



(12) 发明专利

(10) 授权公告号 CN 111968624 B

(45) 授权公告日 2024. 02. 09

(21) 申请号 202010859302.X

G10L 15/183 (2013.01)

(22) 申请日 2020.08.24

G10L 15/22 (2006.01)

(65) 同一申请的已公布的文献号

G10L 15/26 (2006.01)

申请公布号 CN 111968624 A

G10L 25/87 (2013.01)

G06N 3/0464 (2023.01)

(43) 申请公布日 2020.11.20

G06F 40/216 (2020.01)

G06F 9/30 (2018.01)

(73) 专利权人 平安科技(深圳)有限公司

地址 518000 广东省深圳市福田区福田街
道福安社区益田路5033号平安金融中
心23楼

(56) 对比文件

CN 108763068 A, 2018.11.06

CN 111444723 A, 2020.07.24

CN 107657950 A, 2018.02.02

CN 107783960 A, 2018.03.09

CN 110276054 A, 2019.09.24

CN 110784603 A, 2020.02.11

CN 111209729 A, 2020.05.29

US 2004261016 A1, 2004.12.23

US 2010254521 A1, 2010.10.07

(72) 发明人 张涛 姚璐娉 郝正鸿 叶洋甫
石佳

(74) 专利代理机构 深圳市赛恩倍吉知识产权代
理有限公司 44334

专利代理人 刘丽华

审查员 王佳玉

(51) Int. Cl.

G10L 15/02 (2006.01)

G10L 15/16 (2006.01)

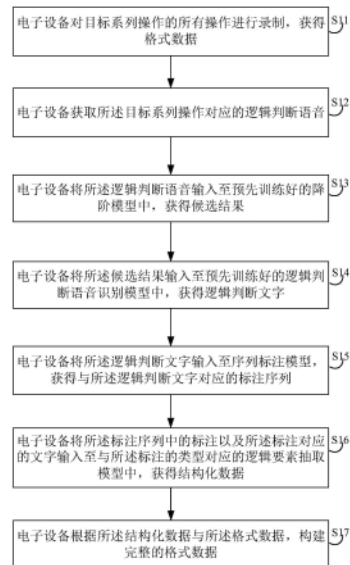
权利要求书2页 说明书12页 附图2页

(54) 发明名称

数据构建方法、装置、电子设备及存储介质

(57) 摘要

本申请涉及人工智能领域,提供一种数据构建方法,所述方法包括:对目标系列操作的所有操作进行录制,获得格式数据;获取目标系列操作对应的逻辑判断语音;将逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果;将候选结果输入至预先训练好的逻辑判断语音识别模型中,获得逻辑判断文字;将逻辑判断文字输入至序列标注模型,获得与逻辑判断文字对应的标注序列;将标注序列中的标注以及标注对应的文字输入至与标注的类型对应的逻辑要素抽取模型中,获得结构化数据;根据结构化数据与格式数据,构建完整的格式数据。本发明还提供一种数据构建装置、电子设备及存储介质。本发明能确保人工智能的操作的准确性。



1. 一种数据构建方法,其特征在于,所述方法包括:

获取逻辑判断场景的话术模板;

根据所述话术模板的关键词,对所述话术模板进行拓展,获得多个话术语料;

将所述多个话术语料确定为专用语言样本集;

获取自然语言样本集;

使用所述自然语言样本集进行训练,获得通用场景语音识别模型,以及使用所述专用语言样本集进行训练,获得操作场景语音识别模型;

根据预设测试集,对所述通用场景语音识别模型以及所述操作场景语音识别模型进行迭代,获得混合比例;

根据所述混合比例,将所述通用场景语音识别模型以及所述操作场景语音识别模型进行混合,获得逻辑判断语音识别模型;

对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应;

获取所述目标系列操作对应的逻辑判断语音;

对所述逻辑判断语音进行预处理,获得待识别语音;

对所述待识别语音进行端点检测,获得所述待识别语音的语音内容;

对所述语音内容进行分帧,以提取语音特征;将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果,包括:将所述语音特征输入至预先训练好的降阶模型中,获得候选结果;

将所述候选结果输入至所述逻辑判断语音识别模型中,获得逻辑判断文字;

将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列,所述序列标注模型包括双向长短期记忆循环网络以及条件随机场;

将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据,所述逻辑要素抽取模型包括双向长短期记忆循环网络以及条件随机场;

根据所述结构化数据与所述格式数据,构建完整的格式数据,包括:根据预设的映射规则,将所述结构化数据映射为对应的逻辑判断数据;将所述逻辑判断数据与所述格式数据进行组合,获得完整的格式数据。

2. 根据权利要求1所述的数据构建方法,其特征在于,所述将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果之前,所述数据构建方法还包括:

对所述逻辑判断语音识别模型进行裁剪,获得裁剪模型;

对所述裁剪模型进行降阶,获得所述降阶模型。

3. 一种数据构建装置,其特征在于,所述数据构建装置包括:

获取模块,用于获取逻辑判断场景的话术模板;

拓展模块,用于根据所述话术模板的关键词,对所述话术模板进行拓展,获得多个话术语料;

确定模块,用于将所述多个话术语料确定为专用语言样本集;

所述获取模块,还用于获取自然语言样本集;

训练模块,用于使用所述自然语言样本集进行训练,获得通用场景语音识别模型,以及

使用所述专用语言样本集进行训练,获得操作场景语音识别模型;

迭代模块,用于根据预设测试集,对所述通用场景语音识别模型以及所述操作场景语音识别模型进行迭代,获得混合比例;

混合模块,用于根据所述混合比例,将所述通用场景语音识别模型以及所述操作场景语音识别模型进行混合,获得逻辑判断语音识别模型;

录制模块,用于对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应;

所述获取模块,还用于获取所述目标系列操作对应的逻辑判断语音;

预处理模块,用于对所述逻辑判断语音进行预处理,获得待识别语音;

检测模块,用于对所述待识别语音进行端点检测,获得所述待识别语音的语音内容;

分帧模块,用于对所述语音内容进行分帧,以提取语音特征;

输入模块,用于将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果,包括:将所述语音特征输入至预先训练好的降阶模型中,获得候选结果;

所述输入模块,还用于将所述候选结果输入至所述逻辑判断语音识别模型中,获得逻辑判断文字;

所述输入模块,还用于将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列,所述序列标注模型包括双向长短期记忆循环网络以及条件随机场;

所述输入模块,还用于将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据,所述逻辑要素抽取模型包括双向长短期记忆循环网络以及条件随机场;

构建模块,用于根据所述结构化数据与所述格式数据,构建完整的格式数据,包括:根据预设的映射规则,将所述结构化数据映射为对应的逻辑判断数据;将所述逻辑判断数据与所述格式数据进行组合,获得完整的格式数据。

4. 一种电子设备,其特征在于,所述电子设备包括处理器和存储器,所述处理器用于执行存储器中存储的计算机程序以实现如权利要求1至2中任意一项所述的数据构建方法。

5. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有至少一个指令,所述至少一个指令被处理器执行时实现如权利要求1至2中任意一项所述的数据构建方法。

数据构建方法、装置、电子设备及存储介质

技术领域

[0001] 本发明涉及人工智能技术领域,尤其涉及一种数据构建方法、装置、电子设备及存储介质。

背景技术

[0002] 目前,随着人工智能技术的发展,可以通过录制人为的操作并生成计算机能够识别的格式数据,根据格式数据,通过程序自动执行这些操作,从而可以实现自动化地完成一部分操作,使得机器可以胜任部分原本由人来处理的工作。但在实践中发现,在需要进行逻辑判断的操作中,录制工具无法生成关于逻辑判断的格式数据,进而无法实现逻辑判断的功能,使得操作出现错误,也就无法完成需要进行逻辑判断的工作任务。

[0003] 因此,如何构建完整的格式数据以确保人工智能的操作的准确性是一个亟需解决的技术问题。

发明内容

[0004] 鉴于以上内容,有必要提供一种数据构建方法、装置、电子设备及存储介质,能够构建完整的格式数据以确保人工智能的操作的准确性。

[0005] 本发明的第一方面提供一种数据构建方法,所述方法包括:

[0006] 对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应;

[0007] 获取所述目标系列操作对应的逻辑判断语音;

[0008] 将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果;

[0009] 将所述候选结果输入至预先训练好的逻辑判断语音识别模型中,获得逻辑判断文字,其中,所述逻辑判断语音识别模型由通用场景语音识别模型以及操作场景语音识别模型混合获得;

[0010] 将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列;

[0011] 将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据;

[0012] 根据所述结构化数据与所述格式数据,构建完整的格式数据。

[0013] 在一种可能的实现方式中,所述将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果之前,所述方法还包括:

[0014] 对所述逻辑判断语音识别模型进行裁剪,获得裁剪模型;

[0015] 对所述裁剪模型进行降阶,获得所述降阶模型。

[0016] 在一种可能的实现方式中,所述对目标系列操作的所有操作进行录制,获得格式数据之前,所述方法还包括:

[0017] 获取自然语言样本集以及获取专用语言样本集;

- [0018] 使用自然语言样本进行训练,获得通用场景语音识别模型,以及使用专用语言样本进行训练,获得操作场景语音识别模型;
- [0019] 根据预设测试集,对所述通用场景语音识别模型以及所述操作场景语音识别模型进行迭代,获得混合比例;
- [0020] 根据所述混合比例,将所述通用场景语音识别模型以及所述操作场景语音识别模型进行混合,获得所述逻辑判断语音识别模型。
- [0021] 在一种可能的实现方式中,所述获取自然语言样本集以及获取专用语言样本集之前,所述方法还包括:
- [0022] 获取逻辑判断场景的话术模板;
- [0023] 根据所述话术模板的关键词,对所述话术模板进行拓展,获得多个话术语料;
- [0024] 将所述多个话术语料确定为所述专用语言样本集。
- [0025] 在一种可能的实现方式中,所述获取所述目标系列操作对应的逻辑判断语音之后,所述方法还包括:
- [0026] 对所述逻辑判断语音进行预处理,获得待识别语音;
- [0027] 对所述待识别语音进行端点检测,获得所述待识别语音的语音内容;
- [0028] 对所述语音内容进行分帧,以提取语音特征;
- [0029] 所述将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果包括:
- [0030] 将所述语音特征输入至预先训练好的降阶模型中,获得候选结果。
- [0031] 在一种可能的实现方式中,所述根据所述结构化数据与所述格式数据,构建完整的格式数据包括:
- [0032] 根据预设的映射规则,将所述结构化数据映射为对应的逻辑判断数据;
- [0033] 将所述逻辑判断数据与所述格式数据进行组合,获得完整的格式数据。
- [0034] 在一种可能的实现方式中,所述序列标注模型包括双向长短期记忆循环网络以及条件随机场,所述逻辑要素抽取模型包括双向长短期记忆循环网络以及条件随机场。
- [0035] 本发明的第二方面提供一种数据构建装置,所述装置包括:
- [0036] 录制模块,用于对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应;
- [0037] 获取模块,用于获取所述目标系列操作对应的逻辑判断语音;
- [0038] 输入模块,用于将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果;
- [0039] 所述输入模块,还用于将所述候选结果输入至预先训练好的逻辑判断语音识别模型中,获得逻辑判断文字,其中,所述逻辑判断语音识别模型由通用场景语音识别模型以及操作场景语音识别模型混合获得;
- [0040] 所述输入模块,还用于将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列;
- [0041] 所述输入模块,还用于将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据;
- [0042] 构建模块,用于根据所述结构化数据与所述格式数据,构建完整的格式数据。
- [0043] 本发明的第三方面提供一种电子设备,所述电子设备包括处理器和存储器,所述

处理器用于执行所述存储器中存储的计算机程序时实现所述的数据构建方法。

[0044] 本发明的第四方面提供一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器执行时实现所述的数据构建方法。

[0045] 由以上技术方案,本发明中,可以通过逻辑判断语音识别模型,将逻辑判断语音转换为文本,经过列标注模型以及逻辑要素抽取模型将所述文本转换为结构数据,结构数据可以映射为对应的逻辑判断信息,可以和提前录制好的格式数据组合构建完整的工作格式数据,可以确保人工智能的操作的准确性。

附图说明

[0046] 图1是本发明公开的一种数据构建方法的较佳实施例的流程图。

[0047] 图2是本发明公开的一种数据构建装置的较佳实施例的功能模块图。

[0048] 图3是本发明实现数据构建方法的较佳实施例的电子设备的结构示意图。

具体实施方式

[0049] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0050] 除非另有定义,本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。本文中在本发明的说明书中所使用的术语只是为了描述具体的实施例的目的,不是旨在于限制本发明。

[0051] 本发明实施例的数据构建方法应用在电子设备中,也可以应用在电子设备和通过网络与所述电子设备进行连接的服务器所构成的硬件环境中,由服务器和电子设备共同执行。网络包括但不限于:广域网、城域网或局域网。

[0052] 其中,服务器可以是指能对网络中其它设备(如电子设备)提供服务的计算机系统。如果一个个人电脑能够对外提供文件传输协议(File Transfer Protocol,简称FTP)服务,也可以叫服务器。从狭义范围上讲,服务器专指某些高性能计算机,能通过网络,对外提供服务,其相对于普通的个人电脑来说,稳定性、安全性、性能等方面都要求更高,因此在CPU、芯片组、内存、磁盘系统、网络等硬件和普通的个人电脑有所不同。

[0053] 所述电子设备是一种能够按照事先设定或存储的指令,自动进行数值计算和/或信息处理的设备,其硬件包括但不限于微处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、数字处理器(DSP)、嵌入式设备等。所述电子设备还可包括网络设备和/或用户设备。其中,所述网络设备包括但不限于单个网络设备、多个网络设备组成的服务器组或基于云计算(Cloud Computing)的由大量主机或网络设备构成的云,其中,云计算是分布式计算的一种,由一群松散耦合的计算机集组成的一个超级虚拟计算机。所述用户设备包括但不限于任何一种可与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互的电子产品,例如,个人计算机、平板电脑、智能手机、个人数字助理PDA等。

[0054] 请参见图1,图1是本发明公开的一种数据构建方法的较佳实施例的流程图。其中,根据不同的需求,该流程图中步骤的顺序可以改变,某些步骤可以省略。

[0055] S11、电子设备对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应。

[0056] 其中,所述格式数据可以通过AI-RPA(Robotic Process Automation,机器人流程自动化)软件录制获得,所述格式数据能够被计算机识别,可以为可编译的计算机语言数据。

[0057] 其中,所述目标系列操作可以包括多个操作,比如,鼠标点击文本框,在文本框输入文本,鼠标点击登录按钮等。

[0058] 本发明实施例中,可以将机器人流程自动化中需要用到的操作都进行录制,每个操作经过录制后,会生成对应的格式数据,能够被识别、执行以实现录制的操作,使得机器能够完成这部分被录制的由人进行的操作。

[0059] S12、电子设备获取所述目标系列操作对应的逻辑判断语音。

[0060] 其中,可以针对需要进行逻辑判断的操作录制语音信息,在使用系统时,有些地方需要根据不同的情况,做出不同的操作。比如:当……时,执行……,否则执行……;当……时,循环执行……;当xxx的时候,搜索“xxx”,否则,关闭浏览器等。

[0061] 本发明实施例中,这些需要进行逻辑判断的地方无法录制对应的格式数据,因为需要进行逻辑判断的地方存在分支,即需要选择进行不同的操作。只能录制进行选择后的操作,而逻辑判断本身不是操作的一部分,无法录制,可以录音把逻辑判断语音录制下来,比如,若a大于b,执行操作1,否则,执行操作2。所述逻辑判断语音可以所述目标系列操作的录制工作完成后进行录制。

[0062] S13、电子设备将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果。

[0063] 其中,所述降阶模型是在预先训练好的逻辑判断语音识别模型的基础上获得的网络复杂度较小,处理速度较快的模型。所述候选结果可以是所述降阶模型判定的概率较大的多个语句。

[0064] 作为一种可选的实施方式,所述将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果之前,所述方法还包括:

[0065] 对所述逻辑判断语音识别模型进行裁剪,获得裁剪模型;

[0066] 对所述裁剪模型进行降阶,获得所述降阶模型。

[0067] 在该可选的实施方式中,可以对所述逻辑判断语音识别模型进行相对熵裁剪,获得相对于所述逻辑判断语音识别模型较小的模型,即所述裁剪模型,在基于相对熵裁剪的基础上,同时降低模型的阶数,将模型从4-gram降为2-gram,从而可以减少最终解码网络的复杂度,保证第一遍解码的效率。即提高了获得所述候选结果的速度。

[0068] S14、电子设备将所述候选结果输入至预先训练好的逻辑判断语音识别模型中,获得逻辑判断文字,其中,所述逻辑判断语音识别模型由通用场景语音识别模型以及操作场景语音识别模型混合获得。

[0069] 其中,所述通用语音识别模型可以使用自然语言样本训练得到,所述操作场景语音识别模型可以使用机器人流程自动化操作场景的专业术语对应的语音样本训练得到,语言的模型的原理是计算某个词在其前面各个词出现的条件下出现的概率,然后选取各个位置上概率最大的词组成句子。

[0070] 本发明实施例中,在获得了候选结果后,可以使用完整的逻辑判断语音识别模型对所述候选结果进行重新打分排序,输出最优的识别结果,该方法可以在不降低语音识别的效率的情况下,提高识别的准确性,提高了人工智能处理语音信息的准确率。

[0071] 其中,为了保证逻辑判断语音识别模型具有一定的通用识别能力,即当用户没有严格按照设定话术表达时的语音识别能力,将通用场景语音识别模型以及操作场景语音识别模型按一定的比例插值混合后生成最终的逻辑判断语音识别语言模型,提高了逻辑判断语音识别模型的鲁棒性。

[0072] 作为一种可选的实施方式,所述获取所述目标系列操作对应的逻辑判断语音之后,所述方法还包括:

[0073] 对所述逻辑判断语音进行预处理,获得待识别语音;

[0074] 对所述待识别语音进行端点检测,获得所述待识别语音的语音内容;

[0075] 对所述语音内容进行分帧,以提取语音特征;

[0076] 所述将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果包括:

[0077] 将所述语音特征输入至预先训练好的降阶模型中,获得候选结果。

[0078] 在该可选的实施方式中,在获得逻辑判断语音后,可以对所述逻辑判断语音进行滤波降噪、预加重等预处理,识别所述逻辑判断语音达到被识别的要求,然后通过端点检测,过滤掉非语言部分,再对剩下的语音部分进行分帧,提取除语音特征,最后将通过所述预先训练好的降阶模型,获得与所述语音特征对应的候选结果,即对应所述语音特征,模型判定概率比较大的一些文字。

[0079] 作为一种可选的实施方式,所述对目标系列操作的所有操作进行录制,获得格式数据之前,所述方法还包括:

[0080] 获取自然语言样本集以及获取专用语言样本集;

[0081] 使用自然语言样本进行训练,获得通用场景语音识别模型,以及使用专用语言样本进行训练,获得操作场景语音识别模型;

[0082] 根据预设测试集,对所述通用场景语音识别模型以及所述操作场景语音识别模型进行迭代,获得混合比例;

[0083] 根据所述混合比例,将所述通用场景语音识别模型以及所述操作场景语音识别模型进行混合,获得所述逻辑判断语音识别模型。

[0084] 在该可选的实施方式中,所述通用场景语音识别模型和所述操作场景语音识别模型的混合比例不是采用等比例混合或是使用经验值,而是通过在预设测试集(比如包含机器人流程自动化专用语料和通用语料的测试集)上不断迭代的方法,找出一个可以使混合后的模型在目标测试集上困惑度(perplexity)最小的比例,这样可以保证模型在机器人流程自动化场景和通用场景均有较好的识别性能,提高语音识别的准确率。

[0085] 作为一种可选的实施方式,所述获取自然语言样本集以及获取专用语言样本集之前,所述方法还包括:

[0086] 获取逻辑判断场景的话术模板;

[0087] 根据所述话术模板的关键词,对所述话术模板进行拓展,获得多个话术语料;

[0088] 将所述多个话术语料确定为所述专用语言样本集。

[0089] 其中,所述话术模板可以是话术1:增加流程开关,假如 $a > b$,则执行步骤1,否则执

行步骤2;话术2:增加流程开关,如 $a>b$,是则执行步骤1,否则执行步骤2等话术,针对每句话术,可以对其中的关键词的排列组合以及相似词语的替换来拓展生成多句话,即生成所述话术语料。

[0090] 本发明实施例中,可以预先收集一些带有专业术语的语句来作为话术模板,然后,针对话术模板中的每句话术,可以对其中的关键词的排列组合以及相似词语的替换来拓展生成多句话,即话术语料,比如“增加流程开关,假如 $a>b$,则执行步骤1,否则执行步骤2”可以拓展出“新增流程,若 $a<b$,则执行步骤2,否则执行步骤1”。

[0091] S15、电子设备将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列。

[0092] 其中,所述序列标注模型包括双向长短期记忆循环网络(Bi-directional Long Short-Term Memory,BILSTM)以及条件随机场(conditional random field,CRF)。

[0093] 其中,双向长短期记忆循环网络(BILSTM)是由前向的长短期记忆网络(LSTM,Long Short-Term Memory)和后向的长短期记忆网络结合成。所述双向长短期记忆循环网络能够在原来的正向运算上,把输入的序列反转,重新计算,实现反向运算,然后将正向运算的结果与反向运算的结果简单堆叠,可以实现考虑上下文信息。

[0094] 其中,所述长短期记忆网络是一种特殊的循环神经网络(Recurrent Neural Network,RNN),主要是为了解决长序列训练过程中的梯度消失和梯度爆炸问题,即相比于普通的循环神经网络,所述长短期记忆网络能够在更长的序列中有更好的表现。

[0095] 其中,循环神经网络是一类是一类以序列数据为输入,在序列的演进方向进行递归且所有节点按链式连接的递归神经网络,可以用于处理序列数据。

[0096] 其中,所述条件随机场(conditional random field,CRF)是一种鉴别式机率模型,是随机场的一种,可以用于标注或分析序列资料,如自然语言文字或是生物序列。

[0097] 本发明实施例中,所述标注序列是为了让计算机可以理解所述逻辑判断文字,将一些特定的字符按照预先规定的规则进行组合而获得的字符串,能够用来表示所述逻辑判断文字。比如:增加开关,假如 a 大于 b ,是则执行步骤一,否则执行步骤二。转换成标注序列:0,0,0,0,0,0,0,B-CONDITION,I-CONDITION,I-CONDITION,E-CONDITION,0,0,0,0,0,B-THENDO,I-THENDO,E-THENDO,0,0,0,0,0,B-ELSEDO,I-ELSEDO,E-ELSEDO。

[0098] S16、电子设备将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据。

[0099] 其中,所述逻辑要素抽取模型包括双向长短期记忆循环网络以及条件随机场,所述结构化数据也是由一个个标注构成。

[0100] 其中,双向长短期记忆循环网络(BILSTM)是由前向的长短期记忆网络(LSTM,Long Short-Term Memory)和后向的长短期记忆网络结合成。

[0101] 其中,比如标注序列:B-CONDITION,I-CONDITION,I-CONDITION,E-CONDITION,0,0,0,0,0,B-THENDO,I-THENDO,E-THENDO,0,0,0,0,0,B-ELSEDO,I-ELSEDO,E-ELSEDO以及对应的文字:增加开关,假如 a 大于 b ,是则执行步骤一,否则执行步骤二,转换成的结构化数据为:{'if_else':{'condition':' $a>b$ ','then_do':['1'],'else_do':['2']},'type':['1']}}。

[0102] S17、电子设备根据所述结构化数据与所述格式数据,构建完整的格式数据。

[0103] 具体的,所述根据所述结构化数据与所述所有格式数据,构建完整的格式数据包括:

[0104] 根据预设的映射规则,将所述结构化数据映射为对应的逻辑判断数据;

[0105] 将所述逻辑判断数据与所述格式数据进行组合,获得完整的格式数据。

[0106] 在该可选的实施方式中,可以预先实现好各种各样的逻辑判断数据,比如条件判断逻辑、循环执行逻辑等,根据结构化数据,确定与所述结构化数据对应的逻辑判断数据,比如结构化数据'if_else'对应条件判断逻辑。将所述逻辑判断数据与所述格式数据进行组合,获得的完整的格式数据,可以实现系统的自动化工作,并可以按照逻辑判断语音所设置的逻辑进行操作,使得人工智能可以完成更复杂的工作。

[0107] 在图1所描述的方法流程中,可以通过逻辑判断语音识别模型,将逻辑判断语音转换为文本,经过列标注模型以及逻辑要素抽取模型将所述文本转换为结构数据,结构数据可以映射为对应的逻辑判断信息,可以和提前录制好的格式数据组合构建完整的工作格式数据,可以确保人工智能的操作的准确性。

[0108] 图2是本发明公开的一种数据构建装置的较佳实施例的功能模块图。

[0109] 请参见图2,所述数据构建装置20运行于电子设备中。所述数据构建装置20可以包括多个由程序代码段所组成的功能模块。所述数据构建装置20中的各个程序段的程序代码可以存储于存储器中,并由至少一个处理器所执行,以执行图1所描述的数据构建方法中的部分或全部步骤。

[0110] 本实施例中,所述数据构建装置20根据其所执行的功能,可以被划分为多个功能模块。所述功能模块可以包括:录制模块201、获取模块202、输入模块203及构建模块204。本发明所称的模块是指一种能够被至少一个处理器所执行并且能够完成固定功能的一系列计算机程序段,其存储在存储器中。

[0111] 录制模块201,用于对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应。

[0112] 其中,所述格式数据可以通过AI-RPA(Robotic Process Automation,机器人流程自动化)软件录制获得,所述格式数据能够被计算机识别,可以为可编译的计算机语言数据。

[0113] 其中,所述目标系列操作可以包括多个操作,比如,鼠标点击文本框,在文本框输入文本,鼠标点击登录按钮等。

[0114] 本发明实施例中,可以将机器人流程自动化中需要用到的操作都进行录制,每个操作经过录制后,会生成对应的格式数据,能够被识别、执行以实现录制的操作,使得机器能够完成这部分被录制的由人进行的操作。

[0115] 获取模块202,用于获取所述目标系列操作对应的逻辑判断语音。

[0116] 其中,可以针对需要进行逻辑判断的操作录制语音信息,在使用系统时,有些地方需要根据不同的情况,做出不同的操作。比如:当……时,执行……,否则执行……;当……时,循环执行……;当xxx的时候,搜索“xxx”,否则,关闭浏览器等。

[0117] 本发明实施例中,这些需要进行逻辑判断的地方无法录制对应的格式数据,因为需要进行逻辑判断的地方存在分支,即需要选择进行不同的操作。只能录制进行选择后的操作,而逻辑判断本身不是操作的一部分,无法录制,可以录音把逻辑判断语音录制下来,

比如,若a大于b,执行操作1,否则,执行操作2。所述逻辑判断语音可以所述目标系列操作的录制工作完成后进行录制。

[0118] 输入模块203,用于将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果。

[0119] 其中,所述降阶模型是在预先训练好的逻辑判断语音识别模型的基础上获得的网络复杂度较小,处理速度较快的模型。所述候选结果可以是所述降阶模型判定的概率较大的多个语句。

[0120] 所述输入模块203,还用于将所述候选结果输入至预先训练好的逻辑判断语音识别模型中,获得逻辑判断文字,其中,所述逻辑判断语音识别模型由通用场景语音识别模型以及操作场景语音识别模型混合获得。

[0121] 其中,所述通用语音识别模型可以使用自然语言样本训练得到,所述操作场景语音识别模型可以使用机器人流程自动化操作场景的专业术语对应的语音样本训练得到,语言的模型的原理是计算某个词在其前面各个词出现的条件下出现的概率,然后选取各个位置上概率最大的词组成句子。

[0122] 本发明实施例中,在获得了候选结果后,可以使用完整的逻辑判断语音识别模型对所述候选结果进行重新打分排序,输出最优的识别结果,该方法可以在不降低语音识别的效率的情况下,提高识别的准确性,提高了人工智能处理语音信息的准确率。

[0123] 其中,为了保证逻辑判断语音识别模型具有一定的通用识别能力,即当用户没有严格按照设定话术表达时的语音识别能力,将通用场景语音识别模型以及操作场景语音识别模型按一定的比例插值混合后生成最终的逻辑判断语音识别语言模型,提高了逻辑判断语音识别模型的鲁棒性。

[0124] 所述输入模块203,还用于将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列。

[0125] 其中,所述序列标注模型包括双向长短期记忆循环网络(Bi-directional Long Short-Term Memory,BILSTM)以及条件随机场(conditional random field,CRF)。

[0126] 其中,双向长短期记忆循环网络(BILSTM)是由前向的长短期记忆网络(LSTM,Long Short-Term Memory)和后向的长短期记忆网络结合成。所述双向长短期记忆循环网络能够在原来的正向运算上,把输入的序列反转,重新计算,实现反向运算,然后将正向运算的结果与反向运算的结果简单堆叠,可以实现考虑上下文信息。

[0127] 其中,所述长短期记忆网络是一种特殊的循环神经网络(Recurrent Neural Network,RNN),主要是为了解决长序列训练过程中的梯度消失和梯度爆炸问题,即相比于普通的循环神经网络,所述长短期记忆网络能够在更长的序列中有更好的表现。

[0128] 其中,循环神经网络是是一类是一类以序列数据为输入,在序列的演进方向进行递归且所有节点按链式连接的递归神经网络,可以用于处理序列数据。

[0129] 其中,所述条件随机场(conditional random field,CRF)是一种鉴别式机率模型,是随机场的一种,可以用于标注或分析序列资料,如自然语言文字或是生物序列。

[0130] 本发明实施例中,所述标注序列是为了让计算机可以理解所述逻辑判断文字,将一些特定的字符按照预先规定的规则进行组合而获得的字符串,能够用来表示所述逻辑判断文字。比如:增加开关,假如a大于b,是则执行步骤一,否则执行步骤二。转换成标注序列:

0,0,0,0,0,0,0,0,B-CONDITION,I-CONDITION,I-CONDITION,E-CONDITION,0,0,0,0,0,0,B-THENDO,I-THENDO,E-THENDO,0,0,0,0,0,0,B-ELSEDO,I-ELSEDO,E-ELSEDO。

[0131] 所述输入模块203,还用于将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据。

[0132] 其中,所述逻辑要素抽取模型包括双向长短期记忆循环网络以及条件随机场,所述结构化数据也是由一个个标注构成。

[0133] 其中,双向长短期记忆循环网络(BILSTM)是由前向的长短期记忆网络(LSTM,Long Short-Term Memory)和后向的长短期记忆网络结合成。

[0134] 其中,比如标注序列:B-CONDITION,I-CONDITION,I-CONDITION,E-CONDITION,0,0,0,0,0,0,B-THENDO,I-THENDO,E-THENDO,0,0,0,0,0,0,B-ELSEDO,I-ELSEDO,E-ELSEDO以及对应的文字:增加开关,假如a大于b,是则执行步骤一,否则执行步骤二,转换成的结构化数据为:{'if_else':{'condition':'a>b,', 'then_do':['1'],'else_do':['2']},'type':[1]}。

[0135] 构建模块204,用于根据所述结构化数据与所述格式数据,构建完整的格式数据。

[0136] 作为一种可选的实施方式,所述数据构建装置20还可以包括:

[0137] 裁剪模块,用于所述输入模块203将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果之前,对所述逻辑判断语音识别模型进行裁剪,获得裁剪模型;

[0138] 降阶模块,用于对所述裁剪模型进行降阶,获得所述降阶模型。

[0139] 在该可选的实施方式中,可以对所述逻辑判断语音识别模型进行相对熵裁剪,获得相对于所述逻辑判断语音识别模型较小的模型,即所述裁剪模型,在基于相对熵裁剪的基础上,同时降低所述裁剪模型的阶数,将模型从4-gram降为2-gram,从而可以减少最终解码网络的复杂度,保证第一遍解码的效率。即提高了获得所述候选结果的速度。

[0140] 作为一种可选的实施方式,所述获取模块201,还用于所述录制模块201对目标系列操作的所有操作进行录制,获得格式数据之前,获取自然语言样本集以及获取专用语言样本集;

[0141] 所述数据构建装置20还可以包括:

[0142] 训练模块,用于使用自然语言样本进行训练,获得通用场景语音识别模型,以及使用专用语言样本进行训练,获得操作场景语音识别模型;

[0143] 迭代模块,用于根据预设测试集,对所述通用场景语音识别模型以及所述操作场景语音识别模型进行迭代,获得混合比例;

[0144] 混合模块,用于根据所述混合比例,将所述通用场景语音识别模型以及所述操作场景语音识别模型进行混合,获得所述逻辑判断语音识别模型。

[0145] 在该可选的实施方式中,所述通用场景语音识别模型和所述操作场景语音识别模型的混合比例不是采用等比例混合或是使用经验值,而是通过在预设测试集(比如包含机器人流程自动化专用语料和通用语料的测试集)上不断迭代的方法,找出一个可以使混合后的模型在目标测试集上困惑度(perplexity)最小的比例,这样可以保证模型在机器人流程自动化场景和通用场景均有较好的识别性能,提高语音识别的准确率。

[0146] 作为一种可选的实施方式,所述获取模块202,还用于获取逻辑判断场景的话术模板;

[0147] 所述数据构建装置20还可以包括：

[0148] 拓展模块，用于根据所述话术模板的关键词，对所述话术模板进行拓展，获得多个话术语料；

[0149] 确定模块，用于将所述多个话术语料确定为所述专用语言样本集。

[0150] 其中，所述话术模板可以是话术1：增加流程开关，假如 $a > b$ ，则执行步骤1，否则执行步骤2；话术2：增加流程开关，如 $a > b$ ，是则执行步骤1，否则执行步骤2等话术，针对每句话术，可以对其中的关键词的排列组合以及相似词语的替换来拓展生成多句话，即生成所述话术语料。

[0151] 本发明实施例中，可以预先收集一些带有专业术语的语句来作为话术模板，然后，针对话术模板中的每句话术，可以对其中的关键词的排列组合以及相似词语的替换来拓展生成多句话，即话术语料，比如“增加流程开关，假如 $a > b$ ，则执行步骤1，否则执行步骤2”可以拓展出“新增流程，若 $a < b$ ，则执行步骤2，否则执行步骤1”。

[0152] 作为一种可选的实施方式，所述数据构建装置20还可以包括：

[0153] 预处理模块，用于所述获取模块202获取所述目标系列操作对应的逻辑判断语音之后，对所述逻辑判断语音进行预处理，获得待识别语音；

[0154] 检测模块，用于对所述待识别语音进行端点检测，获得所述待识别语音的语音内容；

[0155] 分帧模块，用于对所述语音内容进行分帧，以提取语音特征；

[0156] 所述输入模块203将所述逻辑判断语音输入至预先训练好的降阶模型中，获得候选结果的方式具体为：

[0157] 将所述语音特征输入至预先训练好的降阶模型中，获得候选结果。

[0158] 在该可选的实施方式中，在获得逻辑判断语音后，可以对所述逻辑判断语音进行滤波降噪、预加重等预处理，识别所述逻辑判断语音达到被识别的要求，然后通过端点检测，过滤掉非语言部分，再对剩下的语音部分进行分帧，提取除语音特征，最后将通过所述预先训练好的降阶模型，获得与所述语音特征对应的候选结果，即对应所述语音特征，模型判定概率比较大的一些文字。

[0159] 作为一种可选的实施方式，所述构建模块204根据所述结构化数据与所述格式数据，构建完整的格式数据的方式具体为：

[0160] 根据预设的映射规则，将所述结构化数据映射为对应的逻辑判断数据；

[0161] 将所述逻辑判断数据与所述格式数据进行组合，获得完整的格式数据。

[0162] 在该可选的实施方式中，可以预先实现好各种各样的逻辑判断数据，比如条件判断逻辑、循环执行逻辑等，根据结构化数据，确定与所述结构化数据对应的逻辑判断数据，比如结构化数据‘if_else’对应条件判断逻辑。将所述逻辑判断数据与所述格式数据进行组合，获得的完整的格式数据，可以实现系统的自动化工作，并可以按照逻辑判断语音所设置的逻辑进行操作，使得人工智能可以完成更复杂的工作。

[0163] 在图2所描述的数据构建装置20中，可以通过逻辑判断语音识别模型，将逻辑判断语音转换为文本，经过列标注模型以及逻辑要素抽取模型将所述文本转换为结构数据，结构数据可以映射为对应的逻辑判断信息，可以和提前录制好的格式数据组合构建完整的工作格式数据，可以确保人工智能的操作的准确性。

[0164] 如图3所示,图3是本发明实现数据构建方法的较佳实施例的电子设备的结构示意图。所述电子设备3包括存储器31、至少一个处理器32、存储在所述存储器31中并可在所述至少一个处理器32上运行的计算机程序33及至少一条通讯总线34。

[0165] 本领域技术人员可以理解,图3所示的示意图仅仅是所述电子设备3的示例,并不构成对所述电子设备3的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件,例如所述电子设备3还可以包括输入输出设备、网络接入设备等。

[0166] 所述电子设备3还包括但不限于任何一种可与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互的电子产品,例如,个人计算机、平板电脑、智能手机、个人数字助理(Personal Digital Assistant,PDA)、游戏机、交互式网络电视(Internet Protocol Television,IPTV)、智能式穿戴式设备等。所述电子设备3所处的网络包括但不限于互联网、广域网、城域网、局域网、虚拟专用网络(Virtual Private Network,VPN)等。

[0167] 所述至少一个处理器32可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、晶体管逻辑器件、分立硬件组件等。该处理器32可以是微处理器或者该处理器32也可以是任何常规的处理器等,所述处理器32是所述电子设备3的控制中心,利用各种接口和线路连接整个电子设备3的各个部分。

[0168] 所述存储器31可用于存储所述计算机程序33和/或模块/单元,所述处理器32通过运行或执行存储在所述存储器31内的计算机程序和/或模块/单元,以及调用存储在存储器31内的数据,实现所述电子设备3的各种功能。所述存储器31可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等等);存储数据区可存储根据电子设备3的使用所创建的数据等。此外,存储器31可以包括易失性和非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件等。

[0169] 结合图1,所述电子设备3中的所述存储器31存储多个指令以实现一种数据构建方法,所述处理器32可执行所述多个指令从而实现:

[0170] 对目标系列操作的所有操作进行录制,获得格式数据,其中,所述格式数据与所述操作一一对应;

[0171] 获取所述目标系列操作对应的逻辑判断语音;

[0172] 将所述逻辑判断语音输入至预先训练好的降阶模型中,获得候选结果;

[0173] 将所述候选结果输入至预先训练好的逻辑判断语音识别模型中,获得逻辑判断文字,其中,所述逻辑判断语音识别模型由通用场景语音识别模型以及操作场景语音识别模型混合获得;

[0174] 将所述逻辑判断文字输入至序列标注模型,获得与所述逻辑判断文字对应的标注序列;

[0175] 将所述标注序列中的标注以及所述标注对应的文字输入至与所述标注的类型对应的逻辑要素抽取模型中,获得结构化数据;

[0176] 根据所述结构化数据与所述格式数据,构建完整的格式数据。

[0177] 具体地,所述处理器32对上述指令的具体实现方法可参考图1对应实施例中相关步骤的描述,在此不赘述。

[0178] 在图3所描述的电子设备3中,可以通过逻辑判断语音识别模型,将逻辑判断语音转换为文本,经过列标注模型以及逻辑要素抽取模型将所述文本转换为结构数据,结构数据可以映射为对应的逻辑判断信息,可以和提前录制好的格式数据组合构建完整的工作格式数据,可以确保人工智能的操作的准确性。

[0179] 所述电子设备3集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明实现上述实施例方法中的全部或部分流程,也可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一计算机可读存储介质中,该计算机程序在被处理器执行时,可实现上述各个方法实施例的步骤。其中,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存储器(RAM,Random Access Memory)等。

[0180] 在本发明所提供的几个实施例中,应该理解到,所揭露的系统,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述模块的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式。

[0181] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0182] 另外,在本发明各个实施例中的各功能模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能模块的形式实现。

[0183] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外,显然“包括”一词不排除其他单元或步骤,单数不排除复数。系统权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第二等词语用来表示名称,而并不表示任何特定的顺序。

[0184] 最后应说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或等同替换,而不脱离本发明技术方案的精神和范围。

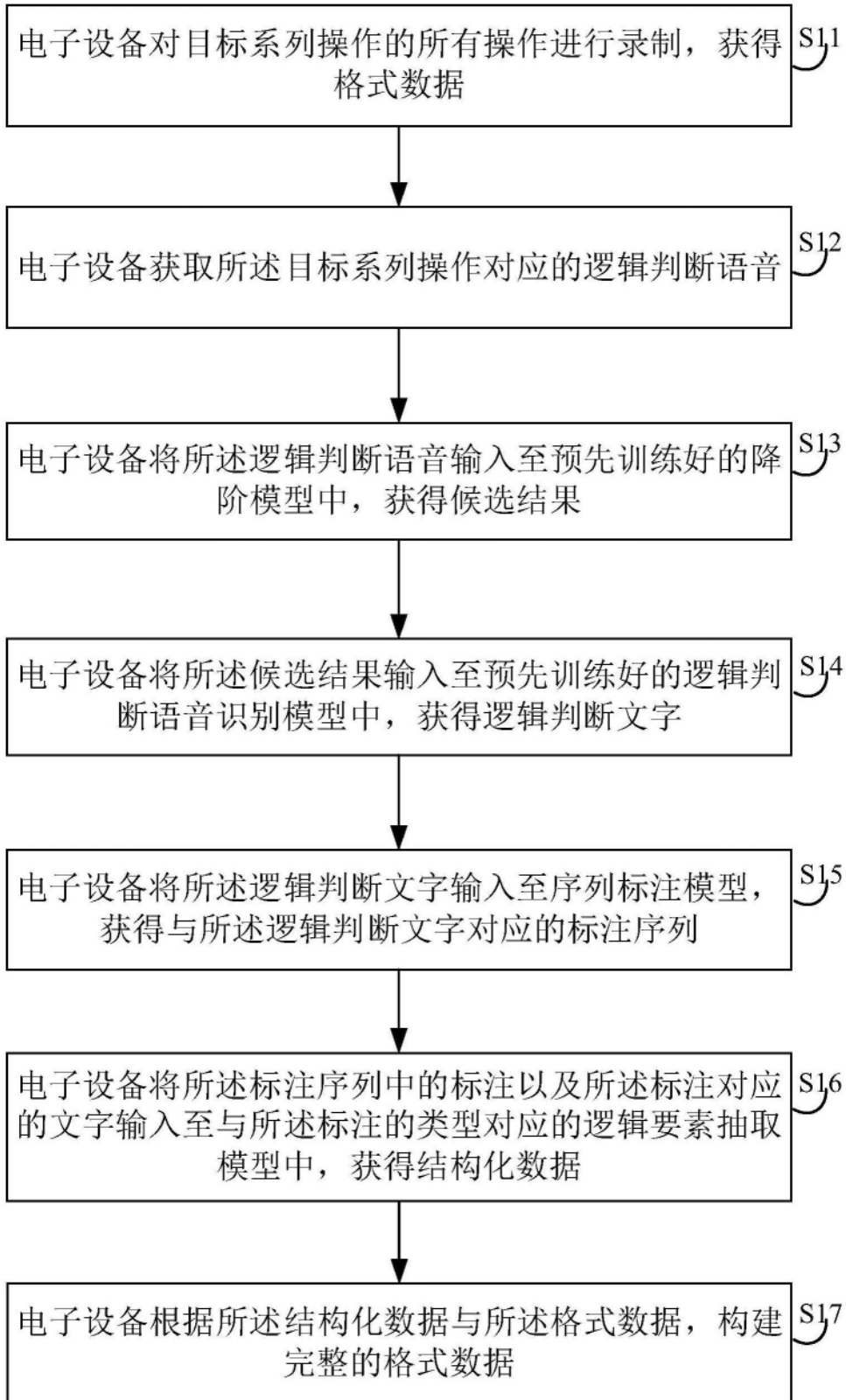


图1

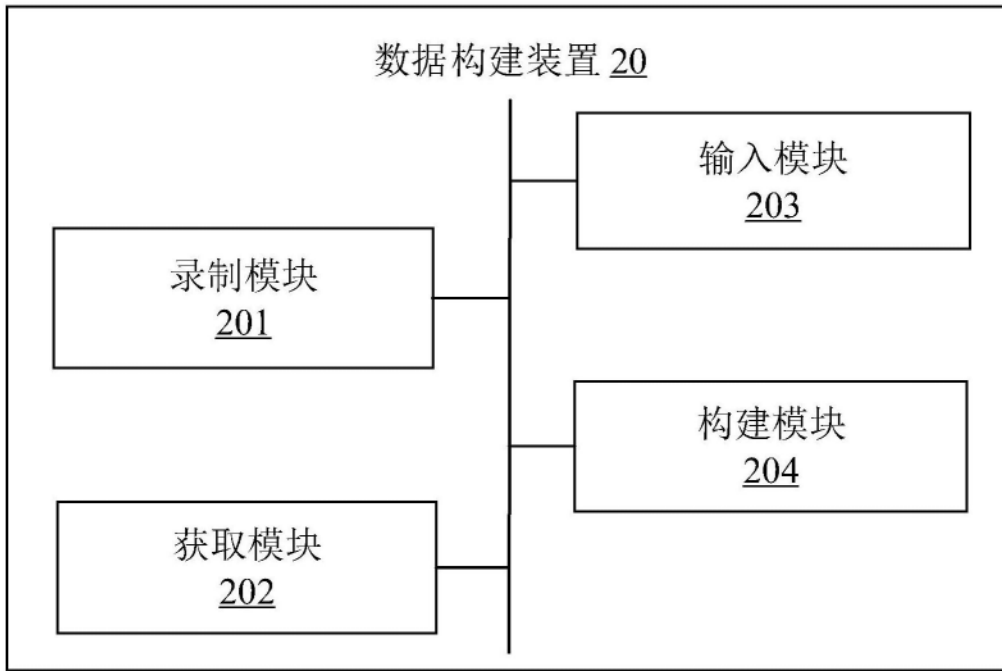


图2

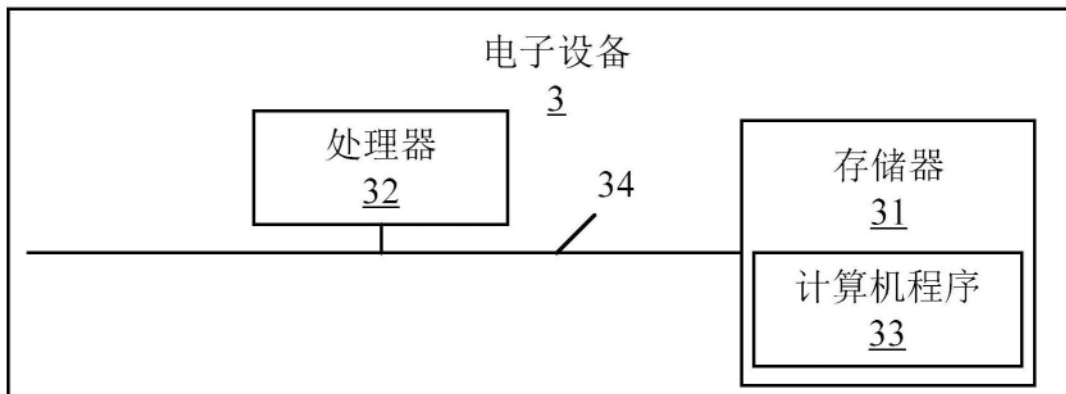


图3