

(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
2024年8月22日 (22.08.2024)



(10) 国际公布号
WO 2024/168917 A1

- (51) 国际专利分类号:
H04L 67/54 (2022.01)
- (21) 国际申请号: PCT/CN2023/076988
- (22) 国际申请日: 2023年2月17日 (17.02.2023)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 北京小米移动软件有限公司 (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) [CN/CN]; 中国北京市海淀区西二旗中路33号院6号楼8层018号, Beijing 100085 (CN)。
- (72) 发明人: 牟勤 (MU, Qin); 中国北京市海淀区西二旗中路33号院6号楼8层018号, Beijing 100085 (CN)。
- (74) 代理人: 北京法胜知识产权代理有限公司 (FASHENG INTELLECTUAL PROPERTY COMPANY, LTD.); 中国北京市海淀区北洼路45号1号楼2层202室, Beijing 100142 (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,

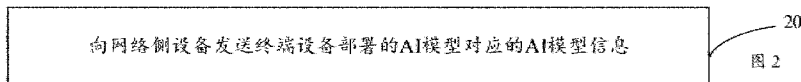
CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, 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))。

(54) Title: AI MODEL REGISTRATION METHOD, AND APPARATUS, DEVICE AND STORAGE MEDIUM

(54) 发明名称: 一种AI模型的注册方法/装置/设备及存储介质



201 Send, to a network-side device, AI model information corresponding to an AI model which is deployed on a terminal device

(57) Abstract: The present disclosure belongs to the technical field of communications. Provided are an AI model registration method, and an apparatus, a device and a storage medium. The method comprises: a terminal device sending, to a network-side device, AI model information corresponding to an AI model which is deployed on the terminal device. Therefore, the present disclosure provides the AI model registration method, such that an AI model which is deployed on a terminal device is registered on a network-side device, and the network-side device then knows model information of the AI model, which is deployed on the terminal device, thereby facilitating subsequent management or monitoring of the AI model, which is deployed on the terminal device, by means of the network-side device.

(57) 摘要: 本公开提出一种AI模型的注册方法/装置/设备/存储介质, 属于通信技术领域。终端设备会向网络侧设备发送终端设备部署的AI模型对应的AI模型信息。由此可知, 本公开提供了一种AI模型的注册方法, 以使得终端设备所部署AI模型在网络侧设备上注册, 从而使得网络侧设备知晓终端设备部署的AI模型的模型信息, 进而便于网络侧设备后续对部署在终端设备上的AI模型进行管理或监测。

WO 2024/168917 A1

一种 AI 模型的注册方法/装置/设备及存储介质

技术领域

本公开涉及通信技术领域，尤其涉及一种 AI 模型的注册方法/装置/设备及存储介质。

5

背景技术

随着 AI (Artificial Intelligent, 人工智能) 技术的不断发展, AI 模型在多个领域得到广泛应用。其中, 将通过终端设备或终端设备的外部服务器训练得到的 AI 模型部署在终端设备上。相关技术中, 需要终端设备向网络侧设备进行 AI 模型的注册, 以使得网络侧设备知晓终端设备的 AI 模型信息, 从而便于网络侧设备对部署在终端设备上的 AI 模型进行管理或监测。

10

因此, 亟需一种 AI 模型的注册方法。

发明内容

本公开提出的 AI 模型的注册方法/装置/设备及存储介质, 用于 AI 模型的注册。

15

第一方面, 本公开实施例提供一种 AI 模型的注册方法, 应用于终端设备, 所述方法包括: 向网络侧设备发送所述终端设备部署的 AI 模型对应的 AI 模型信息。

第二方面, 本公开实施例提供一种 AI 模型的注册方法, 应用于网络侧设备, 所述方法包括: 接收终端设备发送的所述终端设备部署的 AI 模型对应的 AI 模型信息。

20

第三方面, 本公开实施例提供一种通信装置, 该装置被配置在终端设备中, 包括:

发送模块, 用于向网络侧设备发送所述终端设备部署的 AI 模型对应的 AI 模型信息。

第四方面, 本公开实施例提供一种通信装置, 该装置被配置在网络侧设备中, 包括:

接收模块, 用于接收终端设备发送的所述终端设备部署的 AI 模型对应的 AI 模型信息。

25

第五方面, 本公开实施例提供一种通信装置, 该通信装置包括处理器, 当该处理器调用存储器中的计算机程序时, 执行上述第一方面所述的方法。

第六方面, 本公开实施例提供一种通信装置, 该通信装置包括处理器, 当该处理器调用存储器中的计算机程序时, 执行上述第二方面所述的方法。

第七方面, 本公开实施例提供一种通信装置, 该通信装置包括处理器和存储器, 该存储器中存储有计算机程序; 所述处理器执行该存储器所存储的计算机程序, 以使该通信装置执行上述第一方面所述的方法。

30

第八方面, 本公开实施例提供一种通信装置, 该通信装置包括处理器和存储器, 该存储器中存储有计算机程序; 所述处理器执行该存储器所存储的计算机程序, 以使该通信装置执行上述第二方面所述的方法。

第九方面, 本公开实施例提供一种通信装置, 该装置包括处理器和接口电路, 该接口电路用于接收代码指令并传输至该处理器, 该处理器用于运行所述代码指令以使该装置执行上述第一方面所述的方法。

35

第十方面, 本公开实施例提供一种通信装置, 该装置包括处理器和接口电路, 该接口电路用于接收代码指令并传输至该处理器, 该处理器用于运行所述代码指令以使该装置执行上述第二方面所述的方法。

第十一方面, 本公开实施例提供一种通信系统, 该系统包括第三方面所述的通信装置至第四方面所述的通信装置, 或者, 该系统包括第五方面所述的通信装置至第六方面所述的通信装置, 或者, 该系统包括第七方面所述的通信装置至第八方面所述的通信装置, 或者, 该系统包括第九方面所述的通信装置至第十方面所述的通信装置。

40

第十二方面, 本发明实施例提供一种计算机可读存储介质, 用于储存为上述网络设备所用的指令, 当所述指令被执行时, 使所述终端设备执行上述第一方面至第三方面的任一方面所述的方法。

第十三方面, 本公开还提供一种包括计算机程序的计算机程序产品, 当其在计算机上运行时, 使得计算机执行上述第一方面至第二方面的任一方面所述的方法。

45

第十四方面, 本公开提供一种芯片系统, 该芯片系统包括至少一个处理器和接口, 用于支持网络设

备实现第一方面至第二方面的任一方面所述的方法所涉及的功能，例如，确定或处理上述方法中所涉及的数据和信息中的至少一种。在一种可能的设计中，所述芯片系统还包括存储器，所述存储器，用于保存源辅节点必要的计算机程序和数据。该芯片系统，可以由芯片构成，也可以包括芯片和其他分立器件。

第十五方面，本公开提供一种计算机程序，当其在计算机上运行时，使得计算机执行上述第一方面至第二方面的任一方面所述的方法。

附图说明

本公开上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

- 10 图 1 为本公开实施例提供的一种通信系统的架构示意图；
- 图 2 为本公开另一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 3 为本公开再一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 4a 为本公开又一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 4b 为本公开又一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 15 图 5 为本公开另一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 6 为本公开再一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 7 为本公开又一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 8 为本公开一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 9 为本公开另一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 20 图 10 为本公开再一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 11a 为本公开又一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 11b 为本公开又一个实施例所提供的 AI 模型的注册方法的流程示意图；
- 图 12 为本公开一个实施例所提供的通信装置的结构示意图；
- 图 13 为本公开另一个实施例所提供的通信装置的结构示意图；
- 25 图 14 是本公开一个实施例所提供的一种用户设备的框图；
- 图 15 为本公开一个实施例所提供的一种网络侧设备的框图。

具体实施方式

这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本公开实施例相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本公开实施例的一些方面相一致的装置和方法的例子。

在本公开实施例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实施例。在本公开实施例和所附权利要求书中所使用的单数形式的“一种”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，本文中使用的术语“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，尽管在本公开实施例可能采用术语第一、第二、第三等来描述各种信息，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，在不脱离本公开实施例范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”及“若”可以被解释成为“在……时”或“当……时”或“响应于确定”。

下面详细描述本公开的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的要素。下面通过参考附图描述的实施例是示例性的，旨在用于解释本公开，而不能理解为对本公开的限制。

为了便于理解，首先介绍本申请涉及的术语。

- 45 1、人工智能（Artificial Intelligence, AI）

AI是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。

本公开实施例中涉及到的各种网元/功能，其既可以是一个独立的硬件设备，也可以是在硬件设备内的通过计算机代码实现的功能，本公开实施例中并不对此做出限定。

5 请参见图 1，图 1 为本公开实施例提供的一种通信系统的架构示意图。该通信系统可包括但不限于一个网络设备和一个终端设备，图 1 所示的设备数量和形态仅用于举例并不构成对本公开实施例的限定，实际应用中可以包括两个或两个以上的网络设备，两个或两个以上的终端设备。图 1 所示的通信系统以包括一个网络设备 11、一个终端设备 12 为例。

需要说明的是，本公开实施例的技术方案可以应用于各种通信系统。例如：长期演进（long term evolution, LTE）系统、第五代（5th generation, 5G）移动通信系统、5G 新空口（new radio, NR）系统，或者其他未来的新型移动通信系统等。

15 本公开实施例中的网络设备 11 是网络侧的一种用于发射或接收信号的实体。例如，网络设备 11 可以为演进型基站（evolved NodeB, eNB）、发送接收点（transmission reception point, TRP）、NR 系统中的下一代基站（next generation NodeB, gNB）、其他未来移动通信系统中的基站或无线保真（wireless fidelity, WiFi）系统中的接入节点等。本公开的实施例对网络设备所采用的具体技术和具体设备形态不做限定。本公开实施例提供的网络设备可以由集中单元（central unit, CU）与分布式单元（distributed unit, DU）组成的，其中，CU 也可以称为控制单元（control unit），采用 CU-DU 的结构可以将网络设备，例如基站的协议层拆分开，部分协议层的功能放在 CU 集中控制，剩下部分或全部协议层的功能分布在 DU 中，由 CU 集中控制 DU。

20 本公开实施例中的终端设备 12 是用户侧的一种用于接收或发射信号的实体，如手机。终端设备也可以称为终端设备（terminal）、用户设备（user equipment, UE）、移动台（mobile station, MS）、移动终端设备（mobile terminal, MT）等。终端设备可以是具备通信功能的汽车、智能汽车、手机（mobile phone）、穿戴式设备、平板电脑（Pad）、带无线收发功能的电脑、虚拟现实（virtual reality, VR）终端设备、增强现实（augmented reality, AR）终端设备、工业控制（industrial control）中的无线终端设备、无人驾驶（self-driving）中的无线终端设备、远程手术（remote medical surgery）中的无线终端设备、智能电网（smart grid）中的无线终端设备、运输安全（transportation safety）中的无线终端设备、智慧城市（smart city）中的无线终端设备、智慧家庭（smart home）中的无线终端设备等等。本公开的实施例对终端设备所采用的具体技术和具体设备形态不做限定。

30 可以理解的是，本公开实施例描述的通信系统是为了更加清楚的说明本公开实施例的技术方案，并不构成对于本公开实施例提供的技术方案的限制，本领域普通技术人员可知，随着系统架构的演变和新业务场景的出现，本公开实施例提供的技术方案对于类似的技术问题，同样适用。

下面参考附图对本公开实施例所提供的 AI 模型的注册方法/装置/设备及存储介质进行详细描述。

图 2 为本公开另一个实施例所提供的一种 AI 模型的注册方法的流程示意图，其中，该方法由终端设备执行，如图 2 所示，该 AI 模型的注册方法可以包括以下步骤：

35 步骤 201、向网络侧设备发送终端设备部署的 AI 模型对应的 AI 模型信息。

其中，在本公开的一个实施例之中，上述 AI 模型信息可以包括 AI 模型对应的标识。

具体地，在本公开的一个实施例之中，可以预定义至少一个 AI 模型，每个预定义的 AI 模型均有对应的预定义的模型标识，不同的 AI 模型可能对应不同的 AI 模型信息，终端设备可以通过向网络侧设备发送所部署的 AI 模型对应的标识，使得网络侧设备可以通过接收到的 AI 模型的标识确定对应的预定义 AI 模型，并基于预定义 AI 模型确定对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册。其中，在本公开的一个实施例之中，上述 AI 模型对应的模型标识可以由终端设备产生。例如，可以通过终端服务器与网络侧设备协商预定义一系列的 AI 模型或者基于协议标准预定义一系列的 AI 模型，并分配给每个 AI 模型一个模型标识，终端通过上报模型标识来上报 AI 模型信息。在一些实施例中，响应于终端向网络侧设备发送该终端设备的 AI 模型的模型信息，网络设备会为终端设备上报的 AI 模型分配模型 ID（Identity Document，序号）。在一些实施中，终端设备也会向网络侧设备上报一个 AI 模型的模型标识，此时网络侧设备可以不对该 AI 模型进行模型 ID 的分配，也可以为该 AI 模型分配另外一个模型 ID。

其中，在本公开的一个实施例之中，可以通过终端设备与网络侧设备预定义至少一个 AI 模型。在本公开的另一个实施例之中，可以基于协议标准预定义至少一个 AI 模型。

以及，在本公开的一个实施例之中，上述预定义 AI 模型的模型信息可以包括以下信息中的至少一种：

- AI 模型的功能信息；
- AI 模型的应用范围信息；
- 5 AI 模型的输入或输出信息；
- AI 模型能达到的性能信息；
- AI 模型的结构信息。

需要说明的是，在本公开的一个实施例之中，上述网络侧设备可以为基站或核心网节点。

综上所述，在本公开实施例提供的 AI 模型的注册方法之中，终端设备会向网络侧设备发送终端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

图 3 为本公开再一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由终端设备执行，如图 3 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 301、向网络侧设备发送 AI 模型的描述信息。

其中，在本公开的一个实施例之中，上述 AI 模型的描述信息可以包括以下的至少一种：

- AI 模型的功能信息；
- AI 模型的应用范围信息；
- 20 AI 模型的输入或输出信息；
- AI 模型能达到的性能信息；
- AI 模型的结构信息。

以及，在本公开的一个实施例之中，终端设备可以通过向网络侧设备发送 AI 模型的描述信息，以使得终端设备完成该终端设备所部署 AI 模型在网络侧设备上的注册，从而使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而使得网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

综上所述，在本公开实施例提供的 AI 模型的注册方法之中，终端设备会向网络侧设备发送终端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

图 4a 为本公开又一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由终端设备执行，如图 4a 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 401a、响应于终端设备部署了 AI 模型，终端设备主动向网络侧设备上部署的 AI 模型对应的 AI 模型信息。

其中，在本公开的一个实施例之中，若终端设备部署了 AI 模型，此时终端设备可以主动向网络侧设备发送终端设备所部署的 AI 模型对应的 AI 模型信息，以使得终端设备完成该终端设备所部署 AI 模型在网络侧设备上的注册，从而使得网络侧设备知晓终端设备部署 AI 模型的 AI 模型信息。

关于 AI 模型信息的详细介绍可以参考上述实施例中的相关介绍，本公开实施例在此不做赘述。

综上所述，在本公开实施例提供的 AI 模型的注册方法之中，终端设备会向网络侧设备发送终端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

图 4b 为本公开又一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由终端设备

执行，如图 4b 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 401b、接收到网络侧设备发送的上报消息。

步骤 402b、向网络侧设备上部署的 AI 模型对应的 AI 模型信息。

其中，在本公开的一个实施例之中，响应于终端设备接收到网络侧设备发送的上报消息，终端设备
5 会向网络侧设备上部署的 AI 模型对应的 AI 模型信息。

具体地，在本公开的一个实施例之中，若终端设备接收到网络侧设备发送的上报消息，此时终端设备
会向网络侧设备上部署的 AI 模型对应的 AI 模型信息，以使得终端设备完成该终端设备
所部署 AI 模型在网络侧设备上的注册，从而使得网络侧设备知晓终端设备部署 AI 模型的 AI 模型信息，
进而使得网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

10 关于 AI 模型信息的详细介绍可以参考上述实施例中的相关介绍，本公开实施例在此不做赘述。

综上所述，在本公开实施例提供的 AI 模型的注册方法之中，终端设备会向网络侧设备发送终端设备
所部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备
所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得
网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的
15 AI 模型进行管理或监测。

图 5 为本公开另一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由终端设备执
行，如图 5 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 501、接收网络侧设备发送的网络侧设备所支持的 AI 操作。

示例的，在本公开的一个实施例之中，终端设备接收到网络侧设备发送的网络侧设备所支持的 AI
20 操作包括基于 AI 的 CSI (Channel State Information, 信道状态信息) 增强和基于 AI 的波束管理。

步骤 502、根据网络侧设备所支持的 AI 操作，确定向网络侧设备上部署该网络侧设备所支持的 AI
操作所对应的 AI 模型。

其中，在本公开的一个实施例之中，终端设备通过步骤 501 接收网络侧设备发送的网络侧设备所支
持的 AI 操作后，可以根据该网络侧设备所支持的 AI 操作确定终端设备部署的 AI 模型中支持该 AI 操
作的 AI 模型，并向网络侧设备上部署支持该 AI 操作所对应的 AI 模型的 AI 模型信息。
25

示例的，在本公开的一个实施例之中，终端设备接收到网络侧设备发送的网络侧设备所支持的 AI
操作包括基于 AI 的 CSI 增强和基于 AI 的波束管理，则终端设备向网络侧设备上部署该终端设备部署
的 AI 模型中支持进行 AI-based CSI 和 AI-based 波束管理的 AI 模型对应的 AI 模型信息。

关于实施例其它内容的详细介绍可以参考上述实施例中的相关介绍，本公开实施例在此不做赘述。

30 综上所述，在本公开实施例提供的 AI 模型的注册方法之中，终端设备会向网络侧设备发送终端设备
所部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备
所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得
网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的
AI 模型进行管理或监测。

35 图 6 为本公开再一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由终端设备执
行，如图 6 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 601、接收网络侧设备根据 AI 模型信息分配的模型标识。

其中，在本公开的一个实施例之中，当终端设备向网络侧设备发送终端设备所部署的 AI 模型对
应的 AI 模型信息后，终端设备可以接收网络侧设备根据 AI 模型信息分配的模型标识。其中，在本公开
40 的一个实施例之中，上述模型标识可以为模型 ID。

以及，在本公开的一个实施例之中，当终端设备向网络侧设备发送终端设备所部署的 AI 模型对
应的 AI 模型信息包括 AI 模型对应的标识时，此时终端设备可能接收不到网络侧设备分配的模型标识，
或者终端设备接收到网络侧设备为该 AI 模型分配的另一个不同的模型标识，以进行区分。

关于实施例其它内容的详细介绍可以参考上述实施例中的相关介绍，本公开实施例在此不做赘述。

45 综上所述，在本公开实施例提供的 AI 模型的注册方法之中，终端设备会向网络侧设备发送终端设

备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

5 图 7 为本公开又一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由网络侧设备执行，如图 7 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 701、接收终端设备发送的终端设备部署的 AI 模型对应的 AI 模型信息。

其中，在本公开的一个实施例之中，上述 AI 模型信息可以包括 AI 模型对应的标识。

10 具体地，在本公开的一个实施例之中，可以预定义至少一个 AI 模型，每个预定义的 AI 模型均有对应的预定义的模型标识，不同的 AI 模型可能对应不同的 AI 模型信息，网络侧设备可以通过接收终端设备发送的该终端设备所部署的 AI 模型对应的标识，确定对应的预定义 AI 模型，并基于预定义 AI 模型确定对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册。其中，在本公开的一个实施例之中，上述 AI 模型对应的标识可以由终端设备产生。

15 其中，在本公开的一个实施例之中，可以通过终端设备与网络侧设备预定义至少一个 AI 模型。在本公开的另一个实施例之中，可以基于协议标准预定义至少一个 AI 模型。

以及，在本公开的一个实施例之中，上述预定义 AI 模型的模型信息可以包括以下信息中的至少一种：

AI 模型的功能信息；

AI 模型的应用范围信息；

20 AI 模型的输入或输出信息；

AI 模型能达到的性能信息；

AI 模型的结构信息。

需要说明的是，在本公开的一个实施例之中，上述网络侧设备可以为基站或核心网节点。

25 关于本实施例中其它内容的详细介绍可以参考上述实施例中的相关介绍，本公开实施例在此不做赘述。

综上所述，在本公开实施例提供的 AI 模型的注册方法之中，网络侧设备会接收终端设备发送的终端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

30 图 8 为本公开一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由网络侧设备执行，如图 8 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 801、接收终端设备发送的 AI 模型的结构信息。

其中，在本公开的一个实施例之中，上述 AI 模型的结构信息可以包括以下的至少一种：

35 AI 模型的功能信息；

AI 模型的应用范围信息；

AI 模型的输入或输出信息；

AI 模型能达到的性能信息；

AI 模型的结构信息。

40 以及，在本公开的一个实施例之中，网络侧设备可以通过接收终端设备发送的 AI 模型的结构信息知晓终端设备部署的 AI 模型的模型信息，从而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

关于本实施例中其它内容的详细介绍可以参考上述实施例中的相关介绍，本公开的实施例在此不做赘述。

45 综上所述，在本公开实施例提供的 AI 模型的注册方法之中，网络侧设备会接收终端设备发送的终

端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

5 图 9 为本公开另一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由网络侧设备执行，如图 9 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 901、向终端设备发送上报消息，上报消息用于指示终端设备上报 AI 模型对应的 AI 模型信息。

其中，在本公开的一个实施例之中，网络侧设备可以通过向终端发送上报消息，以指示终端设备上报 AI 模型对应的 AI 模型信息。以及，当终端设备接收到网络侧设备发送的上报消息后，终端设备会向网络侧设备发送终端设备部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

关于本实施例中其它内容的详细介绍可以参考上述实施例中的相关介绍，本公开的实施例在此不做赘述。

15 综上所述，在本公开实施例提供的 AI 模型的注册方法之中，网络侧设备会接收终端设备发送的终端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

20 图 10 为本公开再一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由网络侧设备执行，如图 10 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 1001、向终端设备发送网络侧设备所支持的 AI 操作。

示例的，在本公开的一个实施例之中，网络侧设备向终端设备发送的网络侧设备所支持的 AI 操作包括基于 AI 的 CSI (Channel State Information, 信道状态信息) 增强和基于 AI 的波束管理。

25 以及，在本公开的一个实施例之中，网络侧设备接收终端设备上报的 AI 模型是根据网络侧设备所支持的 AI 操作生成。

具体地，在本公开的一个实施例之中，网络侧设备向终端设备发送网络侧设备所支持的 AI 操作后，终端设备可以根据该网络侧设备所支持的 AI 操作确定终端设备部署的 AI 模型中支持该 AI 操作的 AI 模型，并向网络侧设备上报支持该 AI 操作所对应的 AI 模型的 AI 模型信息。

30 关于本实施例中其它内容的详细介绍可以参考上述实施例中的相关介绍，本公开的实施例在此不做赘述。

综上所述，在本公开实施例提供的 AI 模型的注册方法之中，网络侧设备会接收终端设备发送的终端设备部署的 AI 模型对应的 AI 模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的 AI 模型对应的 AI 模型信息，从而完成终端设备所部署 AI 模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的 AI 模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的 AI 模型进行管理或监测。

图 11a 为本公开又一个实施例所提供的一种 AI 模型的注册方法的流程示意图，该方法由网络侧设备执行，如图 11a 所示，该 AI 模型的注册方法可以包括以下步骤：

步骤 1101a、向终端设备发送根据 AI 模型信息分配的模型标识。

40 其中，在本公开的一个实施例之中，当终端设备向网络侧设备发送终端设备所部署的 AI 模型对应的 AI 模型信息后，网络侧设备可以根据终端设备发送的 AI 模型的 AI 模型信息为该 AI 模型分配模型标识。其中，在本公开的一个实施例之中，上述模型标识可以为模型 ID。

以及，在本公开的一个实施例之中，当终端设备向网络侧设备发送终端设备所部署的 AI 模型对应的 AI 模型信息包括 AI 模型对应的标识时，此时网络侧设备可以不向该 AI 模型分配模型标识，或者网络侧设备为该 AI 模型分配的另一个不同的模型标识，以进行区分。

关于本实施例中其它内容的详细介绍可以参考上述实施例中的相关介绍,本公开的实施例在此不做赘述。

综上所述,在本公开实施例提供的AI模型的注册方法之中,网络侧设备会接收终端设备发送的终端设备部署的AI模型对应的AI模型信息。也即是,本公开中终端设备通过向网络侧设备发送该终端设备部署的AI模型对应的AI模型信息,从而完成终端设备所部署AI模型在网络侧设备上的注册,使得网络侧设备知晓终端设备部署的AI模型的模型信息,进而便于网络侧设备后续对部署在终端设备上的AI模型进行管理或监测。

图11b为本公开又一个实施例所提供的一种AI模型的注册方法的流程示意图,该方法由网络侧设备执行,如图11b所示,该AI模型的注册方法可以包括以下步骤:

步骤1101b、当终端设备发生小区切换时,目标基站获取终端设备所支持AI模型的AI模型信息。

其中,在本公开的一个实施例之中,如果终端设备已经在网络侧设备完成AI模型的注册,当终端设备发生小区切换时,终端设备切换至的目标小区所对应的目标基站需要获取终端设备所支持AI模型的AI模型信息,以完成终端设备部署的AI模型在目标基站上的注册,从而不用重新与目标基站进行交互完成注册。

以及,在本公开的一个实施例之中,终端设备完成AI模型的注册的网络侧设备不同时,目标基站获取终端设备所支持AI模型的AI模型信息的方法也有所不同。

具体地,在本公开的一个实施例之中,若终端设备在源小区对应的源基站完成AI模型的注册,则当终端设备发生小区切换时,源基站可以将该终端设备所部署AI模型的AI模型信息发送至终端设备的目标小区所对应的目标基站,目标基站获取终端设备所支持AI模型的AI模型信息的方法可以包括:

目标基站接收源基站发送的终端设备所支持AI模型的AI模型信息。

以及,在本公开的另一个实施例之中,若终端设备在核心网节点完成AI模型的注册,则当终端设备发生小区切换时,目标基站获取终端设备所支持AI模型的AI模型信息的方法可以包括:目标基站从核心网节点获取终端设备所支持AI模型的AI模型信息。

关于本实施例中其它内容的详细介绍可以参考上述实施例中的相关介绍,本公开的实施例在此不做赘述。

综上所述,在本公开实施例提供的AI模型的注册方法之中,网络侧设备会接收终端设备发送的终端设备部署的AI模型对应的AI模型信息。也即是,本公开中终端设备通过向网络侧设备发送该终端设备部署的AI模型对应的AI模型信息,从而完成终端设备所部署AI模型在网络侧设备上的注册,使得网络侧设备知晓终端设备部署的AI模型的模型信息,进而便于网络侧设备后续对部署在终端设备上的AI模型进行管理或监测。

图12为本公开一个实施例所提供的一种通信装置的结构示意图,如图12所示,装置可以包括:

发送模块1201,用于向网络侧设备发送终端设备部署的AI模型对应的AI模型信息。

综上所述,在本公开实施例提供的通信装置之中,终端设备会向网络侧设备发送终端设备部署的AI模型对应的AI模型信息。也即是,本公开中终端设备通过向网络侧设备发送该终端设备所部署的AI模型对应的AI模型信息,从而完成终端设备所部署AI模型在网络侧设备上的注册,使得网络侧设备知晓终端设备部署的AI模型的模型信息,进而便于网络侧设备后续对部署在终端设备上的AI模型进行管理或监测。

可选的,在本公开的一个实施例之中,AI模型信息包括AI模型对应的标识。

可选的,在本公开的一个实施例之中,预定义至少一个AI模型,每个预定义的AI模型有对应的预定义模型标识,预定义AI模型的模型信息包括至少以下信息中的至少一种:

AI模型的功能信息;

AI模型的应用范围信息;

AI模型的输入或输出信息;

AI模型能达到的性能信息;

AI模型的结构信息。

可选的，在本公开的一个实施例之中，AI模型对应的标识由终端设备产生。

可选的，在本公开的一个实施例之中，上述装置还用于：

向网络侧设备发送AI模型的描述信息，其中，AI模型的描述信息包括以下的至少一种：

AI模型的功能信息；

5 AI模型的应用范围信息；

AI模型的输入或输出信息；

AI模型能达到的性能信息；

AI模型的结构信息。

10 可选的，在本公开的一个实施例之中，响应于终端设备部署了AI模型，终端设备主动向网络侧设备上部署的AI模型对应的AI模型信息。

可选的，在本公开的一个实施例之中，响应于接收到网络侧设备发送的上报消息，终端设备向网络侧设备上部署的AI模型对应的AI模型信息。

可选的，在本公开的一个实施例之中，上述装置还用于：

接收网络侧设备发送的网络侧设备所支持的AI操作；

15 根据网络侧设备所支持的AI操作，确定向网络侧设备上部署网络侧设备所支持的AI操作对应的AI模型。

可选的，在本公开的一个实施例之中，上述装置还用于：

接收网络侧设备根据AI模型信息分配的模型标识。

20 图13为本公开另一个实施例所提供的一种通信装置的结构示意图，如图13所示，装置可以包括：接收模块1301，用于接收终端设备发送的终端设备部署的AI模型对应的AI模型信息。

综上所述，在本公开实施例提供的通信装置之中，网络侧设备会接收终端设备发送的终端设备部署的AI模型对应的AI模型信息。也即是，本公开中终端设备通过向网络侧设备发送该终端设备所部署的AI模型对应的AI模型信息，从而完成终端设备所部署AI模型在网络侧设备上的注册，使得网络侧设备知晓终端设备部署的AI模型的模型信息，进而便于网络侧设备后续对部署在终端设备上的AI模型进行管理或监测。

25 可选的，在本公开的一个实施例之中，AI模型信息包括AI模型对应的标识。

可选的，在本公开的一个实施例之中，预定义至少一个AI模型，每个预定义的AI模型有对应的预定义模型标识，预定义AI模型的模型信息包括至少以下信息中的至少一种：

AI模型的功能信息；

30 AI模型的应用范围信息；

AI模型的输入或输出信息；

AI模型能达到的性能信息；

AI模型的结构信息。

可选的，在本公开的一个实施例之中，上述装置，还用于：

35 接收终端设备发送的AI模型的描述信息，其中，AI模型的描述信息包括以下的至少一种：

AI模型的功能信息；

AI模型的应用范围信息；

AI模型的输入或输出信息；

AI模型能达到的性能信息；

40 AI模型的结构信息。

可选的，在本公开的一个实施例之中，上述装置，还用于：

向终端设备发送上报消息，其中，上报消息用于指示终端设备上部署AI模型对应的AI模型信息。

可选的，在本公开的一个实施例之中，上述装置还用于：

45 向终端设备发送网络侧设备所支持的AI操作，其中，终端设备上部署的AI模型根据所述网络侧设备所支持的AI操作生成。

可选的，在本公开的一个实施例之中，上述装置还用于：

向终端设备发送根据 AI 模型信息分配的模型标识。

可选的，在本公开的一个实施例之中，上述装置还用于：

当终端设备发生小区切换时，目标基站获取终端设备所支持 AI 模型的 AI 模型信息。

5 可选的，在本公开的一个实施例之中，上述装置还用于：

目标基站接收源基站发送的终端设备所支持 AI 模型的 AI 模型信息。

可选的，在本公开的一个实施例之中，上述装置还用于：

目标基站从核心网节点获取终端设备所支持 AI 模型的 AI 模型信息。

10 请参见图 14，图 14 是本申请实施例提供的一种通信装置 1400 的结构示意图。通信装置 1400 可以是网络设备，也可以是终端设备，也可以是支持网络设备实现上述方法的芯片、芯片系统、或处理器等，还可以是支持终端设备实现上述方法的芯片、芯片系统、或处理器等。该装置可用于实现上述方法实施例中描述的方法，具体可以参见上述方法实施例中的说明。

15 通信装置 1400 可以包括一个或多个处理器 1401。处理器 1401 可以是通用处理器或者专用处理器等。例如可以是基带处理器或中央处理器。基带处理器可以用于对通信协议以及通信数据进行处理，中央处理器可以用于对通信装置（如，基站、基带芯片，终端设备、终端设备芯片，DU 或 CU 等）进行控制，执行计算机程序，处理计算机程序的数据。

20 可选的，通信装置 1400 中还可以包括一个或多个存储器 1402，其上可以存有计算机程序 1404，处理器 1401 执行所述计算机程序 1404，以使得通信装置 1400 执行上述方法实施例中描述的方法。可选的，所述存储器 1402 中还可以存储有数据。通信装置 1400 和存储器 1402 可以单独设置，也可以集成在一起。

可选的，通信装置 1400 还可以包括收发器 1405、天线 1406。收发器 1405 可以称为收发单元、收发机、或收发电路等，用于实现收发功能。收发器 1405 可以包括接收器和发送器，接收器可以称为接收机或接收电路等，用于实现接收功能；发送器可以称为发送机或发送电路等，用于实现发送功能。

25 可选的，通信装置 1400 中还可以包括一个或多个接口电路 1407。接口电路 1407 用于接收代码指令并传输至处理器 1401。处理器 1401 运行所述代码指令以使通信装置 1400 执行上述方法实施例中描述的方法。

30 通信装置 1400 为终端设备：收发器 1405 用于执行图 3 中的步骤 301-步骤 302；图 4 中的步骤 401 至步骤 402；图 5 中的步骤 501 至步骤 503；图 6a 中的步骤 601a 至步骤 603a；图 6b 中的步骤 601b 至步骤 602b。处理器 1401 用于执行图 2 中的步骤 201；图 3 中的步骤 303-步骤 304；图 4 中的步骤 403；图 5 中的步骤 504 和步骤 505；图 6a 中的步骤 604a；图 6b 中的步骤 603b；图 7 中的步骤 701-步骤 702。

通信装置 1400 为网络设备：收发器 1405 用于执行图 11a 中的步骤 1102a。处理器 1401 用于执行图 8 中的步骤 801；图 9 中的步骤 901 至步骤 904；图 10 中的步骤 1001 至步骤 1005；图 11a 中的步骤 1101a；图 11b 中的步骤 1101b 和步骤 1102b。

35 在一种实现方式中，处理器 1401 中可以包括用于实现接收和发送功能的收发器。例如该收发器可以是收发电路，或者是接口，或者是接口电路。用于实现接收和发送功能的收发电路、接口或接口电路可以是分开的，也可以集成在一起。上述收发电路、接口或接口电路可以用于代码/数据的读写，或者，上述收发电路、接口或接口电路可以用于信号的传输或传递。

40 在一种实现方式中，处理器 1401 可以存有计算机程序 1403，计算机程序 1403 在处理器 1401 上运行，可使得通信装置 1400 执行上述方法实施例中描述的方法。计算机程序 1403 可能固化在处理器 1401 中，该种情况下，处理器 1401 可能由硬件实现。

45 在一种实现方式中，通信装置 1400 可以包括电路，所述电路可以实现前述方法实施例中发送或接收或者通信的功能。本申请中描述的处理器和收发器可实现在集成电路（integrated circuit, IC）、模拟 IC、射频集成电路 RFIC、混合信号 IC、专用集成电路（application specific integrated circuit, ASIC）、印刷电路板（printed circuit board, PCB）、电子设备等上。该处理器和收发器也可以用各种 IC 工艺技术来制造，例如互补金属氧化物半导体（complementary metal oxide semiconductor, CMOS）、N 型金属

氧化物半导体(nMetal-oxide-semiconductor, NMOS)、P型金属氧化物半导体(positive channel metal oxide semiconductor, PMOS)、双极结型晶体管(bipolar junction transistor, BJT)、双极CMOS(BiCMOS)、硅锗(SiGe)、砷化镓(GaAs)等。

5 以上实施例描述中的通信装置可以是网络设备或者终端设备,但本申请中描述的通信装置的范围并不限于此,而且通信装置的结构可以不受图14的限制。通信装置可以是独立的设备或者可以是较大设备的一部分。例如所述通信装置可以是:

(1) 独立的集成电路IC,或芯片,或,芯片系统或子系统;

(2) 具有一个或多个IC的集合,可选的,该IC集合也可以包括用于存储数据,计算机程序的存储部件;

10 (3) ASIC,例如调制解调器(Modem);

(4) 可嵌入在其他设备内的模块;

(5) 接收机、终端设备、智能终端设备、蜂窝电话、无线设备、手持机、移动单元、车载设备、网络设备、云设备、人工智能设备等等;

(6) 其他等等。

15 对于通信装置可以是芯片或芯片系统的情况,可参见图15所示的芯片的结构示意图。图15所示的芯片包括处理器1501和接口1502。其中,处理器1501的数量可以是一个或多个,接口1502的数量可以是多个。

可选的,芯片还包括存储器1503,存储器1503用于存储必要的计算机程序和数据。

本领域技术人员还可以了解到本申请实施例列出的各种说明性逻辑块(illustrative logical block)和步骤(step)可以通过电子硬件、电脑软件,或两者的结合进行实现。这样的功能是通过硬件还是软件来实现取决于特定的应用和整个系统的设计要求。本领域技术人员可以对于每种特定的应用,可以使用各种方法实现所述的功能,但这种实现不应被理解为超出本申请实施例保护的范围。

本申请还提供一种可读存储介质,其上存储有指令,该指令被计算机执行时实现上述任一方法实施例的功能。

25 本申请还提供一种计算机程序产品,该计算机程序产品被计算机执行时实现上述任一方法实施例的功能。

在上述实施例中,可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机程序。在计算机上加载和执行所述计算机程序时,全部或部分地产生按照本申请实施例所述的流程或功能。30 所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机程序可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机程序可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(digital subscriber line, DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用35 介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质(例如,软盘、硬盘、磁带)、光介质(例如,高密度数字视频光盘(digital video disc, DVD))、或者半导体介质(例如,固态硬盘(solid state disk, SSD))等。

本领域普通技术人员可以理解:本申请中涉及的第一、第二等各种数字编号仅为描述方便进行的区分,并不用来限制本申请实施例的范围,也表示先后顺序。

40 本申请中的至少一个还可以描述为一个或多个,多个可以是两个、三个、四个或者更多个,本申请不做限制。在本申请实施例中,对于一种技术特征,通过“第一”、“第二”、“第三”、“A”、“B”、“C”和“D”等区分该种技术特征中的技术特征,该“第一”、“第二”、“第三”、“A”、“B”、“C”和“D”描述的技术特征间无先后顺序或者大小顺序。

本申请中各表所示的对应关系可以被配置,也可以是预定义的。各表中的信息的取值仅仅是举例,45 可以配置为其他值,本申请并不限定。在配置信息与各参数的对应关系时,并不一定要求必须配置各表

5 中示意出的所有对应关系。例如，本申请中的表格中，某些行示出的对应关系也可以不配置。又例如，可以基于上述表格做适当的变形调整，例如，拆分，合并等等。上述各表中标题示出参数的名称也可以采用通信装置可理解的其他名称，其参数的取值或表示方式也可以通信装置可理解的其他取值或表示方式。上述各表在实现时，也可以采用其他的数据结构，例如可以采用数组、队列、容器、栈、线性表、指针、链表、树、图、结构体、类、堆、散列表或哈希表等。

本申请中的预定义可以理解为定义、预先定义、存储、预存储、预协商、预配置、固化、或预烧制。

10 本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。

15 因此，本申请的保护范围应以所述权利要求的保护范围为准。

权利要求书

- 1、一种人工智能 AI 模型的注册方法，其特征在于，应用于终端设备，所述方法包括：
向网络侧设备发送所述终端设备部署的 AI 模型对应的 AI 模型信息。
- 2、如权利要求 1 所述的方法，其特征在于，所述 AI 模型信息包括所述 AI 模型对应的标识。
- 5 3、如权利要求 2 所述的方法，预定义至少一个 AI 模型，每个预定义的 AI 模型有对应的预定义模型标识，所述预定义 AI 模型的模型信息包括至少以下信息中的至少一种：
所述 AI 模型的功能信息；
所述 AI 模型的应用范围信息；
所述 AI 模型的输入或输出信息；
10 所述 AI 模型能达到的性能信息；
所述 AI 模型的结构信息。
- 4、如权利要求 2 所述方法，其特征在于，所述 AI 模型对应的标识由终端设备产生。
- 5、如权利要求 1 所述的方法，其特征在于，还包括：
向所述网络侧设备发送所述 AI 模型的描述信息，其中，所述 AI 模型的描述信息包括以下的至少
15 一种：
所述 AI 模型的功能信息；
所述 AI 模型的应用范围信息；
所述 AI 模型的输入或输出信息；
所述 AI 模型能达到的性能信息；
20 所述 AI 模型的结构信息。
- 6、如权利要求 1 所述的方法，其特征在于，响应于所述终端设备部署了 AI 模型，所述终端设备主动向所述网络侧设备上部署的 AI 模型对应的 AI 模型信息。
- 7、如权利要求 1 所述的方法，其特征在于，响应于接收到所述网络侧设备发送的上报消息，所述终端设备向所述网络侧设备上部署的 AI 模型对应的 AI 模型信息。
- 25 8、如权利要求 1 所述的方法，其特征在于，还包括：
接收所述网络侧设备发送的所述网络侧设备所支持的 AI 操作；
根据所述网络侧设备所支持的 AI 操作，确定向所述网络侧设备上部署所述网络侧设备所支持的 AI 操作对应的 AI 模型。
- 9、如权利要求 1 或 8 所述的方法，其特征在于，还包括：
30 接收所述网络侧设备根据所述 AI 模型信息分配的模型标识。
- 10、一种人工智能 AI 模型的注册方法，其特征在于，应用于网络侧设备，所述方法包括：
接收终端设备发送的所述终端设备部署的 AI 模型对应的 AI 模型信息。
- 11、如权利要求 10 所述的方法，其特征在于，所述 AI 模型信息包括所述 AI 模型对应的标识。
- 12、如权利要求 11 所述的方法，预定义至少一个 AI 模型，每个预定义的 AI 模型有对应的预定义模型标识，所述预定义 AI 模型的模型信息包括至少以下信息中的至少一种：
35 所述 AI 模型的功能信息；
所述 AI 模型的应用范围信息；
所述 AI 模型的输入或输出信息；
所述 AI 模型能达到的性能信息；
40 所述 AI 模型的结构信息。
- 13、如权利要求 10 所述的方法，其特征在于，还包括：
接收所述终端设备发送的所述 AI 模型的描述信息，其中，所述 AI 模型的描述信息包括以下的至少一种：
45 所述 AI 模型的功能信息；
所述 AI 模型的应用范围信息；

所述 AI 模型的输入或输出信息；

所述 AI 模型能达到的性能信息；

所述 AI 模型的结构信息。

14、如权利要求 10 所述的方法，其特征在于，还包括：

5 向所述终端设备发送上报消息，其中，所述上报消息用于指示所述终端设备上报所述 AI 模型对应的 AI 模型信息。

15、如权利要求 10 所述的方法，其特征在于，还包括：

向所述终端设备发送所述网络侧设备所支持的 AI 操作，其中，所述终端设备上报的 AI 模型根据所述网络侧设备所支持的 AI 操作生成。

10 16、如权利要求 10 或 15 所述的方法，其特征在于，还包括：

向所述终端设备发送根据所述 AI 模型信息分配的模型标识。

17、如权利要求 10 所述的方法，其特征在于，还包括：

当所述终端设备发生小区切换时，目标基站获取所述终端设备所支持 AI 模型的 AI 模型信息。

15 18、如权利要求 17 所述的方法，其特征在于，所述目标基站获取所述终端设备所支持 AI 模型的 AI 模型信息，包括：

目标基站接收源基站发送的所述终端设备所支持 AI 模型的 AI 模型信息。

19、如权利要求 17 所述的方法，其特征在于，所述目标基站获取所述终端设备所支持 AI 模型的 AI 模型信息，包括：

目标基站从核心网节点获取所述终端设备所支持 AI 模型的 AI 模型信息。

20 20、一种通信装置，被配置在终端设备中，包括：

发送模块，用于向网络侧设备发送所述终端设备部署的 AI 模型对应的 AI 模型信息。

21、一种通信装置，被配置在网络侧设备中，包括：

接收模块，用于接收终端设备发送的所述终端设备部署的 AI 模型对应的 AI 模型信息。

25 22、一种通信装置，其特征在于，所述装置包括处理器和存储器，其中，所述存储器中存储有计算机程序，所述处理器执行所述存储器中存储的计算机程序，以使所述装置执行如权利要求 1 至 9 中任一项所述的方法，或所述处理器执行所述存储器中存储的计算机程序，以使所述装置执行如权利要求 10 至 19 中任一项所述的方法。

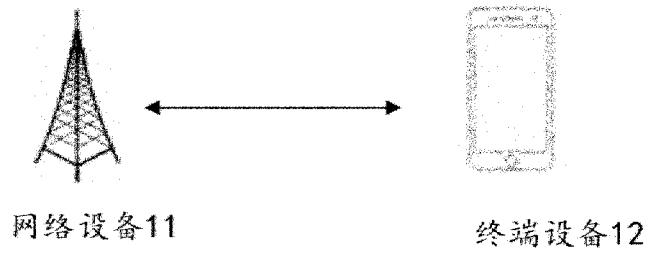


图 1

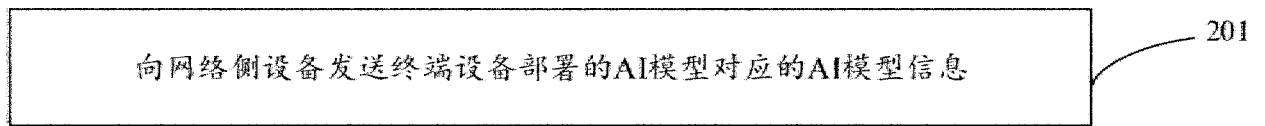


图 2

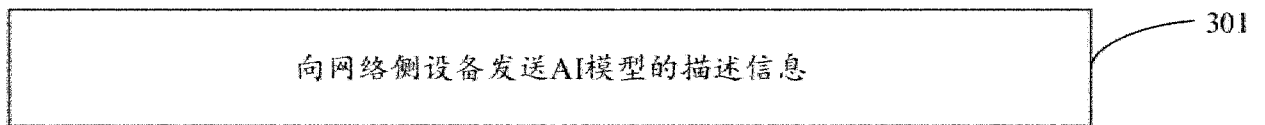


图 3

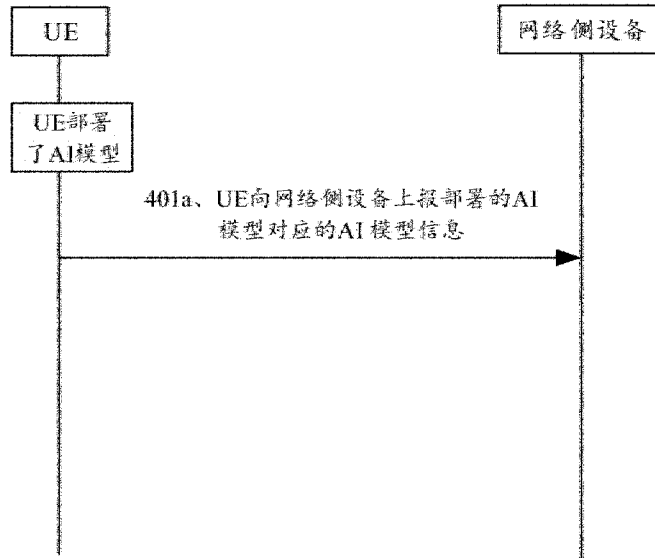


图 4a

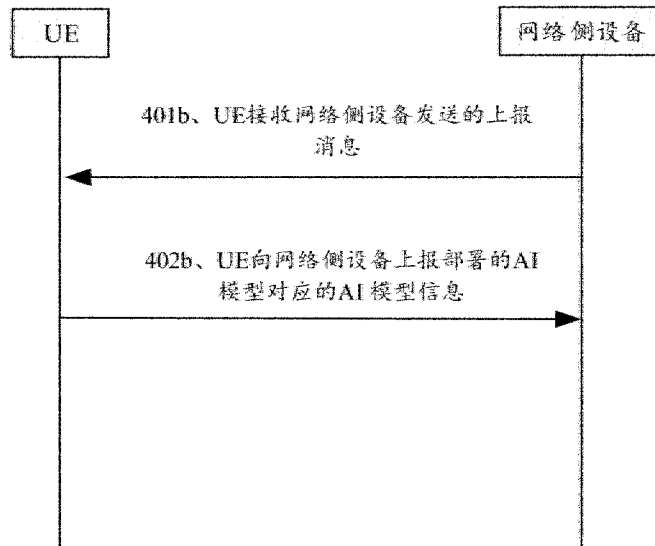


图 4b

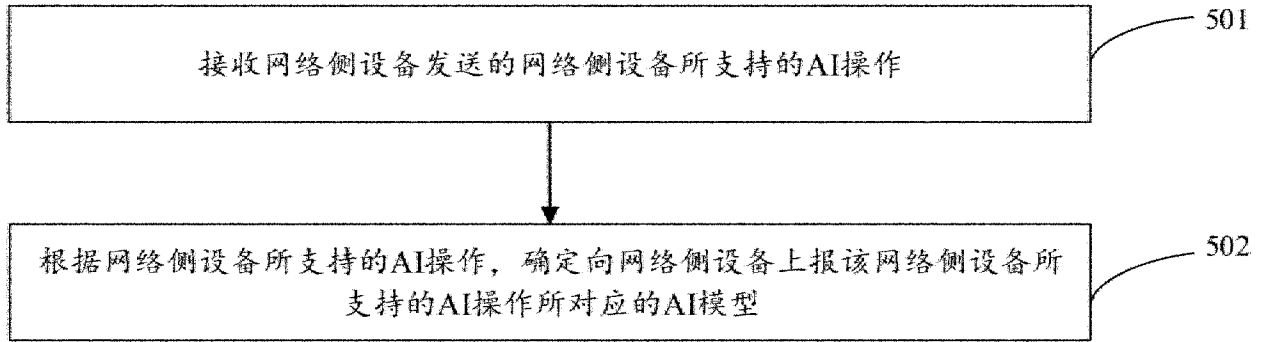


图 5

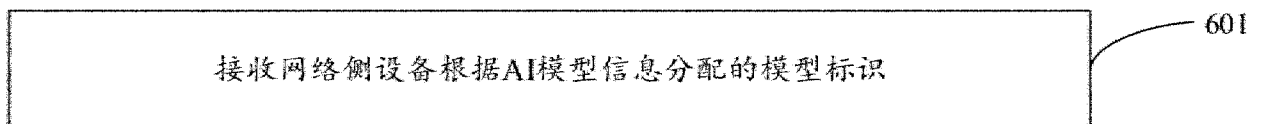


图 6

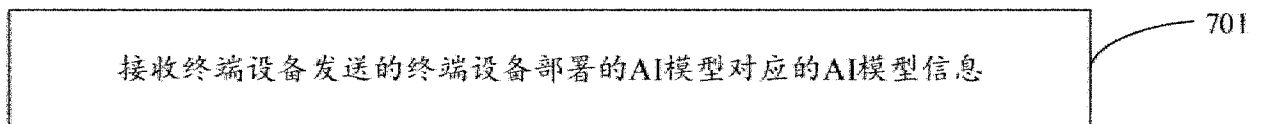


图 7

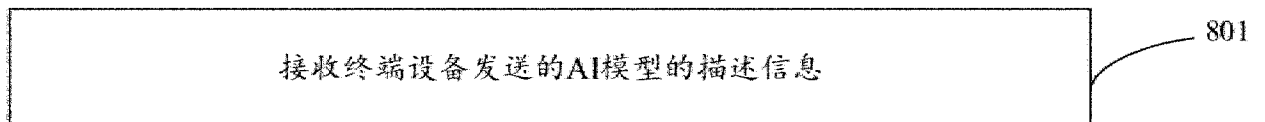


图 8

向终端设备发送上报消息，上报消息用于指示终端设备上报AI模型对应的AI模型信息 901

图 9

向终端设备发送网络侧设备所支持的AI操作 1001

图 10

向终端设备发送根据AI模型信息分配的模型标识 1101a

图 11a

当终端设备发生小区切换时，目标基站获取终端设备所支持AI模型的AI模型信息 1101b

图 11b

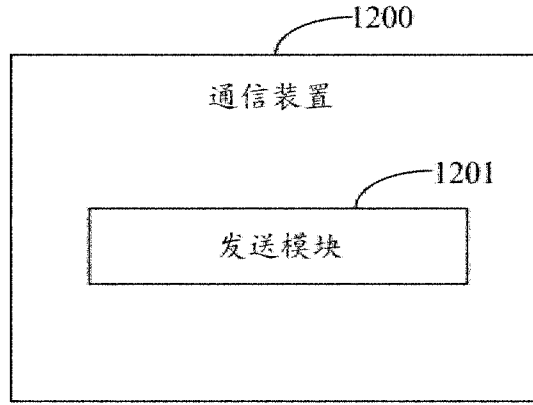


图 12

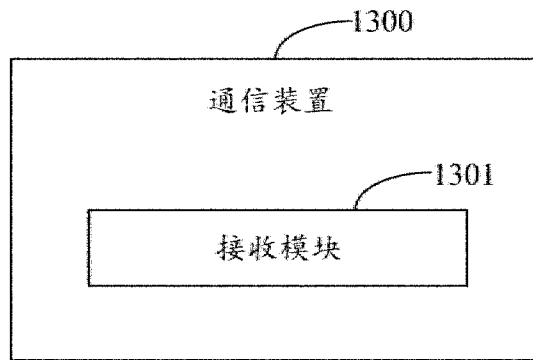


图 13

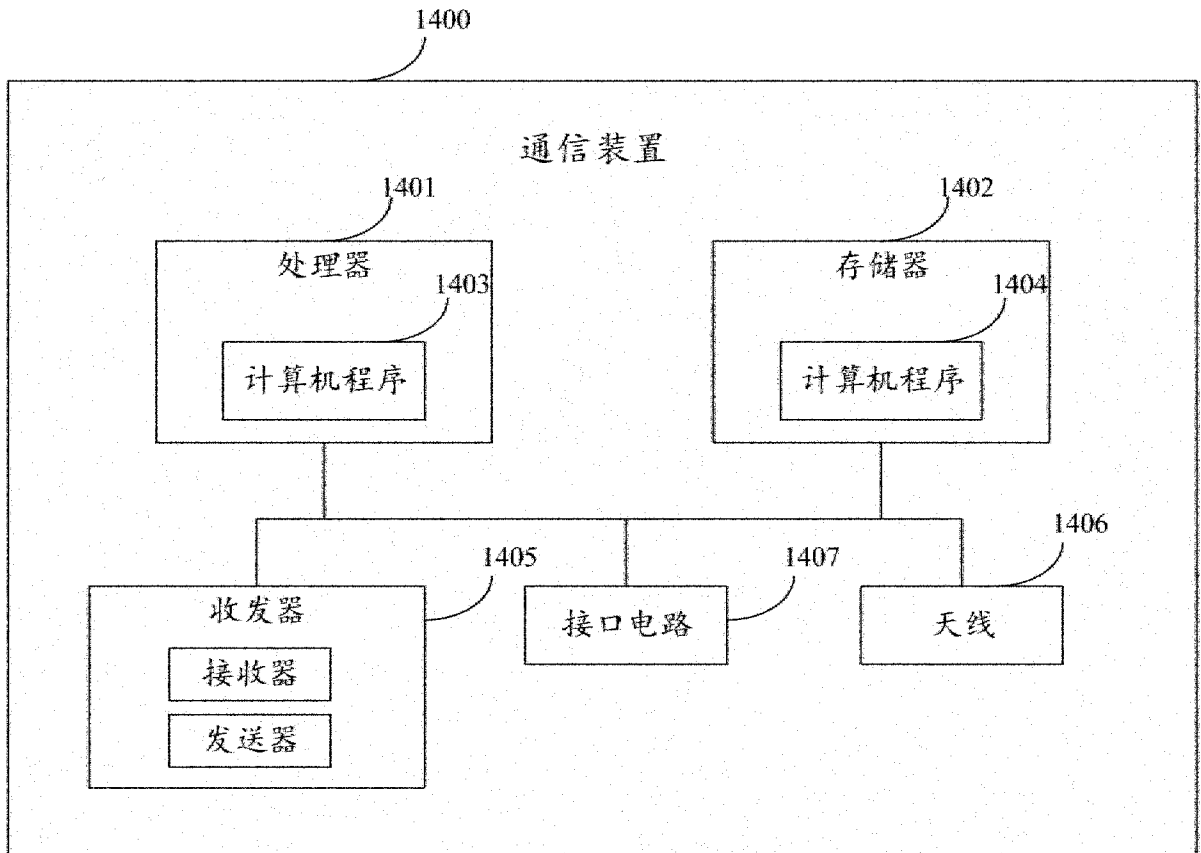


图 14

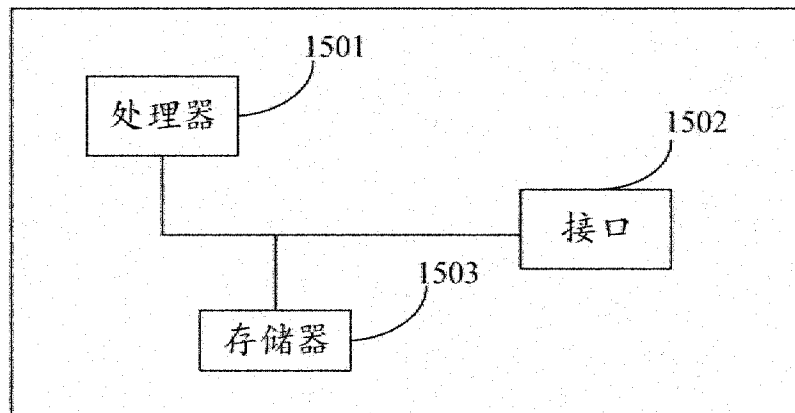


图 15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/076988

A. CLASSIFICATION OF SUBJECT MATTER		
H04L 67/54(2022.01)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: H04W		
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)		
CNKI; CJFD; ENTXT; ENTXTC; CNTXT; DWPI: 终端, UE, 发送, 传输, 配置, AI, 人工智能, 机器学习, 深度学习, 神经网络, 模型, config+, learning, machine, deep, model, neural, network, transmit		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021163895 A1 (GUANGDONG OPPO MOBILE COMMUNICATIONS CO., LTD.) 26 August 2021 (2021-08-26) description, line 3, line 11 to page 10, line 50, and figures 3-8	1-22
X	CN 111837425 A (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 27 October 2020 (2020-10-27) claims 1-26, and description, paragraphs 96-143	1-22
A	CN 114071484 A (HUAWEI TECHNOLOGIES CO., LTD.) 18 February 2022 (2022-02-18) entire document	1-22
A	CN 115392332 A (GONGDADI INNOVATION TECHNOLOGY (SHENZHEN) CO., LTD.) 25 November 2022 (2022-11-25) entire document	1-22
A	US 2022225126 A1 (SAMSUNG ELECTRONICS CO., LTD.) 14 July 2022 (2022-07-14) entire document	1-22
A	WO 2022007579 A1 (SPREADTRUM COMMUNICATIONS (SHANGHAI) CO., LTD.) 13 January 2022 (2022-01-13) entire document	1-22
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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
17 October 2023		23 October 2023
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

International application No.

PCT/CN2023/076988

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2022236840 A1 (GUANGDONG OPPO MOBILE COMMUNICATIONS CO., LTD.) 17 November 2022 (2022-11-17) entire document	1-22

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/076988

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2021163895	A1	26 August 2021	CN	114787793	A	22 July 2022
CN	111837425	A	27 October 2020	WO	2021248371	A1	16 December 2021
				US	2023217366	A1	06 July 2023
				EP	4167645	A1	19 April 2023
				IN	202247074647	A	30 December 2022
CN	114071484	A	18 February 2022	WO	2022022334	A1	03 February 2022
				US	2023179490	A1	08 June 2023
				EP	4181556	A1	17 May 2023
CN	115392332	A	25 November 2022	None			
US	2022225126	A1	14 July 2022	WO	2022154505	A1	21 July 2022
				CN	114765789	A	19 July 2022
WO	2022007579	A1	13 January 2022	EP	4181572	A1	17 May 2023
				US	2023247503	A1	03 August 2023
				CN	113923730	A	11 January 2022
WO	2022236840	A1	17 November 2022	None			

<p>A. 主题的分类</p> <p>H04L 67/54(2022.01)i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																										
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>IPC: H04W</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNKI;CJFD;ENTXT;ENTXTC;CNTXT;DWPI:终端, UE, 发送, 传输, 配置, AI, 人工智能, 机器学习, 深度学习, 神经网络, 模型, config+, learning, machine, deep, model, neural, network, transmit</p>																										
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>WO 2021163895 A1 (OPPO广东移动通信有限公司) 2021年8月26日 (2021 - 08 - 26) 说明书第3行第11行至第10页第50行, 附图3-8</td> <td>1-22</td> </tr> <tr> <td>X</td> <td>CN 111837425 A (北京小米移动软件有限公司) 2020年10月27日 (2020 - 10 - 27) 权利要求1-26, 说明书第96-143</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>CN 114071484 A (华为技术有限公司) 2022年2月18日 (2022 - 02 - 18) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>CN 115392332 A (共达地创新技术(深圳)有限公司) 2022年11月25日 (2022 - 11 - 25) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>US 2022225126 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2022年7月14日 (2022 - 07 - 14) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>WO 2022007579 A1 (展讯通信(上海)有限公司) 2022年1月13日 (2022 - 01 - 13) 全文</td> <td>1-22</td> </tr> <tr> <td>A</td> <td>WO 2022236840 A1 (OPPO广东移动通信有限公司) 2022年11月17日 (2022 - 11 - 17) 全文</td> <td>1-22</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “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</p>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	WO 2021163895 A1 (OPPO广东移动通信有限公司) 2021年8月26日 (2021 - 08 - 26) 说明书第3行第11行至第10页第50行, 附图3-8	1-22	X	CN 111837425 A (北京小米移动软件有限公司) 2020年10月27日 (2020 - 10 - 27) 权利要求1-26, 说明书第96-143	1-22	A	CN 114071484 A (华为技术有限公司) 2022年2月18日 (2022 - 02 - 18) 全文	1-22	A	CN 115392332 A (共达地创新技术(深圳)有限公司) 2022年11月25日 (2022 - 11 - 25) 全文	1-22	A	US 2022225126 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2022年7月14日 (2022 - 07 - 14) 全文	1-22	A	WO 2022007579 A1 (展讯通信(上海)有限公司) 2022年1月13日 (2022 - 01 - 13) 全文	1-22	A	WO 2022236840 A1 (OPPO广东移动通信有限公司) 2022年11月17日 (2022 - 11 - 17) 全文	1-22
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
X	WO 2021163895 A1 (OPPO广东移动通信有限公司) 2021年8月26日 (2021 - 08 - 26) 说明书第3行第11行至第10页第50行, 附图3-8	1-22																								
X	CN 111837425 A (北京小米移动软件有限公司) 2020年10月27日 (2020 - 10 - 27) 权利要求1-26, 说明书第96-143	1-22																								
A	CN 114071484 A (华为技术有限公司) 2022年2月18日 (2022 - 02 - 18) 全文	1-22																								
A	CN 115392332 A (共达地创新技术(深圳)有限公司) 2022年11月25日 (2022 - 11 - 25) 全文	1-22																								
A	US 2022225126 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2022年7月14日 (2022 - 07 - 14) 全文	1-22																								
A	WO 2022007579 A1 (展讯通信(上海)有限公司) 2022年1月13日 (2022 - 01 - 13) 全文	1-22																								
A	WO 2022236840 A1 (OPPO广东移动通信有限公司) 2022年11月17日 (2022 - 11 - 17) 全文	1-22																								
国际检索实际完成的日期	2023年10月17日	国际检索报告邮寄日期	2023年10月23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	授权官员	熊金安 电话号码 (+86) 010-53961789																							

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/076988

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)
WO	2021163895	A1	2021年8月26日	CN	114787793	A	2022年7月22日
CN	111837425	A	2020年10月27日	WO	2021248371	A1	2021年12月16日
				US	2023217366	A1	2023年7月6日
				EP	4167645	A1	2023年4月19日
				IN	202247074647	A	2022年12月30日
CN	114071484	A	2022年2月18日	WO	2022022334	A1	2022年2月3日
				US	2023179490	A1	2023年6月8日
				EP	4181556	A1	2023年5月17日
CN	115392332	A	2022年11月25日	无			
US	202225126	A1	2022年7月14日	WO	2022154505	A1	2022年7月21日
				CN	114765789	A	2022年7月19日
WO	2022007579	A1	2022年1月13日	EP	4181572	A1	2023年5月17日
				US	2023247503	A1	2023年8月3日
				CN	113923730	A	2022年1月11日
WO	2022236840	A1	2022年11月17日	无			