(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2025 年 9 月 18 日 (18.09.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/189918 A1

(51) 国际专利分类号:
G06N 5/04 (2023.01)

(21) 国际申请号: PCT/CN2024/144172

(22) 国际申请日: 2024 年 12 月 31 日 (31.12.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202410278910.X 2024年3月12日 (12.03.2024) CN
202410650697.0 2024年5月21日 (21.05.2024) CN

(71) 申请人: 华为云计算技术有限公司 (**HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.**) [CN/CN]; 中国贵州省贵阳市贵安新区黔中大道交兴功路华为云数据中心 550025 (CN)。

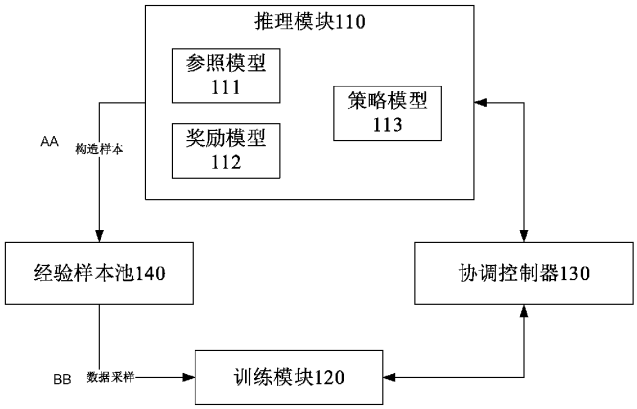
(72) 发明人: 程圣军 (**CHENG, Shengjun**); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。潘小亮 (**PAN, Xiaoliang**); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。徐云昆 (**XU, Yunkun**); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。杜斌 (**DU, Bin**); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。

(74) 代理人: 北京汇思诚业知识产权代理有限公司 (**UNI-INTEL PATENT AND TRADEMARK LAW FIRM**); 中国北京市朝阳区建国门外大街永安东里甲3号通用国际中心A座3层 100022 (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,

(54) **Title:** MODEL TRAINING SYSTEM AND METHOD, AND ELECTRONIC DEVICE

(54) 发明名称: 一种模型训练系统、方法和电子设备



110 Inference module
111 Reference model
112 Reward model
113 Strategy model
120 Training module
130 Coordination controller
140 Experience sample pool
AA Sample construction
BB Data sampling

(57) **Abstract:** Provided in the embodiments of the present application are a model training system and method, and an electronic device. The model training system comprises: an inference module, which is used for generating training samples for model training; an experience sample pool, which is used for storing the training samples; and a training module, which is used for performing model training on the basis of the training samples stored in the experience sample pool, wherein the inference module and the training module asynchronously run in parallel. In the model training system in the embodiments of the present application, by using an asynchronous architecture in which a training module and an inference module are separated, training and inference processes are decoupled, such

WO 2025/189918 A1

[见续页]

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, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

that a good expandability is realized, and the problem of a low sampling efficiency can be solved, thereby improving the efficiency of model training.

(57) 摘要: 本申请实施例提供一种模型训练系统、方法以及电子设备。模型训练系统包括：推理模块，其用于生成模型训练使用的训练样本；经验样本池，其用于保存所述训练样本；训练模块，其用于根据所述经验样本池中保存的所述训练样本，进行模型的训练，其中，所述推理模块与所述训练模块异步并行运行。根据本申请实施例的模型训练系统，采用训练模块与推理模块分离的异步架构，使得训练与推理过程解耦，具备良好的可扩展性，可以解决采样效率低下的问题，提升模型训练的训练效率。