

(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 2025/152868 A1

(43) 国际公布日
2025 年 7 月 24 日 (24.07.2025)

(51) 国际专利分类号:
G06V 20/40 (2022.01)

LTD.) [CN/CN]; 中国北京市大兴区北京经济技术开发区科创十一街18号C座2层221室 100176 (CN)。

(21) 国际申请号: PCT/CN2025/071813

(72) 发明人: 刘武(**LIU, Wu**); 中国北京市大兴区北京经济技术开发区科创十一街18号院2号楼4层A402室 100176 (CN)。 刘鑫辰(**LIU, Xincheng**); 中国北京市大兴区北京经济技术开发区科创十一街18号院2号楼4层A402室 100176 (CN)。 刘鲲(**LIU, Kun**); 中国北京市大兴区北京经济技术开发区科创十一街18号院2号楼4层A402室 100176 (CN)。 何晓冬(**HE, Xiaodong**); 中国北京市大兴区北京经济技术开发区科创十一街18号院2号楼4层A402室 100176 (CN)。

(22) 国际申请日: 2025 年 1 月 10 日 (10.01.2025)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202410057649.0 2024年1月15日 (15.01.2024) CN

(71) 申请人: 北京沃东天骏信息技术有限公司 (**BEIJING WODONG TIANJUN INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市大兴区北京经济技术开发区科创十一街18号院2号楼4层A402室 100176 (CN)。 京东科技控股股份有限公司 (**JINGDONG TECHNOLOGY HOLDING CO.,**

(74) 代理人: 中国贸促会专利商标事务所有限公司 (**CCPIT PATENT AND TRADEMARK LAW OFFICE**); 中国北京市西城区复兴门内大街158号远洋大厦F10层 100031 (CN)。

(54) **Title:** TEMPORAL ACTION SEGMENTATION METHOD AND APPARATUS, MODEL TRAINING METHOD AND APPARATUS, AND STORAGE MEDIUM

(54) 发明名称: 时序动作分割方法、装置及模型训练方法、装置和存储介质

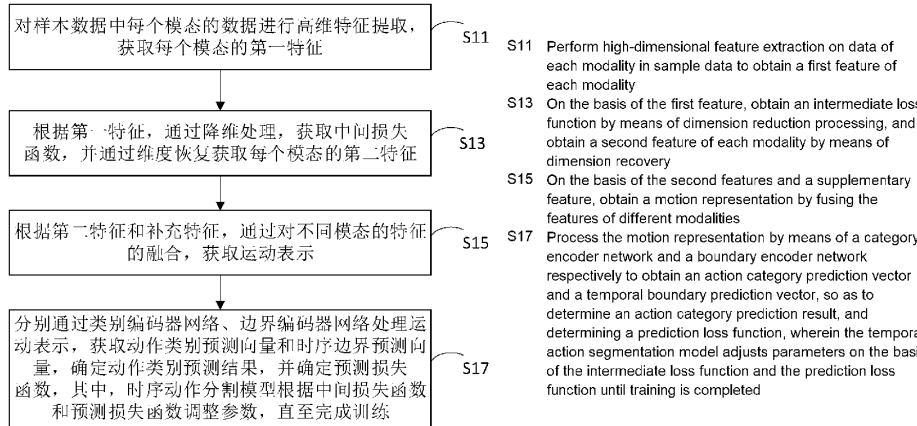


图 1

(57) **Abstract:** The present disclosure relates to the technical field of artificial intelligence, and provides a temporal action segmentation method and apparatus, a model training method and apparatus, and a storage medium. The training method for a temporal action segmentation model of the present disclosure comprises: performing high-dimensional feature extraction on data of each modality in sample data to obtain a first feature of each modality; on the basis of the first feature, obtaining an intermediate loss function by means of dimension reduction processing, and obtaining a second feature of each modality by means of dimension recovery; on the basis of the second features and a supplementary feature, obtaining a motion representation by fusing the features of different modalities; processing the motion representation by means of a category encoder network and a boundary encoder network respectively to obtain an action category prediction vector and a temporal boundary prediction vector, so as to determine an action category prediction result, and determining a prediction loss function, wherein the temporal action segmentation model adjusts parameters on the basis of the intermediate loss function and the prediction loss function until training is completed.

[见续页]

(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, 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, UY, 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)。

根据细则4.17的声明:
— 关于申请人有权要求在先申请的优先权(细则4.17(iii))

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本公开提出一种时序动作分割方法、装置及模型训练方法、装置和存储介质, 涉及人工智能技术领域。本公开的一种时序动作分割模型的训练方法包括: 对样本数据中每个模态的数据进行高维特征提取, 获取每个模态的第一特征; 根据第一特征, 通过降维处理, 获取中间损失函数, 并通过维度恢复获取每个模态的第二特征; 根据第二特征和补充特征, 通过对不同模态的特征的融合, 获取运动表示; 分别通过类别编码器网络、边界编码器网络处理运动表示, 获取动作类别预测向量和时序边界预测向量, 确定动作类别预测结果, 并确定预测损失函数, 其中, 时序动作分割模型根据中间损失函数和预测损失函数调整参数, 直至完成训练。