

(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 2023/056889 A1

(43) 国际公布日
2023 年 4 月 13 日 (13.04.2023)

- (51) 国际专利分类号:
G06N 3/04 (2006.01) *G06V 10/82* (2022.01)
- (21) 国际申请号: PCT/CN2022/123011
- (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:
202111174534.2 2021 年 10 月 9 日 (09.10.2021) CN
- (71) 申请人: 百果园技术 (新加坡) 有限公司
(**BIGO TECHNOLOGY PTE. LTD.**) [SG/SG]; 新加坡新加坡市巴西班让路枫树商业城 30 号楼 15 层 31A, Singapore 117440 (SG)。
- (72) 发明人: 及
- (71) 申请人 (仅对 US): 罗雄文 (**LUO, Xiongwen**) [CN/CN]; 中国广东省广州市番禺区市桥街兴泰路 274 号 C 栋西塔 5-13 层, Guangdong 511402 (CN)。
- (72) 发明人: 卢江虎 (**LU, Jianghu**); 中国广东省广州市番禺区市桥街兴泰路 274 号 C 栋西塔 5-13 层, Guangdong 511402 (CN)。 项伟 (**XIANG, Wei**); 中国广东省广州市番禺区市桥街兴泰路 274 号 C 栋西塔 5-13 层, Guangdong 511402 (CN)。
- (74) 代理人: 北京品源专利代理有限公司 (**BEYOND ATTORNEYS AT LAW**); 中国北京市海淀区莲花池东路 39 号西金大厦 6 层, Beijing 100036 (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 AND SCENE RECOGNITION METHOD AND APPARATUS, DEVICE, AND MEDIUM

(54) 发明名称: 模型训练和场景识别方法、装置、设备及介质

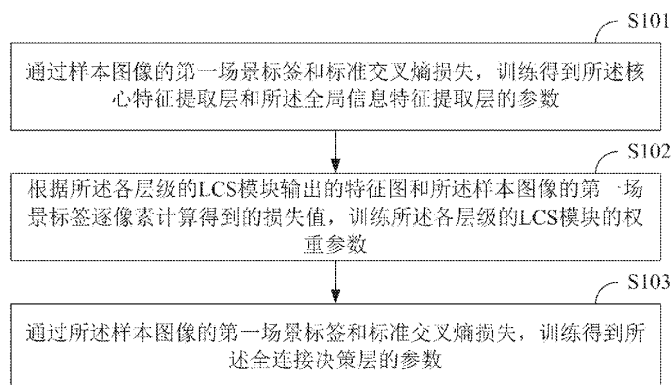


图 1

- S101 Train, by means of a first scene label and a standard cross-entropy loss of a sample image, to obtain parameters of a core feature extraction layer and a global information feature extraction layer
- S102 Train a weight parameter of an LCS module at each level according to a feature map outputted by the LCS module at each level and a loss value obtained by pixel-by-pixel calculation of the first scene label of the sample image
- S103 Train, by means of the first scene label and the standard cross-entropy loss of the sample image, to obtain a parameter of a fully-connected decision-making layer

(57) **Abstract:** The present application discloses a model training and scene recognition method and apparatus, a device, and a medium. When a scene recognition model is trained, training is first performed by means of a first scene label and a standard cross-entropy loss of a sample image to obtain parameters of a core feature extraction layer and a global information feature extraction layer; then, according to a feature map outputted by a local supervised learning (LCS) module having an attention mechanism at each level and a loss value obtained by pixel-by-pixel calculation of the first scene label of the sample image, a weight parameter of the LCS module at each level is trained; and finally, a parameter of a fully-connected decision-making layer of the scene recognition model is obtained by training.

[见续页]

WO 2023/056889 A1

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)。

根据细则4.17的声明:

- 关于申请人有权申请并被授予专利 (细则4.17(ii))
- 关于申请人有权要求在先申请的优先权 (细则4.17(iii))
- 发明人资格 (细则4.17(iv))

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 本申请公开了一种模型训练和场景识别方法、装置、设备及介质, 在训练场景识别模型时, 首先通过样本图像的第一场景标签和标准交叉熵损失, 训练得到核心特征提取层和全局信息特征提取层的参数, 然后根据每个层级的带注意力机制的局部监督学习LCS模块输出的特征图和所述样本图像的第一场景标签逐像素计算得到的损失值, 训练每个层级的LCS模块的权重参数, 最后训练得到场景识别模型的全连接决策层的参数。