

(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



(10) 国际公布号
WO 2022/268203 A1

(43) 国际公布日
2022 年 12 月 29 日 (29.12.2022)

(51) 国际专利分类号:
G06N 3/02 (2006.01)

(21) 国际申请号: PCT/CN2022/101128

(22) 国际申请日: 2022 年 6 月 24 日 (24.06.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110704382.6 2021年6月24日 (24.06.2021) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 罗咪 (LUO, Mi); 新加坡新加坡国立大学工程3路4号E4座8-24室视觉与机器学习实验室, Singapore 117583 (SG)。 陈飞 (CHEN, Fei);

中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 李震国 (LI, Zhenguo); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 冯佳时 (FENG, Jiashi); 新加坡新加坡国立大学工程3路4号E4座8-24室视觉与机器学习实验室, Singapore 117583 (SG)。

(74) 代理人: 北京市金杜律师事务所 (KING & WOOD MALLESONS); 中国北京市朝阳区东三环中路1号环球金融中心办公楼东楼20层, Beijing 100020 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN,

(54) Title: METHOD, APPARATUS AND SYSTEM FOR GENERATING NEURAL NETWORK MODEL, DEVICES, MEDIUM AND PROGRAM PRODUCT

(54) 发明名称: 用于生成神经网络模型的方法、装置、系统、设备、介质和程序产品

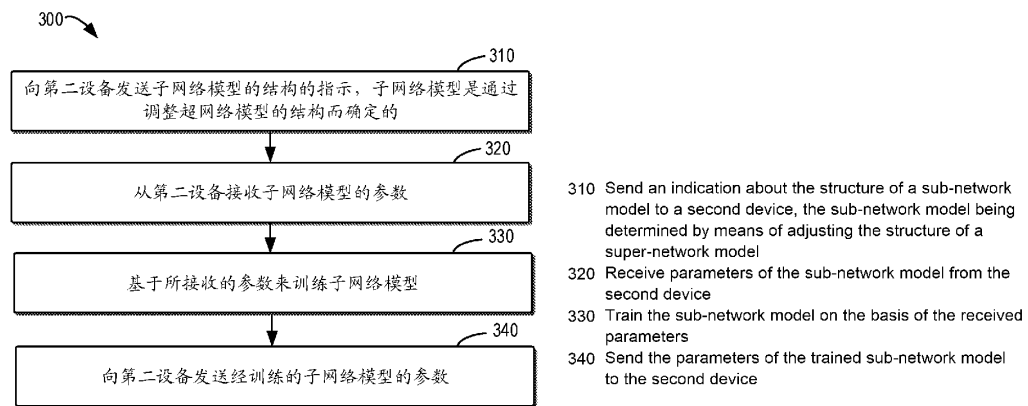


图 3

(57) Abstract: Embodiments of the present disclosure provide a method, apparatus and system for generating a neural network model, devices, a medium, and a program product. The method comprises: a first device sends an indication about the structure of a sub-network model to a second device, the sub-network model being determined by means of adjusting the structure of a super-network model; the first device receives parameters of the sub-network model from the second device, the parameters of the sub-network model being determined by the second device on the basis of the indication and the super-network model; the first device trains the sub-network model on the basis of the received parameters of the sub-network model; and the first device sends to the second device the parameters of the trained sub-network model for the second device to update the super-network model. By means of the foregoing manner, an efficient federated learning solution between a plurality of devices is provided, and communication costs required for a federated learning process and computing costs of the devices are also reduced while improving the accuracy of the model.

[见续页]

WO 2022/268203 A1

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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 本公开的实施例提供了用于生成神经网络模型的方法、装置、系统、设备、介质和程序产品。方法包括第一设备向第二设备发送关于子网络模型的结构指示, 子网络模型是通过调整超网络模型的结构来确定的; 第一设备从第二设备接收子网络模型的参数, 子网络模型的参数由第二设备基于该指示和超网络模型而确定; 第一设备基于所接收的子网络模型的参数来训练子网络模型; 以及第一设备向第二设备发送经训练的子网络模型的参数, 以用于第二设备更新超网络模型。通过上述方式, 提供了多个设备之间的高效的联邦学习方案, 在提升模型精度的同时, 还减少了联邦学习过程所需的通信成本、以及设备的计算成本。