

(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
2022 年 11 月 17 日 (17.11.2022)



(10) 国际公布号
WO 2022/237157 A1

- (51) 国际专利分类号:
G06F 16/78 (2019.01)
- (21) 国际申请号: PCT/CN2021/137579
- (22) 国际申请日: 2021 年 12 月 13 日 (13.12.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110505869.1 2021年5月10日 (10.05.2021) CN
- (71) 申请人: 中国科学院深圳先进技术研究院 (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(74) 代理人: 北京中巡通大知识产权代理有限公司 (BEIJING ZHONG XUN TONG DA INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市西城区榆树馆胡同东口 1 层 106, Beijing 100195 (CN)。

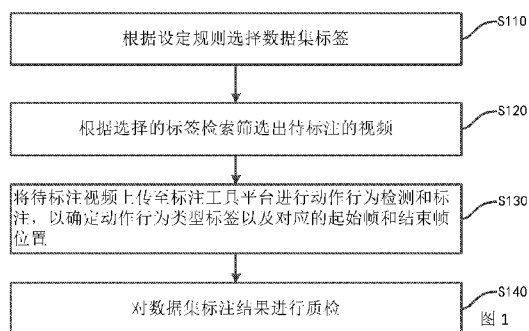
(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,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

(72) 发明人: 马筱(MA, Xiao); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。 乔宇(QIAO, Yu); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。 王利民(WANG, Limin); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。

(54) Title: VIDEO DATA SET LABELING METHOD AND APPARATUS

(54) 发明名称: 一种视频数据集标注方法及装置



S110 Select a data set label according to a set rule
S120 Search for and choose, according to the selected label, a video to be labeled
S130 Upload said video to a labeling tool platform for action behavior detection and labeling, so as to determine an action behavior type label and corresponding start frame and end frame positions
S140 Perform quality inspection on a data set labeling result

(57) Abstract: Disclosed are a video data set labeling method and apparatus. The method comprises: determining a data set label according to a set action category selection rule, the data set label representing short-time instantaneous action and cyclic action types; choosing, according to the data set label, a matched video to be labeled; and uploading the video to be labeled to a labeling tool platform for action behavior detection and labeling, so as to determine an action behavior type label and corresponding start frame and end frame positions. According to the present invention, the boundary definition of the action behavior is more accurate, and both the labeling efficiency and the labeling quality are remarkably improved.

(57) 摘要: 本发明公开了一种视频数据集标注方法及装置。该方法包括: 根据设定的动作类别选择规则确定数据集标签, 该数据集标签表征短时间的瞬时动作和循环动作类型; 根据所述数据集标签筛选出匹配的待标注视频; 将待标注视频上传至标注工具平台进行动作行为检测和标注, 以确定动作行为类型标签以及对应的起始帧和结束帧位置。本发明对于动作行为的边界定义更加精确, 并且标注效率和标注质量均有显著提高。

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))。

一种视频数据集标注方法及装置

技术领域

[0001] 本发明涉及计算机视觉技术领域，更具体地，涉及一种视频数据集标注方法及装置。

背景技术

[0002] 近年来，视频理解在视频内容分析，智能监控，人机交互等领域得到了广泛应用。在视频行为理解中，基于深度学习，有两类比较重要的任务，一类是视频行为分类，其主要是针对修剪过的视频，按照其中的人类行为进行分类。另一类是视频行为检测，目的是在长视频中定位一个动作的开始时间和结束时间。视频行为检测作为视频理解的重要部分，在计算机视觉界已得到了广泛研究。

[0003] 相比于行为分类，行为检测难度更高，现有的行为检测方法通常是首先生成可能存在动作的片段提案，然后再对其进行分类。然而，因为边界的定义较为模糊，统一视频中可能还会存在多个动作同时进行的情况，导致对动作的准确检测具有巨大挑战。不同于行为识别，行为检测要求进行精确的动作片段检测，而对于真实场景下的动作产生，往往边界不是十分确定，尤其是动作的终止，并且对动作完整性的判断也相对困难。由于视频本身边界不明确、以及现有时序检测标注工具相对简陋，导致现有的视频时序检测数据集大多是弱标定方式，这也导致了目前行为检测平均精准度偏低。

[0004] 经分析，现有的视频标注方案主要存在以下缺陷：

[0005] 1)、相关视频时序检测数据集标签定义较为粗粒度，不同标签的时序时长相差较大并且不同标签的边界定义不明晰，无法更为直观地明晰起止边界定义。

[0006] 2)、现有的视频标注工具主要针对目标检测工作，而非对一段未修剪的视频进行行为段的起止时间标注。此外，现有的视频标注工具功能相对简单，界面相对简陋。对于大量的未修剪的数据，没有一个相对便捷，功能齐全的标注工具，所需人工成本昂贵。由于现实视频的复杂性，现有的视频标注工具浏览一次视频大多只能针对一类标签进行标注，然而现实场景中，往往是多个行为同

时出现，因此需要对一段时间的多标签情况也要进行考虑。另外，现有的标注工具往往不能清晰地体现已标注标签的时间片段，容易导致漏标，重复标，错标等情况，并且无法很好地体现同一行为段的起止关联性，也不方便二次质检，因此对于已标注行为段的显示也较为重要。

发明概述

技术问题

[0007] 本发明的目的是克服上述现有技术的缺陷，提供一种视频数据集标注方法及装置。

问题的解决方案

技术解决方案

[0008] 根据本发明的第一方面，提供一种视频数据集标注方法。该方法包括以下步骤：

[0009] 步骤S1：根据设定的动作类别选择规则确定数据集标签，该数据集标签表征短时间的瞬时动作和循环动作类型；

[0010] 步骤S2：根据所述数据集标签筛选出匹配的待标注视频；

[0011] 步骤S3：将待标注视频上传至标注工具平台进行动作行为检测和标注，以确定动作行为类型标签以及对应的起始帧和结束帧位置。

[0012] 步骤S4：对标注结果进行抽样可视化质检并通过行为识别模型对背景样例以及行为段样例进行识别，通过这种方式检测标注质量可以在大幅节省人工成本的同时提高精度。

[0013] 根据本发明的第二方面，提供一种视频数据集标注装置。该装置包括：

[0014] 标签选择模块：用于根据设定的动作类别选择规则确定数据集标签，该数据集标签表征短时间的瞬时动作和循环动作类型；

[0015] 视频检索模块：用于根据所述数据集标签筛选出匹配的待标注视频；

[0016] 数据集标注模块：将待标注视频上传至标注工具平台进行动作行为检测和标注，以确定动作行为类型标签以及对应的起始帧和结束帧位置。

发明的有益效果

有益效果

[0017] 与现有技术相比，本发明的优点在于，提供了一种用于深度学习行为时序检测的数据集的标注技术方案，首先对收集来的视频进行视频查重。为了更好地统一边界定义指标，针对现有的视频数据集的标签根据一定规则进行筛选，通过拆解，合并等方式将标签结构化，并按照一定规则进行了排除，更加精确不同行为的起止边界。此外，还设计了一个视频时序标注的工具，针对一段输入视频进行不同标签起止帧的选取，使得多种标签的时间序列更好地体现，并且对已标注行为段都可以更加直观的展现。

[0018] 通过以下参照附图对本发明的示例性实施例的详细描述，本发明的其它特征及其优点将会变得清楚。

对附图的简要说明

附图说明

[0019] 被结合在说明书中并构成说明书的一部分的附图示出了本发明的实施例，并且连同其说明一起用于解释本发明的原理。

[0020] 图1是根据本发明一个实施例的视频数据集标注方法的流程图；

[0021] 图2是根据本发明一个实施例的视频数据集标注方法的总体过程示意图；

[0022] 图3是根据本发明一个实施例的视频标注工具标注示意图；

[0023] 图4是根据本发明一个实施例的视频标注工具流程示意图。

发明实施例

本发明的实施方式

[0024] 现在将参照附图来详细描述本发明的各种示例性实施例。应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本发明的范围。

[0025] 以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本发明及其应用或使用的任何限制。

[0026] 对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为说明书的一部分。

[0027] 在这里示出和讨论的所有例子中，任何具体值应被解释为仅仅是示例性的，而

不是作为限制。因此，示例性实施例的其它例子可以具有不同的值。

[0028] 应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

[0029] 结合图1和图2所示，所提供的视频数据集标注方法包括以下步骤：

[0030] 步骤S110，根据设定规则选择数据集标签。

[0031] 在此步骤中，通过一定规则定义数据集标签，例如，选择通用场景下的常见标签，而非特定场景特定人群类的标签；排除动作定义较宽泛的标签；排除主要通过交互对象的差异性区分而非人类姿态变化的不同而分类的标签；排除在每个动作中都通用的基础身体状态标签；将可以分为原子动作的动作进行拆分。

[0032] 在一个实施例中，将标签主要分为两类：短时间的瞬时动作和循环动作对动作的过程周期进行划分。所选取数据可通过现有行为识别方法进行筛选。

[0033] 通过定义动作类别中的“动作”并进行筛选和划分，是标签更为合理，粗粒度动作可以划分为更多的细微动作。

[0034] 步骤S120，根据选择的标签检索筛选出待标注的视频。

[0035] 针对标签搜集相关视频，并对其进行查重和筛选。在一个实施例中，查重过程包括：对于待处理视频，在视频库中进行近邻检索，筛选出与待处理视频相似的候选视频，得到候选视频集合；计算每一候选视频与待处理视频的相似性，得到相似度结果；根据相似度结果确定待处理视频是否通过查重检测。其中，可通过对视频第一帧和最后一帧的哈希值的汉明距离来计算相似性。

[0036] 步骤S130，将待标注视频上传至标注工具平台进行动作行为检测和标注，以确定动作行为类型标签以及对应的起始帧和结束帧位置。

[0037] 具体地，参见图3所示，将视频输入至标注工具平台进行视频标注，其中左上方为实际标注起始帧与结束帧的选择区域；右上方为视频选择区域，可批量进行选取（已标注视频会显示已标注结果）；视频选择下方为标签选择菜单，delete（删除），label（标注），save（保存）按钮分别具备删除误操作框，标注操作框以及保存现有结果的功能。

[0038] 利用图3的标注工具平台，能够实现同时标注多标签的功能，更具备现实意义。菜单选择下方为实际标注结果的timeline（时间线）显示区域，便于标注人

员检查错标、漏标以及二次质检；左下方为视频播放区域，可通过滑动条快速浏览视频。并且，实际标注时可设置键盘快捷键，以便于标注人员更快速地进行标注。

[0039] 具体参见图4所示，首先，输入待标注视频并点击播放键，也可通过滑动条快速浏览整条视频。在显示以帧数框中的帧为起始的24帧图像的区域选择待标注动作的起始帧和结束帧（例如，两个框显示的分别是标定的起始帧和标定的结束帧）；通过菜单栏添加标签，并选定标签的开始和结束，然后通过时间轴检测已标注行为片段，这种方式有利于标注人员检查错标、漏标以及二次质检，通过保存，删除菜单对标签标注结果进行选择；结束标注。

[0040] 综上，所设计的标注工具平台相较于现有的标注工具，功能更加丰富，操作更加便捷，界面更加直观，并且利用上述过程进行标注有利于确定更明晰的边界，并实现对一段时间的多标签进行标注，体现同一行为段的起止关联性。此外，标注规则和流程更加精确，减少了标注偏差和时间动作定位的边界不确定性。

[0041] 步骤S140，对数据集标注结果进行质检。

[0042] 在获得利用视频标注工具标注的结果后，可对其进行抽样可视化质检并通过模型对背景样例以及行为段样例进行识别，通过这种识别检测标注质量可以在大幅节省人工成本的同时提高精度。例如，通过TSN（Temporal Segment Networks）方法来对其进行识别。该方法主要由空间流卷积网络和时间流卷积网络构成。但不同于two-stream采用单帧或者单堆帧，TSN使用从整个视频中稀疏地采样一系列短片段，每个片段都将给出其本身对于行为类别的初步预测，从这些片段的“共识”来得到视频级的预测结果。在学习过程中，通过迭代更新模型参数来优化视频级预测的损失值（loss value）。结果进行统计，根据质检结果可确定所标注的数据集是否符合预期。

[0043] 为验证本发明的效果，采用不同方法在数据集上进行了比较，结果参见下表1。

[0044] 表1：不同方法在数据集上的比较

[0045]

Method	Modality	IoU [0.05:0.05:0.5]			IoU [0.5:0.05:0.95]		
		AR@10	AR@100	AUC	AR@10	AR@100	AUC
BMN	rgb	29.51	63.90	48.89	12.02	24.15	18.62
	flow	24.14	65.02	51.04	13.91	26.05	20.60
	rgb+flow	32.13	64.71	50.40	14.12	25.98	20.64
DBG	rgb	26.22	61.96	46.20	9.83	22.59	16.53
	flow	29.75	64.17	49.37	11.91	25.15	19.08
	rgb+flow	28.93	63.72	48.75	11.41	24.61	18.58
G-TAD	rgb	27.71	56.80	43.61	11.23	22.54	17.25
	flow	29.23	58.38	45.02	12.53	23.90	18.55
	rgb+flow	30.03	59.44	46.23	13.05	24.62	19.23

[0046] 表1中，通过三种目前较为先进的时序检测方法BMN、DBG、G-TAD分别通过双流3D卷积（I3D）模型提取RGB、光流、RGB+光流特征在数据集上进行实验。其中，边界匹配网络（BMN）通过一种新的时序提名置信度评估机制-边界匹配机制，以及基于边界匹配机制的边界匹配网络。BMN网络能够同时生成一维边界概率序列，以及二维的BM置信度图来密集的评估所有可能存在的时序提名的置信度分数。稠密边界动作生成器网络（DBG）通过一种快速的、端到端的稠密边界动作生成器对所有的动作提名估计出稠密的边界置信度图。用于时序动作检测的子图定位模型（G-TAD）通过自适应地融合多级语义文本信息，将时序动作检测问题转化为子图定位问题。评价指标主要用AR@AN表示，即通过测量AR与提案的平均数（AN）之间的关系来进行判断。并计算了AR与AN曲线下的面积（AUC），作为ActivityNet-1.3数据集上的另一个评估指标，其中AN的范围是0到100。由表1可以看出，基于本发明的这种标签筛选方法构建的数据集由于标注结果更加精确，使得现有的视频时序检测方法的性能远低于其他数据集。

[0047] 相应地，本发明还提供一种视频数据集标注装置，用于实现上述方法的一个方面或多个方面。例如，该装置包括：标签选择模块：用于根据设定的动作类别选择规则确定数据集标签，该数据集标签表征短时间的瞬时动作和循环动作类型；视频检索模块：用于根据所述数据集标签筛选出匹配的待标注视频；数据

集标注模块：将待标注视频上传至标注工具平台进行动作行为检测和标注，以确定动作行为类型标签以及对应的起始帧和结束帧位置。

[0048] 在计算机视觉领域，除了公开的数据集之外，对很多应用场景都需要专门的数据集做迁移学习或者端到端的训练，这种情况需要大量的训练数据。而本发明可应用于多种领域的视频数据集标注，例如，用于视频辅助裁判方面。因为本发明对视频动作快慢时间不敏感，因此可以普适于多种体育运动场景中，如动作慢的瑜伽和动作变化迅速的体操等。通过更加精确的标注方法，可以使得不同动作之间的边界判断更明确。例如，用于智能视频审核，在移动端即可完成异常动作识别和研判，可以通过开始动作界限来判断是否将要发生该异常行为；还可以运用到其他识别中，如管道的异常检测等。又如，应用于智能安防，可以在计算资源受限的智能终端如智能眼镜、无人机、智能摄像头等上直接进行动作识别，直接反馈异常行为，提高巡防等的即时性和准确性。

[0049] 综上所述，相对于现有技术，本发明的优点主要体现在以下方面：

[0050] 1)、视频时序检测对学术圈和工业界都有着广泛的应用价值，但由于没有明显的边界定义以及人工标注成本损耗大，现有的视频数据集在标注时都存在一定程度的高成本和弱标注的特点，动作标签的某些定义相对而言比较粗糙，不适合定义准确的动作行为边界。另外，由于不同的动作有不同的动作过程周期，不同粒度的人类动作也会带来检测上的困难。本发明更为精确的标注粒度，参考现有行为相关数据集的标注特点，根据动作类别选择中的标准进行排除，筛选，拆分。

[0051] 2)、现有视频时序标注工具功能较少，界面相对简陋。本发明为了帮助标注员有效且一致地标注操作视频段，设计了视频时序检测的工具。其中设置视频浏览区域，帮助标注员快速预览整个视频，并提供了快进等功能；起止帧选择区域，通过不同的操作来注释动作的开始和结束，通过不同的标明代表起止帧；标签选择区域，选择不同段中的标签类别，为了方便多标签的标注，对类别进行了归类及菜单选项；操作菜单区域，用以添加，删除和修改标签以及标签显示区域，以帮助标注员查看结果，避免漏标，重标，错标。此外，为了尽可能避免人主观意识带来的偏差，保证数据一致性，还设计了标注指南，通过维

基百科及相关运动指南通过文字和图片清楚地定义每个标签的边界。标注效率相比纯手工而言提升了效率，并提高了标注质量。

[0052] 3)、现有技术主要通过人工检验。本发明在人工二次质检的基础上，通过现有行为识别模型对标注结果的背景段和行为段进行了样本分类。对比现有数据集，边界定义更加精确。

[0053] 本发明可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质，其上载有用于使处理器实现本发明的各个方面的计算机可读程序指令。

[0054] 计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是一——但不限于——电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM或闪存）、静态随机存取存储器（SRAM）、便携式压缩盘只读存储器（CD-ROM）、数字多功能盘（DVD）、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身，诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波（例如，通过光纤电缆的光脉冲）、或者通过电线传输的电信号。

[0055] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备，或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令，并转发该计算机可读程序指令，以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0056] 用于执行本发明操作的计算机程序指令可以是汇编指令、指令集架构（ISA）指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以

一种或多种编程语言的任意组合编写的源代码或目标代码，所述编程语言包括面向对象的编程语言—诸如Smalltalk、C++、Python等，以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。在一些实施例中，通过利用计算机可读程序指令的状态信息来个性化定制电子电路，例如可编程逻辑电路、现场可编程门阵列(FPGA)或可编程逻辑阵列(PLA)，该电子电路可以执行计算机可读程序指令，从而实现本发明的各个方面。

[0057] 这里参照根据本发明实施例的方法、装置（系统）和计算机程序产品的流程图和/或框图描述了本发明的各个方面。应当理解，流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

[0058] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，这些指令使得计算机、可编程数据处理装置和/或其它设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0059] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上，使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0060] 附图中的流程图和框图显示了根据本发明的多个实施例的系统、方法和计算机

程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分，所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。对于本领域技术人员来说公知的是，通过硬件方式实现、通过软件方式实现以及通过软件和硬件结合的方式实现都是等价的。

[0061] 以上已经描述了本发明的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术改进，或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。本发明的范围由所附权利要求来限定。

权利要求书

- [权利要求 1] 一种视频数据集标注方法，包括以下步骤：
- 步骤S1：根据设定的动作类别选择规则确定数据集标签，该数据集标签表征短时间的瞬时动作和循环动作类型；
- 步骤S2：根据所述数据集标签筛选出匹配的待标注视频；
- 步骤S3：将待标注视频上传至标注工具平台进行动作行为检测和标注，以确定动作行为类型标签以及对应的起始帧和结束帧位置。
- [权利要求 2] 根据权利要求1所述的方法，其特征在于，所述根据设定的动作类别选择规则确定数据集标签包括：
- 选择通用场景下的常见标签，并排除非特定场景特定人群类的标签；
- 排除动作定义宽泛的标签；
- 排除通过交互对象的差异性区分而非人类姿态变化的不同而分类的标签；
- 排除在每个动作中都通用的基础身体状态标签；
- 对于可拆分的动作进行拆分，以获得细粒度标签。
- [权利要求 3] 根据权利要求1所述的方法，其特征在于，步骤S2包括：
- 根据所述数据集标签搜集相关视频，并进行查重和筛选，其中查重通过对视频第一帧和最后一帧的哈希值的汉明距离来计算相似性；
- 根据相似性结果确定待处理视频是否通过查重检测。
- [权利要求 4] 根据权利要求1所述的方法，其特征在于，所述标注工具平台设有起始帧选择区域、视频选择区域、标签选择区域、结果显示区域以及视频播放区域，其中，所述起始帧选择区域用于供用户标注起始帧与结束帧；所述视频选择区域用于供用户选择待标注的一个或多个视频；所述标签选择区域用于供用户标注动作行为标签；所述结果显示区域用于向用户显示标注起始时间；所述视频播放区域用于向用户显示待标注视频的连续多帧图像，以供标注动作行为的起始帧。
- [权利要求 5] 根据权利要求4所述的方法，其特征在于，步骤S3包括：
- 输入待标注视频并点击播放键或通过滑动条浏览待标注视频；

在所述视频选择区域，显示连续的24帧图像，以供用户选择待标注动作行为的起始帧和结束帧；

通过所述标签选择区域的菜单栏添加标签，并选定标签的开始和结束；

通过所述结果显示区域的时间轴检测已标注行为片段，以供用户检查错标、漏标以及二次质检；

通过设置在所述标签选择区域的保存，删除菜单对标注结果进行选择。

[权利要求 6] 根据权利要求1所述的方法，其特征在于，还包括：

步骤S4：对标注结果进行抽样可视化质检并通过行为识别模型对背景样例以及行为段样例进行识别。

[权利要求 7] 根据权利要求6所述的方法，其特征在于，步骤S4包括：

采用TSN行为识别模型对背景视频段以及行为视频段进行识别，给出其本身对于行为类别的预测类别得分，从而检验标注结果质量。

[权利要求 8] 一种视频数据集标注装置，包括：

标签选择模块：用于根据设定的动作类别选择规则确定数据集标签，该数据集标签表征短时间的瞬时动作和循环动作类型；

视频检索模块：用于根据所述数据集标签筛选出匹配的待标注视频；

数据集标注模块：将待标注视频上传至标注工具平台进行动作行为检测和标注，以确定动作行为类型标签以及对应的起始帧和结束帧位置。

[权利要求 9] 一种计算机可读存储介质，其上存储有计算机程序，其中，该程序被处理器执行时实现根据权利要求1至7中任一项所述方法的步骤。

[权利要求 10] 一种计算机设备，包括存储器和处理器，在所述存储器上存储有能够在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现根据权利要求1至7中任一项所述的方法的步骤。

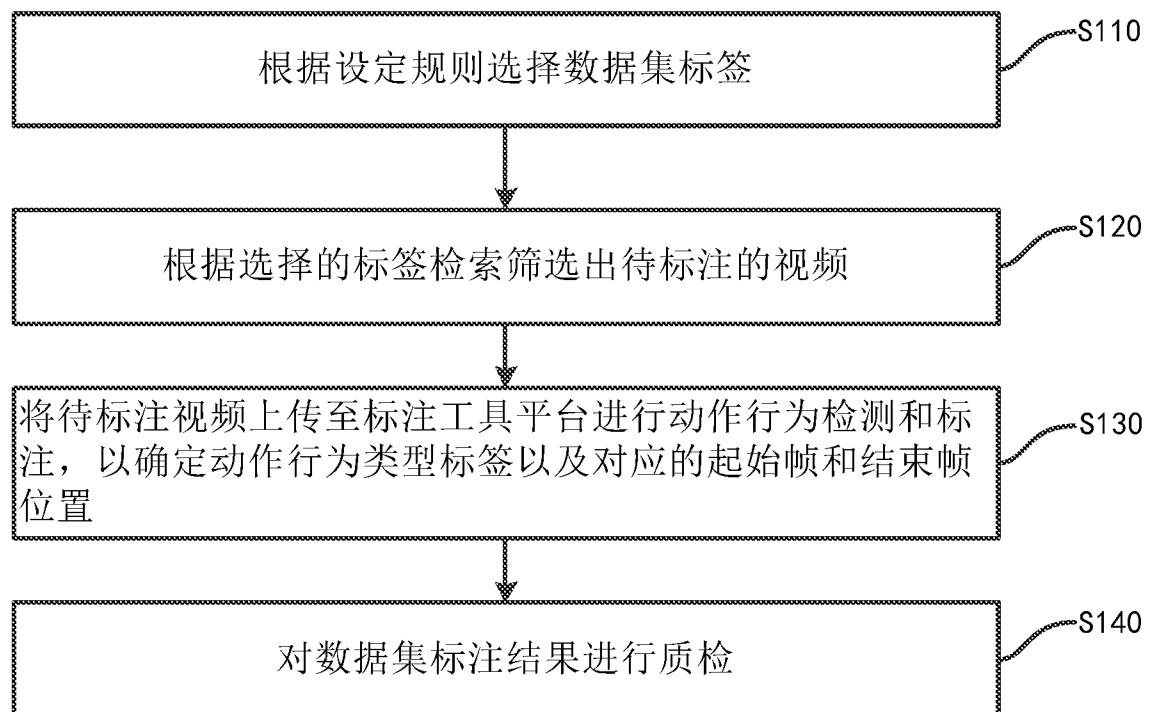


图 1

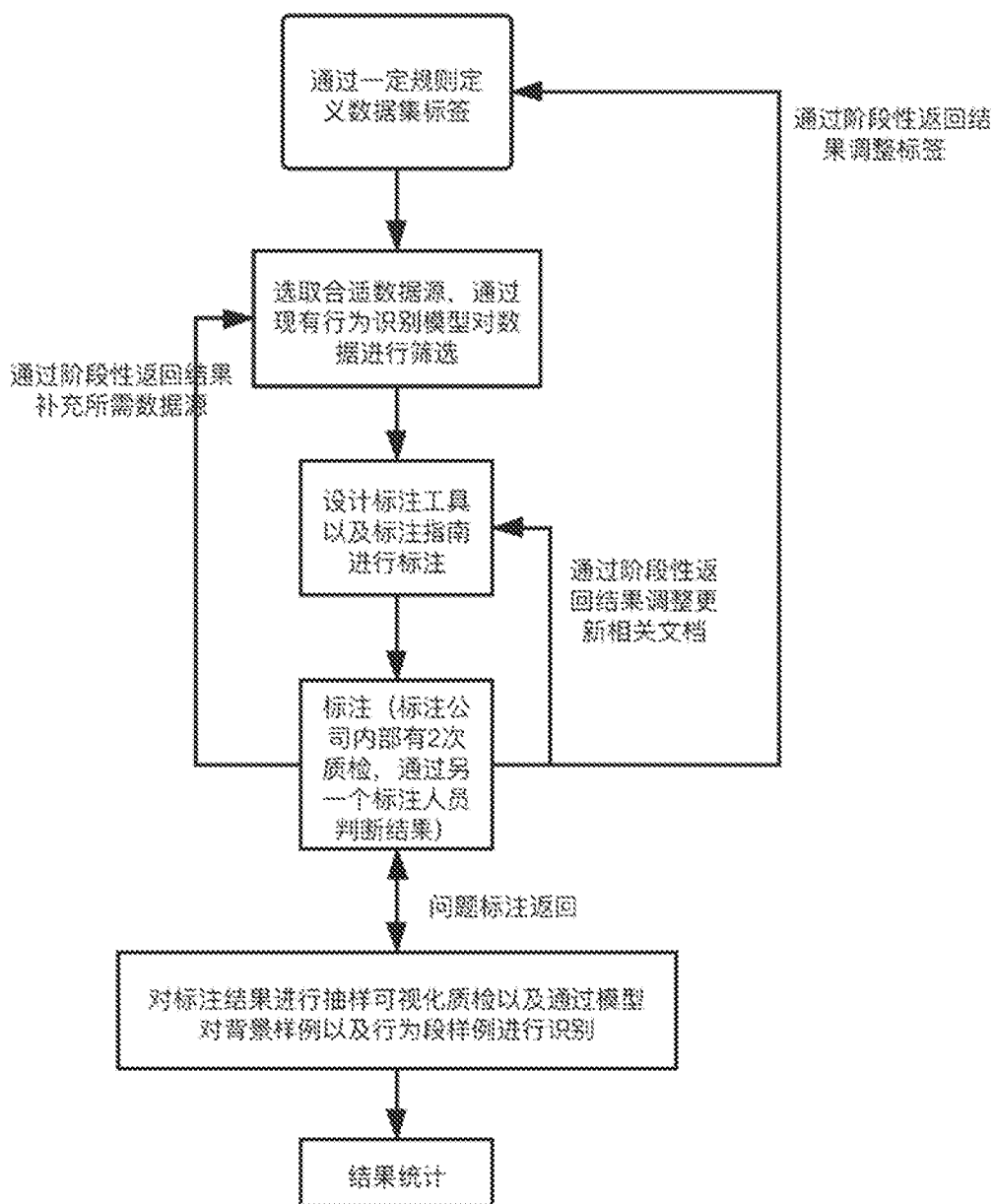


图 2



图 3

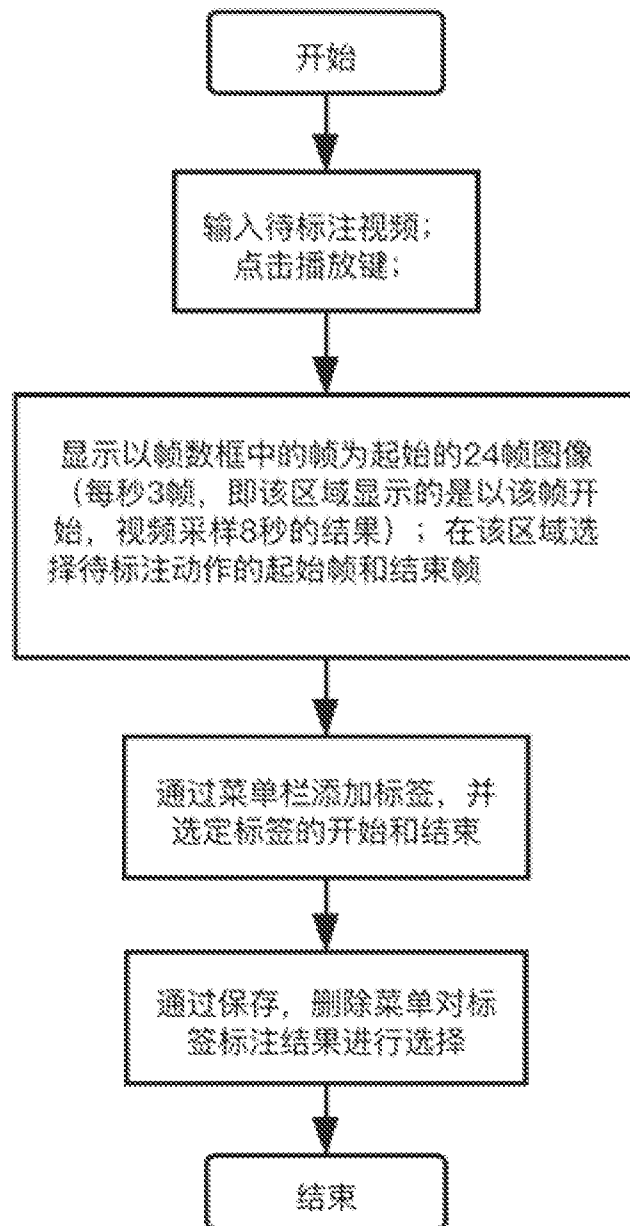


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/137579

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/78(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTEXT; CNABS; DWPI; SIPOABS; CNKI: 视频, 标签, 标注, 打标, 工具, 平台, 软件, 人工, 手动, 打标, 动作, 行为, 运动, 开始, 起始, 结束, 停止, 边界, 起止, 众包, 外包, video, annotat+, label+, tag+, software, platform, manual, artificial, action, behavior, movement, start, begin, end, finish, stop, boundary, crowdsourcing, epiboly

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113139096 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 20 July 2021 (2021-07-20) claims 1-10	1-10
X	CN 112101297 A (HANGZHOU HIKVISION DIGITAL TECHNOLOGY CO., LTD.) 18 December 2020 (2020-12-18) description, paragraphs 123-145, 155-167, 212-245, and 260-287	1-10
A	CN 110457494 A (XINHUA ZHIYUN TECHNOLOGY CO., LTD.) 15 November 2019 (2019-11-15) description, paragraphs 3-56 and 67-207	1-10
A	CN 110996138 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 10 April 2020 (2020-04-10) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 “&” document member of the same patent family

Date of the actual completion of the international search

25 February 2022

Date of mailing of the international search report

07 March 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/137579

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113139096	A	20 July 2021	None			
CN	112101297	A	18 December 2020	None			
CN	110457494	A	15 November 2019	None			
CN	110996138	A	10 April 2020	WO	2021120814	A1	24 June 2021

国际检索报告

国际申请号

PCT/CN2021/137579

A. 主题的分类 G06F 16/78 (2019.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																	
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNTXT; CNABS; DWPI; SIPOABS; CNKI: 视频, 标签, 标注, 打标, 工具, 平台, 软件, 人工, 手动, 打标, 动作, 行为, 运动, 开始, 起始, 结束, 停止, 边界, 起止, 众包, 外包, video, annotat+, label+, tag+, software, platform, manual, artificial, action, behavior, movement, start, begin, end, finish, stop, boundary, crowdsourcing, epiboly																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 113139096 A (中国科学院深圳先进技术研究院) 2021年7月20日 (2021 - 07 - 20) 权利要求1-10</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>CN 112101297 A (杭州海康威视数字技术股份有限公司) 2020年12月18日 (2020 - 12 - 18) 说明书第123-145, 155-167, 212-245, 260-287段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 110457494 A (新华智云科技有限公司) 2019年11月15日 (2019 - 11 - 15) 说明书第3-56, 67-207段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 110996138 A (腾讯科技深圳有限公司) 2020年4月10日 (2020 - 04 - 10) 全文</td> <td>1-10</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 113139096 A (中国科学院深圳先进技术研究院) 2021年7月20日 (2021 - 07 - 20) 权利要求1-10	1-10	X	CN 112101297 A (杭州海康威视数字技术股份有限公司) 2020年12月18日 (2020 - 12 - 18) 说明书第123-145, 155-167, 212-245, 260-287段	1-10	A	CN 110457494 A (新华智云科技有限公司) 2019年11月15日 (2019 - 11 - 15) 说明书第3-56, 67-207段	1-10	A	CN 110996138 A (腾讯科技深圳有限公司) 2020年4月10日 (2020 - 04 - 10) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
PX	CN 113139096 A (中国科学院深圳先进技术研究院) 2021年7月20日 (2021 - 07 - 20) 权利要求1-10	1-10															
X	CN 112101297 A (杭州海康威视数字技术股份有限公司) 2020年12月18日 (2020 - 12 - 18) 说明书第123-145, 155-167, 212-245, 260-287段	1-10															
A	CN 110457494 A (新华智云科技有限公司) 2019年11月15日 (2019 - 11 - 15) 说明书第3-56, 67-207段	1-10															
A	CN 110996138 A (腾讯科技深圳有限公司) 2020年4月10日 (2020 - 04 - 10) 全文	1-10															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件													
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																
国际检索实际完成的日期		国际检索报告邮寄日期															
2022年2月25日		2022年3月7日															
ISA/CN的名称和邮寄地址		受权官员															
中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		徐琳 电话号码 (86-10)62089885															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/137579

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	113139096	A	2021年7月20日	无			
CN	112101297	A	2020年12月18日	无			
CN	110457494	A	2019年11月15日	无			
CN	110996138	A	2020年4月10日	WO	2021120814	A1	2021年6月24日