



(12) 发明专利

(10) 授权公告号 CN 116820895 B

(45) 授权公告日 2024.01.05

(21) 申请号 202311093994.1

(22) 申请日 2023.08.29

(65) 同一申请的已公布的文献号  
申请公布号 CN 116820895 A

(43) 申请公布日 2023.09.29

(73) 专利权人 深圳市广和通无线通信软件有限公司

地址 518000 广东省深圳市南山区西丽街道西丽社区打石一路深圳国际创新谷六栋A座1001

(72) 发明人 吴豪

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

专利代理师 刘静莉

(51) Int.Cl.

G06F 11/30 (2006.01)

G06N 20/00 (2019.01)

(56) 对比文件

CN 111930592 A, 2020.11.13

CN 115576805 A, 2023.01.06

CN 111628967 A, 2020.09.04

CN 111611021 A, 2020.09.01

审查员 阳枝俄

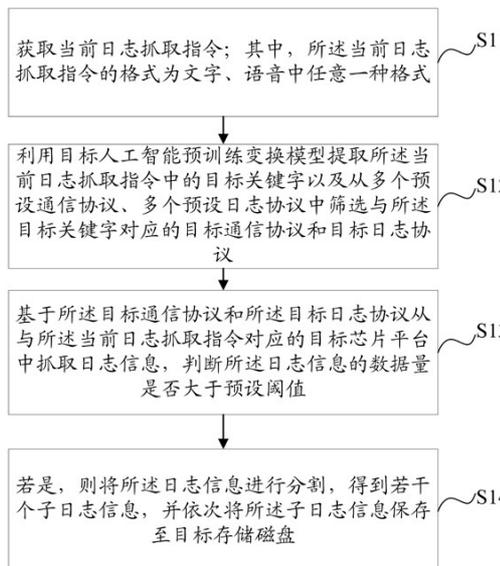
权利要求书3页 说明书11页 附图6页

(54) 发明名称

基于人工智能的日志抓取方法、装置、设备及介质

(57) 摘要

本申请公开了一种基于人工智能的日志抓取方法、装置、设备及介质,所述方法包括:获取当前日志抓取指令;其中,当前日志抓取指令的格式为文字、语音中任意一种格式;利用目标人工智能预训练变换模型提取当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与目标关键字对应的目标通信协议和目标日志协议;基于目标通信协议和目标日志协议从与当前日志抓取指令对应的目标芯片平台中抓取日志信息,若日志信息的数据量大于预设阈值,则将日志信息进行分割,得到若干个子日志信息,并依次将子日志信息保存至目标存储磁盘。降低日志抓取的难度和复杂度,并提高日志抓取效率。



1. 一种基于人工智能的日志抓取方法,其特征在于,应用于部署于主机中的预设通用日志抓取工具,包括:

获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;

利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;

基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;

若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘;

所述利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议,包括:

利用目标人工智能预训练变换模型提取所述当前日志抓取指令中目标关键字;其中,所述目标关键字包括与所述当前日志抓取指令对应的芯片类型、所属项目;基于所述目标关键字生成对应的第一提示信息,以使用户基于所述第一提示信息判断是否执行与所述当前日志抓取指令对应的抓取任务,并将判定结果反馈至所述预设通用日志抓取工具;若所述判定结果表征执行所述抓取任务,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议;

所述利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议,包括:

基于所述芯片类型、所述所属项目确定与所述当前日志抓取指令对应的目标芯片平台,并判断所述目标芯片平台是否与所述主机连接;若所述目标芯片平台与所述主机连接,则判断所述主机中是否存在所述目标芯片平台与所述主机连接的目标端口;如果所述主机中存在所述目标端口,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议。

2. 根据权利要求1所述的基于人工智能的日志抓取方法,其特征在于,所述利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字,包括:

利用目标人工智能预训练变换模型对所述当前日志抓取指令进行识别解析,以将所述当前日志抓取指令的语言形式从自然语言形式转换为机器语言形式,得到转换后日志抓取指令;

提取所述转换后日志抓取指令中的目标关键字。

3. 根据权利要求1或2所述的基于人工智能的日志抓取方法,其特征在于,所述依次将所述子日志信息保存至目标存储磁盘,包括:

确定用于存储所述日志信息的目标存储磁盘,并判断所述目标存储磁盘的当前可存储空间是否小于所述日志信息的所需存储空间;

若所述目标存储磁盘的当前可存储空间不小于所述日志信息的所需存储空间,则在预设用户界面中显示所述目标存储磁盘的存储路径,并依次将所述子日志信息保存至目标存储磁盘。

4. 根据权利要求1所述的基于人工智能的日志抓取方法,其特征在于,所述获取当前日志抓取指令之前,还包括:

采集训练数据;其中,所述训练数据包括历史日志类型、历史芯片类型、历史日志采集方式、历史日志分割方式、预设日志协议、预设通信协议、文字格式的历史日志抓取指令、语音格式的历史日志抓取指令;

利用所述训练数据对初始人工智能预训练变换模型进行训练,以得到符合预期训练结束条件的目标人工智能预训练变换模型。

5. 根据权利要求1所述的基于人工智能的日志抓取方法,其特征在于,还包括:

确定当前时刻对应的当前抓取执行阶段;

生成与所述当前抓取执行阶段对应的第二提示信息,并在预设用户界面中显示所述第二提示信息;

确定下一时刻对应的下一抓取执行阶段,并将所述下一抓取执行阶段作为当前抓取执行阶段,并重新跳转至所述生成与所述当前抓取执行阶段对应的第二提示信息,在预设用户界面中显示所述第二提示信息的步骤。

6. 一种基于人工智能的日志抓取装置,其特征在于,应用于部署于主机中的预设通用日志抓取工具,包括:

指令获取模块,用于获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;

协议筛选模块,用于利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;

阈值判断模块,用于基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;

日志保存模块,用于若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘;

所述协议筛选模块,具体用于:

利用目标人工智能预训练变换模型提取所述当前日志抓取指令中目标关键字;其中,所述目标关键字包括与所述当前日志抓取指令对应的芯片类型、所属项目;基于所述目标关键字生成对应的第一提示信息,以使用户基于所述第一提示信息判断是否执行与所述当前日志抓取指令对应的抓取任务,并将判定结果反馈至所述预设通用日志抓取工具;若所述判定结果表征执行所述抓取任务,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议;

所述协议筛选模块,具体用于:

基于所述芯片类型、所述所属项目确定与所述当前日志抓取指令对应的目标芯片平

台,并判断所述目标芯片平台是否与所述主机连接;若所述目标芯片平台与所述主机连接,则判断所述主机中是否存在所述目标芯片平台与所述主机连接的目标端口;如果所述主机中存在所述目标端口,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议。

7.一种电子设备,其特征在于,包括:

存储器,用于保存计算机程序;

处理器,用于执行所述计算机程序,以实现如权利要求1至5任一项所述的基于人工智能的日志抓取方法的步骤。

8.一种计算机可读存储介质,其特征在于,用于存储计算机程序;其中,所述计算机程序被处理器执行时实现如权利要求1至5任一项所述的基于人工智能的日志抓取方法的步骤。

## 基于人工智能的日志抓取方法、装置、设备及介质

### 技术领域

[0001] 本发明涉及人工智能技术领域,特别涉及基于人工智能的日志抓取方法、装置、设备及介质。

### 背景技术

[0002] 在进行日志数据抓取时,针对不同芯片平台下的无线模块其用于日志抓取的通信协议、日志协议也各不相同,也就是说,用户在抓取不同芯片平台下的无线模块侧相关日志数据时需要依赖特定的通信协议、日志协议的日志抓取工具,如此一来,不同芯片平台对应的通信协议、日志协议无法进行相互适配、复用和通用,从而导致针对不同芯片平台下的日志抓取需要独立开发、维护相关代码,增加了日志抓取的复杂度,并且,现有的日志抓取工具是机械化的、固化式的,即需要专业人员根据指导手册进行日志抓取,使得日志抓取的难度提高,效率降低。

[0003] 综上所述,如何降低日志抓取的难度和复杂度,并提高日志抓取效率是本领域有待解决的问题。

### 发明内容

[0004] 有鉴于此,本发明的目的在于提供一种基于人工智能的日志抓取方法、装置、设备及介质,降低日志抓取的难度和复杂度,并提高日志抓取效率。其具体方案如下:

[0005] 第一方面,本申请公开了一种基于人工智能的日志抓取方法,应用于部署于主机中的预设通用日志抓取工具,包括:

[0006] 获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任何一种格式;

[0007] 利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;

[0008] 基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;

[0009] 若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。

[0010] 可选的,所述利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字,包括:

[0011] 利用目标人工智能预训练变换模型对所述当前日志抓取指令进行识别解析,以将所述当前日志抓取指令的语言形式从自然语言形式转换为机器语言形式,得到转换后日志抓取指令;

[0012] 提取所述转换后日志抓取指令中的目标关键字。

[0013] 可选的,所述利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的

目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议,包括:

[0014] 利用目标人工智能预训练变换模型提取所述当前日志抓取指令中目标关键字;其中,所述目标关键字包括与所述当前日志抓取指令对应的芯片类型、所属项目;

[0015] 基于所述目标关键字生成对应的第一提示信息,以使用户基于所述第一提示信息判断是否执行与所述当前日志抓取指令对应的抓取任务,并将判定结果反馈至所述预设通用日志抓取工具;

[0016] 若所述判定结果表征执行所述抓取任务,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议。

[0017] 可选的,所述利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议,包括:

[0018] 基于所述芯片类型、所述所属项目确定与所述当前日志抓取指令对应的目标芯片平台,并判断所述目标芯片平台是否与所述主机连接;

[0019] 若所述目标芯片平台与所述主机连接,则判断所述主机中是否存在所述目标芯片平台与所述主机连接的目标端口;

[0020] 如果所述主机中存在所述目标端口,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议。

[0021] 可选的,所述依次将所述子日志信息保存至目标存储磁盘,包括:

[0022] 确定用于存储所述日志信息的目标存储磁盘,并判断所述目标存储磁盘的当前可存储空间是否小于所述日志信息的所需存储空间;

[0023] 若所述目标存储磁盘的当前可存储空间不小于所述日志信息的所需存储空间,则在预设用户界面中显示所述目标存储磁盘的存储路径,并依次将所述子日志信息保存至目标存储磁盘。

[0024] 可选的,所述获取当前日志抓取指令之前,还包括:

[0025] 采集训练数据;其中,所述训练数据包括历史日志类型、历史芯片类型、历史日志采集方式、历史日志分割方式、预设日志协议、预设通信协议、文字格式的历史日志抓取指令、语音格式的历史日志抓取指令;

[0026] 利用所述训练数据对初始人工智能预训练变换模型进行训练,以得到符合预期训练结束条件的目标人工智能预训练变换模型。

[0027] 可选的,所述基于人工智能的日志抓取方法,还包括:

[0028] 确定当前时刻对应的当前抓取执行阶段;

[0029] 生成与所述当前抓取执行阶段对应的第二提示信息,并在预设用户界面中显示所述第二提示信息;

[0030] 确定下一时刻对应的下一抓取执行阶段,并将所述下一抓取执行阶段作为当前抓取执行阶段,并重新跳转至所述生成与所述当前抓取执行阶段对应的第二提示信息,在预设用户界面中显示所述第二提示信息的步骤。

[0031] 第二方面,本申请公开了一种基于人工智能的日志抓取装置,应用于部署于主机中的预设通用日志抓取工具,包括:

[0032] 指令获取模块,用于获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;

[0033] 协议筛选模块,用于利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;

[0034] 阈值判断模块,用于基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;

[0035] 日志保存模块,用于若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。

[0036] 第三方面,本申请公开了一种电子设备,包括:

[0037] 存储器,用于保存计算机程序;

[0038] 处理器,用于执行所述计算机程序,以实现前述公开的基于人工智能的日志抓取方法的步骤。

[0039] 第四方面,本申请公开了一种计算机可读存储介质,用于存储计算机程序;其中,所述计算机程序被处理器执行时实现前述公开的基于人工智能的日志抓取方法的步骤。

[0040] 本申请有益效果为:本申请应用于部署于主机中的预设通用日志抓取工具,包括:获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。由此可见,本申请获取的当前日志抓取指令的格式可以是文字、语音中任意一种格式,然后利用目标人工智能预训练变换模型提取当前日志抓取指令中的目标关键字,也就是说,以人机交互的方式完成日志抓取,对用户输入的当前日志抓取指令的格式不做要求,无需根据指导手册进行操作,降低对用户的专业度要求,方便用户进行日志抓取,进一步的,利用目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与目标关键字对应的目标通信协议和目标日志协议,即预设通用日志抓取工具集成了多个预设通信协议、预设日志协议,以实现不同的芯片平台、日志类型完成日志抓取,通用性更高,不会因为芯片平台、日志类型等不同的原因无法完成日志抓取,也即降低日志抓取的难度。

## 附图说明

[0041] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据

提供的附图获得其他的附图。

- [0042] 图1为本申请公开的一种基于人工智能的日志抓取方法流程图；
- [0043] 图2为本申请公开的一种具体的基于人工智能的日志抓取方法流程图；
- [0044] 图3为本申请公开的另一具体的基于人工智能的日志抓取方法流程图；
- [0045] 图4为本申请公开的一种具体的日志抓取架构图；
- [0046] 图5为本申请公开的一种具体的日志采集示意图；
- [0047] 图6为本申请公开的一种日志抓取装置结构示意图；
- [0048] 图7为本申请公开的一种电子设备结构图。

### 具体实施方式

[0049] 下面将结合本申请实施例中的附图,对本申请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] 在进行日志数据抓取时,针对不同芯片平台下的无线模块其用于日志抓取的通信协议、日志协议也各不相同,也就是说,用户在抓取不同芯片平台下的无线模块侧相关日志数据时需要依赖特定的通信协议、日志协议的日志抓取工具,如此一来,不同芯片平台对应的通信协议、日志协议无法进行相互适配、复用和通用,从而导致针对不同芯片平台下的日志抓取需要独立开发、维护相关代码,增加了日志抓取的复杂度,并且,现有的日志抓取工具是机械化的、固化式的,即需要专业人员根据指导手册进行日志抓取,使得日志抓取的难度提高,效率降低。

[0051] 为此本申请相应的提供了一种日志抓取方案,降低日志抓取的难度和复杂度,并提高日志抓取效率。

[0052] 参见图1所示,本申请实施例公开了一种基于人工智能的日志抓取方法,应用于部署于主机中的预设通用日志抓取工具,包括:

[0053] 步骤S11:获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式。

[0054] 本实施例中,所述获取当前日志抓取指令之前,还包括:采集训练数据;其中,所述训练数据包括历史日志类型、历史芯片类型、历史日志采集方式、历史日志分割方式、预设日志协议、预设通信协议、文字格式的历史日志抓取指令、语音格式的历史日志抓取指令;利用所述训练数据对初始人工智能预训练变换模型进行训练,以得到符合预期训练结束条件的目标人工智能预训练变换模型。可以理解的是,需要对初始人工智能预训练变换模型(Generative Pre-Trained Transformer,即GPT)进行训练,使得训练后得到的目标人工智能预训练变换模型可以用于后续用户输入的各种格式的日志抓取指令进行识别,即采集包括历史日志类型、历史芯片类型、历史日志采集方式、历史日志分割方式、预设日志协议、预设通信协议、文字格式的历史日志抓取指令、语音格式的历史日志抓取指令的训练数据,训练数据可以以文本格式存储,例如TXT文件(Text File,即纯文本文件),需要将训练的数据导入到本地环境或云平台环境中,再将训练数据加载到GPT模型中,并选择合适的训练参数,然后用训练数据对初始人工智能预训练变换模型进行训练,直至符合预期训练结束条

件,其中,预期训练结束条件可以为训练过程中的迭代次数,也可以在训练过程中使用可视化工具进行监控,即监控每一次训练后得到的损失函数值、学习率,当损失函数值、学习率满足预设条件后,将满足预期训练结束条件的目标人工智能预训练变换模型导出,关于基于GPT模型对数据训练环境部署,个人或企业环境部署可通过采购特定算力型号配置要求的硬件和深度学习软件环境来搭建,也可使用开源的云平台或开源库来实现。

[0055] 可以在预设用户界面提供当前日志抓取指令的输入接口,输入接口可以为用于键入文字格式的第一输入框,还可以为用于录入语音格式的第二输入框,使得用户可以根据具体情况选择当前日志抓取指令的格式,预设通用日志抓取工具可以设置语音播报功能,与用户进行语音交互,还可以提供语音转文字、文字转语音功能,将当前日志抓取指令进行格式转换。目标人工智能预训练变换模型可以面向不同芯片平台、日志类型以及不同格式的日志抓取指令,以完成日志抓取任务,获取的当前日志抓取指令例如为文字格式的“我需要抓取芯片平台A中B项目的medemlog(调制解调器日志)”,其中,目标关键字为“芯片平台A”以及“B项目”,也就是说,当前日志抓取指令中无需包含具体实现细节,只需要包含相应的要求即可;需要注意的是,当前日志抓取指令中还可以包含其他的目标关键字,例如通信协议、日志协议,还可以包含后续如何对日志信息进行分割的关键字。

[0056] 步骤S12:利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议。

[0057] 本实施例中,所述利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字,包括:利用目标人工智能预训练变换模型对所述当前日志抓取指令进行识别解析,以将所述当前日志抓取指令的语言形式从自然语言形式转换为机器语言形式,得到转换后日志抓取指令;提取所述转换后日志抓取指令中的目标关键字。可以理解的是,需要利用目标人工智能预训练变换模型对当前日志抓取指令进行识别解析,以将当前日志抓取指令的语言形式从自然语言形式转换为机器语言形式,如此一来,才能够提取转换后日志抓取指令中的目标关键字。

[0058] 目标关键字可以包括与当前日志抓取指令对应的芯片类型、所属项目,进一步的,目标人工智能预训练变换模型根据芯片类型、所属项目从多个预设通信协议、多个预设日志协议中筛选出对应的目标通信协议和目标日志协议;需要注意的是,如果当前日志抓取指令中不仅仅包含了芯片类型、所属项目,还包含了日志抓取时所需的通信协议、日志协议,那么可以将当前日志抓取指令中的通信协议、日志协议直接作为目标通信协议和目标日志协议。

[0059] 步骤S13:基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值。

[0060] 确定出目标通信协议和目标日志协议后,就可以基于目标通信协议和目标日志协议从与当前日志抓取指令对应的目标芯片平台中抓取日志信息,并且还需要判断日志信息的数据量是否大于预设阈值,如果日志信息的数据量不大于预设阈值,即若日志信息的数据量较小,则无需对日志信息进行分割,可以直接将日志信息存储至对应的目标存储磁盘。

[0061] 步骤S14:若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。

[0062] 如果日志信息的数据量较大,且大于预设阈值,则可以将日志信息进行分割,将分割后得到的若干个子日志信息依次传输至目标存储磁盘,并进行保存,其中,用户可以预先设置预设阈值,也可以在当前日志抓取指令中包含。

[0063] 本实施例中,还包括:确定当前时刻对应的当前抓取执行阶段;生成与所述当前抓取执行阶段对应的第二提示信息,并在预设用户界面中显示所述第二提示信息;确定下一时刻对应的下一抓取执行阶段,并将所述下一抓取执行阶段作为当前抓取执行阶段,并重新跳转至所述生成与所述当前抓取执行阶段对应的第二提示信息,在预设用户界面中显示所述第二提示信息的步骤。为了更好的给用户提供人机互动,还可以实时显示用于表征抓取进度的提示信息,即确定当前时刻对应的当前抓取执行阶段,在预设用户界面中将用于表征当前抓取执行阶段的第二提示信息进行显示,可以在当前抓取执行阶段为抓取任务结束阶段时,停止更新当前抓取执行阶段。

[0064] 本申请有益效果为:本申请应用于部署于主机中的预设通用日志抓取工具,包括:获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。由此可见,本申请获取的当前日志抓取指令的格式可以是文字、语音中任意一种格式,然后利用目标人工智能预训练变换模型提取当前日志抓取指令中的目标关键字,也就是说,以人机交互的方式完成日志抓取,对用户输入的当前日志抓取指令的格式不做要求,无需根据指导手册进行操作,降低对用户的专业度要求,方便用户进行日志抓取,进一步的,利用目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与目标关键字对应的目标通信协议和目标日志协议,即预设通用日志抓取工具集成了多个预设通信协议、预设日志协议,以实现不同的芯片平台、日志类型完成日志抓取,通用性更高,不会因为芯片平台、日志类型等不同的原因无法完成日志抓取,也即降低日志抓取的难度。

[0065] 参见图2所示,本申请实施例公开了一种具体的基于人工智能的日志抓取方法,应用于部署于主机中的预设通用日志抓取工具,包括:

[0066] 步骤S21:获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式。

[0067] 步骤S22:利用目标人工智能预训练变换模型提取所述当前日志抓取指令中目标关键字;其中,所述目标关键字包括与所述当前日志抓取指令对应的芯片类型、所属项目。

[0068] 步骤S23:基于所述目标关键字生成对应的第一提示信息,以便用户基于所述第一提示信息判断是否执行与所述当前日志抓取指令对应的抓取任务,并将判定结果反馈至所述预设通用日志抓取工具。

[0069] 在第一种具体的实施例中,根据目标关键字生成的第一提示信息可以表征当前已经进行与目标关键字对应的日志抓取任务,那么用户基于该第一提示信息判断是否还需要再执行与当前日志抓取指令对应的抓取任务,以生成判定结果,预设通用日志抓取工具获

取该判定结果。

[0070] 在第二种具体实施例中,根据目标关键字生成的第一提示信息可以表征当前准备执行的抓取任务,用户可以判断当前准备执行的抓取任务是否与当前日志抓取指令对应的抓取任务一致,以生成判定结果,预设通用日志抓取工具获取该判定结果。

[0071] 步骤S24:若所述判定结果表征执行所述抓取任务,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议。

[0072] 本实施例中,所述利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议,包括:基于所述芯片类型、所述所属项目确定与所述当前日志抓取指令对应的目标芯片平台,并判断所述目标芯片平台是否与所述主机连接;若所述目标芯片平台与所述主机连接,则判断所述主机中是否存在所述目标芯片平台与所述主机连接的目标端口;如果所述主机中存在所述目标端口,则利用所述目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与所述芯片类型、所述所属项目对应的目标通信协议和目标日志协议。可以理解的是,在执行抓取任务之前,需要判断所要进行日志抓取的目标芯片平台是否与主机连接,如果连接还需要判断主机中是否存在目标芯片平台与主机连接的目标端口,只有主机中存在目标芯片平台与主机连接的目标端口时,才能进行后续的日志抓取任务。

[0073] 步骤S25:基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值。

[0074] 步骤S26:若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。

[0075] 由此可见,本申请支持多种格式的日志抓取指令,为用户提供多种输入当前日志抓取指令的方式,提高用户体验感,且无需用户关系具体实现抓取细节,降低对用户专业度的要求以及难度;本申请在进行日志抓取前,判断主机是否与目标芯片平台相连接以及是否存在与目标芯片相连接的端口,以保障后续能够完成任务抓取,避免在主机与目标芯片连接不成功的情况下进行任务抓取,浪费工作程序和时间;预设通用日志抓取工具中预先集成多种预设通信协议、预设日志协议,因此针对不同的芯片平台、项目均可以完成日志抓取,通用性更高,更加稳定。

[0076] 参见图3所示,本申请实施例公开了一种基于人工智能的日志抓取方法,应用于部署于主机中的预设通用日志抓取工具,包括:

[0077] 步骤S31:获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式。

[0078] 步骤S32:利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议。

[0079] 本实施例中,一方面提供不同芯片平台下不同项目的通信机制适配功能,其主要支持COM(Communication)通信、Socket(套接字)通信、WebSocket通信、MBIM(Mobile Broadband Interface Model,即移动宽带接口模块)通信、IPC通信(Inter-Process

Communication,即网络进程间通信)、RPC(Remote Procedure Call Protocol,即远程过程调用协议)通信及其各个通信协议对应的可调用功能接口;另一方面还提供不同芯片平台下不同项目的日志协议适配功能,例如:支持TVL(Tag Length Value)编解码功能、支持高通7E指令集编解码功能等。

[0080] 步骤S33:基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值。

[0081] 步骤S34:若是,则将所述日志信息进行分割,得到若干个子日志信息,并确定用于存储所述日志信息的目标存储磁盘,并判断所述目标存储磁盘的当前可存储空间是否小于所述日志信息的所需存储空间。

[0082] 可以理解的是,在进行日志信息存储之前,需要确保目标存储磁盘可以存储下日志信息,即判断目标存储磁盘的当前可存储空间是否小于日志信息的所需存储空间,若目标存储磁盘的当前可存储空间小于日志信息的所需存储空间,可以生成对应的预警信息,使得用户可以根据该预警信息更改用于保存日志信息的磁盘。

[0083] 步骤S35:若所述目标存储磁盘的当前可存储空间不小于所述日志信息的所需存储空间,则在预设用户界面中显示所述目标存储磁盘的存储路径,并依次将所述子日志信息保存至目标存储磁盘。

[0084] 如果目标存储磁盘的当前可存储空间不小于日志信息的所需存储空间,则生成目标存储磁盘的存储路径,并在预设用户界面中显示,使得用户可以选择是否更高用于保存日志信息的磁盘,如果不需要更改,则可以将所有的子日志信息保存至目标存储磁盘。

[0085] 由此可见,本申请为预设通用日志抓取工具开发提供了一套具有模块化、通用性、重用性、高效性、可裁剪性、松耦合性、高内聚性、易用性、健壮性、鲁棒性设计思路,确保了开发出来的预设通用日志抓取工具是高效的,同时也避免了预设通用日志抓取工具本身的不稳定因素而导致用户使用过程中出现不可估量的错误。另外,预设通用日志抓取工具的预设用户界面设计遵循易操作、易理解、高效简洁的布局原则,目的是使用户便于操作、快速上手,大大降低日志抓取的难度。

[0086] 例如图4所示的一种具体的日志抓取架构图,预设通用日志抓取工具部署于主机中,芯片平台位于模块侧,并且预设通用日志抓取工具中包含实时数据显示组件、通信接口适配组件、目标人工智能预训练变换模型,目标人工智能预训练变换模型包含文字语音识别及解析组件、日志数据采集及存储组件、日志协议适配组件,其中,实时数据显示组件用于实时显示用于表征当前抓取执行阶段的第二提示信息,通信接口适配组件用于利用目标人工智能预训练变换模型确定出对应的目标通信协议,文字语音识别及解析组件用于将文字格式或者语音格式的当前日志抓取指令进行识别解析,进而提取目标关键字,日志数据采集及存储组件用于进行日志的采集,并将采集的日志信息存储至目标磁盘,日志协议适配组件用于筛选出目标日志协议,本申请使用目标人工智能预训练变换模型模块接口、组件化方法和面向对象思想进行分析、抽象、建模等一系列处理后,为适配和满足不同的业务及功能需求,对各个组成的核心功能部分进行了独立模块设计和功能逻辑整合,并基于MVM(Model-View-Model)开发框架模型来保证功能处理逻辑部分和业务需求分开,从而确保了工具的灵活性、适配性和通用性。下面以图5所示的一种具体的日志采集示意图为例,对本申请进行说明:

[0087] 步骤1) 获取用户输入的文件格式的当前日志抓取指令或者语音格式的当前日志抓取指令;其中,当前日志抓取指令包含芯片类型、所属项目;

[0088] 步骤2) 利用目标人工智能预训练变换模型对当前日志抓取指令进行识别解析,以将当前日志抓取指令的语言形式从自然语言形式转换为机器语言形式,得到转换后日志抓取指令;

[0089] 步骤3) 提取转换后日志抓取指令中的目标关键字,即芯片类型、所属项目;

[0090] 步骤4) 基于目标关键字生成对应的第一提示信息,以使用户基于第一提示信息判断是否执行与当前日志抓取指令对应的抓取任务;

[0091] 步骤5) 若获取到的判定结果表征不执行抓取任务,则跳转至步骤1),即可以获取下一日志抓取指令;若获取到的判定结果表征执行抓取任务,则基于芯片类型、所属项目确定与当前日志抓取指令对应的目标芯片平台,并判断目标芯片平台是否与主机连接;

[0092] 步骤6) 若目标芯片平台与主机连接,则判断主机中是否存在目标芯片平台与主机连接的目标端口;

[0093] 步骤7) 如果主机中存在目标端口,则利用目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与芯片类型、所属项目对应的目标通信协议和目标日志协议;

[0094] 步骤8) 基于目标通信协议和目标日志协议从与当前日志抓取指令对应的目标芯片平台中抓取日志信息,当日志抓取完毕时,会接收到结束指令;

[0095] 步骤9) 显示抓取结果以及用于存储日志信息的目标存储磁盘的存储路径,将日志信息存储至目标存储磁盘中。

[0096] 参见图6所示,本申请实施例公开了一种基于人工智能的日志抓取装置,应用于部署于主机中的预设通用日志抓取工具,包括:

[0097] 指令获取模块11,用于获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;

[0098] 协议筛选模块12,用于利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;

[0099] 阈值判断模块13,用于基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;

[0100] 日志保存模块14,用于若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标存储磁盘。

[0101] 本申请有益效果为:本申请应用于部署于主机中的预设通用日志抓取工具,包括:获取当前日志抓取指令;其中,所述当前日志抓取指令的格式为文字、语音中任意一种格式;利用目标人工智能预训练变换模型提取所述当前日志抓取指令中的目标关键字以及从多个预设通信协议、多个预设日志协议中筛选与所述目标关键字对应的目标通信协议和目标日志协议;基于所述目标通信协议和所述目标日志协议从与所述当前日志抓取指令对应的目标芯片平台中抓取日志信息,判断所述日志信息的数据量是否大于预设阈值;若是,则将所述日志信息进行分割,得到若干个子日志信息,并依次将所述子日志信息保存至目标

存储磁盘。由此可见,本申请获取的当前日志抓取指令的格式可以是文字、语音中任意一种格式,然后利用目标人工智能预训练变换模型提取当前日志抓取指令中的目标关键字,也就是说,以人机交互的方式完成日志抓取,对用户输入的当前日志抓取指令的格式不做要求,无需根据指导手册进行操作,降低对用户的专业度要求,方便用户进行日志抓取,进一步的,利用目标人工智能预训练变换模型从多个预设通信协议、多个预设日志协议中筛选与目标关键字对应的目标通信协议和目标日志协议,即预设通用日志抓取工具集成了多个预设通信协议、预设日志协议,以实现不同的芯片平台、日志类型完成日志抓取,通用性更高,不会因为芯片平台、日志类型等不同的原因无法完成日志抓取,也即降低日志抓取的难度。

[0102] 进一步的,本申请实施例还提供了一种电子设备。图7是根据一示例性实施例示出的电子设备20结构图,图中的内容不能认为是对本申请的使用范围的任何限制。

[0103] 图7为本申请实施例提供的一种电子设备的结构示意图。具体可以包括:至少一个处理器21、至少一个存储器22、电源23、通信接口24、输入输出接口25和通信总线26。其中,所述存储器22用于存储计算机程序,所述计算机程序由所述处理器21加载并执行,以实现前述任一实施例公开的由电子设备执行的基于人工智能的日志抓取方法中的相关步骤。

[0104] 本实施例中,电源23用于为电子设备上的各硬件设备提供工作电压;通信接口24能够为电子设备创建与外界设备之间的数据传输通道,其所遵循的通信协议是能够适用于本申请技术方案的任意通信协议,在此不对其进行具体限定;输入输出接口25,用于获取外界输入数据或向外界输出数据,其具体的接口类型可以根据具体应用需要进行选取,在此不进行具体限定。

[0105] 其中,处理器21可以包括一个或多个处理核心,比如4核心处理器、8核心处理器等。处理器21可以采用DSP(Digital Signal Processing,数字信号处理)、FPGA(Field-Programmable Gate Array,现场可编程门阵列)、PLA(Programmable Logic Array,可编程逻辑阵列)中的至少一种硬件形式来实现。处理器21也可以包括主处理器和协处理器,主处理器是用于对在唤醒状态下的数据进行处理的处理器,也称CPU(Central Processing Unit,中央处理器);协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中,处理器21可以在集成有GPU(Graphics Processing Unit,图像处理器),GPU用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中,处理器21还可以包括AI(Artificial Intelligence,人工智能)处理器,该AI处理器用于处理有关机器学习的计算操作。

[0106] 另外,存储器22作为资源存储的载体,可以是只读存储器、随机存储器、磁盘或者光盘等,其上所存储的资源包括操作系统221、计算机程序222及数据223等,存储方式可以是短暂存储或者永久存储。

[0107] 其中,操作系统221用于管理与控制电子设备上的各硬件设备以及计算机程序222,以实现处理器21对存储器22中海量数据223的运算与处理,其可以是Windows、Unix、Linux等。计算机程序222除了包括能够用于完成前述任一实施例公开的由电子设备执行的基于人工智能的日志抓取方法的计算机程序之外,还可以进一步包括能够用于完成其他特定工作的计算机程序。数据223除了可以包括电子设备接收到的由外部设备传输进来的数据,也可以包括由自身输入输出接口25采集到的数据等。

[0108] 进一步的,本申请还公开了一种计算机可读存储介质,用于存储计算机程序;其中,所述计算机程序被处理器执行时实现前述公开的基于人工智能的日志抓取方法。关于该方法的具体步骤可以参考前述实施例中公开的相应内容,在此不再进行赘述。

[0109] 本说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其它实施例的不同之处,各个实施例之间相同或相似部分互相参见即可。对于实施例公开的装置而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。

[0110] 专业人员还可以进一步意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块,或者二者的结合来实施。软件模块可以置于随机存储器(Random Access Memory,即RAM)、内存、只读存储器(Read-Only Memory,即ROM)、电可编程EPROM(Erasable Programmable Read Only Memory)、电可擦除可编程EEPROM(Electrically Erasable Programmable read only memory)、寄存器、硬盘、可移动磁盘、CD-ROM(Compact Disc Read-Only Memory,即紧凑型光盘只读储存器)、或技术领域内所公知的任意其它形式的存储介质中。

[0111] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0112] 以上对本发明所提供的一种基于人工智能的日志抓取方法、装置、设备及介质进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

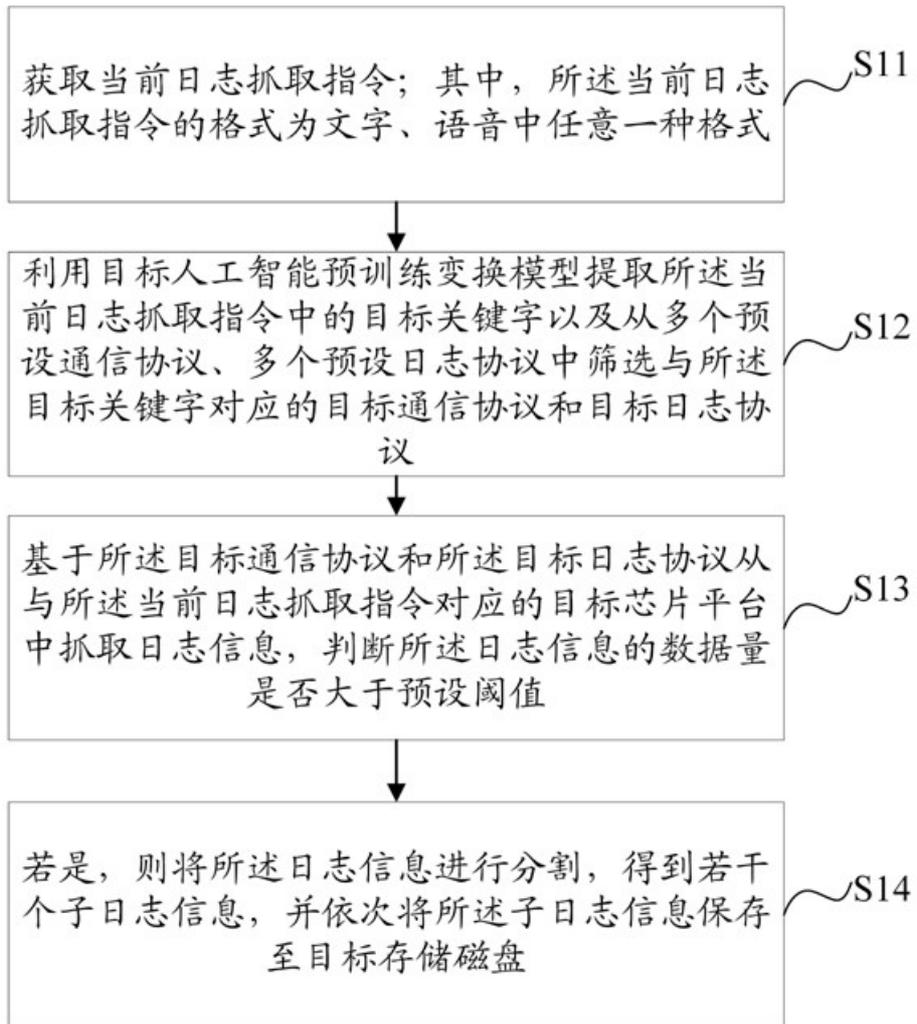


图 1

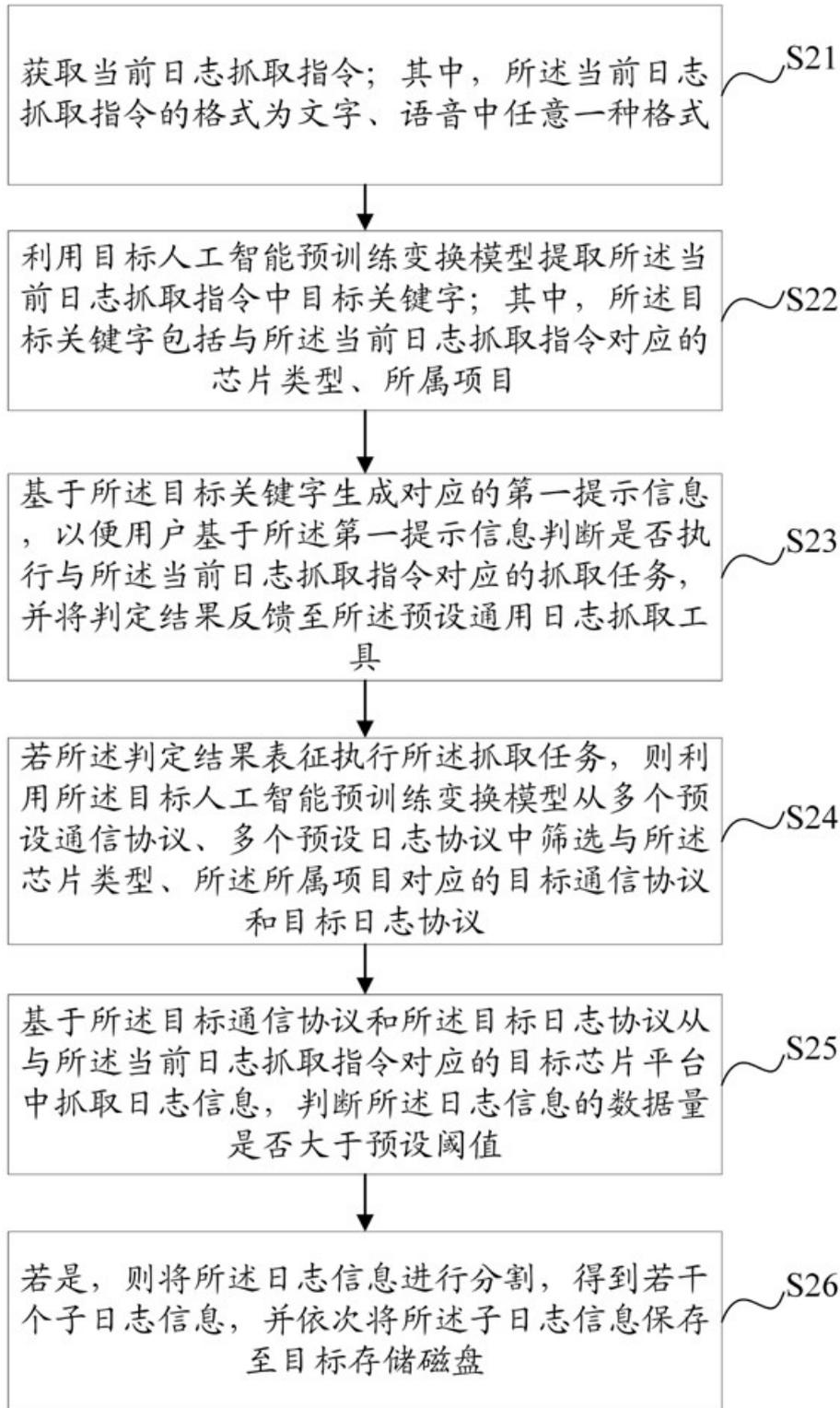


图 2

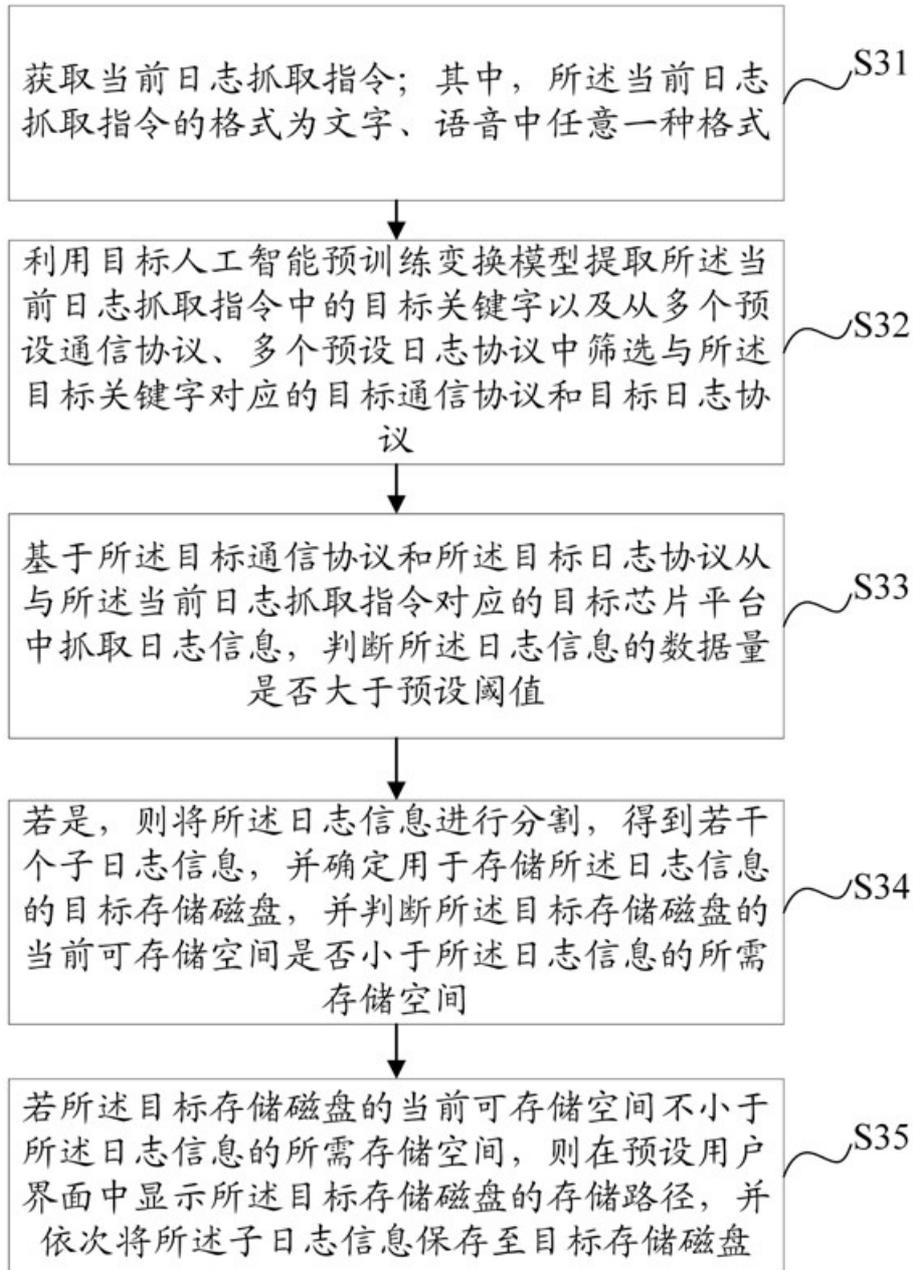


图 3

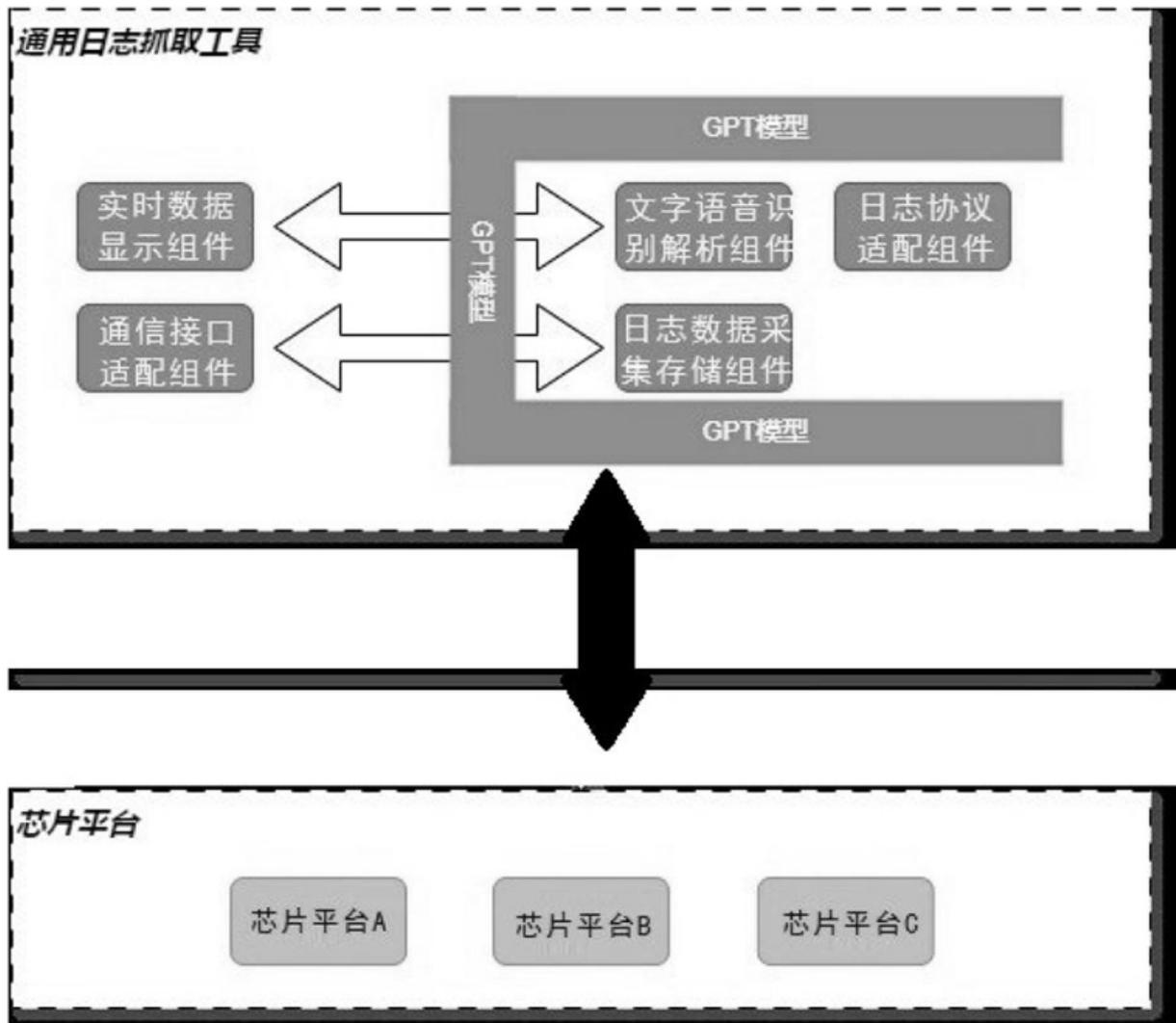


图 4

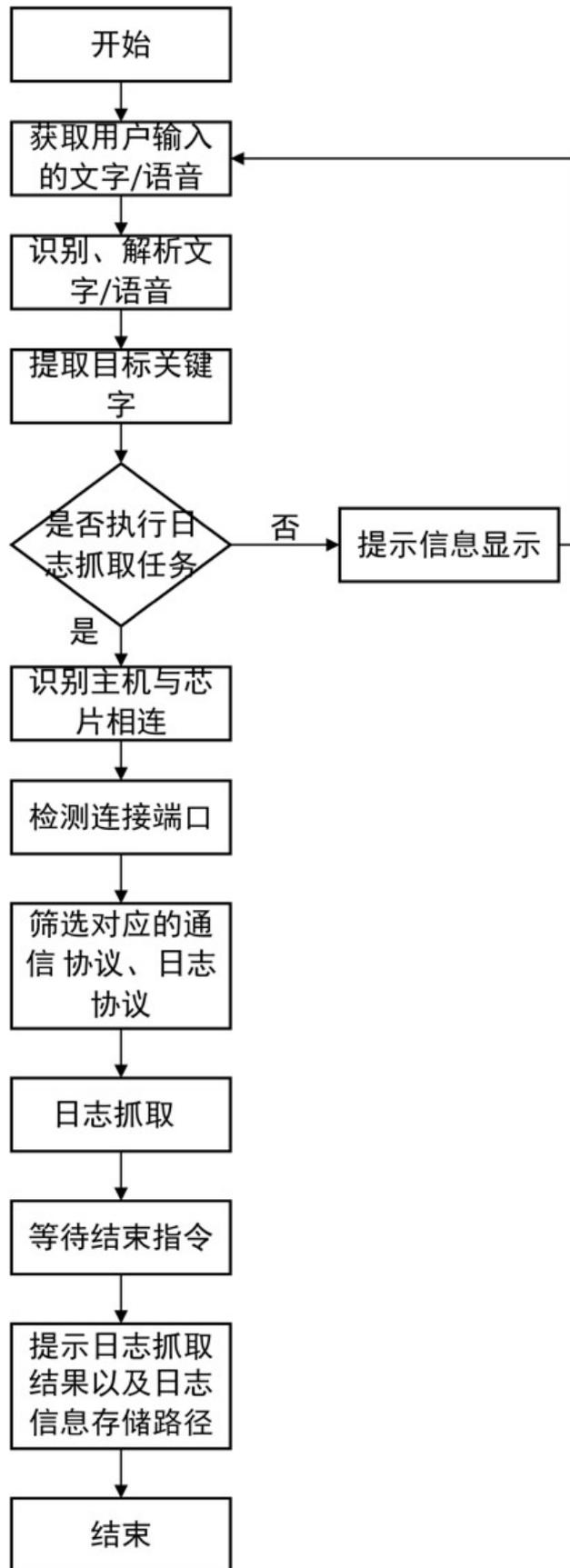


图 5

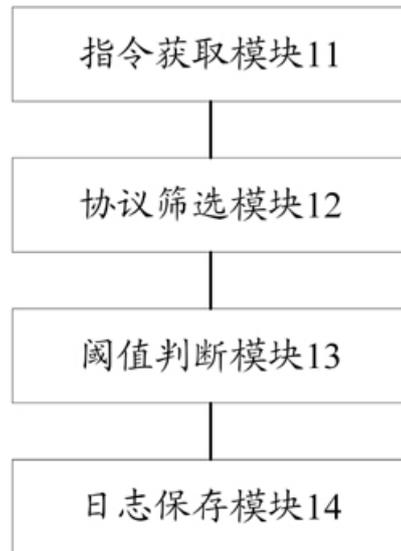


图 6

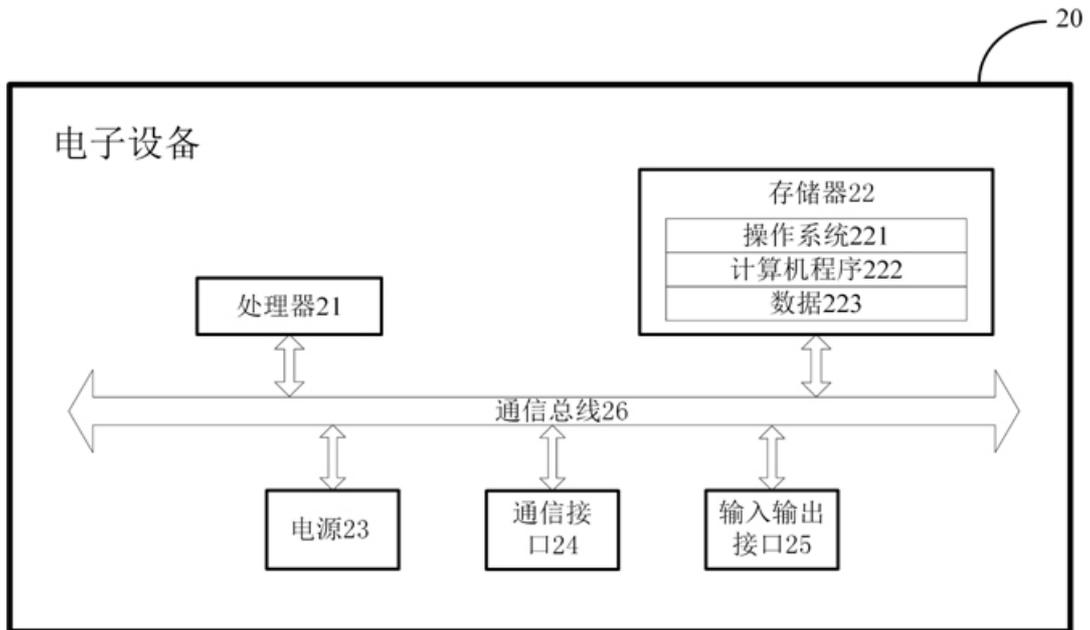


图 7