实现方式与前述第一方面或第二方面中的第一比特位图类似，具体可以参阅前述的相关介绍。

[0043] 基于第三方面或第四方面，一种可能的实现方式中，第二比特位图中的一个比特与第二子模型中的一层神经网络对应；当第二比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第二附加模块；当第二比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第二附加模块；或者，当第二比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第二附加模块；当第二比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第二附加模块。该实现方式与前述第一方面或第二方面中的第一比特位图类似，具体可以参阅前述的相关介绍。

[0044] 基于第三方面或第四方面，一种可能的实现方式中，至少一个第二附加模块包括多个第二附加模块；第十指示信息用于指示在第二子模型的第Y层神经网络与第Y+1层神经网络之间插入多个第二附加模块，Y为大于或等于1且小于第二子模型包括的神经网络层数；其中，多个第二附加模块的输入维度之和等于第Y层神经网络的输出维度，多个第二附加模块的输出维度之和等于第Y+1层神经网络的输入维度；或者，多个第二附加模块的输入维度之和大于第Y层神经网络的输出

维度，第Y层神经网络在输出数据上补零，使得第Y层神经网络的输出维度与多个第二附加模块的输入维度之和相等，多个第二附加模块中的一个或多个第二附加模块在输出数据上降维，使得多个第二附加模块的输出维度之和等于第Y+1层神经网络的输入维度；或者，

- [0045] 第十指示信息用于指示在第二子模型的最外层神经网络插入多个第二附加模块；其中，多个第二附加模块的输入维度之和等于最外层神经网络的输出维度；或者，多个第二附加模块的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与多个第二附加模块的输入维度之和相等。该实现方式与前述第一方面或第二方面中在一层神经网络插入多个第一附加模块的实现方式类似，具体可以参阅前述的相关介绍。
- [0046] 基于第三方面或第四方面，一种可能的实现方式中，该至少一个第二附加模块中各个第二附加模块为神经网络或信号处理模块。
- [0047] 基于第三方面或第四方面，一种可能的实现方式中，第三请求包括以下至少一项：第二子模型的结构、第二超参数、第二损失函数、第二训练停止条件、或第二子模型在多个子模型中的编号；其中，多个子模型是AI任务节点对AI模型节点请求训练的第一模型拆分得到的，第二训练停止条件是第二AI训练节点停止训练第二子模型的条件。该实现方式与前述第一方面或第二方面中的第一请求类似，具体可以参阅前述相关的有益效果介绍。
- [0048] 基于第四方面，一种可能的实现方式中，第二AI训练节点接收来自AI任务节点的第二配置信息之前，方法还包括：第二AI训练节点接收来自第一AI训练节点的信道测量信号；第二AI训练节点向第一AI训练节点发送信道状态信息，信道状态信息是第二AI训练节点测量信道测量信号得到的。从而便于第一AI训练节点向AI任务节点提供该信道状态信息。便于AI任务节点为第二AI训练节点配置合适的第二附加模块。
- [0049] 基于第四方面，一种可能的实现方式中，至少一个第二附加模块包括第二附加模块1；第十指示信息用于指示在第二子模型的第Y层神经网络与第Y+1层神经网络之间插入第二附加模块1；Y为大于或等于1且小于第二子模型包括的神经

网络层数；第二附加模块1的输入维度小于第Y层神经网络的输出维度；方法还包括：第二AI训练节点确定第二附加模块2，第二AI训练节点在第Y层神经网络与第Y+1层神经网络之间插入第二附加模块1和第二附加模块2；其中，第二附加模块1的输入维度与第二附加模块2的输入维度之和等于第Y层神经网络的输出维度，第二附加模块1的输出维度与第二附加模块2的输出维度之和等于第Y+1层神经网络的输入维度；或者，第二附加模块1的输入维度与第二附加模块2的输入维度之和大于第Y层神经网络的输出维度，第Y层神经网络在输出数据上补零，使得第Y层神经网络的输出维度与第二附加模块1的输入维度与第二附加模块2的输入维度之和相等，第二附加模块1和/或第二附加模块2在输出数据上降维，使得第二附加模块1的输出维度与第二附加模块2的输出维度之和等于第Y+1层神经网络的输入维度；或者，

[0050] 至少一个第二附加模块包括第二附加模块1；第十指示信息用于指示在第二子模型的最外层神经网络插入第二附加模块1；方法还包括：第二AI训练节点确定第二附加模块2，第二AI训练节点在最外层神经网络插入第二附加模块1和第二附加模块2；其中，第二附加模块1的输入维度与第二附加模块2的输入维度之和等于最外层神经网络的输出维度；或者，第二附加模块1的输入维度与第二附加模块2的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与第二附加模块1的输入维度与第二附加模块2的输入维度之和相等。

[0051] 基于第四方面，一种可能的实现方式中，方法还包括：第二AI训练节点通过第二冗余处理模块对待输出的符号进行冗余处理，再向第一AI训练节点发送待输出的符号；和/或，第二AI训练节点通过第二冗余处理模块对第二AI训练节点在空口上接收到来自第一AI训练节点的符号进行去冗余处理，再通过至少一个第二附加模块将经过去冗余处理的符号转换为待输入第二AI训练节点的第二子模型的符号。该实现方式与前述第一方面或第二方面中第一AI训练节点采用第一冗余处理模块进行冗余发送或冗余接收的过程类似，具体有益效果可以参阅前述的相关介绍。

- [0052] 基于第四方面，一种可能的实现方式中，第二冗余处理模块的维度等于冗余倍数与至少一个第二附加模块的维度的乘积。从而实现对该至少一个第二附加模块输出的符号进行冗余发送。
- [0053] 基于第四方面，一种可能的实现方式中，方法还包括：第二AI训练节点接收来自第一AI训练节点的信道测量信号；第二AI训练节点向第一AI训练节点发送冗余倍数，冗余倍数是第二AI训练节点通过测量信道测量信号确定的。从而实现第一AI训练节点和第二AI训练节点均获取该冗余倍数。
- [0054] 基于第四方面，一种可能的实现方式中，方法还包括：第二AI训练节点接收来自第一AI训练节点的信道测量信号；第二AI训练节点向第一AI训练节点发送信道状态信息，信道状态信息是第二AI训练节点测量信道测量信号得到的，信道状态信息用于确定所述冗余倍数。从而便于第一AI训练节点确定冗余倍数。
- [0055] 基于第四方面，一种可能的实现方式中，方法还包括：第二AI训练节点接收来自第一AI训练节点的冗余倍数。
- [0056] 基于第四方面，一种可能的实现方式中，第二冗余处理模块为神经网络，或信号处理模块。
- [0057] 本申请第五方面提供一种第一通信装置，包括：
- [0058] 收发模块，用于向第二通信装置发送第一配置信息，第一配置信息用于配置至少一个第一附加模块，至少一个第一附加模块用于将第二通信装置的第一子模型输出的符号映射为第二通信装置在空口上待输出的符号，和/或，至少一个第一附加模块用于将第二通信装置在空口上接收到的符号转换为用于更新第一子模型的更新参数。
- [0059] 基于第五方面，一种可能的实现方式中，收发模块还用于：向第二通信装置发送第一请求，第一请求用于请求第二通信装置对第一子模型进行训练。
- [0060] 本申请第六方面提供一种第二通信装置，包括：
- [0061] 收发模块，用于接收来自第一通信装置的第一配置信息，第一配置信息用于配置至少一个第一附加模块；
- [0062] 处理模块，用于通过至少一个第一附加模块将第二通信装置的第一子模型输出的符号映射为第二通信装置在空口上待输出的符号，并向第二AI训练节点发送

该待输出的符号；和/或，通过至少一个第一附加模块将第二通信装置在空口上接收到来自第二AI训练节点的符号转换为用于更新第一子模型的更新参数；通过更新参数更新第一子模型。

[0063] 基于第六方面，一种可能的实现方式中，收发模块还用于：接收来自第一通信装置的第一请求，第一请求用于请求第二通信装置对第一子模型进行训练。

[0064] 基于第五方面或第六方面，一种可能的实现方式中，第一配置信息包括以下至少一项：至少一个第一附加模块的索引、类型、结构参数、第一指示信息、第二指示信息、第三指示信息、第四指示信息、或第五指示信息；其中，第一指示信息用于指示至少一个第一附加模块是否参与第一子模型的训练；第二指示信息用于指示至少一个第一附加模块是否作为第一子模型的一部分上报给第一通信装置；第三指示信息用于指示至少一个第一附加模块中各个第一附加模块的部署方式，部署方式包括串联部署或并联部署；第四指示信息用于指示是否向第二通信装置指示至少一个第一附加模块的类型、结构参数；第五指示信息用于指示至少一个第一附加模块中各个第一附加模块插入第一子模型的位置。

[0065] 基于第五方面或第六方面，一种可能的实现方式中，第一配置信息还包括第一超网络的索引，第一超网络用于生成至少一个第一附加模块。

[0066] 基于第五方面或第六方面，一种可能的实现方式中，第五指示信息包括第一比特位图，第一比特位图中的比特与第一子模型中的神经网络层对应，第一比特位图中的比特用于指示在该比特对应的神经网络层是否插入第一附加模块。

[0067] 基于第五方面或第六方面，一种可能的实现方式中，第一比特位图中的一个比特与第一子模型中的一层神经网络对应；当第一比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块；当第一比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块；或者，当第一比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块；当第一比特位图中的一个比特的取值为

1, 指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块。

- [0068] 基于第五方面或第六方面, 一种可能的实现方式中, 至少一个第一附加模块包括多个第一附加模块; 第五指示信息用于指示在第一子模型的第X层神经网络与第X+1层神经网络之间插入多个第一附加模块, X为大于或等于1且小于第一子模型包括的神经网络层数; 其中, 多个第一附加模块的输入维度之和等于第X层神经网络的输出维度, 多个第一附加模块的输出维度之和等于第X+1层神经网络的输入维度; 或者, 多个第一附加模块的输入维度之和大于第X层神经网络的输出维度, 第X层神经网络在输出数据上补零, 使得第X层神经网络的输出维度与多个第一附加模块的输入维度之和相等, 多个第一附加模块中的一个或多个第一附加模块在输出数据上降维, 使得多个第一附加模块的输出维度之和等于第X+1层神经网络的输入维度; 或者,
- [0069] 第五指示信息用于指示在第一子模型的最外层神经网络插入多个第一附加模块; 其中, 多个第一附加模块的输入维度之和等于最外层神经网络的输出维度; 或者, 多个第一附加模块的输入维度之和大于最外层神经网络的输出维度, 最外层神经网络在输出数据上补零, 使得最外层神经网络的输出维度与多个第一附加模块的输入维度之和相等。
- [0070] 基于第五方面或第六方面, 一种可能的实现方式中, 至少一个第一附加模块中各个第一附加模块为神经网络, 或者为信号处理模块。
- [0071] 基于第五方面或第六方面, 一种可能的实现方式中, 第一请求包括以下至少一项: 第一子模型的结构、第一超参数、第一损失函数、第一训练停止条件、第一子模型在多个子模型中的编号、或数据集; 其中, 多个子模型是第一通信装置对AI模型节点请求训练的第一模型拆分得到的, 第一训练停止条件是第二通信装置停止训练第一子模型的条件, 数据集用于作为第一子模型的输入数据。
- [0072] 基于第五方面, 一种可能的实现方式中, 收发模块还用于: 向第二通信装置发送第二请求, 第二请求用于请求第二通信装置与第二AI训练节点之间的信道状态信息; 接收来自第二通信装置的信道状态信息; 第一通信装置还包括处理

模块；处理模块，用于根据任务类型和信道状态信息确定至少一个第一附加模块，任务类型是AI模型节点向第一通信装置请求的任务类型。

[0073] 基于第六方面，一种可能的实现方式中，收发模块还用于：接收来自第一通信装置的第二请求，第二请求用于请求第二通信装置与第二AI训练节点之间的信道状态信息；向第二AI训练节点发送信道测量信号；接收来自第二AI训练节点的信道状态信息，信道状态信息是第二AI训练节点测量信道测量信号得到的；向第一通信装置发送信道状态信息。

[0074] 基于第五方面，一种可能的实现方式中，收发模块还用于：向第二AI训练节点发送第二配置信息，第二配置信息用于配置至少一个第二附加模块，至少一个第二附加模块用于将用于更新第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号，和/或，至少一个第二附加模块用于将第二AI训练节点在空口上接收到的符号映射为第二AI训练节点的第二子模型输入的符号。

[0075] 基于第五方面，一种可能的实现方式中，收发模块还用于：向第二AI训练节点发送第三请求，第三请求用于请求第二AI训练节点对第二子模型进行训练。

[0076] 基于第六方面，一种可能的实现方式中，至少一个第一附加模块包括第一附加模块1；第五指示信息用于指示在第一子模型的第X层神经网络与第X+1层神经网络之间插入第一附加模块1，X为大于或等于1且小于第一子模型包括的神经网络层数，第一附加模块1的输入维度小于第X层神经网络的输出维度；处理模块还用于：

[0077] 确定第一附加模块2；在第X层神经网络与第X+1层神经网络之间插入第一附加模块1和第一附加模块2；其中，第一附加模块1的输入维度与第一附加模块2的输入维度之和等于第X层神经网络的输出维度，第一附加模块1的输出维度与第一附加模块2的输出维度之和等于第X+1层神经网络的输入维度；或者，第一附加模块1的输入维度与第一附加模块2的输入维度之和大于第X层神经网络的输出维度，第X层神经网络在输出数据上补零，使得第X层神经网络的输出维度与第一附加模块1的输入维度与第一附加模块2的输入维度之和相等，第一附加模块1和/或第一附加模块2在输出数据上降维，使得第一附加模块1的输出维度与第一附加模块2的输出维度之和等于第X+1层神经网络的输入维度；或者，

- [0078] 至少一个第一附加模块包括第一附加模块1；第五指示信息用于指示在第一子模型的最外层神经网络插入第一附加模块1；处理模块还用于：
- [0079] 确定第一附加模块2；第二通信装置在最外层神经网络插入第一附加模块1和第一附加模块2；其中，第一附加模块1的输入维度与第一附加模块2的输入维度之和等于最外层神经网络的输出维度；或者，第一附加模块1的输入维度与第一附加模块2的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与第一附加模块1的输入维度与第一附加模块2的输入维度之和相等。
- [0080] 基于第六方面，一种可能的实现方式中，处理模块还用于：通过第一冗余处理模块对待输出的符号进行冗余处理，再向第二AI训练节点发送经过冗余处理的待输出的符号；和/或，通过第一冗余处理模块对第二通信装置在空口上接收到来自第二AI训练节点的符号进行去冗余处理，再通过至少一个第一附加模块将过去冗余处理的符号转换为更新参数。
- [0081] 基于第六方面，一种可能的实现方式中，第一冗余处理模块的维度等于冗余倍数与至少一个第一附加模块的维度的乘积。
- [0082] 基于第六方面，一种可能的实现方式中，收发模块还用于：向第二AI训练节点发送信道测量信号；接收来自第二AI训练节点的冗余倍数，冗余倍数是第二AI训练节点通过测量信道测量信号确定的。
- [0083] 基于第六方面，一种可能的实现方式中，收发模块还用于：向第二AI训练节点发送信道测量信号；接收来自第二AI训练节点的信道状态信息，信道状态信息是第二AI训练节点测量信道测量信号得到的；处理模块还用于：根据信道状态信息确定冗余倍数。
- [0084] 基于第六方面，一种可能的实现方式中，第一冗余处理模块为神经网络，或信号处理模块。
- [0085] 本申请第七方面提供一种第一通信装置，包括：
- [0086] 收发模块，用于向第三通信装置发送第二配置信息，第二配置信息用于配置至少一个第二附加模块，至少一个第二附加模块用于将用于更新第一子模型的更新参数转换为第三通信装置在空口上待输出的符号，和/或，至少一个第二附加

模块用于将第三通信装置在空口上接收到的符号映射为第三通信装置的第二子模型输入的符号。

- [0087] 基于第七方面，一种可能的实现方式中，收发模块，还用于向第三通信装置发送第三请求，第三请求用于请求第三通信装置对第二子模型进行训练。
- [0088] 本申请第八方面提供一种第三通信装置，包括：
- [0089] 收发模块，用于接收来自第一通信装置的第二配置信息，第二配置信息用于配置至少一个第二附加模块；
- [0090] 处理模块，用于通过至少一个第二附加模块将用于更新第一AI训练节点的第一子模型的更新参数转换为第三通信装置在空口上待输出的符号，并向第一AI训练节点发送所述待输出的符号；和/或，通过至少一个第二附加模块将第三通信装置在空口上接收到来自第一AI训练节点的符号映射为待输入第三通信装置的第二子模型的符号。
- [0091] 基于第八方面，一种可能的实现方式中，收发模块还用于：接收来自第一通信装置的第三请求，第三请求用于请求第三通信装置对第二子模型进行训练。
- [0092] 关于第七方面或第八方面中的其他可能的实现方式可以参阅前述第三方面或第四方面中的相关介绍，这里不再一一赘述。
- [0093] 基于第七方面或第八方面，一种可能的实现方式中，第二配置信息包括以下至少一项：至少一个第二附加模块的索引、类型、结构参数、第六指示信息、第七指示信息、第八指示信息、第九指示信息、或第十指示信息；其中，第六指示信息用于指示至少一个第二附加模块是否参与第二子模型的训练；第七指示信息用于指示至少一个第二附加模块是否作为第二子模型的一部分上报给第一通信装置；第八指示信息用于指示至少一个第二附加模块中各个第二附加模块的部署方式，部署方式包括串联部署或并联部署；第九指示信息用于指示是否向第三通信装置指示至少一个第二附加模块的类型、结构参数；第十指示信息用于指示至少一个第二附加模块中各个第二附加模块插入第二子模型的位置。
- [0094] 基于第七方面或第八方面，一种可能的实现方式中，第二配置信息还包括第二超网络的索引，第二超网络用于生成至少一个第二附加模块。

- [0095] 基于第七方面或第八方面，一种可能的实现方式中，第十指示信息包括第二比特位图，第二比特位图中的比特与第二子模型中的神经网络层对应，第二比特位图中的比特用于指示在该比特对应的神经网络层是否插入第二附加模块。
- [0096] 基于第七方面或第八方面，一种可能的实现方式中，第二比特位图中的一个比特与第二子模型中的一层神经网络对应；当第二比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第二附加模块；当第二比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第二附加模块；或者，当第二比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第二附加模块；当第二比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第二附加模块。
- [0097] 基于第七方面或第八方面，一种可能的实现方式中，至少一个第二附加模块包括多个第二附加模块；第十指示信息用于指示在第二子模型的第Y层神经网络与第Y+1层神经网络之间插入多个第二附加模块，Y为大于或等于1且小于第二子模型包括的神经网络层数；其中，多个第二附加模块的输入维度之和等于第Y层神经网络的输出维度，多个第二附加模块的输出维度之和等于第Y+1层神经网络的输入维度；或者，多个第二附加模块的输入维度之和大于第Y层神经网络的输出维度，第Y层神经网络在输出数据上补零，使得第Y层神经网络的输出维度与多个第二附加模块的输入维度之和相等，多个第二附加模块中的一个或多个第二附加模块在输出数据上降维，使得多个第二附加模块的输出维度之和等于第Y+1层神经网络的输入维度；或者，
- [0098] 第十指示信息用于指示在第二子模型的最外层神经网络插入多个第二附加模块；其中，多个第二附加模块的输入维度之和等于最外层神经网络的输出维度；或者，多个第二附加模块的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与多个第二附加模块的输入维度之和相等。

- [0099] 基于第七方面或第八方面，一种可能的实现方式中，该至少一个第二附加模块中各个第二附加模块为神经网络或信号处理模块。
- [0100] 基于第七方面或第八方面，一种可能的实现方式中，第三请求包括以下至少一项：第二子模型的结构、第二超参数、第二损失函数、第二训练停止条件、或第二子模型在多个子模型中的编号；其中，多个子模型是第一通信装置对AI模型节点请求训练的第一模型拆分得到的，第二训练停止条件是第三通信装置停止训练第二子模型的条件。
- [0101] 基于第八方面，一种可能的实现方式中，收发模块还用于：接收来自第一AI训练节点的信道测量信号；向第一AI训练节点发送信道状态信息，信道状态信息是第三通信装置测量信道测量信号得到的。
- [0102] 基于第八方面，一种可能的实现方式中，至少一个第二附加模块包括第二附加模块1；第十指示信息用于指示在第二子模型的第Y层神经网络与第Y+1层神经网络之间插入第二附加模块1；Y为大于或等于1且小于第二子模型包括的神经网络层数；第二附加模块1的输入维度小于第Y层神经网络的输出维度；处理模块还用于：确定第二附加模块2；在第Y层神经网络与第Y+1层神经网络之间插入第二附加模块1和第二附加模块2；其中，第二附加模块1的输入维度与第二附加模块2的输入维度之和等于第Y层神经网络的输出维度，第二附加模块1的输出维度与第二附加模块2的输出维度之和等于第Y+1层神经网络的输入维度；或者，第二附加模块1的输入维度与第二附加模块2的输入维度之和大于第Y层神经网络的输出维度，第Y层神经网络在输出数据上补零，使得第Y层神经网络的输出维度与第二附加模块1的输入维度与第二附加模块2的输入维度之和相等，第二附加模块1和/或第二附加模块2在输出数据上降维，使得第二附加模块1的输出维度与第二附加模块2的输出维度之和等于第Y+1层神经网络的输入维度；或者，
- [0103] 至少一个第二附加模块包括第二附加模块1；第十指示信息用于指示在第二子模型的最外层神经网络插入第二附加模块1；处理模块还用于：确定第二附加模块2；在最外层神经网络插入第二附加模块1和第二附加模块2；其中，第二附加模块1的输入维度与第二附加模块2的输入维度之和等于最外层神经网络的输出维度；或者，第二附加模块1的输入维度与第二附加模块2的输入维度之和大于

最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与第二附加模块1的输入维度与第二附加模块2的输入维度之和相等。

- [0104] 基于第八方面，一种可能的实现方式中，处理模块还用于：通过第二冗余处理模块对待输出的符号进行冗余处理，再向第一AI训练节点发送待输出的符号；和/或，通过第二冗余处理模块对第三通信装置在空口上接收到来自第一AI训练节点的符号进行去冗余处理，再通过至少一个第二附加模块将经过去冗余处理的符号转换为待输入第三通信装置的第二子模型的符号。
- [0105] 基于第八方面，一种可能的实现方式中，第二冗余处理模块的维度等于冗余倍数与至少一个第二附加模块的维度的乘积。
- [0106] 基于第八方面，一种可能的实现方式中，收发模块还用于：接收来自第一AI训练节点的信道测量信号；向第一AI训练节点发送冗余倍数，冗余倍数是第三通信装置通过测量信道测量信号确定的。
- [0107] 基于第八方面，一种可能的实现方式中，收发模块还用于：接收来自第一AI训练节点的信道测量信号；向第一AI训练节点发送信道状态信息，信道状态信息是第三通信装置测量信道测量信号得到的，信道状态信息用于确定所述冗余倍数。
- [0108] 基于第八方面，一种可能的实现方式中，收发模块还用于：接收来自第一AI训练节点的冗余倍数。
- [0109] 基于第八方面，一种可能的实现方式中，第二冗余处理模块为神经网络，或信号处理模块。
- [0110] 本申请第九方面提供一种通信装置，该通信装置包括处理器。该处理器用于调用存储起中的计算机程序或计算机指令，使得处理器实现如第一方面至第四方面中任一方面中的任意一种实现方式。
- [0111] 可选的，该通信装置还包括收发器，该处理器用于控制该收发器收发信号。
- [0112] 本申请第十方面提供一种通信装置，该通信装置包括处理器，处理器用于执行如第一方面至第四方面中任一方面的任意一种实现方式。

- [0113] 本申请第十一方面提供一种包括指令的计算机程序产品，其特征在于，当其在计算机上运行时，使得该计算机执行如第一方面至第四方面中任一方面中的任一种的实现方式。
- [0114] 本申请第十二方面提供一种计算机可读存储介质，包括计算机指令，当该指令在计算机上运行时，使得计算机执行如第一方面至第四方面中任一方面中的任一种实现方式。
- [0115] 本申请第十三方面提供一种芯片装置，包括处理器，用于调用该存储器中的计算机程序或计算机指令，以使得该处理器执行上述如第一方面至第四方面中任一方面中的任一种实现方式。
- [0116] 可选的，该处理器通过接口与该存储器耦合。
- [0117] 本申请第十四方面提供一种通信系统，通信系统包括AI任务节点和第一AI训练节点；AI任务节点用于执行如第一方面所示的方法，第一AI训练节点用于执行如第二方面所示的方法。可选的，通信系统还包括第三AI训练节点，AI任务节点用于执行如第三方面所示的方法，第三AI训练节点用于执行如第四方面所示的方法。
- [0118] 本申请第十五方面提供一种通信系统，通信系统包括第一AI训练节点和第二AI训练节点；第一AI训练节点用于执行如第二方面所示的方法，第二AI训练节点用于执行如第四方面所示的方法。可选的，通信系统还包括AI任务节点，AI任务节点用于执行如第一方面或第三方面所示的方法。

附图说明

- [0119] 图1A为本申请实施例通信系统的一个结构示意图；
- [0120] 图1B为本申请实施例通信系统的另一个结构示意图；
- [0121] 图1C为本申请实施例通信系统的再一个结构示意图；
- [0122] 图1D为本申请实施例通信系统的又一个结构示意图；
- [0123] 图2A为本申请实施例通信系统的又一个结构示意图；
- [0124] 图2B为本申请实施例通信系统的又一个结构示意图；
- [0125] 图2C为本申请实施例通信系统的又一个结构示意图；
- [0126] 图2D为本申请实施例神经网络的一个结构示意图；

- [0127] 图2E为本申请实施例中间梯度的一个计算示意图；
- [0128] 图3为本申请实施例模型训练配置方法的一个实施例示意图；
- [0129] 图4为本申请实施例第一AI训练节点与第二AI训练节点之间的交互示意图；
- [0130] 图5A为本申请实施例至少一个第一附加模块插入第一子模型以及至少一个第二附加模块插入第二子模型的一个示意图；
- [0131] 图5B为本申请实施例至少一个第一附加模块插入第一子模型以及至少一个第二附加模块插入第二子模型的另一个示意图；
- [0132] 图5C为本申请实施例至少一个第一附加模块插入第一子模型以及至少一个第二附加模块插入第二子模型的再一个示意图；
- [0133] 图6为本申请实施例AI任务节点下发超网络的一个示意图；
- [0134] 图7为本申请实施例AI任务节点下发至少一个第一附加模块的类型、结构参数和/或索引的一个示意图；
- [0135] 图8为本申请实施例第一附加模块1和第二附加模块2适配第一子模型的最外层神经网络的示意图；
- [0136] 图9为本申请实施例提供的技术方案产生的传输开销相比于采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销的一个效果示意图；
- [0137] 图10为通过本申请实施例提供的技术方案得到的第一模型对图片分类的准确度相比于采用应用层和物理层进行数据和/或参数的传输的方案得到的模型对图片分类的准确度的一个效果示意图；
- [0138] 图11为本申请实施例第一AI训练节点与第二AI训练节点之间进行冗余传输的一个示意图；
- [0139] 图12为本申请实施例第一AI训练节点与第二AI训练节点之间进行冗余传输所带来的增益的一个示意图；
- [0140] 图13为本申请实施例模型训练配置方法的另一个实施例示意图；
- [0141] 图14为本申请实施例包含至少一个第一附加模块或至少一个第二附加模块的第一模型的一个示意图；
- [0142] 图15为本申请实施例去除至少一个第一附加模块和至少一个第二附加模块的第一模型的一个示意图；

- [0143] 图16为本申请实施例通信装置的一个结构示意图；
- [0144] 图17为本申请实施例通信装置的另一个结构示意图；
- [0145] 图18为本申请实施例通信装置的再一个结构示意图；
- [0146] 图19为本申请实施例终端设备的一个结构示意图；
- [0147] 图20为本申请实施例基站的一个结构示意图。

具体实施方式

- [0148] 本申请实施例提供一种模型训练配置方法以及相关装置，用于AI任务节点为第一AI训练节点配置至少一个第一附加模块。有利于实现第一AI训练节点通过该至少一个第一附加模块实现与第二AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。
- [0149] 在本申请中描述的参考“一个实施例”或“一些实施例”等意味着在本申请的一个或多个实施例中包括结合该实施例描述的特定特征、结构或特点。由此，在本说明书中的不同之处出现的语句“在一个实施例中”、“在一些实施例中”、“在其他一些实施例中”、“在另外一些实施例中”等不是必然都参考相同的实施例，而是意味着“一个或多个但不是所有的实施例”，除非是以其他方式另外特别强调。术语“包括”、“包含”、“具有”及它们的变形都意味着“包括但不限于”，除非是以其他方式另外特别强调。
- [0150] 在本申请的描述中，除非另有说明，“/”表示“或”的意思，例如，A/B可以表示A或B。本文中的“和/或”仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A和/或B，可以表示：单独存在A，同时存在A和B，单独存在B这三种情况。此外，“至少一个”是指一个或多个，“多个”是指两个或两个以上。“以下至少一项(个)”或其类似表达，是指的这些项中的任意组合，包括单项(个)或复数项(个)的任意组合。例如，a，b，或c中的至少一项(个)，可以表示：a，b，c，a和b，a和c，b和c，或a和b和c。其中a，b，c可以是单个，也可以是多个。
- [0151] 本申请实施例的技术方案可应用于第五代移动通信(5th-generation, 5G)、卫星通信及短距等无线通信系统中。本申请实施例提及的无线通信系统包括但不限于：窄带物联网系统(Narrow Band-Internet of Things, NB-IoT)、长期演

进系统(Long Term Evolution, LTE)、5G通信系统的三大应用场景增强移动宽带(enhanced mobile broadband, eMBB),超可靠低时延通信(ultra reliable low latency communication, URLLC)和海量机器类通信(massive machine type of communication, mMTC)以及随着通信技术的不断发展,还包括第六代移动通信(6th-generation, 6G)等后续演进的通信系统等。

[0152] 本申请提供的技术方案应用的通信系统包括: AI任务节点和第一AI训练节点。可选的,通信系统还包括第二AI训练节点。可选的,进一步的,通信系统还包括AI模型节点。

[0153] 或者说,本申请提供的技术方案应用的通信系统包括: 第一AI训练节点和第二AI训练节点。可选的,通信系统还包括AI任务节点。可选的,进一步的,通信系统还包括AI模型节点。需要说明的是,上述通信系统还可以包括更多AI训练节点,具体本申请不做限定。例如,通信系统还包括第三AI训练节点。

[0154] 本申请中, AI模型节点用于向AI任务节点发起模型训练请求。该模型训练请求用于AI模型节点请求对该AI模型节点请求的任务类型对应的模型进行训练。AI任务节点用于对该任务类型对应的模型进行拆分得到多个子模型,选择相应的AI训练节点并请求AI训练节点对相应的子模型进行训练。AI训练节点用于接收来自AI任务节点的请求,并对相应的子模型进行训练。

[0155] 本申请中,可选的, AI任务节点可以为网络设备或服务器,具体本申请不做限定。AI训练节点可以为网络设备或终端设备,具体本申请不做限定。AI模型节点可以为网络设备或终端设备,具体本申请不做限定。可选的,网络设备可以为核心网设备或接入网设备。

[0156] 下面结合图1A至图1D介绍本申请的一些可能的应用场景。对于其他应用场景本申请仍适用,具体本申请不做限定。

[0157] 图1A为本申请实施例通信系统的一个结构示意图。请参阅图1A,通信系统包括网络设备、终端设备1、终端设备2和终端设备3。AI任务节点可以为接入网设备。第一AI训练节点为终端设备1,第二AI训练节点为终端设备2。AI模型节点为终端设备3。接入网设备与终端设备可以执行本申请的技术方案。

- [0158] 图1A所示的设备数量和形态用于举例并不构成对本申请实施例的限定。实际应用中，图1A所述的通信系统可以包括更多的接入网设备和更多的终端设备。
- [0159] 图1B为本申请实施例通信系统的另一个结构示意图。请参阅图1B，通信系统包括服务器、接入网设备1、接入网设备2和终端设备1。AI任务节点可以为服务器，第一AI训练节点为接入网设备1，第二AI训练节点为终端设备1。AI模型节点为接入网设备2。服务器、接入网设备和终端设备可以执行本申请的技术方案。
- [0160] 图1B所示的设备数量和形态用于举例并不构成对本申请实施例的限定。实际应用中，图1B所述的通信系统可以包括更多的服务器、更多的接入网设备和更多的终端设备。
- [0161] 图1C为本申请实施例通信系统的再一个结构示意图。请参阅图1C，通信系统包括核心网设备、接入网设备1、接入网设备2和接入网设备3。AI任务节点可以为核心网设备，第一AI训练节点可以为接入网设备1，第二AI训练节点可以为接入网设备2。AI模型节点可以为接入网设备3。核心网设备与接入网设备可以执行本申请的技术方案。
- [0162] 图1C所示的设备数量和形态用于举例并不构成对本申请实施例的限定。实际应用中，图1C所述的通信系统可以包括更多的核心网设备和更多的接入网设备。
- [0163] 图1D为本申请实施例通信系统的又一个结构示意图。请参阅图1D，通信系统包括服务器、接入网设备1、接入网设备2和接入网设备3。AI任务节点可以为服务器，第一AI训练节点可以为接入网设备1，第二AI训练节点可以为接入网设备2。AI模型节点可以为接入网设备3。服务器与接入网设备可以执行本申请的技术方案。
- [0164] 图1D所示的设备数量和形态用于举例并不构成对本申请实施例的限定。实际应用中，图1D所述的通信系统可以包括更多的服务器和更多的接入网设备。
- [0165] 下面介绍本申请涉及的终端设备、接入网设备、核心网设备和服务器。
- [0166] 终端设备，又称之为用户设备(user equipment, UE)、移动台(mobile station, MS)、移动终端(mobile terminal, MT)、客户前置设备(customer premise equipment, CPE)等。终端设备是包括无线通信功能(向用户提供语音/数据连通性)的设备。例如，具有无线连接功能的手持式设备、车载设

备、机器类型通信(machine type communication, MTC)终端等。目前, 终端设备可以包括: 手机(mobile phone)、平板电脑、笔记本电脑、掌上电脑、移动互联网设备(mobile internet device, MID)、可穿戴设备, 虚拟现实(virtual reality, VR)设备、增强现实(augmented reality, AR)设备、工业控制(industrial control)中的无线终端、无人驾驶(self driving)中的无线终端、远程手术(remote medical surgery)中的无线终端、智能电网(smart grid)中的无线终端、运输安全(transportation safety)中的无线终端、智慧城市(smart city)中的无线终端、或智慧家庭(smart home)中的无线终端等。例如, 无人驾驶中的无线终端可以为无人机、直升机、或飞机等。例如, 车联网中的无线终端可以为车载设备、整车设备、车载模块、车辆、或轮船等。工业控制中的无线终端可以为摄像头、机器人、或机械臂等。智慧家庭中的无线终端可以为电视、空调、扫地机、音箱、或机顶盒等。

[0167] 需要说明的是, 终端设备可以是带有芯片的设备或装置, 或者集成有电路的设备或装置, 或者是上述示出的设备或装置中的芯片、模块或控制单元, 具体本申请不做限定。

[0168] 接入网设备是一种部署在无线接入网中为终端设备提供无线通信功能的设备。接入网设备又可以称为接入网(radio access network, RAN)实体、接入节点、网络节点、网络设备、或通信装置等。

[0169] 具体的, 接入网设备可以为第三代合作伙伴计划(3rd generation partnership project, 3GPP)相关的蜂窝系统的接入网设备。例如, 第四代(fourth-generation, 4G)移动通信系统, 或5G移动通信系统。接入网设备还可以是开放式接入网(open RAN, O-RAN或ORAN)或云无线接入网络(cloud radio access network, CRAN)中的接入网设备。或者, 接入网设备还可以是以上两种或两种以上的通信系统融合得到的通信系统中的接入网设备。

[0170] 接入网设备包括但不限于: 演进型节点B(evolved Node B, eNB)、无线网络控制器(radio network controller, RNC)、节点B(Node B, NB)、基站控制器(base station controller, BSC)、基站收发台(base transceiver station, BTS)、家庭基站(例如, home evolved NodeB, 或

home Node B, HNB)、基带单元(baseband unit, BBU), 无线保真(wireless fidelity, WIFI)系统中的接入点(access point, AP)、宏基站、微基站、无线中继节点、施主节点、CRAN场景下的无线控制器、无线回传节点、传输点(transmission point, TP)或者传输接收点(transmission and receiving point, TRP)。接入网设备还可以为5G移动通信系统中的接入网设备。例如, 新空口(new radio, NR)系统中的下一代基站(next generation NodeB, gNB), TRP, TP, 或者, 5G移动通信系统中的基站的一个或一组(包括多个天线面板)天线面板。或者, 接入网设备还可以为构成gNB或传输点的网络节点。例如, 集中式单元(centralized unit, CU)、分布式单元(distributed unit, DU)、CU-控制面(control plane, CP)、CU-用户面(user plane, UP)、或者无线单元(radio unit, RU)等。CU和DU可以是单独设置, 或者也可以包括在同一网元中。例如, BBU。RU可以包括在射频设备或射频单元中。例如, 在射频拉远单元(remote radio unit, RRU)、有源天线处理单元(active antenna unit, AAU)或远程射频头(remote radio head, RRH)中。或者, 接入网设备还可以是服务器、可穿戴设备、车辆或车载设备等。例如, V2X技术中的接入网设备可以是路侧单元(road side unit, RSU)。

[0171] 需要说明的是, 在不同系统中, CU(或CU-CP和CU-UP)、DU或RU也可以有不同的名称, 但是本领域的技术人员可以理解其含义。例如, 在ORAN系统中, CU也可以称为开放式集中式单元(open centralized unit, O-CU)或者开放式CU, DU也可以称为开放式分布式单元(open distributed unit, O-DU), 集中式单元-控制面(centralized unit control plane, CU-CP)也可以称为开放式集中式单元-控制面(open centralized unit control plane, O-CU-CP)或开放式CU-CP, 集中式单元-用户面(centralized unit user plane, CU-UP)也可以称为开放式集中式单元-用户面(open centralized unit user plane, O-CU-UP)或开放式CU-UP, RU也可以称为开放式无线单元(open radio unit, O-RU), 具体本申请不做限定。本申请中的CU、CU-CP、CU-UP、DU和RU中的任一单元, 可以通过软件模块、硬件模块、或者软件模块与硬件模块结合来一起实现。

[0172] 可选的，对于ORAN系统中的网元，各个网元可以实现如下述表1所示的协议层功能。

[0173] 表1

ORAN 网元	协议层功能
O-CU-CP	无线资源控制 (radio resource control, RRC), 分组数据汇聚层协议控制面 (packet data convergence protocol control plane, PDCP-C)
O-CU-UP	服务数据适配协议 (service data adaptation protocol, SDAP), 分组数据汇聚层协议用户面 (packet data convergence protocol user plane , PDCP-U)
O-DU	无线链路控制 (radio link control, RLC), 媒体接入控制 (media access control, MAC), 物理高层 (PHY-high)
O-RU	物理低层 (PHY-low)

[0174] 需要说明的是，上述表1仅仅是一种示例。实际应用中，各个网元支持的协议层功能并不限定。例如，各个网元可以支持更多的协议层功能，或者具体结合实际配置各个网元支持的协议层功能。具体本申请不做限定。

[0175] 下面介绍接入网设备的CU和DU的架构。接入网设备包括至少一个CU和至少一个DU。可选的，接入网设备还包括至少一个RU。

[0176] 下面以接入网设备包括一个CU和一个DU为例进行介绍。CU具有核心网的部分功能，CU可以包括CU-CP和CU-UP。CU和DU可以根据其实现的无线网络的协议层功能进行配置。例如，CU被配置为用以实现分组数据汇聚层协议 (packet data convergence protocol, PDCP) 层及以上协议层 (例如，RRC层和/或SDAP层的功能。DU被配置为用以实现PDCP层以下协议层 (例如，RLC层、MAC层、和/或物理 (physical, PHY) 层) 的功能。又例如，CU被配置为用以实现PDCP层以上协议层 (如RRC层和/或SDAP层) 的功能，DU被配置为用以实现PDCP层及以下协议层 (例如RLC层、MAC层、和/或PHY层等) 的功能。

- [0177] 当CU包括CU-CP和CU-UP时，CU-CP用于实现CU的控制面功能，CU-UP用于实现CU的用户面功能。例如CU被配置为用以实现PDCP层、RRC层和SDAP层的功能时，CU-CP用于实现RRC层功能和PDCP层的控制面功能，CU-UP用于实现SDAP层功能和PDCP层的用户面功能。
- [0178] CU-CP可以与核心网中用于实现控制面功能的网元交互。核心网中用于实现控制面功能的网元可以是接入和移动性功能网元，例如5G移动通信系统中的接入和移动功能(access and mobility management function, AMF)。所述接入和移动性功能网元用于负责移动网络中的移动性管理，如终端设备的位置更新、终端设备的注册网络、终端设备的切换等。CU-UP可以与核心网中用于实现用户面功能的网元交互。核心网中用于实现用户面功能的网元，例如，5G移动通信系统中的用户功能(user plane function, UPF)，用于负责终端设备中数据的转发和接收。
- [0179] 以上CU，DU的配置仅仅是一种举例，也可以根据需要配置CU，DU具有的功能。例如，可以将CU或者DU配置为具有更多协议层的功能，或者将CU或DU配置为具有协议层的部分处理功能。例如，将RLC层的部分功能和RLC层以上的协议层的功能设置在CU，将RLC层的剩余功能和RLC层以下的协议层的功能设置在DU。再例如，可以按照业务类型或者其他系统需求对CU或者DU的功能进行划分。例如，按时延划分，将处理时间需要满足较小时延要求的功能设置在DU，不需要满足该时延要求的功能设置在CU。
- [0180] DU和RU可以合作共同实现PHY层的功能。一个DU可以和一个或多个RU相连。DU和RU所具有的功能可以根据设计被配置为多种方式。例如，DU被配置用于实现基带功能，RU被配置用于实现中射频功能。再例如，DU被配置为用以实现PHY层中的高层功能，RU被配置为实现PHY层中的低层功能或者实现该低层功能和射频功能。物理层中的高层功能可以包括物理层的一部分功能，该部分功能更加靠近MAC层，物理层中的低层功能可以包括物理层的另一部分功能，该部分功能更加靠近中射频侧。
- [0181] 需要说明的是，接入网设备可以是上述示出的设备或装置，也可以是上述示出的设备或装置中的部件(例如，芯片)、模块、或单元，具体本申请不做限定。

- [0182] 核心网设备负责终端设备接入网络的接入控制、注册管理、业务管理、移动性管理等。例如，核心网设备为接入与移动性管理功能(access and mobility management function, AMF)。
- [0183] 服务器用于为特定应用提供业务功能。例如，服务器可以是特定应用的服务器。
- [0184] 本申请提供的通信系统中可以引入AI网元来实现部分或全部AI相关操作。AI网元也可以称为AI节点、AI设备、AI实体、AI模块、AI模型、或AI单元等。所述AI网元可以是内置在通信系统的网元中。例如，AI网元可以是内置在：接入网设备、核心网设备、云服务器、或网管(operation, administration and maintenance, OAM)中的AI模块，用以实现AI相关的功能。所述OAM可以是作为核心网设备网管和/或作为接入网设备的网管。或者，所述AI网元也可以是通信系统中独立设置的网元。可选的，终端或终端内置的芯片中也可以包括AI实体，用于实现AI相关的功能。下面结合图2A介绍本申请实施例提供的另一种通信系统。
- [0185] 图2A为本申请实施例通信系统的又一个结构示意图。如图2A所示，通信系统中网元之间通过接口(例如，NG接口，或Xn接口)，或空口相连。这些网元节点(例如，核心网设备、接入网节点(RAN节点)、终端或OAM中的一个或多个设备)中设置有一个或多个AI模块(为清楚起见，图2A中仅示出1个)。接入网节点可以作为单独的RAN节点，也可以包括多个RAN节点，例如，包括CU和DU。CU和/或DU也可以设置有一个或多个AI模块。可选的，CU还可以被拆分为CU-CP和CU-UP。CU-CP和/或CU-UP中设置有一个或多个AI模型。
- [0186] AI模块用以实现相应的AI功能。不同网元中部署的AI模块可以相同或不同。AI模块的模型根据不同的参数配置，AI模块可以实现不同的功能。AI模块的模型可以是基于以下一项或多项参数配置的：结构参数(例如神经网络层数、神经网络宽度、层间的连接关系、神经元的权值、神经元的激活函数、或激活函数中的偏置中的至少一项)、输入参数(例如输入参数的类型和/或输入参数的维度)、或输出参数(例如输出参数的类型和/或输出参数的维度)。其中，激活函数中的偏置还可以称为神经网络的偏置。

- [0187] 一个AI模块可以具有一个或多个模型。一个模型可以推理得到一个输出，该输出包括一个参数或者多个参数。不同模型的学习过程、训练过程、或推理过程可以部署在不同的节点或设备中，或者可以部署在相同的节点或设备中。
- [0188] 需要说明的是，上述图2A所示的接入网节点可以作为一个整体，或者上述图2A所示的接入网节点采用CU和DU分离的架构。如图2B示出了接入网节点作为一个整体，接入网节点与终端设备连接的一个通信系统示意图。如图2C示出了接入网节点采用CU和DU分离的架构的一个通信系统示意图。
- [0189] 本申请中，可选的，AI任务节点、第一AI训练节点、第二AI训练节点和AI模型节点均可以是上述所述的AI网元。
- [0190] AI与网络的结合将是未来研究的一个重要方向。AI技术应用于通信设备中。例如，第一通信设备与第二通信设备是两个AI训练节点。第一通信设备与第二通信设备之间的信道是未知的，第一通信设备可以将数据映射到空口资源上进行传输，即将数据转换为空口上待发送的信号，从而实现对数据的反向传输。实现第一通信设备与第二通信设备之间的联合训练。第一通信设备与第二通信设备无需采用物理层进行信道编码、符号调制，再生成信号发送。由此可知，如何实现第一通信设备将数据映射到空口资源上进行传输，即如何实现第一通信设备将数据转换为空口上待发送的信号，是值得考虑的问题。具体可以参阅后文实施例的相关介绍。
- [0191] 本申请提供的技术方案适用于分布式学习的通信系统。分布式学习是实现联合训练的一种学习方法。分布式学习包括拆分学习(split learning)与联邦学习(federated learning)。本申请的技术方案主要采用拆分学习的方式实现联合训练。具体的，本申请中，AI任务节点可以将AI模型节点请求的第一模型拆分为第一子模型和第二子模型。然后，AI任务节点请求第一AI训练节点对第一子模型进行训练，以及请求第二AI训练节点对第二子模型进行训练。第一AI训练节点与第二AI训练节点之间可以通过本申请的技术方案实现数据和/或参数的传输。实现第一AI训练节点与第二AI训练节点之间的联合训练。
- [0192] 下面介绍本申请涉及的一些技术术语。

[0193] 1、神经网络：神经网络可以由神经元组成的，神经元可以是指以 x_s 和截距 b 为输入的运算单元，该运算单元的输出可以为：

$$h_{w,b}(x) = f(W^T x) = f\left(\sum_{s=1}^n W_s x_s + b\right) \quad (1)$$

[0194] 其中， $s=1, 2, \dots, n$ ， n 为大于1的自然数， W_s 为 x_s 的权重。需要说明的是，可选的， x_s 的权重也可以通过权重梯度加上该神经元上一次使用的权重计算得到。 b 为神经元的偏置。 f 为神经元的激活函数(activation functions)，用于将非线性特性引入神经网络中，来将神经元中的输入信号转换为输出信号。也就是在一个神经元中输入参数，该神经元可输出相应的参数。神经网络是将许多个上述单一的神经元联结在一起形成的网络，即一个神经元的输出可以是另一个神经元的输入。

[0195] 神经网络可以具有多层神经元，下面以深度神经网络(deep neural network, DNN)为例进行介绍。深度神经网络是具备很多层隐含层的神经网络。我们常说的多层神经网络和深度神经网络其本质上是同一个东西。从DNN按不同层的位置划分，DNN内部的神经网络可以分为三类：输入层，隐含层，输出层。一般来说第一层是输入层，最后一层是输出层，中间的层数都是隐含层。层与层之间是全连接的，也就是说，第 a 层的任意一个神经元一定与第 $a+1$ 层的任意一个神经元相连。在深度神经网络中，更多的隐含层让网络更能够刻画现实世界中的复杂情形。理论上而言，模型参数越多的模型复杂度越高，“容量”也就越大，也就意味着它能完成更复杂的学习任务。下面结合图2D介绍神经网络包括的多层神经网络层。如图2D所示的神经网络，该神经网络包括六个神经网络层，具体为神经网络层(neural network, NN1)至NN6。NN1包括两个输出维度，NN2包括三个输出维度，NN3包括四个输出维度，NN4包括四个输出维度，NN5包括三个输出维度，NN6包括两个输出维度。需要说明的是，上述图2D仅仅是为了方便示例，仅示出了每个神经网络层包括的输出维度。实际上，每个神经网络层包括该神经网络层的输出维度以及该神经网络层与上一个神经网络层之间的连接结构。例如，NN2包括NN2的三个输出维度以及NN2的三个输出维度与

NN1的两个输出维度之间的连接结构。对于其他神经网络层同样类似，这里不一一说明。

[0196] 2、中间梯度：是神经网络参数的梯度表达式中的一项或多项，或者是多项的乘积。当中间梯度是神经网络参数的梯度表达式中的多项时，该多项是分别回传给前一AI训练节点的。

[0197] 例如，如图2E所示的通信系统包括神经网络#1、神经网络#2以及神经网络#3，其各神经网络对应的参数分别为 θ_1 、 θ_2 、 θ_3 。 A_0 、 A_1 、 A_2 分别为神经网络#1、神经网络#2、神经网络#3的输入， σ 为一个或多个对数据进行处理函数。其中， $Z_i = \theta_i g A_{i-1}$ ， $A_{i-1} = \sigma g(Z_{i-1})$ ， i 可取1、2、3。

[0198] 神经网络#3参数、神经网络#2参数以及神经网络#1参数的梯度分别为：

$$\frac{\partial l}{\partial \theta_3} = \frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial \theta_3} \quad (2)$$

$$\frac{\partial l}{\partial \theta_2} = \frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial A_2} \frac{\partial A_2}{\partial Z_2} \frac{\partial Z_2}{\partial \theta_2} \quad (3)$$

$$\frac{\partial l}{\partial \theta_1} = \frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial A_2} \frac{\partial A_2}{\partial Z_2} \frac{\partial Z_2}{\partial A_1} \frac{\partial A_1}{\partial Z_1} \frac{\partial Z_1}{\partial \theta_1} \quad (4)$$

[0199] 其中， l 为损失函数。从而，神经网络#3参数的中间梯度为公式(2)中的一项或多项，或者多项的乘积。例如，神经网络#3参数的中间梯度为： $\frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3}$ ，再

例如，神经网络#3参数的中间梯度为 $\frac{\partial l}{\partial A_3}$ 和 $\frac{\partial A_3}{\partial Z_3}$ 。神经网络#2参数的中间梯度为

公式(3)中的一项或多项，或者多项的乘积。例如，神经网络#2参数的中间梯度为 $\frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial A_2}$ ，再例如，神经网络#2参数的中间梯度为 $\frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial A_2} \frac{\partial A_2}{\partial Z_2}$ ，又例如，

神经网络#2参数的中间梯度为 $\frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3}$ 。神经网络 #1参数的中间梯度为公式(4

)中的一项或多项,或者多项的乘积。例如,神经网络#1参数的中间梯度为

$\frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial A_2} \frac{\partial A_2}{\partial Z_2} \frac{\partial Z_2}{\partial A_1}$,再例如,神经网络#1参数的中间梯度为

$$\frac{\partial l}{\partial A_3} \frac{\partial A_3}{\partial Z_3} \frac{\partial Z_3}{\partial A_2} \frac{\partial A_2}{\partial Z_2} \frac{\partial Z_2}{\partial A_1} \frac{\partial A_1}{\partial Z_1}。$$

- [0200] 3、损失函数:用于度量模型的预测值 $f(x)$ 与真实值 Y 的差异程度的运算函数。损失函数是一个非负实值函数,通常使用 $L(Y, f(x))$ 来表示,损失函数越小,模型的鲁棒性越好。
- [0201] 4、超网络:表示用于产生较大规模网络参数的小规模网络。在这过程中,主网络的作用与任一神经网络一样,将输入样本映射到对应的目标值,而超网络的作用则是接收一系列包含主网络参数结构信息的值作为输入然后产生主网络的某一层的参数。
- [0202] 5、超参数:用于估计模型的参数。例如,深度神经网络的学习率。
- [0203] 本申请中,第一附加模块也可以称为第一附加层神经网络、第一附加神经网络、第一附加算法、第一附加变换、第一附加匹配、第一附加适配、或第一附加均衡,具体本申请对第一附加模块的名称不做限定。对于第二附加模块同样类似,例如,第二附加模块也可以称为第二附加层神经网络、第二附加神经网络、第二附加算法、第二附加变换、第二附加匹配、第二附加适配、或第二附加均衡,具体本申请对第二附加模块的名称不做限定。
- [0204] 下面结合具体实施例介绍本申请的技术方案。
- [0205] 图3为本申请实施例模型训练配置方法的一个实施例示意图。请参阅图3,方法包括:
- [0206] 301、AI任务节点向第一AI训练节点发送第一请求。第一请求用于请求第一AI训练节点对第一子模型进行训练。相应的,第一AI训练节点接收来自AI任务节点的第一请求。
- [0207] 其中,第一子模型是第一模型经过拆分得到的多个子模型中的其中一个子模型。第一模型是AI模型节点请求训练的模型。例如,如图4所示,第一模型是AI模型节点请求训练的模型。AI任务节点将第一模型拆分为第一子模型和第二子模型。然后,AI任务节点向第一AI训练节点发送第一请求,以请求第一AI训练

节点对第一子模型进行训练。如图4所示，第一模型为神经网络，第一子模型是将神经网络进行拆分得到的第一部分神经网络，第二子模型是将神经网络进行拆分得到的第二部分神经网络。需要说明的是，图4仅仅是一种示例。实际应用中，AI任务节点也可以结合实际情况将第一模型拆分为更多个子模型，具体本申请不做限定。

[0208] 可选的，第一请求包括以下至少一项：第一子模型的结构、第一超参数、第一损失函数、第一训练停止条件、第一子模型在多个子模型中的编号、或数据集。

[0209] 关于第一超参数、第一损失函数的相关功能请参阅前述技术术语中的介绍。

[0210] 例如，第一子模型为神经网络，第一子模型的结构包括该神经网络的层数、每一层采用的神经网络的类型(例如，卷积神经网络(convolutional neural networks, CNN)、或循环神经网络(recurrent neural networks, RNN))、每层神经网络中的神经元数量等。

[0211] 第一训练停止条件是第一AI训练节点停止训练第一子模型的条件。例如，训练次数大于预设阈值；或者，第一子模型的训练时间超过预设阈值；或者，第一模型对应的测试性能大于预设阈值。第一模型对应的测试性能是AI模型节点通过第一模型执行AI模型节点请求的任务得到任务执行性能。

[0212] AI任务节点可以对该多个子模型进行排序编号。例如，该多个子模型中，包含的神经网络层的层序号越大的子模型的排序编号越大，包含的神经网络层的层序号越小的子模型的排序编号越小。或者，该多个子模型中，包含的神经网络层的层序号越小的子模型的排序编号越小，包含的神经网络层的层序号越大的子模型的排序编号越大。或者，该多个子模型中，执行时间越早的子模型的排序编号越小，执行时间越晚的子模型的排序编号越大。或者，该多个子模型中，执行时间越晚的子模型的排序编号越小，执行时间越早的子模型的排序编号越大。子模型的执行时间是AI任务节点规定的该子模型所在的设备接收相应的数据和/或参数，并对该子模型进行训练的时间。因此，第一子模型在多个子模型中的编号指按照子模型包含的神经网络层的层序号从小到大的顺序或从大到小的顺序对多个子模型进行排序编号得到的该第一子模型的编号。或者，第

一子模型在多个子模型中的编号指按照子模型的执行时间从早到晚的顺序或从晚到早的顺序对多个子模型进行排序编号得到的该第一子模型的编号。例如，如图4所示，AI任务节点将第一模型拆分为第一子模型和第二子模型。第一子模型包括神经网络层1至神经网络层2，第二子模型包括神经网络层3至神经网络层4。因此可知，AI任务节点按照子模型包含的神经网络层的层序号从小到大的顺序对第一子模型和第二子模型进行排序编号，得到第一子模型的编号为1以及第二子模型的编号为2。

[0213] 数据集用于作为第一子模型的输入数据。具体的，第一AI训练节点可以将数据集输入第一子模型，得到第一子模型的输出数据。

[0214] 可选的，图3所示的实施例还包括步骤301a。步骤301a可以在步骤301之前执行。

[0215] 301a、AI模型节点向AI任务节点发送第四请求。第四请求用于请求对第一模型进行训练。相应的，AI任务节点接收来自AI模型节点的第四请求。

[0216] 可选的，第四请求包括以下至少一项：任务类型、第一模型的结构、训练复杂度、或时间限制需求。

[0217] 任务类型是指AI模型节点请求的任务类型，该任务类型对应第一模型。例如，图片分类任务，第一模型用于支持AI模型节点执行图片分类任务。

[0218] 例如，第一模型为神经网络。第一模型的结构包括该神经网络的层数、每一层采用的神经网络的类型(例如，CNN、或RNN)、每层神经网络中的神经元数量等。

[0219] 训练复杂度包括计算复杂度和/或存储复杂度。计算复杂度表示训练第一模型的复杂度。存储复杂度表示存储训练得到的第一模型的参数的复杂度。第一模型参数越多，存储复杂度越高。

[0220] 时间限制需求表示第一模型的训练时间。即表示在该时间限制需求指示的时间内完成该第一模型的训练。

[0221] 302、AI任务节点向第一AI训练节点发送第一配置信息。第一配置信息用于配置至少一个第一附加模块。相应的，第一AI训练节点接收来自第一AI训练节点的第一配置信息。

- [0222] 至少一个第一附加模块用于将第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号，和/或，至少一个第一附加模块用于将第一AI训练节点在空口上接收到的符号转换为用于更新第一子模型的更新参数。其中，该第一子模型输出的符号可以是第一AI训练节点将数据集输入第一子模型得到的第一子模型输出的符号；或者，该第一子模型输出的符号可以是第一子模型输出的数据再经过转换得到的符号。第一子模型输出的数据是第一AI训练节点将数据集输入第一子模型得到的第一子模型输出的数据。
- [0223] 可选的，更新参数包括用于更新第一子模型的中间梯度或权重参数。本文主要以中间梯度为例进行介绍。
- [0224] 可选的，该至少一个第一附加模块中各个第一附加模块为神经网络，或信号处理模块。信号处理模块可以是软件模块，或硬件模块，或软硬结合模块。该信号处理模块支持相应的信号处理算法。下面主要以各个第一附加模块为神经网络为例进行介绍。
- [0225] 例如，如图4所示，该至少一个第一附加模块包括一个第一附加模块。第一AI训练节点将数据集输入第一子模型，得到第一子模型输出的符号。或者，第一AI训练节点将数据集输入第一子模型，得到第一子模型输出的数据，再将该第一子模型输出的数据转换为第一子模型输出的符号。也就是第一子模型可以直接输出符号，或者，第一子模型也可以输出数据，第一AI训练节点再将数据转换为符号。然后，第一AI训练节点通过第一附加模块将第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号。第一AI训练节点通过该第一附加模块将在空口上接收到的符号转换为用于更新第一子模型的中间梯度。然后，第一AI训练节点通过该中间梯度更新第一子模型。需要说明的是，图4所示的第一附加模块包括四个输出维度，该四个输出维度与第一子模型之间的连接结构可以理解为第一附加模块的结构。图4所示的示例仅仅是为了方便示图，实际上，第一附加模块包括第一附加模块的四个输出维度与该第一子模型之间的连接结构以及该四个输出维度。需要说明的是，图4所示的第一附加模块包括一个神经网络层。上述图4所示的第一附加模块仅仅是一种示例。实际应用中，本

申请中，每个第一附加模块包括一个或多个神经网络层，具体本申请不做限定。

- [0226] 需要说明的是，第一AI训练节点通过至少一个附加模块将第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号的过程与申请号为202110412895.X的专利文件中第一通信装置将第一中间梯度映射到空口资源上，生成第一梯度信号的过程类似，具体可以参阅该专利文件中的相关介绍。第一AI训练节点将在空口上接收到的符号通过该至少一个第一附加模块转换为用于更新第一子模型的更新参数的过程与申请号为202110412895.X的专利文件中第二通信装置根据第二梯度信号，确定第二中间梯度的过程类似，具体可以参阅该专利文件中的相关介绍。
- [0227] 可选的，第一配置信息包括以下至少一项：该至少一个第一附加模块的索引、类型、结构参数、第一指示信息、第二指示信息、第三指示信息、第四指示信息、或第五指示信息。
- [0228] 具体的，该第一配置信息包括该至少一个第一附加模块中各个第一附加模块的索引、类型、结构参数。例如，各个第一附加模块为神经网络，各个第一附加模块的类型包括该各个神经网络中包括的层数、每层神经网络包括的神经元数量、以及每层神经网络的类型。各个第一附加模块的结构参数包括各个神经网络中每层神经网络中神经元的参数。
- [0229] 需要说明的是，上述示出了第一配置信息包括该至少一个第一附加模块的具体配置的实现方式。可选的，第一配置信息也可以包括可执行文件。该可执行文件用于第一AI训练节点实现该至少一个第一附加模块的功能。也就是第一配置信息中不显式指示该至少一个第一附加模块的具体配置，而是配置可执行文件。第一AI训练节点执行该可执行文件可以实现该至少一个第一附加模块的功能。从而避免公开该至少一个第一附加模块的具体配置。
- [0230] 第一指示信息用于指示至少一个第一附加模块是否参与第一子模型的训练。也就是说第一AI训练节点训练第一子模型时，是否将该至少一个第一附加模块也作为模型训练的一部分，对该至少一个第一附加模块进行训练。第一指示信息的长度可以1比特、2比特或3比特等，具体本申请不做限定。例如，第一指示信

息为1比特。该第一指示信息的取值为0时，指示该至少一个第一附加模块不参与第一子模型的训练；该第一指示信息的取值为1时，指示该至少一个第一附加模块参与第一子模型的训练。或者，该第一指示信息的取值为0时，指示该至少一个第一附加模块参与第一子模型的训练；该第一指示信息的取值为1时，指示该至少一个第一附加模块不参与第一子模型的训练。例如，如图4所示，如果第一指示信息指示至少一个第一附加模块参与第一子模型的训练，第一AI训练节点可以将第一子模型和第一附加模块作为一个整体进行模型训练。

[0231] 第二指示信息用于指示至少一个第一附加模块是否作为第一子模型的一部分上报给AI任务节点。也就是说第一AI训练节点上报经过训练得到的第一子模型时，是否将该至少一个第一附加模块作为该经过训练得到的第一子模型的一部分上报给AI任务节点。第二指示信息的长度可以1比特、2比特或3比特等，具体本申请不做限定。例如，第二指示信息为1比特，该第二指示信息的取值为0时，指示该至少一个第一附加模块不作为第一子模型的一部分上报给AI任务节点；该第二指示信息的取值为1时，指示该至少一个第一附加模块作为第一子模型的一部分上报给AI任务节点。或者，该第二指示信息的取值为0时，指示该至少一个第一附加模块作为第一子模型的一部分上报给AI任务节点；该第二指示信息的取值为1时，指示该至少一个第一附加模块不作为第一子模型的一部分上报给AI任务节点。例如，如图4所示，如果第二指示信息指示至少一个第一附加模块作为第一子模型的一部分上报给AI任务节点，第一AI训练节点可以将经过训练得到的第一子模型和第一附加模块作为一个整体上报给AI任务节点。

[0232] 第三指示信息用于指示至少一个第一附加模块中各个第一附加模块的部署方式，部署方式包括串联部署或并联部署。第三指示信息的长度可以1比特、2比特或3比特等，具体本申请不做限定。例如，第三指示信息为1比特。第三指示信息的取值为0时，指示该至少一个第一附加模块中各个第一附加模块的部署方式为串联部署；第三指示信息的取值为1时，指示该至少一个第一附加模块中各个第一附加模块的部署方式为并联部署。或者，第三指示信息的取值为0时，指示该至少一个第一附加模块中各个第一附加模块的部署方式为并联部署；第三

指示信息的取值为1时，指示该至少一个第一附加模块中各个第一附加模块的部署方式为串联部署。

[0233] 例如，如图5A所示，该至少一个第一附加模块包括第一附加模块1和第一附加模块2。第三指示信息用于指示第一附加模块1的部署方式为串联部署，第一附加模块2的部署方式为串联部署。第一附加模块1串联在第一子模型的NN1与第一子模型的NN2之间，第一附加模块2串联在第一子模型的NN2上。具体如图5A所示。在该实现方式中，第一附加模块1的输入维度与NN1的输出维度相等，第一附加模块1的输出维度与NN2的输入维度相等。第一附加模块2的输入维度与NN2的输出维度相等。

[0234] 再例如，如图5B所示，该至少一个第一附加模块包括第一附加模块1和第一附加模块2。第三指示信息用于指示第一附加模块1的部署方式为串联部署，第一附加模块2的部署方式为串联部署。第一附加模块1串联在第一子模型的NN1上，第一附加模块2串联在第一子模型的NN1与第一子模型的NN2之间。具体图5B所示。在该实现方式中，第一附加模块1的输出维度与NN1的输入维度相等。第一附加模块2的输入维度与NN1的输出维度相等，第一附加模块2的输出维度与NN2的输入维度相等。

[0235] 再例如，如图5C所示，该至少一个第一附加模块包括第一附加模块1和第一附加模块2。第三指示信息用于指示第一附加模块1和第一附加模块2分别对应的部署方式都为并联部署。因此可知，第一附加模块1并联在第一子模型的NN1上，第一附加模块2并联在第一子模型的NN2上。具体如图5C所示。在该实现方式中，第一附加模块1的输入维度与NN1的输入维度相等，第一附加模块1的输出维度与NN1的输出维度相等。第一附加模块2的输入维度与NN2的输入维度相等，第一附加模块2的输出维度与NN2的输出维度相等。

[0236] 第四指示信息用于指示是否向第一AI训练节点指示至少一个第一附加模块的类型、结构参数。第四指示信息的长度可以1比特、2比特或3比特等，具体本申请不做限定。例如，第四指示信息为1比特。第四指示信息的取值为0时，指示不向第一AI训练节点指示至少一个第一附加模块的类型、结构参数；第四指示信息的取值为1时，指示向第一AI训练节点指示至少一个第一附加模块的类型、结

构参数。或者，第四指示信息的取值为0时，指示向第一AI训练节点指示至少一个第一附加模块的类型、结构参数；第四指示信息的取值为1时，指示不向第一AI训练节点指示至少一个第一附加模块的类型、结构参数。

[0237] 一种可能的实现方式中，该第一配置信息还包括超网络的索引。该超网络用于确定该至少一个第一附加模块。

[0238] 在该实现方式中，第四指示信息指示不向第一AI训练节点指示该至少一个第一附加模块的类型、结构参数。例如，如图6所示，AI任务节点向第一AI训练节点发送超网络的索引。第一AI训练节点可以第一附加模块i的序号i作为超网络的输入参数输入超网络，从而得到第一附加模块i。i为大于或等于1且小于或等于该至少一个第一附加模块包括的第一附加模块的数量。

[0239] 需要说明的是，可选的，若该至少一个第一附加模块中，第一附加模块的类型不相同，第一AI训练节点还应当进一步将第一附加模块i的类型作为超网络的输入参数输入超网络，从而得到第一附加模块i。也就是超网络的输入参数包括第一附加模块i的序号i和类型。

[0240] 另一种可能的实现方式中，该第一配置信息包括该至少一个第一附加模块的索引，和/或，该至少一个第一附加模块的类型、结构参数。

[0241] 在该实现方式中，第四指示信息指示向第一AI训练节点指示该至少一个第一附加模块的类型、结构参数，AI任务节点可以指示该至少一个第一附加模块的索引，和/或，AI任务节点采用约定格式向第一AI训练节点下发该至少一个第一附加模块的类型、结构参数。例如，如图7所示，AI任务节点向第一AI训练节点发送该至少一个第一附加模块的类型、结构参数。

[0242] 第五指示信息用于指示至少一个第一附加模块中各个第一附加模块插入第一子模型的位置。例如，第一子模型为神经网络，第五指示信息用于指示该至少一个第一附加模块中各个第一附加模块分别插入神经网络中的哪个神经网络层上。

[0243] 可选的，第五指示信包括第一比特位图，第一比特位图中的比特与第一子模型中的神经网络层对应，第一比特位图中的比特用于指示在该比特对应的神经网络层是否插入第一附加模块。

- [0244] 第一比特位图中一个比特对应第一子模型中的一层神经网络或多层神经网络。或者，第一比特位图中的多个比特对应第一子模型中的一层神经网络。需要说明的是，可选的，当第一比特位图中的一个比特对应第一子模型中的多层神经网络时，第一配置信息还包括第十一指示信息，该第十一指示信息用于指示第一比特位图中的各个比特分别对应的神经网络层。下面以第一比特位图中的一个比特对应第一子模型中的一层神经网络为例进行介绍。
- [0245] 可选的，第一比特位图中的一个比特与第一子模型中的一层神经网络对应。当第一比特位图中的一个比特的取值为1，指示在比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块；当该第一比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块。或者，当该第一比特位图中的一个比特的取值为0，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块；当该第一比特位图中的一个比特的取值为1，指示在该比特对应的一层神经网络与该比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块。
- [0246] 可选的，第一比特位图中的1比特与该至少一个第一附加模块中的第一附加模块一一对应。例如，该至少一个第一附加模块按照第一附加模块的配置顺序与第一比特位图中的1比特一一对应。例如，第一子模型包括四层神经网络，因此第一比特位图包括四个比特，第一比特位图为1001。第一比特位图中，第一个比特对应四层神经网络中的第一层神经网络，第二个比特对应四层神经网络中的第二层神经网络，第三个比特对应四层神经网络中的第三层神经网络，第四个比特对应四层神经网络中的第四层神经网络。该至少一个第一附加模块包括第一附加模块1和第二附加模块2。第一比特位图中的第一个比特的取值为1，指示在第一层神经网络与第二层神经网络之间插入第一附加模块1。第一比特位图的第三个比特的取值为1，指示在第三层神经网络与第四层神经网络之间插入第一附加模块2。第一比特位图中的其他比特的取值都为0。
- [0247] 303、第一AI训练节点通过至少一个第一附加模块将第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号，并向第二AI训练节点发送待输出

的符号；和/或，通过该至少一个第一附加模块将第一AI训练节点在空口上接收来自第二AI训练节点的符号转换为用于更新第一子模型的更新参数，通过更新参数更新第一子模型。

[0248] 具体的，第一AI训练节点接收到第一请求之后，第一AI训练节点确定该第一子模型。第一AI训练节点接收到第一配置信息之后，第一AI训练节点确定该至少一个第一附加模块。第一AI训练节点可以执行上述步骤303。例如，如图4所示，第一AI训练节点将第一子模型输出的符号通过第一附加模块转换为第一AI训练节点在空口上待输出的符号，并向第二AI训练节点发送待输出的符号。或者，第一AI训练节点在空口上接收来自第二AI训练节点的符号，再通过该至少一个第一附加模块将第一AI训练节点在空口上接收的来自第二AI训练节点的符号转换为用于更新第一子模型的中间梯度。第一AI训练节点通过中间梯度更新第一子模型。

[0249] 由此可知，第一AI训练节点不是将第一子模型输出的符号在应用层做信源编码，再以比特流发回物理层，由物理层进行信道编码、符号调制，再生成信号发送。而是通过至少一个第一附加模块将第一子模型输出的符号转换为第一AI训练节点在空口上待输出的符号，从而在无线空口上发送第一AI训练节点在空口上待输出的符号。实现向第二AI训练节点发送第一子模型输出的符号。第一AI训练节点不是将第一AI训练节点在空口上接收到的符号通过物理层进行信道解码、符号解调，再恢复为比特流，并通过应用层进行信源解码。而是第一AI训练节点通过至少一个第一附加模块将第一AI训练节点在空口上接收的符号转换为用于更新第一子模型的中间梯度。从而实现接收第二AI训练节点在空口上的符号。实现第一AI训练节点与第二AI训练节点的联合训练。

[0250] 本申请中，第一AI训练节点通过上述步骤303的方式与第二AI训练节点进行数据和/或参数的传输。在保持模型训练精度不变的情况，相比于采用应用层和物理层进行数据和/或参数的传输的方案来说，本申请的技术方案能够显著降低传输开销。也就是第一AI训练节点与第二AI训练节点之间的传输开销得到显著降低。本申请中，相比于采用应用层和物理层进行数据和/或参数的传输的方案来说，第一AI训练节点与第二AI训练节点之间的传输可以理解为是一种有损传

输，即数据和/或参数的传输是有误差的。但是由实验证明，相比于采用应用层和物理层进行数据和/或参数的传输的方案来说，本申请的方案中第一AI训练节点与第二AI训练节点之间的传输性能并没有大幅度下降。下面示出了本申请中研发人员经过实验得到的一些实验数据。

[0251] 如图9所示，在第一模型的训练精度不变的情况下，纵坐标表示第一差值与采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销的比值。第一差值为采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销与本申请的技术方案产生的传输开销的差。因此该比值越大，本申请的技术方案产生的传输开销越小。如图9可知，在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为16比特的情况下，示出了信道的信噪比为10dB(分贝)、0dB、和-10dB，第一差值与采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销的比值分别为80%、90%、100%。在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为8比特的情况下，示出了信道的信噪比为10dB、0dB、和-10dB，第一差值与采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销的比值分别为60%、88%、98%。在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为4比特的情况下，示出了信道的信噪比为10dB、0dB、和-10dB，第一差值与采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销的比值分别为40%、80%、95%。可知，在相同的信噪比的情况下，本申请的技术方案产生的传输开销远小于采用应用层和物理层进行数据和/或参数的传输的方案产生的传输开销。

[0252] 例如，AI模型节点向AI任务节点请求的是图片分类任务。如图10中，在第一模型的训练精度不变的情况下，纵坐标表示第二差值与采用应用层和物理层进行数据和/或参数的传输的方案对图片进行分类的分类准确度之间的比值。第二差值为本申请的技术方案对图片进行分类的分类准确度与采用应用层和物理层进行数据和/或参数的传输的方案对图片进行分类的分类准确度之差。由此可知，比值越大，表示本申请的技术方案对图片进行分类的分类准确度越高。

- [0253] 如图10可知,在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为16比特的情况下,示出了信道的信噪比为10dB(分贝)、0dB、和-10dB,纵坐标的取值分别为-0.28%、-0.38%、-0.68%。在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为8比特的情况下,示出了信道的信噪比为10dB、0dB、和-10dB,纵坐标的取值分别为-0.2%、-0.3%、-0.6%。在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为4比特的情况下,示出了信道的信噪比为10dB、0dB、和-10dB,纵坐标的取值分别为0.23%、0.08%、-0.2%。可知,在相同的信噪比的情况下,本申请的技术方案对图片进行分类的分类准确度并没有大幅降低。反而在采用应用层和物理层进行数据和/或参数的传输的方案中应用层将数据量化为4比特,信噪比为10dB或0dB的情况下,本申请的技术方案对图片进行分类的分类准确度高于采用应用层和物理层进行数据和/或参数的传输的方案对图片进行分类的分类准确度。
- [0254] 一种可能的实现方式中,至少一个第一附加模块包括多个第一附加模块。第五指示信息用于指示在该第一子模型的第X层神经网络与第X+1层神经网络之间插入该多个第一附加模块。X为大于或等于1且小于第一子模型包括的神经网络层数。其中,多个第一附加模块的输入维度之和等于第X层神经网络的输出维度,多个第一附加模块的输出维度之和等于第X+1层神经网络的输入维度;或者,多个第一附加模块的输入维度之和大于第X层神经网络的输出维度,第X层神经网络在输出数据上补零,使得第X层神经网络的输出维度与多个第一附加模块的输入维度之和相等。多个第一附加模块中的一个或多个第一附加模块在输出数据上降维,使得多个第一附加模块的输出维度之和等于第X+1层神经网络的输入维度。在该实现方式中,可选的,第一配置信息包括该多个第一附加模块的索引。
- [0255] 另一种可能的实现方式中,第五指示信息用于指示在第一子模型的最外层神经网络插入多个第一附加模块。其中,多个第一附加模块的输入维度之和等于最外层神经网络的输出维度。或者,多个第一附加模块的输入维度之和大于最外层神经网络的输出维度,最外层神经网络在输出数据上补零,使得最外层神经

网络的输出维度与多个第一附加模块的输入维度之和相等。在该实现方式中，可选的，第一配置信息包括该多个第一附加模块的索引。

[0256] 本申请中，该至少一个第一附加模块取代传统物理层，因此可根据信道特性生成该至少一个第一附加模块。通常来说，第一附加模块的维度与无线的时频资源相关。例如，如图8所示，一个第一附加模块的维度映射到一个资源块(resource block, RB)或一个时隙的无线时频资源。也就是说一个第一附加模块的维度映射到单位无线时频资源上。而第一子模型中各层神经网络的维度通常与神经网络的设计相关。因此，第一附加模块的维度与神经网络的维度不一定匹配。因此可知，AI任务节点可以指示在同一神经网络层中插入多个第一附加模块。如图8所示，AI任务节点指示在第一子模型的最外层神经网络插入第一附加模块1和第一附加模块2。第一附加模块1的输入维度与第一附加模块2的输入维度大于最外层神经网络的输出维度。因此最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与多个第一附加模块的输入维度之和相等。从而使得多个第一附加模块的维度与最外层神经网络的维度适配。上述技术方案避免第一AI训练节点存储大量不同维度的第一附加模块。避免导致第一AI训练节点存储第一附加模块的存储压力问题。需要说明的是，图8所示的第一附加模块1包括五个输出维度，该五个输出维度与第一子模型之间的连接结构可以理解为第一附加模块1的结构。图8所示的第一附加模块2包括四个输出维度，该四个输出维度与第一子模型之间的连接结构可以理解为第一附加模块2的结构。图8所示的示例仅仅是为了方便示图。实际上，第一附加模块1包括第一附加模块1与第一子模型之间的连接结构和该五个输出维度。第一附加模块2包括第一附加模块2与第一子模型之间的连接结构以及该四个输出维度。

[0257] 再一种可能的实现方式中，至少一个第一附加模块包括第一附加模块1。第五指示信息用于指示在第一子模型的第X层神经网络与第X+1层神经网络之间插入第一附加模块1。X为大于或等于1且小于所述第一子模型包括的神经网络层数。第一附加模块1的输入维度小于第X层神经网络的输出维度。在该实现方式中，第一AI训练节点还可以执行以下操作：第一AI训练节点确定第一附加模块2，第一AI训练节点在第X层神经网络与第X+1层神经网络之间插入第一附加模块1和第

一附加模块2。其中，第一附加模块1的输入维度与第一附加模块2的输入维度之和等于第X层神经网络的输出维度，第一附加模块1的输出维度与第一附加模块2的输出维度之和等于第X+1层神经网络的输入维度。或者，第一附加模块1的输入维度与第一附加模块2的输入维度之和大于第X层神经网络的输出维度，第X层神经网络在输出数据上补零，使得第X层神经网络的输出维度与第一附加模块1的输入维度与第一附加模块2的输入维度之和相等，第一附加模块1和/或第一附加模块2在输出数据上降维，使得第一附加模块1的输出维度与第一附加模块2的输出维度之和等于第X+1层神经网络的输入维度。

[0258] 在该实现方式中，AI任务节点可以只向第一AI训练节点指示第一附加模块1的索引，以及通过第五指示信息指示在第一子模型的第X层神经网络与第X+1层神经网络之间插入第一附加模块1。第一AI训练节点确定第一附加模块1之后，确定第一附加模块1的输入维度小于第X层神经网络的输入维度。因此，第一AI训练节点可以从预设的第一附加模块中选择第二附加模块2，并将第一附加模块1和第一附加模块2插入第一子模型的第X层神经网络与第X+1层神经网络之间。从而实现第一附加模块1的输入维度和第二附加模块2的输入维度之和与第一子模型的第X层神经网络的输出维度适配。第一附加模块1的输出维度和第一附加模块2的输出维度与第X+1层神经网络的输入维度适配。

[0259] 再一种可能的实现方式中，至少一个第一附加模块包括第一附加模块1。第五指示信息用于指示在第一子模型的最外层神经网络插入第一附加模块1。在该实现方式中，第一AI训练节点还可以执行以下操作：第一AI训练节点确定第一附加模块2，第一AI训练节点在最外层神经网络插入第一附加模块1和第一附加模块2。其中，第一附加模块1的输入维度与第一附加模块2的输入维度之和等于最外层神经网络的输出维度。或者，第一附加模块1的输入维度与第一附加模块2的输入维度之和大于最外层神经网络的输出维度，最外层神经网络在输出数据上补零，使得最外层神经网络的输出维度与第一附加模块1的输入维度与第一附加模块2的输入维度之和相等。

[0260] 在该实现方式中，AI任务节点可以只向第一AI训练节点指示第一附加模块1的索引，以及通过第五指示信息指示在第一子模型的最外层神经网络插入第一附

加模块1。第一AI训练节点确定第一附加模块1之后，确定第一附加模块1的输入维度小于最外层神经网络的输出维度。因此，第一AI训练节点可以从预设的第一附加模块中选择第二附加模块2，并将第一附加模块1和第一附加模块2插入第一子模型的第X层神经网络与第X+1层神经网络之间。从而实现第一附加模块1的输入维度和第二附加模块2的输入维度之和与最外层神经网络的输出维度适配。

[0261] 需要说明的是，上述几种实现方式是以在一层神经网络插入两个第一附加模块的示例进行介绍。实际应用中，在一层神经网络可以插入更多第一附加模块，从而使得第一附加模块的维度与所插入的神经网络层的维度适配。

[0262] 可选的，上述步骤303具体包括：第一AI训练节点通过第一冗余处理模块对待输出的符号进行冗余处理，再向第二AI训练节点发送经过冗余处理的待输出的符号；和/或，第一AI训练节点通过第一冗余处理模块对第一AI训练节点在空口上接收到来自第二AI训练节点的符号进行去冗余处理，再通过至少一个第一附加模块将经过去冗余处理的符号转换为更新参数。

[0263] 可选的，第一冗余处理模块为神经网络或信号处理模块，具体本申请不做限定。关于信号处理模块可以参阅前述的相关介绍。

[0264] 可选的，第一冗余处理模块的维度等于冗余倍数与至少一个第一附加模块的维度的乘积。例如， β 为冗余倍数。即第一AI训练节点进行冗余发送对应的冗余倍数。同理，对于第二AI训练节点， β 是第二AI训练节点进行冗余发送对应的冗余倍数。

[0265] 例如，在AI模型节点请求部分场景回归任务的场景下，第一AI训练节点与第二AI训练节点之间的有损传输会造成模型训练性能受限。第一AI训练节点可以采用冗余传输方式提升该场景下的模型训练性能。如图11所示， β 等于2。第一AI训练节点对通过第一附加模块将第一子模型输出的符号转换为第一AI训练节点在空口上待输出的符号。第一AI训练节点通过第一冗余处理模块对第一AI训练节点在空口上待输出的符号进行冗余处理，并向第二AI训练节点发送经过冗余处理的待输出的符号。相应的，第二AI训练节点也执行相应的冗余传输，因此第一AI训练节点通过第一冗余处理模块对第一AI训练节点在空口上接收到来自第二AI训练节点的符号进行去冗余处理，再通过第一附加模块将经过去冗余

处理的符号转换为更新参数。然后，第一AI训练节点通过该更新参数更新第一子模型。需要说明的是，图11所示的第一附加模块包括四个输出维度，该四个输出维度与第一子模型之间的连接结构可以理解为第一附加模块的结构。图11所示的示例仅仅是为了方便示图，实际上，第一附加模块包括第一附加模块的四个输出维度与该第一子模型之间的连接结构以及该四个输出维度。图11所示的第一冗余处理模块包括八个输出维度，该八个输出维度与第一子模型之间的连接结构可以理解为第一冗余处理模块的结构。图11所示的示例仅仅是为了方便示图，实际上，第一冗余处理模块包括第一冗余处理模块的八个输出维度与该第一子模型之间的连接结构以及该八个输出维度。

[0266] 图11所示的第二附加模块包括四个输出维度，该四个输出维度与第二子模型之间的连接结构可以理解为第二附加模块的结构。图11所示的示例仅仅是为了方便示图，实际上，第二附加模块包括第二附加模块的四个输出维度与该第二子模型之间的连接结构以及该四个输出维度。图11所示的第二冗余处理模块包括八个输出维度，该八个输出维度与第二子模型之间的连接结构可以理解为第二冗余处理模块的结构。图11所示的示例仅仅是为了方便示图，实际上，第二冗余处理模块包括第二冗余处理模块的八个输出维度与该第二子模型之间的连接结构以及该八个输出维度。

[0267] 一种可能的实现方式中，冗余倍数是根据当前信道的信噪比和目标信噪比确定的。例如，冗余倍数 β 可以表示为如下公式5：

$$\beta = \begin{cases} 10^{(s_t - snr)/10} & , snr < s_t \\ 1 & , \text{其他} \end{cases} \quad (5)$$

[0268] 其中，snr表示当前信道的信噪比， s_t 表示目标信噪比，即AI模型节点请求训练的第一模型对应的目标性能对应的信噪比。第一模型对应的目标性能是指AI模型节点通过第一模型执行相应的任务对应所需的任务执行性能。例如，对于图片分类任务，第一模型对应的目标性能可以是AI模型节点通过经过训练的第一模型对图片进行分类时所需达到的分类准确度。

- [0269] 关于第一AI训练节点获取冗余倍数的方式有多种，具体可以参阅后文图13所示的实施例中的相关介绍，这里不再赘述。
- [0270] 本申请中，第一AI训练节点与第二AI训练节点通过冗余传输方式可以实现信号功率的增益，同时还带来了数据增强的增益。如图12所示，曲线1是在没有采用冗余传输方式的情况下，本申请的技术方案在不同信噪比下对图片的分类准确度的大小。由图12可知，随着信噪比的降低，分类准确度下降。图12中，x标记表示采用冗余传输方式执行本申请的技术方案，可知当前信道的信噪比 (signal-to-noise ratio, SNR) = 12dB, $\beta = 2$ ，能够得到的当前信道的等效信噪比为15dB。因此对应的分类准确度也相应较大。因此，本申请的技术方案中，第一AI训练节点与第二AI训练节点之间进行冗余传输，有利于提升信号功率增益，还带来数据增强的增益。即分类准确度的提升。
- [0271] 需要说明的是，上述步骤301和步骤302之间没有固定的执行顺序。可以先执行步骤 301，再执行步骤302；或者，先执行步骤302，再执行步骤301；或者，依据情况同时执行步骤301和步骤302，具体本申请不做限定。
- [0272] 可选的，图3所示的实施例还包括301b、步骤302a和步骤303a。
- [0273] 301b、AI任务节点向第二AI训练节点发送第三请求。第三请求用于请求第二AI训练节点对第二子模型进行训练。相应的，第二AI训练节点接收来自AI任务节点的第三请求。
- [0274] 第三请求与第一请求类似，关于第三请求可以参阅前述第一请求的相关介绍。例如，如图4所示，第一模型是AI模型节点请求训练的模型。AI任务节点将第一模型拆分为第一子模型和第二子模型。然后，AI任务节点向第二AI训练节点发送第三请求，以请求第二AI训练节点对第二子模型进行训练。
- [0275] 关于第三请求包括的内容与第一请求包括的内容类似，具体可以参阅前述第一请求包括的内容的相关介绍，这里不再赘述。
- [0276] 302a、AI任务节点向第二AI训练节点发送第二配置信息。第二配置信息用于配置至少一个第二附加模块。相应的，第二AI训练节点接收来自AI任务节点的第二配置信息。

- [0277] 该至少一个第二附加模块用于将用于更新第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号；和/或，该至少一个第二附加模块用于将第二AI训练节点在空口上接收到的符号映射为第二子模型输入的符号。
- [0278] 关于更新参数请参阅前述步骤302的相关介绍，这里不再赘述。可选的，该至少一个第二附加模块中各个第二附加模块为神经网络，或信号处理模块。关于信号处理模块可以参阅前述步骤302的相关介绍。
- [0279] 例如，如图4所示，该至少一个第二附加模块包括一个第二附加模块。第二AI训练节点将第二AI训练节点在空口接收到的来自第一AI训练节点的符号映射为第二子模型输入的符号，并输入第二子模型得到第二子模型的输出数据。然后，第二AI训练节点根据输出数据和AI任务节点为第二AI训练节点配置的第二损失函数计算得到用于更新第二子模型的权重梯度以及用于更新第一子模型的中间梯度。关于中间梯度的生成过程可以参阅申请号为202110412895.X的专利文件中的相关介绍。然后，第二AI训练节点通过该权重梯度更新该第二子模型。第二AI训练节点通过该至少一个第二附加模块将该中间梯度转换为第二AI训练节点在空口上待输出的符号，并向第一AI训练节点发送该待输出的符号。需要说明的是，图4所示的第二附加模块包括四个输出维度。该四个输出维度与第二子模型之间的连接结构可以理解为第二附加模块的结构。图4所示的示例仅仅是为了方便示图。实际上，第二附加模块包括第二附加模块的四个输出维度与第二子模型之间的连接结构以及该四个输出维度。
- [0280] 需要说明的是，第二AI训练节点通过该至少一个第二附加模块将用于更新第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号的过程与申请号为202110412895.X的专利文件中第一通信装置将第一中间梯度映射到空口资源上，生成第一梯度信号的过程类似，具体可以参阅该专利文件中的相关介绍。第二AI训练节点通过该至少一个第二附加模块将第二AI训练节点在空口上接收到的符号映射为第二子模型输入的符号的过程与申请号为202110412895.X的专利文件中第二通信装置根据第二梯度信号，确定第二中间梯度的过程类似，具体可以参阅该专利文件中的相关介绍。

- [0281] 第二配置信息与第一配置信息类似，具体可以参阅前述第一配置信息的相关介绍，这里不再赘述。
- [0282] 需要说明的是，上述步骤301b与步骤302a之间没有固定的执行顺序。可以先执行步骤301b，再执行步骤302a；或者，先执行步骤302a，再执行步骤301b；或者，依据情况同时执行步骤301b和步骤302a，具体本申请不做限定。
- [0283] 303a、第二AI训练节点通过至少一个第二附加模块将用于更新第一AI训练节点的第一子模型的更新参数转换为第二AI训练节点在空口上待输出的符号，并向第一AI训练节点发送该待输出的符号；和/或，通过至少一个第二附加模块将第二AI训练节点在空口上接收到的来自第一AI训练节点的符号映射为待输入第二AI训练节点的第二子模型的符号。
- [0284] 具体的，第二AI训练节点接收到第三请求之后，第二AI训练节点确定该第二子模型。第二AI训练节点接收到第二配置信息之后，第二AI训练节点确定该至少一个第二附加模块。然后，第二AI训练节点执行上述步骤303a。例如，如图4所示的，第二AI训练节点确定用于更新第一子模型的中间梯度。然后，第二AI训练节点通过该至少一个第二附加模块将该中间梯度转换为第二AI训练节点在空口上待输出的符号，并向第一AI训练节点发送该待输出的符号。和/或，第二AI训练节点将第二AI训练节点在空口上接收到的来自第一AI训练节点的符号映射为待输入第二AI训练节点的第二子模型的符号。第二AI训练节点将待输入第二AI训练节点的第二子模型的符号输入第二子模型得到第二子模型的输出数据。然后，第二AI训练节点根据输出数据和AI任务节点为第二AI训练节点配置的第二损失函数计算得到用于更新第二子模型的权重梯度以及用于更新第一子模型的中间梯度。第二AI训练节点根据权重梯度更新第二子模型。
- [0285] 可选的，上述步骤303a具体包括：第二AI训练节点通过第二冗余处理模块对待输出的符号进行冗余处理，再向第一AI训练节点发送经过冗余处理的待输出的符号；和/或，第二AI训练节点通过第二冗余处理模块对第二AI训练节点在空口上接收到来自第一AI训练节点的符号进行去冗余处理，再通过至少一个第二附加模块将经过去冗余处理的符号转换为待输入第二AI训练节点的第二子模型的符号。该实现方式与前述步骤303中第一AI训练节点进行冗余传输的过程类似，

具体可以参阅前述的相关介绍。可选的，第二冗余处理模块的维度等于冗余倍数与至少一个第二附加模块的维度的乘积。关于冗余倍数 β 请参阅前述的相关介绍。

[0286] 如图11所示， β 等于2。第二AI训练节点通过第二冗余处理模块对待输出的符号进行冗余处理，再向第一AI训练节点发送经过冗余处理的待输出的符号。相应的，第一AI训练节点执行了冗余发送。因此第二AI训练节点通过第二冗余处理模块对第二AI训练节点在空口上接收到来自第一AI训练节点的符号进行去冗余处理，再通过至少一个第二附加模块将经过去冗余处理的符号转换为待输入第二AI训练节点的第二子模型的符号。

[0287] 可选的，第二冗余处理模块的输入维度大于或等于至少一个第二附加模块的输出维度，第二冗余处理模块的输出维度小于或等于第一冗余处理模块的输出维度。通常情况下，第二冗余处理模块的输出维度等于第一冗余处理模块的输出维度。例如，如图11所示，第一冗余处理模块的输出维度为8，第二附加模块的输出维度为4，第二冗余处理模块的输入维度和输出维度均为8。

[0288] 关于第二AI训练节点获取冗余倍数的方式有多种，具体可以参阅后文图13所示的实施例中的相关介绍，这里不再赘述。

[0289] 需要说明的是，步骤301b、步骤302a和步骤303a与前述步骤301至步骤303之间没有固定的执行顺序。可以先执行步骤301b、步骤302a和步骤303a，再执行步骤301至步骤303；或者，先执行步骤301至步骤303，再执行步骤301b、步骤302a和步骤303a；或者，依据情况同时执行步骤301b、步骤302a和步骤303a以及步骤301至步骤303，具体本申请不做限定。

[0290] 可选的，图3所示的实施例还包括步骤301c至步骤301f。

[0291] 301c、AI任务节点向第一AI训练节点发送第二请求。第二请求用于请求第一AI训练节点与第二AI训练节点之间的信道状态信息。相应的，第一AI训练节点接收来自AI任务节点的第二请求。

[0292] 301d、第一AI训练节点向第二AI训练节点发送信道测量信号。相应的，第二AI训练节点接收来自第一AI训练节点的信道测量信号。

- [0293] 301e、第二AI训练节点向第一AI训练节点发送信道状态信息。相应的，第一AI训练节点接收来自第二AI训练节点的信道状态信息。
- [0294] 其中，信道状态信息是第二AI训练节点测量该信道测量信号得到的。
- [0295] 301f、第一AI训练节点向AI任务节点发送信道状态信息。相应的，AI任务节点接收来自第一AI训练节点的信道状态信息。
- [0296] 基于上述步骤301f，可选的，图3所示的实施例还包括：AI任务节点根据信道状态信息和任务类型确定该至少一个第一附加模块。可选的，AI任务节点根据信道状态信息和任务确定该至少一个第二附加模块。从而实现AI任务节点结合第一AI训练节点和第二AI训练节点之间的信道状态为第一AI训练节点和第二AI训练节点配置合适的附加模块。从而抵抗信道的干扰。
- [0297] 需要说明的是，可选的，当AI任务节点通过该信道状态信息确定第一AI训练节点与第二AI训练节点之间的信道状态较差，AI任务节点也可以重新选择其他AI训练节点以及对第一模型进行重新拆分，并请求重新选择的AI训练节点进行对拆分的子模型进行训练。
- [0298] 可选的，图3所示的实施例还包括步骤304至步骤306。步骤304可以在步骤303之后执行。步骤305可以在步骤303a之后执行。步骤306可以在步骤304和步骤305之后执行。
- [0299] 304、第一AI训练节点向AI任务节点发送经过训练的第一子模型的结构参数。相应的，AI任务节点接收来自第一AI训练节点的经过训练的第一子模型的结构参数。
- [0300] 具体的，第一AI训练节点对第一子模型进行训练后得到经过训练的第一子模型的结构参数。然后，第一AI训练节点可以向AI任务节点发送经过训练的第一子模型的结构参数。
- [0301] 可选的，第一AI训练节点还向AI任务节点发送至少一个第一附加模块的结构参数。在该实现方式中，第一AI训练节点将该至少一个第一附加模块作为第一子模型的一部分上报给AI任务节点。例如，如图4所示，第一AI训练节点将第一子模型和第一附加模块分别对应的结构参数上报给AI任务节点。

- [0302] 305、第二AI训练节点向AI任务节点发送经过训练的第二子模型的结构参数。相应的，AI任务节点接收来自第二AI训练节点的经过训练的第二子模型的结构参数。
- [0303] 具体的，第二AI训练节点对第二子模型进行训练后得到经过训练的第二子模型的结构参数。然后，第二AI训练节点可以向AI任务节点发送经过训练的第二子模型的结构参数。
- [0304] 可选的，第二AI训练节点还向AI任务节点发送至少一个第二附加模块的结构参数。在该实现方式中，第二AI训练节点将该至少一个第二附加模块作为第二子模型的一部分上报给AI任务节点。例如，如图4所示，第二AI训练节点将第二子模型和第二附加模块分别对应的结构参数上报给AI任务节点。
- [0305] 306、AI任务节点向AI模型节点发送经过训练的第一模型的结构参数。相应的，AI模型节点接收来自AI任务节点的经过训练的第一模型的结构参数。
- [0306] 具体的，AI任务节点根据经过训练的第一子模型的结构参数和经过训练的第二子模型的结构参数确定经过训练的第一模型的结构参数。然后，AI任务节点向AI模型节点发送经过训练的第一模型的结构参数。从而实现在AI模型节点上部署该第一模型。
- [0307] 可选的，第一AI训练节点还向AI任务节点发送至少一个第一附加模块的结构参数。第二AI训练节点还向AI任务节点发送至少一个第二附加模块的结构参数。也就是第一AI训练节点将该至少一个第一附加模块作为第一子模型的一部分上报给AI任务节点。第二AI训练节点将该至少一个第二附加模块作为第二子模型的一部分上报给AI任务节点。AI任务节点可以保留该至少一个第一附加模块和该至少一个第二附加模块。例如，AI任务节点保留第一附加模块和第二附加模块得到的第一模型如图14所示。其中，第一附加模块的输出维度和第二附加模块的输出维度合并。需要说明的是，图14所示的示图仅仅是为了方便示图。实际上，图14所示的第一附加模块包括第一附加模块的四个输出维度以及第一附加模块与第一子模型之间的连接结构。图14所示的第二附加模块包括第二附加模块的四个输出维度以及第二附加模块与第二子模型之间的连接结构。第一附加模块和第二附加模块共用四个输出维度。或者AI任务节点也可以去除该至少

一个第一附加模块和该至少一个第二附加模块。例如，AI任务节点去除该至少一个第二附加模块得到的第一模型如图15所示。

[0308] 需要说明的是，上述图3所示的实施例中是以AI任务节点将第一模型拆分为第一子模型和第二子模型，并请求第一AI训练节点训练第一子模型以及请求第二AI训练节点训练第二子模型的过程为例进行介绍。实际应用中，AI任务节点可以将第一模型拆分为多个子模型，然后，AI任务节点请求多个AI训练节点分别训练对应的子模型。例如，AI任务节点可以将第一模型拆分为三个子模型。AI任务节点请求第一AI训练节点训练三个子模型中的第一个子模型，请求第二AI训练节点训练三个子模型中的第二个子模型，以及请求第三AI训练节点训练三个子模型中的第三个子模型。具体本申请不做限定。

[0309] 需要说明的是，可选的，在ORAN系统中，可以通过O-CU或O-DU执行上述图3所示的实施例中AI任务节点执行的操作。如果O-CU划分为O-CU-UP和O-CU-CP，那么可以通过O-CU-UP和/或O-CU-CP执行上述图3所示的实施例中AI任务节点执行的操作。可选的，可以通过O-CU或O-DU执行上述图3所示的实施例中的AI模型节点执行的操作。如果O-CU划分为O-CU-UP和O-CU-CP，那么可以通过O-CU-UP和/或O-CU-CP执行上述图3所示的实施例中AI模型节点执行的操作。可选的，可以通过O-RU、O-DU或O-CU执行上述图3所示的实施例中第一AI训练节点执行的操作。如果O-CU划分为O-CU-UP和O-CU-CP，那么可以通过O-CU-UP和/或O-CU-CP执行上述图3所示的实施例中第一AI训练节点执行的操作。可选的，可以通过O-RU、O-DU或O-CU执行上述图3所示的实施例中第二AI训练节点执行的操作。如果O-CU划分为O-CU-UP和O-CU-CP，那么可以通过O-CU-UP和/或O-CU-CP执行上述图3所示的实施例中第二AI训练节点执行的操作。

[0310] 本申请中，可选的，AI任务节点可以为第一接入网设备，第一AI训练节点可以为第二接入网设备。AI任务节点和第一AI训练节点构成通信系统。第一接入网设备用于执行上述图3所示的实施例中AI任务节点执行的操作。第二接入网设备用于执行上述图3所示的实施例中的第一AI训练节点执行的操作。可选的，通信系统还包括第二AI训练节点，第二AI训练节点可以为第三接入网设备。第三接入网设备用于执行上述图3所示的实施例中第二AI训练节点执行的操作。可选

的，通信系统还包括AI模型节点，AI模型节点为第四接入网设备，第四接入网设备用于执行上述图3所示的实施例中AI模型节点执行的操作。

[0311] 本申请中，可选的，第一AI训练节点为第一终端设备。第二AI训练节点为第二终端设备。第一终端设备和第二终端设备构成通信系统。第一终端设备用于执行上述图3所示的实施例中的第一AI训练节点执行的操作。第二终端设备用于执行上述图3所示的实施例中的第二AI训练节点执行的操作。可选的，通信系统还包括AI任务节点，AI任务节点为第一接入网设备或第三终端设备。第一接入网设备或第三终端设备用于执行上述图3所示的实施例中AI任务节点执行的操作。可选的，通信系统还包括AI模型节点，AI模型节点为第二接入网设备，第二接入网设备用于执行上述图3所示的实施例中AI模型节点执行的操作。

[0312] 本申请中，可选的，第一AI训练节点可以为第一接入网设备。第二AI训练节点可以为第二接入网设备。第一接入网设备和第二接入网设备构成通信系统。第一接入网设备用于执行上述图3所示的实施例中的第一AI训练节点执行的操作。第二接入网设备用于执行上述图3所示的实施例中的第二AI训练节点执行的操作。可选的，通信系统还包括AI任务节点，AI任务节点可以为服务器。服务器用于执行上述图3所示的实施例中AI任务节点执行的操作。可选的，通信系统还包括AI模型节点，AI模型节点可以为第三接入网设备或终端设备，第三接入网设备或终端设备用于执行上述图3所示的实施例中AI模型节点执行的操作。

[0313] 本申请实施例中，AI任务节点向第一AI训练节点发送第一请求。第一请求用于请求第一AI训练节点对第一子模型进行训练。然后，AI任务节点向第一AI训练节点发送第一配置信息。第一配置信息用于配置至少一个第一附加模块，该至少一个第一附加模块用于将第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号，和/或，至少一个第一附加模块用于将第一AI训练节点在空口上接收到的符号转换为用于更新第一子模型的更新参数。由此可知，本申请的技术方案中，AI任务节点可以为第一AI训练节点配置该至少一个第一附加模块。该至少一个第一附加模块用于将第一子模型输出的符号映射为第一AI训练节点在空口上待输出的符号，和/或，至少一个第一附加模块用于将第一AI训练节点在空口上接收到的符号转换为用于更新第一子模型的更新参数。有利于实

现第一AI训练节点通过该至少一个第一附加模块实现与第二AI训练节点进行数据和/或参数等的传输。实现第一AI训练节点与第二AI训练节点进行联合训练。

[0314] 下面结合图13所示的实施例介绍第一AI训练节点和第二AI训练节点分别获取冗余倍数的两种可能的实现方式。对于其他方式本申请仍适用，下述实现方式不属于对本申请的限定。

[0315] 请参阅图13，下面结合步骤1301至步骤1302介绍方式1。

[0316] 1301、第一AI训练节点向第二AI训练节点发送信道测量信号。相应的，第二AI训练节点接收来自第一AI训练节点的信道测量信号。

[0317] 具体的，第二AI训练节点接收到信道测量信号之后，第二AI训练节点测量信道测量信号得到第一AI训练节点与第二AI训练节点之间的信道状态信息。然后，第二AI训练节点通过信道状态信息确定冗余倍数。例如，第二AI训练节点可以通过信道状态信息确定第一AI训练节点与第二AI训练节点之间的当前信道的信噪比。然后，第二AI训练节点根据第一模型对应的目标性能和该当前信道的信噪比确定该冗余倍数。关于第一模型对应的目标性能请参阅前述的相关介绍。例如，目标性能越高，信噪比越低，则冗余倍数越大。

[0318] 1302、第二AI训练节点向第一AI训练节点发送该冗余倍数。相应的，第一AI训练节点接收来自第二AI训练节点的冗余倍数。

[0319] 在上述步骤1301中，第二AI训练节点确定冗余倍数。第二AI训练节点可以向第一AI训练节点发送该冗余倍数。

[0320] 下面结合步骤1303至步骤1305介绍方式2。

[0321] 1303、第一AI训练节点向第二AI训练节点发送信道测量信号。相应的，第二AI训练节点接收来自第一AI训练节点的信道测量信号。

[0322] 1304、第二AI训练节点向第一AI训练节点发送信道状态信息。相应的，第一AI训练节点接收来自第二AI训练节点的信道状态信息。

[0323] 第二AI训练节点接收到信道测量信号之后，第二AI训练节点测量信道测量信号得到第一AI训练节点与第二AI训练节点之间的信道状态信息。

- [0324] 1305、第一AI训练节点向第二AI训练节点发送冗余倍数。相应的，第二AI训练节点接收来自第一AI训练节点的冗余倍数。
- [0325] 第一AI训练节点接收到信道状态信息之后，第一AI训练节点根据该信道状态信息确定冗余倍数。
- [0326] 可选的，在ORAN系统中，可以通过O-RU、O-DU或O-CU执行上述图13所示的实施例中第一AI训练节点执行的操作。如果O-CU划分为O-CU-UP和O-CU-CP，那么可以通过O-CU-UP和/或O-CU-CP执行上述图13所示的实施例中第一AI训练节点执行的操作。可选的，可以通过O-RU、O-DU或O-CU执行上述图13所示的实施例中第二AI训练节点执行的操作。如果O-CU划分为O-CU-UP和O-CU-CP，那么可以通过O-CU-UP和/或O-CU-CP执行上述图13所示的实施例中第二AI训练节点执行的操作。
- [0327] 下面对本申请实施例提供的通信装置进行描述。
- [0328] 图16为本申请实施例通信装置的一个结构示意图。请参阅图16，通信装置可以用于执行图3所示的实施例中AI任务节点执行的过程，具体请参考上述方法实施例中的相关介绍。
- [0329] 通信装置1600包括收发模块1601。可选的，通信装置1600还包括处理模块1602。
- [0330] 处理模块1602用于进行数据处理。收发模块1601可以实现相应的通信功能。收发模块1601还可以称为通信接口或通信模块。
- [0331] 可选地，该通信装置1600还可以包括存储模块，该存储模块可以用于存储指令和/或数据，处理模块1602可以读取存储模块中的指令和/或数据，以使得通信装置实现前述方法实施例。
- [0332] 该通信装置1600可以用于执行上文方法实施例中AI任务节点所执行的动作。该通信装置1600可以为AI任务节点或者可配置于AI任务节点的部件。处理模块1602用于执行上文方法实施例中AI任务节点侧的处理相关的操作。收发模块1601用于执行上文方法实施例中AI任务节点侧的接收相关的操作。
- [0333] 可选的，收发模块1601可以包括发送模块和接收模块。发送模块用于执行上述方法实施例中的发送操作。接收模块用于执行上述方法实施例中的接收操作。

- [0334] 需要说明的是，通信装置1600可以包括发送模块，而不包括接收模块。或者，通信装置1600可以包括接收模块，而不包括发送模块。具体可以视通信装置1600执行的上述方案中是否包括发送动作和接收动作。
- [0335] 可选的，该通信装置1600用于执行上述图3所示的实施例中AI任务节点所执行的动作。具体可以参阅上述图3所示的实施例中的相关介绍，这里不详细展开。
- [0336] 应理解，各模块执行上述相应过程的具体过程在上述方法实施例中已经详细说明，为了简洁，在此不再赘述。
- [0337] 上文实施例中的处理模块1602可以由至少一个处理器或处理器相关电路实现。收发模块1601可以由收发器或收发器相关电路实现。收发模块1601还可称为通信模块或通信接口。存储模块可以通过至少一个存储器实现。
- [0338] 图17为本申请实施例通信装置的另一个结构示意图。请参阅图17，通信装置可以用于执行图3和图13中任一实施例中第一AI训练节点或第二AI训练节点执行的过程，具体请参考上述方法实施例中的相关介绍。
- [0339] 通信装置1700包括收发模块1701和处理模块1702。
- [0340] 处理模块1702用于进行数据处理。收发模块1701可以实现相应的通信功能。收发模块1701还可以称为通信接口或通信模块。
- [0341] 可选地，该通信装置1700还可以包括存储模块，该存储模块可以用于存储指令和/或数据，处理模块1702可以读取存储模块中的指令和/或数据，以使得通信装置实现前述方法实施例。
- [0342] 该通信装置1700可以用于执行上文方法实施例中第一AI训练节点或第二AI训练节点所执行的动作。该通信装置1700可以为第一AI训练节点或第二AI训练节点，或者可配置于第一AI训练节点或第二AI训练节点的部件。处理模块1702用于执行上文方法实施例中第一AI训练节点或第二AI训练节点侧的处理相关的操作。收发模块1701用于执行上文方法实施例中第一AI训练节点或第二AI训练节点侧的发送或接收相关的操作。
- [0343] 可选的，收发模块1701可以包括发送模块和接收模块。发送模块用于执行上述方法实施例中的发送操作。接收模块用于执行上述方法实施例中的接收操作。

- [0344] 需要说明的是，通信装置1700可以包括发送模块，而不包括接收模块。或者，通信装置1700可以包括接收模块，而不包括发送模块。具体可以视通信装置1700执行的上述方案中是否包括发送动作和接收动作。
- [0345] 可选的，该通信装置1700用于执行上述图3和图13中任一实施例中第一AI训练节点或第二AI训练节点所执行的动作。具体可以参阅上述图3和图13中任一实施例中的相关介绍，这里不详细展开。
- [0346] 应理解，各模块执行上述相应过程的具体过程在上述方法实施例中已经详细说明，为了简洁，在此不再赘述。
- [0347] 上文实施例中的处理模块1702可以由至少一个处理器或处理器相关电路实现。收发模块1701可以由收发器或收发器相关电路实现。收发模块1701还可称为通信模块或通信接口。存储模块可以通过至少一个存储器实现。
- [0348] 本申请实施例还提供一种通信装置1800。请参阅图18，该通信装置1800包括处理器1810，处理器1810与存储器1820耦合，存储器1820用于存储计算机程序或指令和/或数据，处理器1810用于执行存储器1820存储的计算机程序或指令和/或数据，使得上文方法实施例中的方法被执行。该通信装置1800用于实现上文方法实施例中由AI任务节点、第一AI训练节点或第二AI训练节点执行的操作。
- [0349] 可选的，该通信装置1800包括的处理器1810为一个或多个。
- [0350] 可选的，如图18所示，该通信装置1800还可以包括存储器1820。
- [0351] 可选的，该通信装置1800包括的存储器1820可以为一个或多个。
- [0352] 可选的，该存储器1820可以与该处理器1810集成在一起，或者分离设置。
- [0353] 可选的，如图18所示，该通信装置1800还可以包括收发器1830，收发器1830用于信号的接收和/或发送。例如，处理器1810用于控制收发器1830进行信号的接收和/或发送。
- [0354] 本申请还提供一种通信装置1900，该通信装置1900可以为终端设备、终端设备中的处理器、或芯片。该通信装置1900可以用于执行上述方法实施例中由AI任务节点、第一AI训练节点或第二AI训练节点所执行的操作。
- [0355] 当该通信装置1900为终端设备时，图19示出了一种简化的终端设备的结构示意图。如图19所示，终端设备包括处理器、存储器、以及收发器。存储器可以存

储计算机程序代码，收发器包括发射机1931、接收机1932、射频电路(图中未示出)、天线1933以及输入输出装置(图中未示出)。

- [0356] 处理器主要用于对通信协议以及通信数据进行处理；对终端设备进行控制，执行软件程序以及处理软件程序的数据等。存储器主要用于存储软件程序和数据。射频电路主要用于基带信号与射频信号的转换以及对射频信号的处理。天线主要用于收发电磁波形式的射频信号。输入输出装置可以包括触摸屏、显示屏，或键盘等。输入输出装置主要用于接收用户输入的数据以及对用户输出数据。需要说明的是，有些种类的终端设备可以不具有输入输出装置。
- [0357] 当需要发送数据时，处理器对待发送的数据进行基带处理后，输出基带信号至射频电路。然后，射频电路将基带信号进行射频处理后将射频信号通过天线以电磁波的形式向外发送。当有数据发送到终端设备时，射频电路通过天线接收到射频信号。射频电路将射频信号转换为基带信号，并将基带信号输出至处理器。处理器将基带信号转换为数据并对该数据进行处理。为便于说明，图19中仅示出了一个存储器、处理器和收发器。在实际的终端设备的产品中，可以存在一个或多个处理器和一个或多个存储器。存储器也可以称为存储介质或者存储设备等。存储器可以是独立于处理器设置，也可以是与处理器集成在一起，本申请实施例对此不做限制。
- [0358] 在本申请实施例中，可以将具有收发功能的天线和射频电路视为终端设备的收发模块，将具有处理功能的处理器视为终端设备的处理模块。
- [0359] 如图19所示，终端设备包括处理器1910、存储器1920和收发器1930。处理器1910也可以称为处理单元，处理单板，处理模块，或处理装置等。收发器1930也可以称为收发单元，收发机，或收发装置等。
- [0360] 可选的，将收发器1930中用于实现接收功能的器件视为接收模块，将收发器1930中用于实现发送功能的器件视为发送模块，即收发器1930包括接收器和发送器。收发器有时也可以称为收发机、收发模块、或收发电路等。接收器有时也可以称为接收机、接收模块、或接收电路等。发送器有时也可以称为发射机、发射模块或者发射电路等。

- [0361] 处理器1901用于执行上述图3所示的实施例中AI任务节点、第一AI训练节点或第二AI训练节点侧的处理动作，或图13所示的实施例中第一AI训练节点或第二AI训练节点侧的处理动作。收发器1930用于执行上述图3所示的实施例中AI任务节点、第一AI训练节点或第二AI训练节点侧的收发动作，或图13所示的实施例中第一AI训练节点或第二AI训练节点侧的收发动作。
- [0362] 应理解，图19仅为示例而非限定，上述包括收发模块和处理模块的终端设备可以不依赖于图16、图17或图19所示的结构。
- [0363] 当该通信装置1900为芯片时，该芯片包括处理器、存储器和收发器。其中，收发器可以是输入输出电路或通信接口。处理器可以为该芯片上集成的处理模块或者微处理器或者集成电路。上述方法实施例中AI任务节点、第一AI训练节点或第二AI训练节点的发送操作可以理解为芯片的输出，上述方法实施例中AI任务节点、第一AI训练节点或第二AI训练节点的接收操作可以理解为芯片的输入。
- [0364] 本申请还提供一种通信装置2000，该通信装置2000可以是网络设备也可以是芯片。该通信装置2000可以用于执行上述图3所示的实施例中AI任务节点、第一AI训练节点或第二AI训练节点所执行的操作，或用于执行上述图13所示的实施例中第一AI训练节点或第二AI训练节点所执行的操作。
- [0365] 当该通信装置2000为网络设备时，例如为基站。图20示出了一种简化的基站结构示意图。基站包括2010部分、2020部分以及2030部分。2010部分主要用于基带处理，对基站进行控制等；2010部分通常是基站的控制中心，通常可以称为处理器，用于控制基站执行上述方法实施例中AI任务节点、第一AI训练节点或第二AI训练节点侧的处理操作。2020部分主要用于存储计算机程序代码和数据。2030部分主要用于射频信号的收发以及射频信号与基带信号的转换；2030部分通常可以称为收发模块、收发机、收发电路、或者收发器等。2030部分的收发模块，也可以称为收发机或收发器等，其包括天线2033和射频电路(图中未示出)，其中射频电路主要用于进行射频处理。可选的，可以将2030部分中用于实现接收功能的器件视为接收机，将用于实现发送功能的器件视为发射机，

即2030部分包括接收机2032和发射机2031。接收机也可以称为接收模块、接收器、或接收电路等，发送机可以称为发射模块、发射器或者发射电路等。

[0366] 2010部分与2020部分可以包括一个或多个单板，每个单板可以包括一个或多个处理器和一个或多个存储器。处理器用于读取和执行存储器中的程序以实现基带处理功能以及对基站的控制。若存在多个单板，各个单板之间可以互联以增强处理能力。作为一种可选的实施方式，也可以是多个单板共用一个或多个处理器，或者是多个单板共用一个或多个存储器，或者是多个单板同时共用一个或多个处理器。

[0367] 例如，在一种实现方式中，2030部分的收发模块用于执行图3所示的实施例中由AI任务节点、第一AI训练节点或第二AI训练节点执行的收发相关的过程。2010部分的处理器用于图3所示的实施例中由AI任务节点、第一AI训练节点或第二AI训练节点执行的收发相关的过程。在另一种实现方式中，2030部分的收发模块用于执行图13所示的实施例中由第一AI训练节点或第二AI训练节点执行的收发相关的过程。2010部分的处理器用于图13所示的实施例中由第一AI训练节点或第二AI训练节点执行的收发相关的过程。

[0368] 应理解，图20仅为示例而非限定，上述包括处理器、存储器以及收发器的网络设备可以不依赖于图16、图17或图20所示的结构。

[0369] 当该通信装置2000为芯片时，该芯片包括收发器、存储器和处理器。其中，收发器可以是输入输出电路、通信接口；处理器为该芯片上集成的处理器、或者微处理器、或者集成电路。上述方法实施例中AI任务节点、第一AI训练节点或第二AI训练节点的发送操作可以理解为芯片的输出，上述方法实施例中AI任务节点、第一AI训练节点或第二AI训练节点的接收操作可以理解为芯片的输入。

[0370] 本申请实施例还提供一种计算机可读存储介质，其上存储有用于实现上述方法实施例中由AI任务节点、第一AI训练节点或第二AI训练节点执行的方法的计算机指令。

[0371] 例如，该计算机程序被计算机执行时，使得该计算机可以实现上述方法实施例中由AI任务节点、第一AI训练节点或第二AI训练节点执行的方法。

- [0372] 本申请实施例还提供一种包含指令的计算机程序产品，该指令被计算机执行时使得该计算机实现上述方法实施例中由AI任务节点、第一AI训练节点或第二AI训练节点执行的方法。
- [0373] 本申请实施例还提供一种通信系统，该通信系统包括上文实施例中的AI任务节点和上文实施例中的第一AI训练节点。可选的，该通信系统还包括上文实施例中的第二AI训练节点。可选的，该通信系统还包括上文实施例中的AI模型节点。
- [0374] 本申请实施例还提供另一种通信系统，该通信系统包括上文实施例中的第一AI训练节点和上文实施例中的第二AI训练节点。可选的，该通信系统还包括上文实施例中的AI任务节点。可选的，该通信系统还包括上文实施例中的AI模型节点。
- [0375] 本申请实施例还提供一种芯片装置，包括处理器，用于调用该存储器中存储的计算机程序或计算机指令，以使得该处理器执行上述图3和图13所示的实施例提供的方法。
- [0376] 一种可能的实现方式中，该芯片装置的输入对应上述图3和图13所示的实施例中任一个实施例中的接收操作，该芯片装置的输出对应上述图3和图13所示的实施例中任一个实施例中的发送操作。
- [0377] 可选的，该处理器通过接口与存储器耦合。
- [0378] 可选的，该芯片装置还包括存储器，该存储器中存储有计算机程序或计算机指令。
- [0379] 其中，上述任一处提到的处理器，可以是一个通用中央处理器，微处理器，特定应用集成电路(application-specific integrated circuit, ASIC)，或一个或多个用于控制上述图3和图13所示的实施例中任一个实施例提供的方法的程序执行的集成电路。上述任一处提到的存储器可以为只读存储器(read-only memory, ROM)或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器(random access memory, RAM)等。

- [0380] 所属领域的技术人员可以清楚地了解到，为描述方便和简洁，上述提供的任何一种通信装置中相关内容的解释及有益效果均可参考上文提供的对应的方法实施例，此处不再赘述。
- [0381] 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。另一点，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，可以是电性，机械或其它的形式。
- [0382] 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实施例方案的目的。
- [0383] 另外，在本申请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。
- [0384] 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，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，本申请的技术方案本质上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)执行本申请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、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。
- [0385] 以上所述，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述 各实施例所记载的技术方案进行修改，或者对其中部分技

术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的范围。

权利要求书

[权利要求 1]

一种模型训练配置方法，其特征在于，所述方法包括：
人工智能AI任务节点向第一AI训练节点发送第一请求，所述第一请求用于请求所述第一AI训练节点对第一子模型进行训练；
所述AI任务节点向所述第一AI训练节点发送第一配置信息，所述第一配置信息用于配置至少一个第一附加模块，所述至少一个第一附加模块用于将所述第一子模型输出的符号映射为所述第一AI训练节点在空口上待输出的符号，和/或，所述至少一个第一附加模块用于将所述第一AI训练节点在空口上接收到的符号转换为用于更新所述第一子模型的更新参数。

[权利要求 2]

一种模型训练配置方法，其特征在于，所述方法包括：
第一人工智能AI训练节点接收来自AI任务节点的第一请求，所述第一请求用于请求所述第一AI训练节点对第一子模型进行训练；
所述第一AI训练节点接收来自所述AI任务节点的第一配置信息，所述第一配置信息用于配置至少一个第一附加模块；
所述第一AI训练节点通过所述至少一个第一附加模块将所述第一子模型输出的符号映射为所述第一AI训练节点在空口上待输出的符号，并向第二AI训练节点发送所述待输出的符号；和/或，所述第一AI训练节点通过所述至少一个第一附加模块将所述第一AI训练节点在空口上接收到来自所述第二AI训练节点的符号转换为用于更新所述第一子模型的更新参数，所述第一AI训练节点通过更新参数更新所述第一子模型。

[权利要求 3]

根据权利要求1或2所述的方法，其特征在于，所述第一配置信息包括以下至少一项：所述至少一个第一附加模块的索引、类型、结构参数、第一指示信息、第二指示信息、第三指示信息、第四指示信息、或第五指示信息；
其中，所述第一指示信息用于指示所述至少一个第一附加模块是否参与所述第一子模型的训练；所述第二指示信息用于指示所述至少

一个第一附加模块是否作为所述第一子模型的一部分上报给所述AI任务节点；所述第三指示信息用于指示所述至少一个第一附加模块中各个第一附加模块的部署方式，所述部署方式包括串联部署或并联部署；所述第四指示信息用于指示是否向所述第一AI训练节点指示所述至少一个第一附加模块的类型、结构参数；所述第五指示信息用于指示所述至少一个第一附加模块中各个第一附加模块插入所述第一子模型的位置。

[权利要求 4] 根据权利要求3所述的方法，其特征在于，所述第一配置信息还包括超网络的索引，所述超网络用于生成所述至少一个第一附加模块。

[权利要求 5] 根据权利要求3或4所述的方法，其特征在于，所述第五指示信息包括比特位图，所述比特位图中的比特与所述第一子模型中的神经网络层对应，所述比特位图中的比特用于指示在所述比特对应的神经网络层是否插入第一附加模块。

[权利要求 6] 根据权利要求5所述的方法，其特征在于，所述比特位图中的一个比特与所述第一子模型中的一层神经网络对应；
当所述比特位图中的一个比特的取值为1，指示在所述比特对应的一层神经网络与所述比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块；当所述比特位图中的一个比特的取值为0，指示在所述比特对应的一层神经网络与所述比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块；
或者，
当所述比特位图中的一个比特的取值为0，指示在所述比特对应的一层神经网络与所述比特对应的一层神经网络的下一层神经网络之间不插入第一附加模块；当所述比特位图中的一个比特的取值为1，指示在所述比特对应的一层神经网络与所述比特对应的一层神经网络的下一层神经网络之间插入相应的第一附加模块。

[权利要求 7]

根据权利要求3至6中任一项所述的方法，其特征在于，所述至少一个第一附加模块包括多个第一附加模块；

所述第五指示信息用于指示在所述第一子模型的第X层神经网络与第X+1层神经网络之间插入所述多个第一附加模块，所述X为大于或等于1且小于所述第一子模型包括的神经网络层数；其中，所述多个第一附加模块的输入维度之和等于所述第X层神经网络的输出维度，所述多个第一附加模块的输出维度之和等于所述第X+1层神经网络的输入维度；或者，所述多个第一附加模块的输入维度之和大于所述第X层神经网络的输出维度，所述第X层神经网络在输出数据上补零，使得所述第X层神经网络的输出维度与所述多个第一附加模块的输入维度之和相等，所述多个第一附加模块中的一个或多个第一附加模块在输出数据上降维，使得所述多个第一附加模块的输出维度之和等于所述第X+1层神经网络的输入维度；

或者，

所述第五指示信息用于指示在所述第一子模型的最外层神经网络插入所述多个第一附加模块；其中，所述多个第一附加模块的输入维度之和等于所述最外层神经网络的输出维度；或者，所述多个第一附加模块的输入维度之和大于所述最外层神经网络的输出维度，所述最外层神经网络在输出数据上补零，使得所述最外层神经网络的输出维度与所述多个第一附加模块的输入维度之和相等。

[权利要求 8]

根据权利要求1至7中任一项所述的方法，其特征在于，所述至少一个第一附加模块中各个第一附加模块为神经网络，或者为信号处理模块。

[权利要求 9]

根据权利要求1至8中任一项所述的方法，其特征在于，所述第一请求包括以下至少一项：所述第一子模型的结构、超参数、损失函数、训练停止条件、所述第一子模型在多个子模型中的编号、或数据集；其中，所述多个子模型是所述AI任务节点对AI模型节点请求训练的第一模型拆分得到的，所述训练停止条件是所述第一AI训练

节点停止训练所述第一子模型的条件，所述数据集用于作为所述第一子模型的输入数据。

[权利要求 10]

根据权利要求1、3至9中任一项所述的方法，其特征在于，在所述AI任务节点向所述第一AI训练节点发送第一配置信息之前，所述方法还包括：

所述AI任务节点向所述第一AI训练节点发送第二请求，所述第二请求用于请求所述第一AI训练节点与第二AI训练节点之间的信道状态信息；

所述AI任务节点接收来自所述第一AI训练节点的所述信道状态信息；

所述AI任务节点根据任务类型和所述信道状态信息确定所述至少一个第一附加模块，所述任务类型是AI模型节点向所述AI任务节点请求的任务类型。

[权利要求 11]

根据权利要求2至9中任一项所述的方法，其特征在于，在所述第一AI训练节点接收来自所述第一AI任务节点的第一配置信息之前，所述方法还包括：

所述第一AI训练节点接收来自所述AI任务节点的第二请求，所述第二请求用于请求所述第一AI训练节点与第二AI训练节点之间的信道状态信息；

所述第一AI训练节点向所述第二AI训练节点发送信道测量信号；

所述第一AI训练节点来自所述第二AI训练节点的信道状态信息，所述信道状态信息是所述第二AI训练节点测量所述信道测量信号得到的；

所述第一AI训练节点向所述AI任务节点发送所述信道状态信息。

[权利要求 12]

根据权利要求1、3至9、10中任一项所述的方法，其特征在于，所述方法还包括：

所述AI任务节点向第二AI训练节点发送第三请求，所述第三请求用于请求所述第二AI训练节点对第二子模型进行训练；

所述AI任务节点向所述第二AI训练节点发送第二配置信息，所述第二配置信息用于配置至少一个第二附加模块，所述至少一个第二附加模块用于将用于更新所述第一子模型的更新参数转换为所述第二AI训练节点在空口上待输出的符号，和/或，所述至少一个第二附加模块用于将所述第二AI训练节点在空口上接收到的符号映射为所述第二子模型输入的符号。

[权利要求 13]

根据权利要求2至9、11中任一项所述的方法，其特征在于，所述至少一个第一附加模块包括第一附加模块1；所述第五指示信息用于指示在所述第一子模型的第X层神经网络与第X+1层神经网络之间插入所述第一附加模块1，所述X为大于或等于1且小于所述第一子模型包括的神经网络层数，所述第一附加模块1的输入维度小于所述第X层神经网络的输出维度；所述方法还包括：

所述第一AI训练节点确定第一附加模块2，所述第一AI训练节点在所述第X层神经网络与所述第X+1层神经网络之间插入所述第一附加模块1和所述第一附加模块2；其中，所述第一附加模块1的输入维度与所述第一附加模块2的输入维度之和等于所述第X层神经网络的输出维度，所述第一附加模块1的输出维度与所述第一附加模块2的输出维度之和等于所述第X+1层神经网络的输入维度；或者，所述第一附加模块1的输入维度与所述第一附加模块2的输入维度之和大于所述第X层神经网络的输出维度，所述第X层神经网络在输出数据上补零，使得所述第X层神经网络的输出维度与所述第一附加模块1的输入维度与所述第一附加模块2的输入维度之和相等，所述第一附加模块1和/或所述第一附加模块2在输出数据上降维，使得所述第一附加模块1的输出维度与所述第一附加模块2的输出维度之和等于所述第X+1层神经网络的输入维度；

或者，

所述至少一个第一附加模块包括第一附加模块1；所述第五指示信息用于指示在所述第一子模型的最外层神经网络插入所述第一附加模块1；所述方法还包括：

所述第一AI训练节点确定第一附加模块2，所述第一AI训练节点在所述最外层神经网络插入所述第一附加模块1和所述第一附加模块2；其中，所述第一附加模块1的输入维度与所述第一附加模块2的输入维度之和等于所述最外层神经网络的输出维度；或者，所述第一附加模块1的输入维度与所述第一附加模块2的输入维度之和大于所述最外层神经网络的输出维度，所述最外层神经网络在输出数据上补零，使得所述最外层神经网络的输出维度与所述第一附加模块1的输入维度与所述第一附加模块2的输入维度之和相等。

[权利要求 14]

根据权利要求2至9、11、和13中任一项所述的方法，其特征在于，所述方法还包括：

所述第一AI训练节点通过第一冗余处理模块对所述待输出的符号进行冗余处理，再向所述第二AI训练节点发送经过冗余处理的待输出的符号；和/或，

所述第一AI训练节点通过所述第一冗余处理模块对所述第一AI训练节点在空口上接收到来自所述第二AI训练节点的符号进行去冗余处理，再通过所述至少一个第一附加模块将经过去冗余处理的符号转换为所述更新参数。

[权利要求 15]

根据权利要求14所述的方法，其特征在于，所述第一冗余处理模块的维度等于冗余倍数与所述至少一个第一附加模块的维度的乘积。

[权利要求 16]

根据权利要求15所述的方法，其特征在于，所述方法还包括：

所述第一AI训练节点向所述第二AI训练节点发送信道测量信号；
所述第一AI训练节点接收来自所述第二AI训练节点的所述冗余倍数，所述冗余倍数是所述第二AI训练节点通过测量所述信道测量信号确定的。

[权利要求 17]

根据权利要求15所述的方法，其特征在于，所述方法还包括：

所述第一AI训练节点向所述第二AI训练节点发送信道测量信号；
所述第一AI训练节点接收来自所述第二AI训练节点的信道状态信息，所述信道状态信息是所述第二AI训练节点测量所述信道测量信号得到的；

所述第一AI训练节点根据所述信道状态信息确定所述冗余倍数。

[权利要求 18] 根据权利要求14至17中任一项所述的方法，其特征在于，所述第一冗余处理模块为神经网络，或信号处理模块。

[权利要求 19] 一种模型训练配置方法，其特征在于，所述方法包括：
第二人工智能AI训练节点接收来自AI任务节点的第三请求，所述第三请求用于请求所述第二AI训练节点对第二子模型进行训练；
所述第二AI训练节点接收来自所述AI任务节点的第二配置信息，所述第二配置信息用于配置至少一个第二附加模块；
所述第二AI训练节点通过所述至少一个第二附加模块将用于更新所述第一AI训练节点的第一子模型的更新参数转换为所述第二AI训练节点在空口上待输出的符号，并向所述第一AI训练节点发送所述待输出的符号；和/或，所述第二AI训练节点通过所述至少一个第二附加模块将所述第二AI训练节点在空口上接收到来自所述第一AI训练节点的符号映射为待输入所述第二AI训练节点的第二子模型的符号。

[权利要求 20] 根据权利要求19所述的方法，其特征在于，所述方法还包括：
所述第二AI训练节点通过第二冗余处理模块对所述待输出的符号进行冗余处理，再向所述第一AI训练节点发送所述待输出的符号；和/或，
所述第二AI训练节点通过所述第二冗余处理模块对所述第二AI训练节点在空口上接收到来自所述第一AI训练节点的符号进行去冗余处理，再通过所述至少一个第二附加模块将经过去冗余处理的符号转换为所述待输入所述第二AI训练节点的第二子模型的符号。

- [权利要求 21] 根据权利要求20所述的方法，其特征在于，所述第二冗余处理模块的维度等于冗余倍数与所述至少一个第二附加模块的维度的乘积。
- [权利要求 22] 根据权利要求21所述的方法，其特征在于，所述方法还包括：
所述第二AI训练节点接收来自所述第一AI训练节点的信道测量信号；
所述第二AI训练节点向所述第一AI训练节点发送所述冗余倍数，所述冗余倍数是所述第二AI训练节点通过测量所述信道测量信号确定的。
- [权利要求 23] 根据权利要求21所述的方法，其特征在于，所述方法还包括：
所述第二AI训练节点接收来自所述第一AI训练节点的信道测量信号；
所述第二AI训练节点向所述第一AI训练节点发送信道状态信息，所述信道状态信息是所述第二AI训练节点测量所述信道测量信号得到的，所述信道状态信息用于确定所述冗余倍数。
- [权利要求 24] 根据权利要求20至23中任一项所述的方法，其特征在于，所述第二冗余处理模块为神经网络，或信号处理模块。
- [权利要求 25] 一种通信装置，其特征在于，所述通信装置包括收发模块，所述收发模块用于执行如权利要求1、3至10、12中任一项所述方法中的收发操作。
- [权利要求 26] 一种通信装置，其特征在于，所述通信装置包括收发模块和处理模块；
所述收发模块用于执行如权利要求2至9、11、13中任一项所述方法中的收发操作，所述处理模块用于执行如权利要求2至9、11、13中任一项所述方法中的处理操作；或者，
所述收发模块用于执行如权利要求19至24中任一项所述方法中的收发操作，所述处理模块用于执行如权利要求19至24中任一项所述方法中的处理操作。

- [权利要求 27] 一种通信装置，其特征在于，所述通信装置包括处理器，所述处理器用于执行存储器中的计算机程序或计算机指令，以执行如权利要求1至24中任一项所述的方法。
- [权利要求 28] 一种计算机可读存储介质，其特征在于，其上存储有计算机程序，所述计算机程序被通信装置执行时，使得所述通信装置执行如权利要求1至24任一项所述的方法。
- [权利要求 29] 一种通信系统，其特征在于，所述通信系统包括AI任务节点和第一AI训练节点；所述AI任务节点用于执行如权利要求1、3至10、12中任一项所述的方法，所述第一AI训练节点用于执行如权利要求2至9、11、13中任一项所述的方法。
- [权利要求 30] 根据权利要求29所述的通信系统，其特征在于，所述通信系统还包括第二AI训练节点，所述第二AI训练节点用于执行19至24中任一项所述方法。