

(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 2023/087953 A1

(43) 国际公布日
2023 年 5 月 25 日 (25.05.2023)

(51) 国际专利分类号:
G06N 3/04 (2006.01) *G06N 3/08* (2006.01)

(21) 国际申请号: PCT/CN2022/123139

(22) 国际申请日: 2022 年 9 月 30 日 (30.09.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202111387843.8 2021 年 11 月 22 日 (22.11.2021) 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: 茹彬鑫 (RU, Binxin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 万星辰 (WAN, Xingchen); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 埃斯佩兰卡·佩德罗 (ESPERANCA, Pedro); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 卡路奇·法比奥·玛利亚 (CARLUCCI, Fabio Maria); 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, Zhenguó); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京亿腾知识产权代理事务所 (普通合伙) (E-TONE INTELLECTUAL PROPERTY FIRM (GENERAL PARTNERSHIP)); 中国北京市

(54) Title: METHOD AND APPARATUS FOR SEARCHING FOR NEURAL NETWORK ENSEMBLE MODEL, AND ELECTRONIC DEVICE

(54) 发明名称: 搜索神经网络集成模型的方法、装置和电子设备

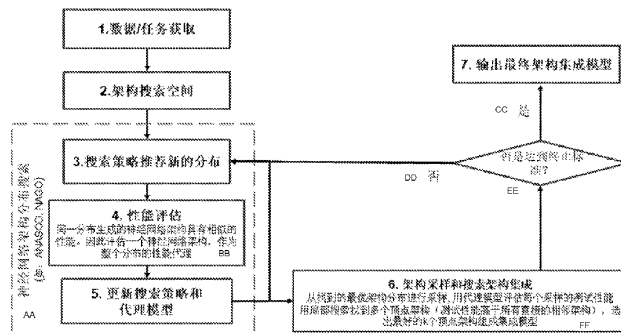


图 5

AA Neural network architecture distribution search (such as ANASOD, NAGO)
BB Neural network architectures generated from the same distribution have similar performance, and therefore, a neural network architecture is evaluated as a performance surrogate for the entire distribution
CC Yes
DD No
EE Is a termination criterion met?
FF Sample from optimal architecture distributions found, evaluate test performance of each sample by using an ensemble model, find multiple vertex architectures (the test performance thereof is higher than all direct adjacent architectures) by means of local search, and select the best k vertex architectures to form an ensemble model

1 Data/task acquisition
2 Architecture search space
3 Recommend a new distribution according to a search policy
4 Performance evaluation
5 Update the search policy and a surrogate model
6 Architecture sampling and found architecture ensemble
7 Output a final architecture ensemble model

(57) Abstract: A method for searching for a neural network architecture ensemble model, comprising: acquiring a data set, the data set comprising samples and labels in a classification task; and using a neural network architecture distribution search algorithm for searching, comprising: determining hyper-parameters of a neural network architecture distribution; sampling a valid neural network architecture in the architecture distribution defined by the hyper-parameters; training and evaluating the neural network architecture on the data set to obtain performance indexes; determining a neural network architecture distribution of shared hyper-parameters according

[见续页]

海淀区中关村紫金数码园3号楼707,
Beijing 100190 (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, ME, 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, 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))。

to the performance indexes to obtain a candidate pool of base learners, the base learners being neural network architectures that meet the requirements of the architecture distribution; and determining a surrogate model, the surrogate model being used for predicting test performance of an unevaluated neural network architecture; and predicting the test performance of the base learners in the candidate pool by means of the surrogate model, and determining k diverse base learners that meet task scenario requirements to form an ensemble model, wherein the size of the ensemble model is k.

(57) 摘要: 一种搜索神经网络架构集成模型的方法, 包括: 获取数据集, 数据集包括分类任务中的样本和标注; 使用神经网络架构分布搜索算法进行搜索, 包括: 确定神经网络架构分布的超参; 在超参定义的架构分布中采样一个有效的(valid)神经网络架构; 对神经网络架构在数据集上训练和评估, 得到性能指标; 根据性能指标确定共享超参的神经网络架构分布, 获得基础学习器的候选池; 基础学习器为符合所述架构分布要求的神经网络架构; 确定代理模型; 代理模型用于预测未评估的神经网络架构的测试性能; 通过代理模型预测所述候选池中基础学习器的测试性能, 确定符合任务场景要求的k个多元化(diverse)基础学习器组成集成模型, 所述集成模型的大小为k。