



(12) 发明专利

(10) 授权公告号 CN 113033543 B

(45) 授权公告日 2024. 04. 05

(21) 申请号 202110461569.8

G06N 3/044 (2023.01)

(22) 申请日 2021.04.27

G06N 3/0464 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 113033543 A

(56) 对比文件

CN 104809436 A, 2015.07.29

CN 111191649 A, 2020.05.22

CN 111767911 A, 2020.10.13

IN 109CHE2011 A, 2014.03.14

JP H07225812 A, 1995.08.22

CN 112686812 A, 2021.04.20

CN 112016315 A, 2020.12.01

CN 105184294 A, 2015.12.23

CN 105678300 A, 2016.06.15

CN 111860682 A, 2020.10.30

CN 112364873 A, 2021.02.12

WO 2020097909 A1, 2020.05.22

WO 2020248471 A1, 2020.12.17

(43) 申请公布日 2021.06.25

(73) 专利权人 中国平安人寿保险股份有限公司

地址 518000 广东省深圳市福田区益田路

5033号平安金融中心14、15、16、41、

44、45、46层

(72) 发明人 易苗 张蓉

(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

专利代理师 刘丽华

审查员 余思妍

(51) Int. Cl.

G06V 30/14 (2022.01)

G06V 30/18 (2022.01)

G06V 10/82 (2022.01)

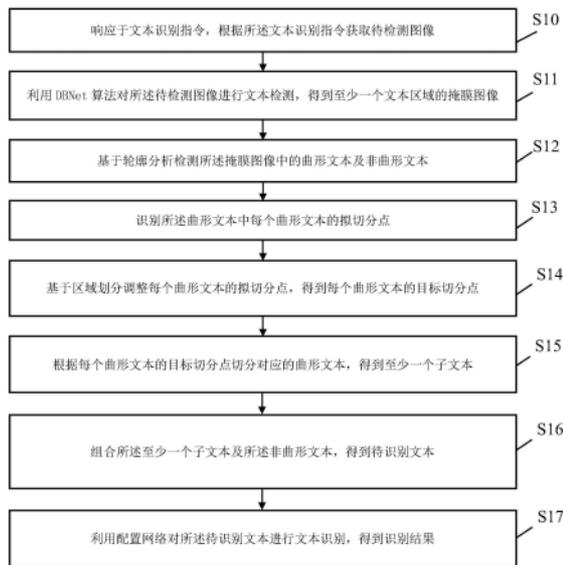
权利要求书2页 说明书14页 附图2页

(54) 发明名称

曲形文本识别方法、装置、设备及介质

(57) 摘要

本发明涉及人工智能领域,提供一种曲形文本识别方法、装置、设备及介质,能够提供精确的文本轮廓区域的掩膜图像,进一步执行曲形文本的判断,进行有针对性的拆分,以减少不必要的计算成本,对于弯曲度最大的拟切分点,对临近区域进行二值化分析,以对拆分点进行微调,并尽量减少对同一字符的切分,得到的所述待识别文本均为非扭曲的正常文本,进而将对难识别的曲形文本的识别问题转化为对多个正常文本的识别问题,通过卷积神经网络学习局部特征信息,再基于循环神经网络学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。此外,本发明还涉及区块链技术,识别结果可存储于区块链节点中。



1. 一种曲形文本识别方法,其特征在于,所述曲形文本识别方法包括:
 - 响应于文本识别指令,根据所述文本识别指令获取待检测图像;
 - 利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像;
 - 基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本,包括:对于所述掩膜图像中的每个文本区域,根据预设间隔建立至少一个点组成每个文本区域的拟合点集,获取每个拟合点集中的初始点及结束点,连接每个拟合点集中的所述初始点及所述结束点,得到每个文本区域的基准线,对于每个文本区域,计算对应拟合点集中的每个点到对应基准线的垂直距离,当每个点到对应基准线的垂直距离都大于预设阈值时,确定对应的文本区域为所述曲形文本;
 - 识别所述曲形文本中每个曲形文本的拟切分点,包括:对于每个曲形文本,按照由大到小的顺序对每个点到对应基准线的垂直距离进行排序,获取排在首位的点作为每个曲形文本的拟切分点,所述排在首位的点的弯曲程度最高;
 - 基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点,包括:将每个拟切分点确定为中心,根据配置延伸范围进行区域划分,得到每个拟切分点对应的临近区域,对每个临近区域进行二值化处理,得到每个临近区域的二值图像,计算每个临近区域的二值图像的垂直投影,根据每个临近区域的垂直投影确定每个曲形文本的目标切分点;
 - 根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本;
 - 组合所述至少一个子文本及所述非曲形文本,得到待识别文本;
 - 利用配置网络对所述待识别文本进行文本识别,得到识别结果。
2. 如权利要求1所述的曲形文本识别方法,其特征在于,所述根据所述文本识别指令获取待检测图像包括:
 - 解析所述文本识别指令的方法体,得到所述文本识别指令所携带的信息;
 - 获取预设标签;
 - 根据所述预设标签构建正则表达式;
 - 利用所述正则表达式在所述文本识别指令所携带的信息中进行遍历,并将遍历到的数据确定为目标地址;
 - 连接至所述目标地址,并获取所述目标地址处存储的数据作为所述待检测图像。
3. 如权利要求1所述的曲形文本识别方法,其特征在于,所述利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像包括:
 - 利用DBNet的主干网提取所述待检测图像的图像特征;
 - 对所述图像特征进行上采样处理,得到与所述待检测图像具有相同尺寸的特征图;
 - 基于DBNet算法,根据所述特征图进行预测,得到概率图及阈值图;
 - 根据所述概率图及所述阈值图进行二值化处理,得到所述至少一个文本区域的掩膜图像。
4. 如权利要求1所述的曲形文本识别方法,其特征在于,所述基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本还包括:
 - 当每个点到对应基准线的垂直距离不都大于所述预设阈值时,确定对应的文本区域为所述非曲形文本。

5. 如权利要求1所述的曲形文本识别方法,其特征在于,所述利用配置网络对所述待识别文本进行文本识别,得到识别结果包括:

利用卷积神经网络对所述待识别文本进行特征提取,得到目标特征;

利用循环神经网络提取所述目标特征的时序特征;

将所述时序特征输入至序列识别层,并获取所述序列识别层的输出作为所述识别结果。

6. 一种曲形文本识别装置,所述装置用于实现如权利要求1至5中任意一项所述的曲形文本识别方法,其特征在于,所述曲形文本识别装置包括:

获取单元,用于响应于文本识别指令,根据所述文本识别指令获取待检测图像;

检测单元,用于利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像;

所述检测单元,还用于基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本;

识别单元,用于识别所述曲形文本中每个曲形文本的拟切分点;

调整单元,用于基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点;

切分单元,用于根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本;

组合单元,用于组合所述至少一个子文本及所述非曲形文本,得到待识别文本;

识别单元,用于利用配置网络对所述待识别文本进行文本识别,得到识别结果。

7. 一种电子设备,其特征在于,所述电子设备包括:

存储器,存储至少一个指令;及

处理器,执行所述存储器中存储的指令以实现如权利要求1至5中任意一项所述的曲形文本识别方法。

8. 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中存储有至少一个指令,所述至少一个指令被电子设备中的处理器执行以实现如权利要求1至5中任意一项所述的曲形文本识别方法。

曲形文本识别方法、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,尤其涉及一种曲形文本识别方法、装置、设备及介质。

背景技术

[0002] 在场景文本识别中,一个具有挑战性的任务是处理扭曲或者不规则布局的文本,曲形文本在自然场景中很常见,改善扭曲文档图像的OCR(Optical Character Recognition,光学字符识别)识别准确率是一项亟需解决的工作。

[0003] 现有的针对扭曲文档的识别方法大多是先将文档进行矫正再识别,矫正的方法通常包括:

[0004] (1) 基于硬件的扭曲文档矫正。

[0005] 该方法通过特质的硬件设备(如结构光源等)扫描纸张的三维形状信息,然后根据三维形状信息对文档图像进行矫正后进行识别。该方法虽然精度高且适用于各张形状,但往往硬件设备昂贵且携带不易。

[0006] (2) 基于3D(three dimensional)模型重建的文档矫正算法。

[0007] 该方法从造成文档扭曲的因素(摆放角度、光源方向等)出发对文档进行3D建模,并利用已有的数学知识对扭曲进行矫正。但是,该方法需要清楚得知道造成扭曲的原因。

[0008] (3) 基于内容切分的文档矫正。

[0009] 该方法是直接通过分析文档图像的倾斜角、文本行特征等的扭曲矫正算法。但是,可矫正文档对象有限,且会大幅度增加额外计算成本高,难以实际部署应用,且在图像矫正的过程中虽然可以一定程度上缓解图片中文本行的扭曲分布,但是在映射的计算过程中也会导致文字的变形,带来新的识别问题。

发明内容

[0010] 鉴于以上内容,有必要提供一种曲形文本识别方法、装置、设备及介质,能够首先通过卷积神经网络学习局部特征信息,再基于循环神经网络想学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。

[0011] 一种曲形文本识别方法,所述曲形文本识别方法包括:

[0012] 响应于文本识别指令,根据所述文本识别指令获取待检测图像;

[0013] 利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像;

[0014] 基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本;

[0015] 识别所述曲形文本中每个曲形文本的拟切分点;

[0016] 基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点;

[0017] 根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本;

[0018] 组合所述至少一个子文本及所述非曲形文本,得到待识别文本;

- [0019] 利用配置网络对所述待识别文本进行文本识别,得到识别结果。
- [0020] 根据本发明优选实施例,所述根据所述文本识别指令获取待检测图像包括:
- [0021] 解析所述文本识别指令的方法体,得到所述文本识别指令所携带的信息;
- [0022] 获取预设标签;
- [0023] 根据所述预设标签构建正则表达式;
- [0024] 利用所述正则表达式在所述文本识别指令所携带的信息中进行遍历,并将遍历到的数据确定为目标地址;
- [0025] 连接至所述目标地址,并获取所述目标地址处存储的数据作为所述待检测图像。
- [0026] 根据本发明优选实施例,所述利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像包括:
- [0027] 利用DBNet的主干网提取所述待检测图像的图像特征;
- [0028] 对所述图像特征进行上采样处理,得到与所述待检测图像具有相同尺寸的特征图;
- [0029] 基于DBNet算法,根据所述特征图进行预测,得到概率图及阈值图;
- [0030] 根据所述概率图及所述阈值图进行二值化处理,得到所述至少一个文本区域的掩膜图像。
- [0031] 根据本发明优选实施例,所述基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本包括:
- [0032] 对于所述掩膜图像中的每个文本区域,根据预设间隔建立至少一个点组成每个文本区域的拟合点集;
- [0033] 获取每个拟合点集中的初始点及结束点;
- [0034] 连接每个拟合点集中的所述初始点及所述结束点,得到每个文本区域的基准线;
- [0035] 对于每个文本区域,计算对应拟合点集中的每个点到对应基准线的垂直距离;
- [0036] 当每个点到对应基准线的垂直距离都大于预设阈值时,确定对应的文本区域为所述曲形文本;或者
- [0037] 当每个点到对应基准线的垂直距离不都大于所述预设阈值时,确定对应的文本区域为所述非曲形文本。
- [0038] 根据本发明优选实施例,所述识别所述曲形文本中每个曲形文本的拟切分点包括:
- [0039] 对于每个曲形文本,按照由大到小的顺序对每个点到对应基准线的垂直距离进行排序;
- [0040] 获取排在首位的点作为每个曲形文本的拟切分点。
- [0041] 根据本发明优选实施例,所述基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点包括:
- [0042] 将每个拟切分点确定为中心,根据配置延伸范围进行区域划分,得到每个拟切分点对应的临近区域;
- [0043] 对每个临近区域进行二值化处理,得到每个临近区域的二值图像;
- [0044] 计算每个临近区域的二值图像的垂直投影;
- [0045] 根据每个临近区域的垂直投影确定每个曲形文本的目标切分点。

[0046] 根据本发明优选实施例,所述利用配置网络对所述待识别文本进行文本识别,得到识别结果包括:

[0047] 利用卷积神经网络对所述待识别文本进行特征提取,得到目标特征;

[0048] 利用循环神经网络提取所述目标特征的时序特征;

[0049] 将所述时序特征输入至序列识别层,并获取所述序列识别层的输出作为所述识别结果。

[0050] 一种曲形文本识别装置,所述曲形文本识别装置包括:

[0051] 获取单元,用于响应于文本识别指令,根据所述文本识别指令获取待检测图像;

[0052] 检测单元,用于利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像;

[0053] 所述检测单元,还用于基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本;

[0054] 识别单元,用于识别所述曲形文本中每个曲形文本的拟切分点;

[0055] 调整单元,用于基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点;

[0056] 切分单元,用于根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本;

[0057] 组合单元,用于组合所述至少一个子文本及所述非曲形文本,得到待识别文本;

[0058] 识别单元,用于利用配置网络对所述待识别文本进行文本识别,得到识别结果。

[0059] 一种电子设备,所述电子设备包括:

[0060] 存储器,存储至少一个指令;及

[0061] 处理器,执行所述存储器中存储的指令以实现所述曲形文本识别方法。

[0062] 一种计算机可读存储介质,所述计算机可读存储介质中存储有至少一个指令,所述至少一个指令被电子设备中的处理器执行以实现所述曲形文本识别方法。

[0063] 由以上技术方案可以看出,本发明能够响应于文本识别指令,根据所述文本识别指令获取待检测图像,利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像,提供精确的文本轮廓区域的掩膜图像,为后续的文本切分提供可靠的数据基础,基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本,能够基于轮廓分析在掩膜图像的基础上进一步执行曲形文本的判断,以便后续进行有针对性的拆分,减少不必要的计算成本,识别所述曲形文本中每个曲形文本的拟切分点,基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点,对于弯曲度最大的拟切分点,分析其临近区域,并对该区域进行二值化分析,以对拆分点进行微调,并尽量减少对同一字符的切分,根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本,组合所述至少一个子文本及所述非曲形文本,得到待识别文本,通过组合所述至少一个子文本及所述非曲形文本,得到的所述待识别文本均为非扭曲的正常文本,进而将对难识别的曲形文本的识别问题转化为对多个正常文本的识别问题,利用配置网络对所述待识别文本进行文本识别,得到识别结果,首先通过卷积神经网络学习局部特征信息,再基于循环神经网络想学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。

附图说明

- [0064] 图1是本发明曲形文本识别方法的较佳实施例的流程图。
- [0065] 图2是本发明曲形文本识别装置的较佳实施例的功能模块图。
- [0066] 图3是本发明实现曲形文本识别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

[0067] 为了使本发明的目的、技术方案和优点更加清楚,下面结合附图和具体实施例对本发明进行详细描述。

[0068] 如图1所示,是本发明曲形文本识别方法的较佳实施例的流程图。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。

[0069] 所述曲形文本识别方法应用于一个或者多个电子设备中,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(Application Specific Integrated Circuit,ASIC)、可编程门阵列(Field-Programmable Gate Array,FPGA)、数字处理器(Digital Signal Processor,DSP)、嵌入式设备等。

[0070] 所述电子设备可以是任何一种可与用户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。

[0071] 所述电子设备还可以包括网络设备和/或用户设备。其中,所述网络设备包括,但不限于单个网络服务器、多个网络服务器组成的服务器组或基于云计算(Cloud Computing)的由大量主机或网络服务器构成的云。

[0072] 所述电子设备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。

[0073] S10,响应于文本识别指令,根据所述文本识别指令获取待检测图像。

[0074] 在本实施例中,所述文本识别指令可以由指定工作人员触发,也可以定期触发,本发明不限制。

[0075] 在本发明的至少一个实施例中,所述根据所述文本识别指令获取待检测图像包括:

[0076] 解析所述文本识别指令的方法体,得到所述文本识别指令所携带的信息;

[0077] 获取预设标签;

[0078] 根据所述预设标签构建正则表达式;

[0079] 利用所述正则表达式在所述文本识别指令所携带的信息中进行遍历,并将遍历到的数据确定为目标地址;

[0080] 连接至所述目标地址,并获取所述目标地址处存储的数据作为所述待检测图像。

[0081] 其中,所述文本识别指令实质上是一条代码,在所述文本识别指令中,根据代码的编写原则,{}之间的内容被称之为所述方法体。

[0082] 所述预设标签可以进行自定义配置,所述预设标签与地址具有一一对应的关系,例如:所述预设标签可以是ADD,进一步以所述预设标签建立正则表达式ADD(),并以ADD()进行遍历。

[0083] 通过上述实施方式,能够基于正则表达式及预设标签快速确定目标地址,并进一步获取所述目标地址处存储的数据作为所述待检测图像,提高了数据获取的效率。

[0084] S11,利用DBNet(Differentiable Binarization Net)算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像。

[0085] 在本发明的至少一个实施例中,所述利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像包括:

[0086] 利用DBNet的主干网提取所述待检测图像的特征;

[0087] 对所述特征图进行上采样处理,得到与待检测图像具有相同尺寸的特征图;

[0088] 基于DBNet算法,根据所述特征图进行预测,得到概率图及阈值图;

[0089] 根据所述概率图及所述阈值图进行二值化处理,得到所述至少一个文本区域的掩膜图像。

[0090] 其中,二值化处理是指将概率图转化为边界框和文字区域,通过与阈值的比较,实现二值化。

[0091] 在本实施例中,DBNet的主干网可以采用resnet18或者resnet50。为了提高网络的特征提取能力,还可以引入变形卷积。在resnet输出的至少一个特征图后,采用标准的FPN(feature pyramid networks,特征金字塔)网络结构,即特征金字塔上采样处理到相同的尺寸,得到的特征图用于基于DBNet的head部分生成概率图及阈值图,并通过设置一个固定阈值将分割网络训练得到的概率图转化为二值图,并将转化后得到的二值图确定为所述至少一个文本区域的掩膜图像。

[0092] 具体地,在DBNet网络结构中包括特征提取模块、上采样融合模块及特征图输出等模块。将图片输入网络后,经过特征提取模块和上采样融合模块,得到特征图,在特征图输出模块使用特征图预测出概率图及阈值图,最后计算出二值图,并输出。

[0093] 其中,可以采用标准二值化算法,也可以采用带有自适应阈值的可微分二值化算法,本发明不限制。

[0094] 通过上述实施方式,首先基于DBNet文本检测算法检测待检测图像中的文本区域,能够提供精确的文本轮廓区域的掩膜图像,为后续的文本切分提供可靠的数据基础。

[0095] S12,基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本。

[0096] 在本发明的至少一个实施例中,所述基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本包括:

[0097] 对于所述掩膜图像中的每个文本区域,根据预设间隔建立至少一个点组成每个文本区域的拟合点集;

[0098] 获取每个拟合点集中的初始点及结束点;

[0099] 连接每个拟合点集中的所述初始点及所述结束点,得到每个文本区域的基准线;

[0100] 对于每个文本区域,计算对应拟合点集中的每个点到对应基准线的垂直距离;

[0101] 当每个点到对应基准线的垂直距离都大于预设阈值时,确定对应的文本区域为所述曲形文本;或者

[0102] 当每个点到对应基准线的垂直距离不都大于所述预设阈值时,确定对应的文本区域为所述非曲形文本。

[0103] 通过上述实施方式,能够基于轮廓分析在掩膜图像的基础上进一步执行曲形文本的判断,以便后续进行有针对性的拆分,减少不必要的计算成本。

[0104] S13,识别所述曲形文本中每个曲形文本的拟切分点。

[0105] 在本发明的至少一个实施例中,所述识别所述曲形文本中每个曲形文本的拟切分点包括:

[0106] 对于每个曲形文本,按照由大到小的顺序对每个点到对应基准线的垂直距离进行排序;

[0107] 获取排在首位的点作为每个曲形文本的拟切分点。

[0108] 可以理解的是,垂直距离最高,代表点的弯曲程度最高,以弯曲程度最高的点作为拟切分点,能够更加准确的执行拆分。

[0109] S14,基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点。

[0110] 在本发明的至少一个实施例中,所述基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点包括:

[0111] 将每个拟切分点确定为中心,根据配置延伸范围进行区域划分,得到每个拟切分点对应的临近区域;

[0112] 对每个临近区域进行二值化处理,得到每个临近区域的二值图像;

[0113] 计算每个临近区域的二值图像的垂直投影;

[0114] 根据每个临近区域的垂直投影确定每个曲形文本的目标切分点。

[0115] 可以理解的是,若直接在距离最大点处对文本进行拆分,可能会将一个字拆分成两部分,影响后续的认识,因此,在上述实施方式中,对于弯曲度最大的拟切分点,分析其临近区域,并对该区域进行二值化分析,以对拆分点进行微调,并尽量减少对同一字符的切分。

[0116] S15,根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本。

[0117] 通过上述实施方式,基于检测到的文本区域信息,对每一个曲形文本行进行分析,仅对弯曲度大的文本行进行切分,减小了计算量,并基于二值化方法调整切分点,以保证字符的完整性。

[0118] S16,组合所述至少一个子文本及所述非曲形文本,得到待识别文本。

[0119] 在本实施例中,所述至少一个子文本是经过上述实施方式矫正后得到的文本,因此属于非变形文本。

[0120] 而可以理解的是,当所述待检测图像为曲形文本时,检测效果必然受到一定影响,检测的准确性不佳,因此,本实施方式进一步根据上述矫正后得到的非变形的子文本及所述待检测图像中原本存在的所述非曲形文本构建数据集,作为所述待识别文本。

[0121] 例如:当所述至少一个子文本为x1、x2、x3,所述非曲形文本为x4时,则得到的所述待识别文本为x1、x2、x3及x4组成的数据集。

[0122] 通过组合所述至少一个子文本及所述非曲形文本,得到的所述待识别文本均为非扭曲的正常文本,进而将对难识别的曲形文本的识别问题转化为对多个正常文本的识别问题。

[0123] S17,利用配置网络对所述待识别文本进行文本识别,得到识别结果。

[0124] 在本实施例中,所述配置网络可以为任意具有文本识别功能的网络,如CNN+CTC (Convolutional Neural Networks+Connectionist Temporal Classification,卷积神经网络+连接时序分类)网络。

[0125] 在本实施例中,所述利用配置网络对所述待识别文本进行文本识别,得到识别结果包括:

[0126] 利用卷积神经网络对所述待识别文本进行特征提取,得到目标特征;

[0127] 利用循环神经网络提取所述目标特征的时序特征;

[0128] 将所述时序特征输入至序列识别层,并获取所述序列识别层的输出作为所述识别结果。

[0129] 其中,所述序列识别层可以为连接时序分类CTC。

[0130] 通过上述实施方式,能够首先通过卷积神经网络学习局部特征信息,再基于循环神经网络想学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。

[0131] 需要说明的是,为了进一步保证数据的安全性,可以将识别结果部署于区块链,以避免数据被恶意篡改。

[0132] 由以上技术方案可以看出,本发明能够响应于文本识别指令,根据所述文本识别指令获取待检测图像,利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像,提供精确的文本轮廓区域的掩膜图像,为后续的文本切分提供可靠的数据基础,基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本,能够基于轮廓分析在掩膜图像的基础上进一步执行曲形文本的判断,以便后续进行有针对性的拆分,减少不必要的计算成本,识别所述曲形文本中每个曲形文本的拟切分点,基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点,对于弯曲度最大的拟切分点,分析其临近区域,并对该区域进行二值化分析,以对拆分点进行微调,并尽量减少对同一字符的切分,根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本,组合所述至少一个子文本及所述非曲形文本,得到待识别文本,通过组合所述至少一个子文本及所述非曲形文本,得到的所述待识别文本均为非扭曲的正常文本,进而将对难识别的曲形文本的识别问题转化为对多个正常文本的识别问题,利用配置网络对所述待识别文本进行文本识别,得到识别结果,首先通过卷积神经网络学习局部特征信息,再基于循环神经网络想学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。

[0133] 如图2所示,是本发明曲形文本识别装置的较佳实施例的功能模块图。所述曲形文本识别装置11包括获取单元110、检测单元111、识别单元112、调整单元113、切分单元114、组合单元115、识别单元116。本发明所称的模块/单元是指一种能够被处理器13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12中。在本实施例中,关于各模块/单元的功能将在后续的实施例中详述。

[0134] 响应于文本识别指令,获取单元110根据所述文本识别指令获取待检测图像。

[0135] 在本实施例中,所述文本识别指令可以由指定工作人员触发,也可以定期触发,本发明不限制。

[0136] 在本发明的至少一个实施例中,所述获取单元110根据所述文本识别指令获取待检测图像包括:

[0137] 解析所述文本识别指令的方法体,得到所述文本识别指令所携带的信息;

[0138] 获取预设标签;

[0139] 根据所述预设标签构建正则表达式;

[0140] 利用所述正则表达式在所述文本识别指令所携带的信息中进行遍历,并将遍历到的数据确定为目标地址;

[0141] 连接至所述目标地址,并获取所述目标地址处存储的数据作为所述待检测图像。

[0142] 其中,所述文本识别指令实质上是一条代码,在所述文本识别指令中,根据代码的编写原则,{}之间的内容被称之为所述方法体。

[0143] 所述预设标签可以进行自定义配置,所述预设标签与地址具有一一对应的关系,例如:所述预设标签可以是ADD,进一步以所述预设标签建立正则表达式ADD(),并以ADD()进行遍历。

[0144] 通过上实施方式,能够基于正则表达式及预设标签快速确定目标地址,并进一步获取所述目标地址处存储的数据作为所述待检测图像,提高了数据获取的效率。

[0145] 检测单元111利用DBNet(Differentiable Binarization Net)算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像。

[0146] 在本发明的至少一个实施例中,所述检测单元111利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像包括:

[0147] 利用DBNet的主干网提取所述待检测图像的图像特征;

[0148] 对所述图像特征进行上采样处理,得到与所述待检测图像具有相同尺寸的特征图;

[0149] 基于DBNet算法,根据所述特征图进行预测,得到概率图及阈值图;

[0150] 根据所述概率图及所述阈值图进行二值化处理,得到所述至少一个文本区域的掩膜图像。

[0151] 其中,二值化处理是指将概率图转化为边界框和文字区域,通过与阈值的比较,实现二值化。

[0152] 在本实施例中,DBNet的主干网可以采用resnet18或者resnet50。为了提高网络的特征提取能力,还可以引入变形卷积。在resnet输出的至少一个特征图后,采用标准的FPN(feature pyramid networks,特征金字塔)网络结构,即特征金字塔上采样处理到相同的尺寸,得到的特征图用于基于DBNet的head部分生成概率图及阈值图,并通过设置一个固定阈值将分割网络训练得到的概率图转化为二值图,并将转化后得到的二值图确定为所述至少一个文本区域的掩膜图像。

[0153] 具体地,在DBNet网络结构中包括特征提取模块、上采样融合模块及特征图输出等模块。将图片输入网络后,经过特征提取模块和上采样融合模块,得到特征图,在特征图输出模块使用特征图预测出概率图及阈值图,最后计算出二值图,并输出。

[0154] 其中,可以采用标准二值化算法,也可以采用带有自适应阈值的可微分二值化算法,本发明不限制。

[0155] 通过上述实施方式,首先基于DBNet文本检测算法检测待检测图像中的文本区域,

能够提供精确的文本轮廓区域的掩膜图像,为后续的文本切分提供可靠的数据基础。

[0156] 所述检测单元111基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本。

[0157] 在本发明的至少一个实施例中,所述检测单元111基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本包括:

[0158] 对于所述掩膜图像中的每个文本区域,根据预设间隔建立至少一个点组成每个文本区域的拟合点集;

[0159] 获取每个拟合点集中的初始点及结束点;

[0160] 连接每个拟合点集中的所述初始点及所述结束点,得到每个文本区域的基准线;

[0161] 对于每个文本区域,计算对应拟合点集中的每个点到对应基准线的垂直距离;

[0162] 当每个点到对应基准线的垂直距离都大于预设阈值时,确定对应的文本区域为所述曲形文本;或者

[0163] 当每个点到对应基准线的垂直距离不都大于所述预设阈值时,确定对应的文本区域为所述非曲形文本。

[0164] 通过上述实施方式,能够基于轮廓分析在掩膜图像的基础上进一步执行曲形文本的判断,以便后续进行有针对性的拆分,减少不必要的计算成本。

[0165] 识别单元112识别所述曲形文本中每个曲形文本的拟切分点。

[0166] 在本发明的至少一个实施例中,所述识别单元112识别所述曲形文本中每个曲形文本的拟切分点包括:

[0167] 对于每个曲形文本,按照由大到小的顺序对每个点到对应基准线的垂直距离进行排序;

[0168] 获取排在首位的点作为每个曲形文本的拟切分点。

[0169] 可以理解的是,垂直距离最高,代表点的弯曲程度最高,以弯曲程度最高的点作为拟切分点,能够更加准确的执行拆分。

[0170] 调整单元113基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点。

[0171] 在本发明的至少一个实施例中,所述调整单元113基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点包括:

[0172] 将每个拟切分点确定为中心,根据配置延伸范围进行区域划分,得到每个拟切分点对应的临近区域;

[0173] 对每个临近区域进行二值化处理,得到每个临近区域的二值图像;

[0174] 计算每个临近区域的二值图像的垂直投影;

[0175] 根据每个临近区域的垂直投影确定每个曲形文本的目标切分点。

[0176] 可以理解的是,若直接在距离最大点处对文本进行拆分,可能会将一个字拆分成两部分,影响后续的认识,因此,在上述实施方式中,对于弯曲度最大的拟切分点,分析其临近区域,并对该区域进行二值化分析,以对拆分点进行微调,并尽量减少对同一字符的切分。

[0177] 切分单元114根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本。

[0178] 通过上述实施方式,基于检测到的文本区域信息,对每一个曲形文本行进行分析,

仅对弯曲度大的文本行进行切分,减小了计算量,并基于二值化方法调整切分点,以保证字符的完整性。

[0179] 组合单元115组合所述至少一个子文本及所述非曲形文本,得到待识别文本。

[0180] 在本实施例中,所述至少一个子文本是经过上述实施方式矫正后得到的文本,因此属于非变形文本。

[0181] 而可以理解的是,当所述待检测图像为曲形文本时,检测效果必然受到一定影响,检测的准确性不佳,因此,本实施方式进一步根据上述矫正后得到的非变形的子文本及所述待检测图像中原本存在的所述非曲形文本构建数据集,作为所述待识别文本。

[0182] 例如:当所述至少一个子文本为x1、x2、x3,所述非曲形文本为x4时,则得到的所述待识别文本为x1、x2、x3及x4组成的数据集。

[0183] 通过组合所述至少一个子文本及所述非曲形文本,得到的所述待识别文本均为非扭曲的正常文本,进而将对难识别的曲形文本的识别问题转化为对多个正常文本的识别问题。

[0184] 识别单元116利用配置网络对所述待识别文本进行文本识别,得到识别结果。

[0185] 在本实施例中,所述配置网络可以为任意具有文本识别功能的网络,如CNN+CTC (Convolutional Neural Networks+Connectionist Temporal Classification,卷积神经网络+连接时序分类)网络。

[0186] 在本实施例中,所述识别单元116利用配置网络对所述待识别文本进行文本识别,得到识别结果包括:

[0187] 利用卷积神经网络对所述待识别文本进行特征提取,得到目标特征;

[0188] 利用循环神经网络提取所述目标特征的时序特征;

[0189] 将所述时序特征输入至序列识别层,并获取所述序列识别层的输出作为所述识别结果。

[0190] 其中,所述序列识别层可以为连接时序分类CTC。

[0191] 通过上述实施方式,能够首先通过卷积神经网络学习局部特征信息,再基于循环神经网络想学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。

[0192] 需要说明的是,为了进一步保证数据的安全性,可以将识别结果部署于区块链,以避免数据被恶意篡改。

[0193] 由以上技术方案可以看出,本发明能够响应于文本识别指令,根据所述文本识别指令获取待检测图像,利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像,提供精确的文本轮廓区域的掩膜图像,为后续的文本切分提供可靠的数据基础,基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本,能够基于轮廓分析在掩膜图像的基础上进一步执行曲形文本的判断,以便后续进行有针对性的拆分,减少不必要的计算成本,识别所述曲形文本中每个曲形文本的拟切分点,基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点,对于弯曲度最大的拟切分点,分析其临近区域,并对该区域进行二值化分析,以对拆分点进行微调,并尽量减少对同一字符的切分,根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本,组合所述至少一个子文本及所述非曲形文本,得到待识别文本,通过组合所述至少一个子文本及

所述非曲形文本,得到的所述待识别文本均为非扭曲的正常文本,进而将对难识别的曲形文本的识别问题转化为对多个正常文本的识别问题,利用配置网络对所述待识别文本进行文本识别,得到识别结果,首先通过卷积神经网络学习局部特征信息,再基于循环神经网络想学习时序特征,最后利用序列识别层端到端的语音识别策略识别文字序列,提升了识别效果。

[0194] 如图3所示,是本发明实现曲形文本识别方法的较佳实施例的电子设备的结构示意图。

[0195] 所述电子设备1可以包括存储器12、处理器13和总线,还可以包括存储在所述存储器12中并可在所述处理器13上运行的计算机程序,例如曲形文本识别程序。

[0196] 本领域技术人员可以理解,所述示意图仅仅是电子设备1的示例,并不构成对电子设备1的限定,所述电子设备1既可以是总线型结构,也可以是星形结构,所述电子设备1还可以包括比图示更多或更少的其他硬件或者软件,或者不同的部件布置,例如所述电子设备1还可以包括输入输出设备、网络接入设备等。

[0197] 需要说明的是,所述电子设备1仅为举例,其他现有的或今后可能出现的电子产品如可适应于本发明,也应包含在本发明的保护范围以内,并以引用方式包含于此。

[0198] 其中,存储器12至少包括一种类型的可读存储介质,所述可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器(例如:SD或DX存储器等)、磁性存储器、磁盘、光盘等。存储器12在一些实施例中可以是电子设备1的内部存储单元,例如该电子设备1的移动硬盘。存储器12在另一些实施例中也可以是电子设备1的外部存储设备,例如电子设备1上配备的插接式移动硬盘、智能存储卡(Smart Media Card,SMC)、安全数字(Secure Digital,SD)卡、闪存卡(Flash Card)等。进一步地,存储器12还可以既包括电子设备1的内部存储单元也包括外部存储设备。存储器12不仅可以用于存储安装于电子设备1的应用软件及各类数据,例如曲形文本识别程序的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0199] 处理器13在一些实施例中可以由集成电路组成,例如可以由单个封装的集成电路所组成,也可以是由多个相同功能或不同功能封装的集成电路所组成,包括一个或者多个中央处理器(Central Processing unit,CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。处理器13是所述电子设备1的控制核心(Control Unit),利用各种接口和线路连接整个电子设备1的各个部件,通过运行或执行存储在所述存储器12内的程序或者模块(例如执行曲形文本识别程序等),以及调用存储在所述存储器12内的数据,以执行电子设备1的各种功能和处理数据。

[0200] 所述处理器13执行所述电子设备1的操作系统以及安装的各类应用程序。所述处理器13执行所述应用程序以实现上述各个曲形文本识别方法实施例中的步骤,例如图1所示的步骤。

[0201] 示例性的,所述计算机程序可以被分割成一个或多个模块/单元,所述一个或者多个模块/单元被存储在所述存储器12中,并由所述处理器13执行,以完成本发明。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述所述计算机程序在所述电子设备1中的执行过程。例如,所述计算机程序可以被分割成获取单元110、检测单元111、识别单元112、调整单元113、切分单元114、组合单元115、识别单元116。

[0202] 上述以软件功能模块的形式实现的集成的单元,可以存储在一个计算机可读取存储介质中。上述软件功能模块存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机、计算机设备,或者网络设备等)或处理器(processor)执行本发明各个实施例所述曲形文本识别方法的部分。

[0203] 所述电子设备1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时,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示相关的硬件设备来完成,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。

[0204] 其中,所述计算机程序包括计算机程序代码,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)、随机存取存储器等。

[0205] 进一步地,计算机可读存储介质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等;存储数据区可存储根据区块链节点的使用所创建的数据等。

[0206] 本发明所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了一批网络交易的信息,用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

[0207] 总线可以是外设部件互连标准(peripheral component interconnect,简称PCI)总线或扩展工业标准结构(extended industry standard architecture,简称EISA)总线等。该总线可以分为地址总线、数据总线、控制总线等。为便于表示,在图3中仅用一根箭头表示,但并不表示仅有一根总线或一种类型的总线。所述总线被设置为实现所述存储器12以及至少一个处理器13等之间的连接通信。

[0208] 尽管未示出,所述电子设备1还可以包括给各个部件供电的电源(比如电池),优选地,电源可以通过电源管理装置与所述至少一个处理器13逻辑相连,从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备1还可以包括多种传感器、蓝牙模块、Wi-Fi模块等,在此不再赘述。

[0209] 进一步地,所述电子设备1还可以包括网络接口,可选地,所述网络接口可以包括有线接口和/或无线接口(如WI-FI接口、蓝牙接口等),通常用于在该电子设备1与其他电子设备之间建立通信连接。

[0210] 可选地,该电子设备1还可以包括用户接口,用户接口可以是显示器(Display)、输入单元(比如键盘(Keyboard)),可选地,用户接口还可以是标准的有线接口、无线接口。可选地,在一些实施例中,显示器可以是LED显示器、液晶显示器、触控式液晶显示器以及OLED(Organic Light-Emitting Diode,有机发光二极管)触摸器等。其中,显示器也可以适当的称为显示屏或显示单元,用于显示在电子设备1中处理的信息以及用于显示可视化的用户

界面。

[0211] 应该了解,所述实施例仅为说明之用,在专利申请范围上并不受此结构的限制。

[0212] 图3仅示出了具有组件12-13的电子设备1,本领域技术人员可以理解的是,图3示出的结构并不构成对所述电子设备1的限定,可以包括比图示更少或者更多的部件,或者组合某些部件,或者不同的部件布置。

[0213] 结合图1,所述电子设备1中的所述存储器12存储多个指令以实现一种曲形文本识别方法,所述处理器13可执行所述多个指令从而实现:

[0214] 响应于文本识别指令,根据所述文本识别指令获取待检测图像;

[0215] 利用DBNet算法对所述待检测图像进行文本检测,得到至少一个文本区域的掩膜图像;

[0216] 基于轮廓分析检测所述掩膜图像中的曲形文本及非曲形文本;

[0217] 识别所述曲形文本中每个曲形文本的拟切分点;

[0218] 基于区域划分调整每个曲形文本的拟切分点,得到每个曲形文本的目标切分点;

[0219] 根据每个曲形文本的目标切分点切分对应的曲形文本,得到至少一个子文本;

[0220] 组合所述至少一个子文本及所述非曲形文本,得到待识别文本;

[0221] 利用配置网络对所述待识别文本进行文本识别,得到识别结果。

[0222] 具体地,所述处理器13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。

[0223] 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。

[0224] 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。

[0225] 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。

[0226] 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。

[0227] 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关联图标记视为限制所涉及的权利要求。

[0228] 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。本发明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。第一、第二等词语用来表示名称,而并不表示任何特定的顺序。

[0229] 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。

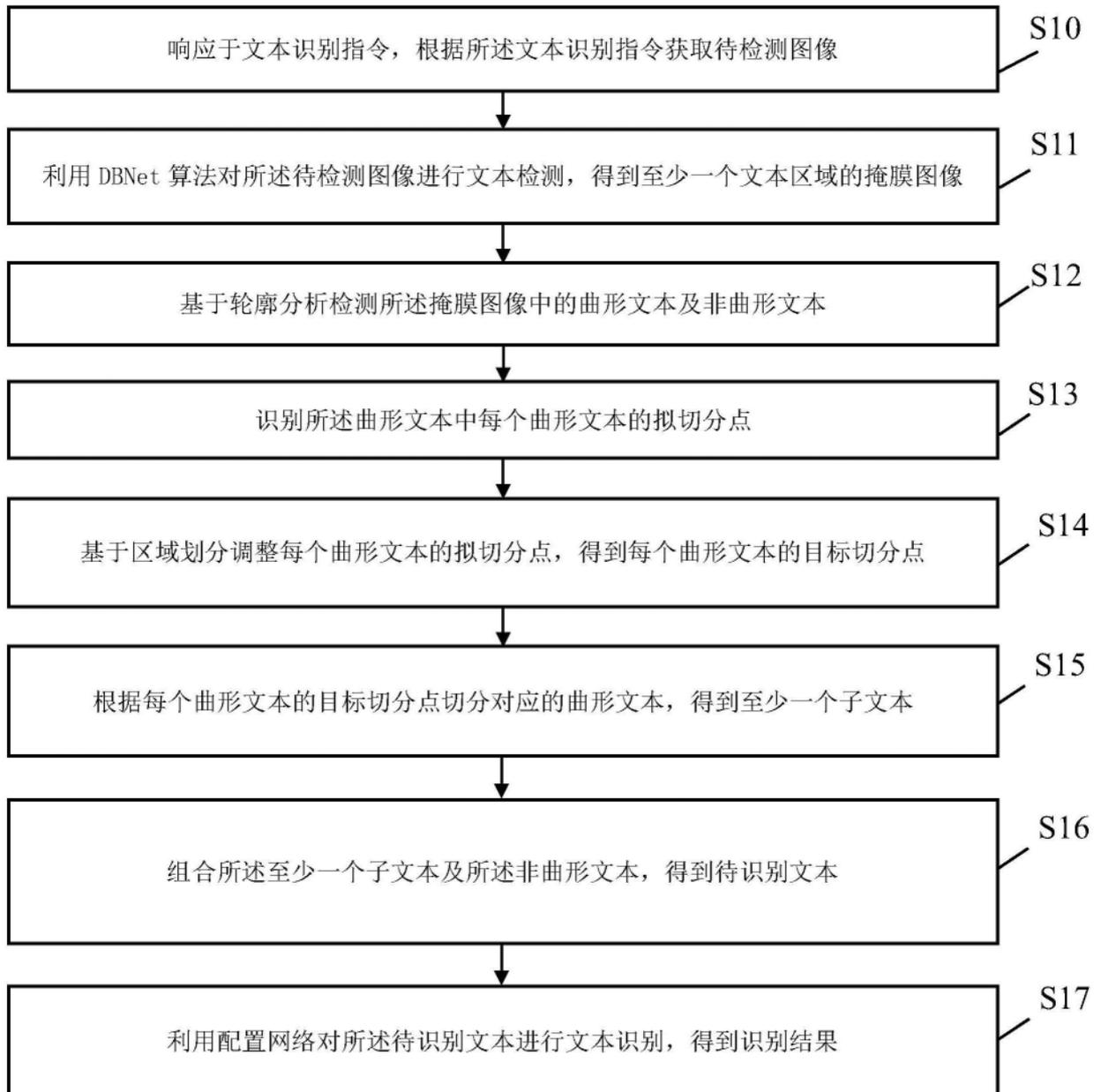


图1

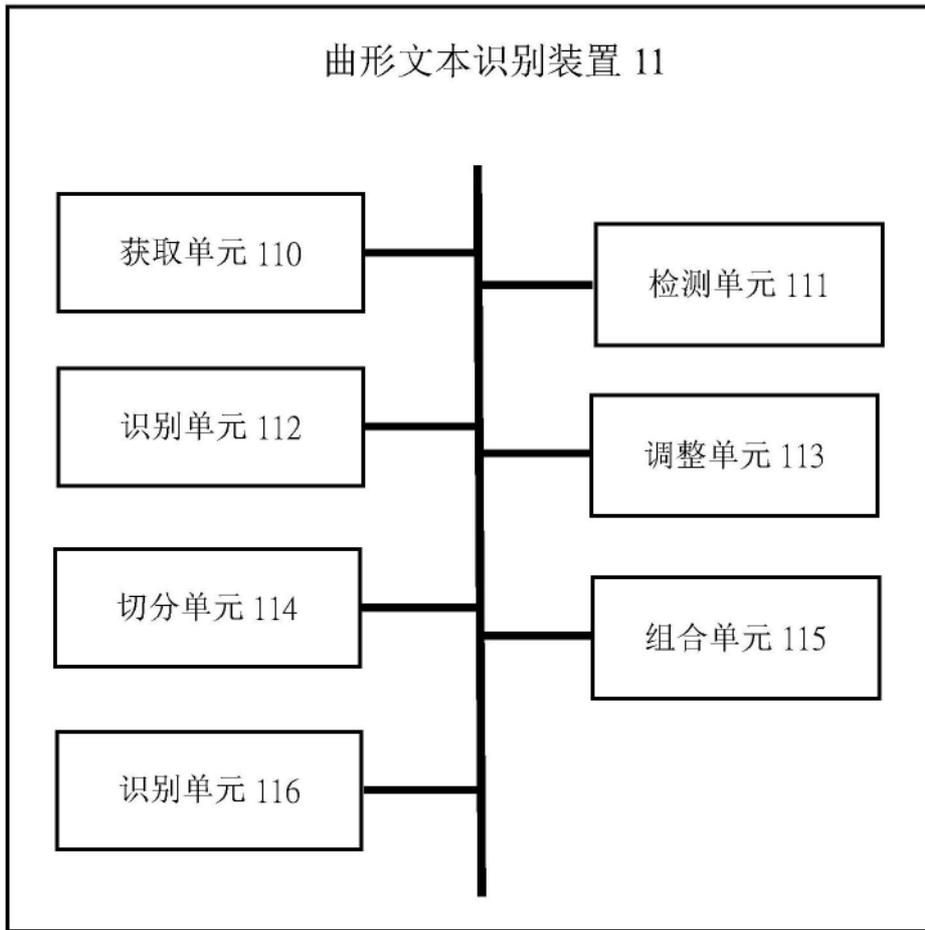


图2

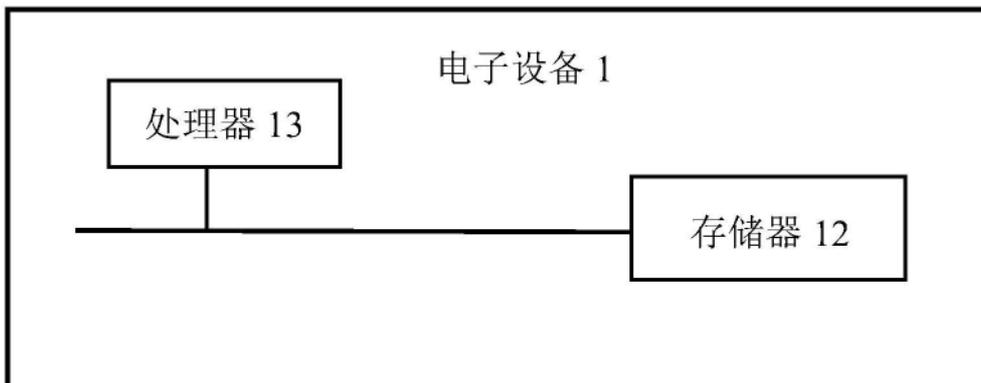


图3