



(12) 发明专利申请

(10) 申请公布号 CN 111767911 A

(43) 申请公布日 2020. 10. 13

(21) 申请号 202010573766.4

G06N 3/04 (2006.01)

(22) 申请日 2020.06.22

G06N 3/08 (2006.01)

(71) 申请人 平安科技(深圳)有限公司

地址 518033 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

(72) 发明人 汤鑫 梁晓云

(74) 专利代理机构 北京鸿元知识产权代理有限  
公司 11327

代理人 李玉琦 曹素云

(51) Int. Cl.

G06K 9/20 (2006.01)

G06K 9/34 (2006.01)

G06K 9/46 (2006.01)

G06K 9/62 (2006.01)

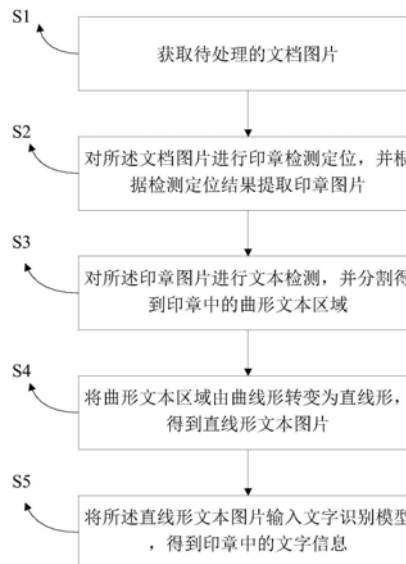
权利要求书2页 说明书12页 附图4页

(54) 发明名称

面向复杂环境的印章文字检测识别方法、装置及介质

(57) 摘要

本发明涉及人工智能,公开了一种面向复杂环境的印章文字检测识别方法、装置及介质,方法包括:获取待处理的文档图片;对文档图片进行印章检测定位,并提取印章图片;对印章图片进行文本检测,并分割得到曲形文本区域;将曲形文本区域由曲线形转变为直线形,得到直线形文本图片;将直线形文本图片输入文字识别模型,得到印章中的文字信息;其中,文字识别模型采用SAR网络进行文字识别,SAR网络包括ResNet模块,用于提取文字特征,获取特征向量;基于LSTM编码器-解码器的框架,框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制。本发明对印章进行端对端的检测识别,对印章的背景文本干扰的鲁棒性较高。



1. 一种面向复杂环境的印章文字检测识别方法,其特征在於,所述方法包括:  
获取待处理的文档图片;

对所述文档图片进行印章检测定位,并根据检测定位结果提取印章图片,其中,所述印章图片为包括印章的最小矩形图片;

对所述印章图片进行文本检测,并分割得到印章中的曲形文本区域;

将曲形文本区域由曲线形转变为直线形,得到直线形文本图片;

将所述直线形文本图片输入文字识别模型,得到印章中的文字信息;

其中,所述文字识别模型采用SAR网络进行文字识别,所述SAR网络包括ResNet模块,用于提取文字特征,并获取特征向量;基于LSTM编码器-解码器的框架,所述框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制;

通过所述ResNet模块获取特征向量,将所述特征向量输入所述LSTM编码器中获得隐藏状态向量;将隐藏状态向量输入施加了注意力机制的解码器中,获得印章中的文字信息。

2. 如权利要求1所述的面向复杂环境的印章文字检测识别方法,其特征在於,对所述文档图片进行印章检测定位的步骤包括:

训练YOLOv3检测模型;

利用训练得到的YOLOv3检测模型获取印章在文档图片中的位置坐标。

3. 如权利要求1所述的面向复杂环境的印章文字检测识别方法,其特征在於,采取PSENet文字检测网络对所述印章图片进行文本检测,检测出印章图片中的各个文本区域。

4. 如权利要求3所述的面向复杂环境的印章文字检测识别方法,其特征在於,对印章图片进行文本检测的步骤包括:

将印章图片输入PSENet文字检测网络,获取与输入的印章图片对应的低维特征图;

对输入的印章图片进行下采样处理,得到高维特征图;

对所述高维特征图进行上采样处理,并与所述低维特征图进行特征融合,得到与输入的印章图片相同尺寸的输出图片;

使用广度优先搜索算法搜索所述输出图片,获取文本连通域,得到印章图片中的文本区域。

5. 如权利要求1所述的面向复杂环境的印章文字检测识别方法,其特征在於,将曲形文本区域由曲线形转变为直线形,得到直线形文本图片的步骤包括:

假设曲形文本区域为圆形区域的一部分,获取所述圆形区域的圆心坐标和圆半径;

根据所述圆心坐标和圆半径估算曲形文本区域对应的圆弧区域,得到所述圆弧区域起点及终点所对应的弧度,并获取所述圆弧区域的最小半径和最大半径;

根据所述圆心坐标、最小半径、最大半径和所述圆弧区域起点及终点所对应的弧度,将直线形文本图片中的坐标对应到印章图片中的坐标,从而将曲形文本区域映射到矩形区域,得到直线形文本图片。

6. 如权利要求5所述的面向复杂环境的印章文字检测识别方法,其特征在於,通过下式将直线形文本图片中的坐标对应到印章图片中的坐标,

$$\begin{cases} x' = c_0 + (r + y)\cos(\alpha - \frac{x}{R}) \\ y' = c_1 - (r + y)\sin(\alpha - \frac{x}{R}) \end{cases}$$

其中,  $(c_0, c_1)$  表示圆心坐标,  $r$  表示圆弧区域的最小半径,  $R$  表示圆弧区域的最大半径,  $\alpha$  表示圆弧区域起点及终点所对应的弧度,  $(x, y)$  表示直线形文本图片中的坐标,  $(x', y')$  表示坐标  $(x, y)$  对应到印章图片中的坐标。

7. 如权利要求5所述的面向复杂环境的印章文字检测识别方法, 其特征在于, 将曲形文本区域由曲线形转变为直线形的步骤之前, 还包括: 判断文本区域是否为曲形, 若文本区域的边缘线的点坐标包围区域的面积与所述包围区域的最小外接矩形面积的比值小于预设阈值, 则所述文本区域为曲形。

8. 一种面向复杂环境的印章文字检测识别装置, 其特征在于, 包括:

图片获取模块, 用于获取待处理的文档图片;

印章提取模块, 用于对所述文档图片进行印章检测定位, 并根据检测定位结果提取印章图片, 其中, 所述印章图片为包括印章的最小矩形图片;

检测分割模块, 用于对所述印章图片进行文本检测, 并分割得到印章中的曲形文本区域;

文本转化模块, 用于将曲形文本区域由曲线形转变为直线形, 得到直线形文本图片;

文字识别模块, 用于将所述直线形文本图片输入文字识别模型, 得到印章中的文字信息;

其中, 所述文字识别模型采用SAR网络进行文字识别, 所述SAR网络包括ResNet模块, 用于提取文字特征, 并获取特征向量; 基于LSTM编码器-解码器的框架, 所述框架包括LSTM编码器和解码器; 注意力模块, 用于向解码器施加注意力机制;

通过所述ResNet模块获取特征向量, 将所述特征向量输入所述LSTM编码器中获得隐藏状态向量; 将隐藏状态向量输入施加了注意力机制的解码器中, 获得印章中的文字信息。

9. 一种电子设备, 其特征在于, 所述电子设备包括:

至少一个处理器; 以及,

与所述至少一个处理器通信连接的存储器; 其中,

所述存储器存储有可被所述至少一个处理器执行的指令, 所述指令被所述至少一个处理器执行, 以使所述至少一个处理器能够执行如权利要求1至7中任一所述的面向复杂环境的印章文字检测识别方法。

10. 一种计算机可读存储介质, 存储有计算机程序, 其特征在于, 所述计算机程序被处理器执行时实现如权利要求1至7中任一所述的面向复杂环境的印章文字检测识别方法。

## 面向复杂环境的印章文字检测识别方法、装置及介质

### 技术领域

[0001] 本发明涉及人工智能,尤其涉及一种面向复杂环境的印章文字检测识别方法、装置、电子设备及计算机可读存储介质。

### 背景技术

[0002] 印章具有权威性,广泛应用于我国的国家机关、团体、企事业单位,盖有印章的文本是具有法律效力的,而印章印文的检验在文件检验中占据了较大比例。目前通常需要采用人工检验的办法来确认文件中所盖印章是否正确,大量的人工校验耗费人力且效率低。面对检验鉴定的工作量日益增长,盖章条件、样本提取质量多样化等情况,印章自动化检测识别具有很大的研究价值和经济效益。但是现有的印章自动识别只能处理没有背景文本干扰的情况,而在真实场景下,由于背景文本的干扰、印章质量的差异等因素,现实场景下印章识别的难度很大。

### 发明内容

[0003] 本发明提供一种面向复杂环境的印章文字检测识别方法、装置、电子设备及计算机可读存储介质,其主要目的在于,在存在背景文本干扰情况下,对印章文字进行检测识别。

[0004] 为了实现上述目的,本发明的第一个方面是提供一种面向复杂环境的印章文字检测识别方法,所述方法包括:

[0005] 获取待处理的文档图片;

[0006] 对所述文档图片进行印章检测定位,并根据检测定位结果提取印章图片,其中,所述印章图片为包括印章的最小矩形图片;

[0007] 对所述印章图片进行文本检测,并分割得到印章中的曲形文本区域;

[0008] 将曲形文本区域由曲线形转变为直线形,得到直线形文本图片;

[0009] 将所述直线形文本图片输入文字识别模型,得到印章中的文字信息;

[0010] 其中,所述文字识别模型采用SAR (Show, Attend and Read) 网络进行文字识别,所述SAR网络包括残差网络 (Residual Network, ResNet) 模块,用于提取文字特征,并获取特征向量;基于长短期记忆网络 (Long Short-Term Memory, LSTM) 编码器-解码器的框架,所述框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制;

[0011] 通过所述ResNet模块获取特征向量,将所述特征向量输入所述LSTM编码器中获得隐藏状态向量;将隐藏状态向量输入施加了注意力机制的解码器中,获得印章中的文字信息。

[0012] 在一个实施例中,对所述文档图片进行印章检测定位的步骤包括:训练 YOLOv3检测模型;利用训练得到的YOLOv3检测模型获取印章在文档图片中的位置坐标。

[0013] 在一个实施例中,采取渐进式扩展网络式 (Progressive Scale Expansion Network, PSENet) 文字检测网络对所述印章图片进行文本检测,检测出印章图片中的各个

文本区域。

[0014] 在一个实施例中,对印章图片进行文本检测的步骤包括:将印章图片输入PSENet文字检测网络,获取与输入的印章图片对应的低维特征图;对输入的印章图片进行下采样处理,得到高维特征图;对所述高维特征图进行上采样处理,并与所述低维特征图进行特征融合,得到与输入的印章图片相同尺寸的输出图片;使用广度优先搜索算法(Breadth First Search,BFS)搜索所述输出图片,获取文本连通域,得到印章图片中的文本区域。

[0015] 在一个实施例中,将曲形文本区域由曲线形转变为直线形,得到直线形文本图片的步骤包括:假设曲形文本区域为圆形区域的一部分,获取所述圆形区域的圆心坐标和圆半径;根据所述圆心坐标和圆半径估算曲形文本区域对应的圆弧区域,得到所述圆弧区域起点及终点所对应的弧度,并获取所述圆弧区域的最小半径和最大半径;根据所述圆心坐标、最小半径、最大半径和所述圆弧区域起点及终点所对应的弧度,将直线形文本图片中的坐标对应到印章图片中的坐标,从而将曲形文本区域映射到矩形区域,得到直线形文本图片。

[0016] 在一个实施例中,通过下式将直线形文本图片中的坐标对应到印章图片中的坐标,

$$[0017] \quad \begin{cases} x' = c_0 + (r + y)\cos(\alpha - \frac{x}{R}) \\ y' = c_1 - (r + y)\sin(\alpha - \frac{x}{R}) \end{cases}$$

[0018] 其中,  $(c_0, c_1)$  表示圆心坐标,  $r$  表示圆弧区域的最小半径,  $R$  表示圆弧区域的最大半径,  $\alpha$  表示圆弧区域起点及终点所对应的弧度,  $(x, y)$  表示直线形文本图片中的坐标,  $(x', y')$  表示坐标  $(x, y)$  对应到印章图片中的坐标。

[0019] 在一个实施例中,将曲形文本区域由曲线形转变为直线形的步骤之前,还包括:判断文本区域是否为曲形,若文本区域的边缘线的点坐标包围区域的面积与所述包围区域的最小外接矩形面积的比值小于预设阈值,则所述文本区域为曲形。

[0020] 为了实现上述目的,本发明的第二个方面是提供一种面向复杂环境的印章文字检测识别装置,包括:

[0021] 图片获取模块,用于获取待处理的文档图片;

[0022] 印章提取模块,用于对所述文档图片进行印章检测定位,并根据检测定位结果提取印章图片,其中,所述印章图片为包括印章的最小矩形图片;

[0023] 检测分割模块,用于对所述印章图片进行文本检测,并分割得到印章中的曲形文本区域;

[0024] 文本转化模块,用于将曲形文本区域由曲线形转变为直线形,得到直线形文本图片;

[0025] 文字识别模块,用于将所述直线形文本图片输入文字识别模型,得到印章中的文字信息;

[0026] 其中,所述文字识别模型采用SAR网络进行文字识别,所述SAR网络包括ResNet模块,用于提取文字特征,并获取特征向量;基于LSTM编码器-解码器的框架,所述框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制;

[0027] 通过所述ResNet模块获取特征向量,将所述特征向量输入所述LSTM编码器中获得

隐藏状态向量;将隐藏状态向量输入施加了注意力机制的解码器中,获得印章中的文字信息。

[0028] 为了实现上述目的,本发明的第三个方面是提供一种电子设备,所述电子设备包括:至少一个处理器;以及与所述至少一个处理器通信连接的存储器;其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如上所述的面向复杂环境的印章文字检测识别方法。

[0029] 为了实现上述目的,本发明的第四个方面是提供一种计算机可读存储介质,存储有计算机程序,所述计算机程序被处理器执行时实现如上所述的面向复杂环境的印章文字检测识别方法。

[0030] 本发明通过文字识别模型对印章进行端对端的检测识别,对印章的背景文本干扰的鲁棒性较高,并且,能够将印章中的曲形文本转化为直线形文本,解决了印章文字难以识别的问题。

[0031] 本发明基于人工智能和图片检测技术自动从文档中提取印章文本内容,且不依赖于印章颜色,对黑白、红色、蓝色等各种颜色的印章都能处理,避免了雇佣大量人力来比对印章内容,节省了人力,提高经济效益。

## 附图说明

[0032] 图1为本发明一实施例提供的印章文字检测识别方法的流程示意图;

[0033] 图2a为本发明一实施例提供的文档图片的示意图;

[0034] 图2b为本发明一实施例提供的印章图片的示意图;

[0035] 图2c为本发明一实施例提供的曲形文本区域的示意图;

[0036] 图2d为本发明一实施例提供的直线形文本图片的示意图;

[0037] 图3为本发明一实施例提供的印章文字检测识别装置的模块示意图;

[0038] 图4为本发明一实施例提供的实现印章文字检测识别方法的电子设备的内部结构示意图;

[0039] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

## 具体实施方式

[0040] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0041] 本发明提供一种面向复杂环境的印章文字检测识别方法。参照图1所示,为本发明一实施例提供的印章文字检测识别方法的流程示意图。该方法可以由一个装置执行,该装置可以由软件和/或硬件实现。

[0042] 在本实施例中,面向复杂环境的印章文字检测识别方法包括:

[0043] 步骤S1,获取待处理的文档图片,参照图2a所示,为本发明一实施例提供的文档图片的示意图,文档图片中具有待识别的印章,并且对印章颜色没有限定,可以是黑白、红色、蓝色等;

[0044] 步骤S2,对所述文档图片进行印章检测定位,并根据检测定位结果提取印章图片,其中,所述印章图片为包括印章的最小矩形图片,参照图2b所示,为本发明一实施例提供的印章图片的示意图,以圆形印章为例,印章图片为包括印章外圆在内的最小矩形;对于输入

的文档,先检测是否存在印章,若存在印章,则需要定位印章位置,根据印章位置提取印章图片;

[0045] 步骤S3,对所述印章图片进行文本检测,并分割得到印章中的曲形文本区域,其中,曲形文本区域指的是包括印章中待识别的文字在内的整体区域呈曲形,可以是椭圆形或圆形,参照图2c所示,为本发明一实施例提供的曲形文本区域的示意图,通过对图2b所示的印章图片进行文本检测分割得到,如图2c所示,在曲形文本区域中,仅包括曲形文本“无锡市XXXXXX会”;

[0046] 步骤S4,将曲形文本区域由曲线形转变为直线形,得到直线形文本图片,参照图2d所示,为本发明一实施例提供的直线形文本图片的示意图,通过对图2c所示的曲形文本区域转化得到,将圆弧形区域转化为矩形区域;

[0047] 步骤S5,将所述直线形文本图片输入文字识别模型,得到印章中的文字信息,例如,获取到印章信息为“无锡市XXXXXX会”;

[0048] 其中,所述文字识别模型采用SAR网络进行文字识别,所述SAR网络包括ResNet模块,用于提取文字特征,并获取特征向量;基于LSTM编码器-解码器的框架,所述框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制;通过所述ResNet模块获取特征向量,将所述特征向量输入所述LSTM编码器中获得隐藏状态向量;将隐藏状态向量输入施加了注意力机制的解码器中,获得印章中的文字信息。

[0049] 本发明的印章文字检测识别可以识别复杂背景下的印章文字,并且,可以识别变形的印章文字,提高印章文字识别的准确率,对印章的背景文本干扰的鲁棒性较好。

[0050] 本发明的文字识别模型利用真实复杂背景图进行训练,所述真实复杂背景图包括公司名称和印章类型字符串数据,作为印章文本内容,以及复杂背景,其中,复杂背景指的是印章图片背景中包括的底色或其他干扰文字等,例如,图2b中的“2019年4月2日”,作为印章图片的背景文字,其不属于印章本身包括的内容,而属于文档的内容。在训练文字识别模型时,将印章类型字符串放置于复杂背景中,并对印章文字做形变以及位置变换处理,位置变换包括旋转、上下左右移动等,使得训练样本具有多样性。

[0051] 在一个实施例中,对所述文档图片进行印章检测定位的步骤包括:训练 YOLOv3检测模型;利用训练得到的YOLOv3检测模型获取印章在文档图片中的位置坐标。

[0052] 其中,YOLOv3检测模型是基于YOLOv2模型改进的目标检测算法,包括多个卷积层,通过所述卷积层获取不同尺度下的特征图,每个特征图均包含了预测目标区域(印章图片)的中心点坐标、目标区域的尺寸以及分类,根据不同尺度下的特征图得到印章在文档图片中的位置坐标。采用这种多尺度特征预测的方法使得预测结果更加精确。

[0053] 其中,训练YOLOv3检测模型的训练样本为带有印章的文本图片,并标注有文本图片中的印章位置坐标。利用这批数据训练模型可以实现印章的精准定位。

[0054] 对于已经检测到的印章,现有的文字识别只能处理横排文字的识别,而印章中可能同时存在曲形文本和矩形文本,因此,需要将印章图片中的所有文本区域检测出来,根据需要分别进行提取。由于印章中各个文字之间的间距比较小,在此通过印章的定位检测得到印章图片之后,将印章图片裁剪下来(参照图2b所示),然后,将裁剪得到的印章图片放大,可以拉开各个文字之间的间隔,方便通过检测算法检测出不同的文本区域。

[0055] 在一个实施例中,采取PSENet (Progressive Scale Expansion Network) 文字检

测网络对所述印章图片进行文本检测,检测出印章图片中的各个文本区域,包括曲形文本区域和/或矩形文本区域。对于矩形文本区域,可以直接输入文字识别模型中进行文字识别,而对于曲形文本区域,则需要将曲形文本区域转化为矩形文本区域,再输入文字识别模型中进行处理。

[0056] 进一步地,对印章图片进行文本检测的步骤包括:

[0057] 将印章图片输入PSENet文字检测网络,获取与输入的印章图片对应的低维特征图,输入图片维度为 $[B, 3, H, W]$ ,其中,B表示批尺寸,H表示图片高度,W表示图片宽度;

[0058] 对输入的印章图片进行下采样处理,得到高维特征图;

[0059] 对所述高维特征图进行上采样处理,并与所述低维特征图进行特征融合,得到与输入的印章图片相同尺寸的输出图片,维度为 $[B, C, H, W]$ ,其中C 为设置的核个数,核根据从小到大表示为 $S_1 \dots S_n$ ,即分割区域;

[0060] 使用广度优先搜索算法(Breadth First Search,BFS)搜索所述输出图片,从 $S_1$ 开始,根据 $S_2$ 加入更多像素来扩展区域,直到 $S_n$ 搜索结束,获取文本连通域,得到印章图片中的文本区域。

[0061] 由于印章中可能存在有曲形文本区域和矩形文本区域,对于曲形文本区域,在印章文字识别时,最关键的就是将曲形文字转为横排直的文本。通过步骤S3可以检测分割得到一块曲形文本区域,其边缘线的点坐标,记为 $\{p_i\}_{i=1}^N$ ,其中, $p_i$ 表示第i个点的坐标 $(x_i, y_i)$ ,这N个点包围的区域记为S。

[0062] 在一个实施例中,将曲形文本区域由曲线形转变为直线形,得到直线形文本图片的步骤包括:

[0063] 步骤S41,假设曲形文本区域S为圆形区域的一部分,获取所述圆形区域的圆心坐标和圆半径;具体地,根据印章的特点,印章曲形文本,主要是圆形和椭圆形。因此,可以假设曲形文本区域S是一个圆形区域的一部分。求解圆形区域的圆心和半径,可以转化为求解下述优化问题:

$$[0064] \quad \min_{r,c} \sum_{i=1}^N (\|p_i - c\|^2 - r^2)^2$$

[0065] 其中, $r$ 为圆半径, $c$ 为圆心坐标, $p_i$ 表示边缘线上第i个点的点坐标,N 为曲形文本区域S的边缘线上点的数量;

[0066] 步骤S42,根据所述圆心坐标和圆半径估算曲形文本区域对应的圆弧区域,得到所述圆弧区域起点及终点所对应的弧度,并获取所述圆弧区域的最小半径和最大半径;具体地,采取试探法,通过边缘线上每个点 $\{p_i\}_{i=1}^N(x_i, y_i)$ 与圆心 $c = (c_0, c_1)$ 的连线,可以求得相应的弧度 $\{\theta_i\}_{i=1}^N$ ,即:

$$[0067] \quad \theta_i = \arctan \frac{c_1 - y_i}{c_0 - x_i}$$

[0068] 其中, $(x_i, y_i)$ 为边缘线上点坐标, $(c_0, c_1)$ 为圆心坐标。

[0069] 现在需要找到文本区域的起始点和终点所对应的弧度值 $\alpha$ 。显然就是找到一个直



线,其与圆心连线对应的弧度值不在 $\{p_i\}_{i=1}^N$ 与圆心连线的弧度值集合中,也即现需要找到一个弧度 $\alpha$ ,使得 $\min_{i \in \{1,2,\dots,N\}} |\theta_i - \alpha| > 0.05$ 。具体地,通过采样的方法,从 $[0, 2\pi]$ 区间中每隔0.01取样,判断是否满足公式 $\min_{i \in \{1,2,\dots,N\}} |\theta_i - \alpha| > 0.05$ ,如果满足,则将弧度值 $\alpha$ 作为曲形文本区域起点及终点所对应的弧度。

[0070] 参照图2c所示,曲形文本区域是由两个同心圆组成的,分别通过下述公式估计最小半径 $r$ 和最大半径 $R$ :

$$[0071] \quad \begin{cases} r = \min_{i=1,2,\dots,N} \|p_i - c\| \\ R = \max_{i=1,2,\dots,N} \|p_i - c\| \end{cases}$$

[0072] 其中, $p_i$ 为边缘线上的第 $i$ 个点坐标, $c$ 为圆心坐标, $r$ 为最小半径, $R$ 为最大半径。

[0073] 为了保证这对同心圆能包括所有的文本,可以将 $R$ 适当的增大以及 $r$ 适当缩小一点,具体调整量根据实际情况确定。

[0074] 步骤S43,根据所述圆心坐标、最小半径、最大半径和所述圆弧区域起点及终点所对应的弧度,将直线形文本图片中的坐标对应到印章图片中的坐标,从而将曲形文本区域映射到矩形区域,得到直线形文本图片。具体地,假设曲形文本区域展开为矩形区域,高为 $R-r$ ,长为 $2\pi R$ 。通过下式将直线形文本图片中的坐标对应到印章图片中的坐标,

$$[0075] \quad \begin{cases} x' = c_0 + (r + y) \cos(\alpha - \frac{x}{R}) \\ y' = c_1 - (r + y) \sin(\alpha - \frac{x}{R}) \end{cases}$$

[0076] 其中, $(c_0, c_1)$ 表示圆心坐标, $r$ 表示圆弧区域的最小半径, $R$ 表示圆弧区域的最大半径, $\alpha$ 表示圆弧区域起点及终点所对应的弧度, $(x, y)$ 表示直线形文本图片中的坐标, $(x', y')$ 表示坐标 $(x, y)$ 对应到印章图片中的坐标。

[0077] 由于印章图片中分割得到的既可能是曲形文本区域,也可能是矩形文本区域。优选地,将曲形文本区域由曲线形转变为直线形的步骤之前,还包括:判断文本区域是否为曲形(包括圆形或椭圆形),若文本区域的边缘线的点坐标包围区域的面积与所述包围区域的最小外接矩形面积的比值小于预设阈值,则所述文本区域为曲形。具体地,利用OPENCV算法,可以计算 $\{p_i\}_{i=1}^N$ 包围区域的面积 $A_1$ ,同时计算 $\{p_i\}_{i=1}^N$ 包围区域的最小外接矩形面积

$A_2$ ,如果 $\frac{A_1}{A_2} < \sigma$ ,则判断出 $\{p_i\}_{i=1}^N$ 包围的区域是曲形区域,其中, $\sigma$ 为设定阈值,其值在 $[0, 1]$ 区间内,优选为 $\sigma = 0.7$ 。若其包围的区域面积 $A_1$ 与最小外接矩形面积 $A_2$ 越接近,比值越接近于1,则文本区域为矩形。

[0078] 本发明通过文字识别模型对印章进行端对端的检测识别,对印章的背景文本干扰的鲁棒性较高,并且,能够将印章中的曲形文本转化为直线形文本,不仅能够识别印章中的曲形文本,也能够识别印章中的矩形文本(直线形文本),解决了印章文字难以识别的问题。

[0079] 如图3所示,是本发明印章文字检测识别装置的功能模块图。

[0080] 本发明所述印章文字检测识别装置100可以安装于电子设备中。根据实现的功能,所述印章文字检测识别装置可以包括图片获取模块101、印章提取模块102、检测分割模块103、文本转化模块104、文字识别模块105。本发所述模块也可以称之为单元,是指一种能够

被电子设备处理器所执行,并且能够完成固定功能的一系列计算机程序段,其存储在电子设备的存储器中。

[0081] 在本实施例中,关于各模块/单元的功能如下:

[0082] 图片获取模块101用于获取待处理的文档图片,其中,文档图片中具有待识别的印章,并且对印章颜色没有限定,可以是黑白、红色、蓝色等;

[0083] 印章提取模块102用于对所述文档图片进行印章检测定位,并根据检测定位结果提取印章图片,其中,所述印章图片为包括印章的最小矩形图片;参照图2b所示,以圆形印章为例,印章图片为包括印章外圆在内的最小矩形;对于输入的文档,先检测是否存在印章,若存在印章,则需要定位印章位置,根据印章位置提取印章图片;

[0084] 检测分割模块103用于对所述印章图片进行文本检测,并分割得到印章中的曲形文本区域;如图2c所示,在曲形文本区域中,仅包括曲形文本“无锡市XXXXXX会”;

[0085] 文本转化模块104用于将曲形文本区域由曲线形转变为直线形,得到直线形文本图片;参照图2d所示,通过对图2c所示的曲形文本区域转化得到,将圆弧形区域转化为矩形区域;

[0086] 文字识别模块105用于将所述直线形文本图片输入文字识别模型,得到印章中的文字信息;

[0087] 其中,所述文字识别模型采用SAR网络进行文字识别,所述SAR网络包括ResNet模块,用于提取文字特征,并获取特征向量;基于LSTM编码器-解码器的框架,所述框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制;

[0088] 通过所述ResNet模块获取特征向量,将所述特征向量输入所述LSTM编码器中获得隐藏状态向量;将隐藏状态向量输入施加了注意力机制的解码器中,获得印章中的文字信息。

[0089] 本发明的印章文字检测识别可以识别复杂背景下的印章文字,并且,可以识别变形的印章文字,提高印章文字识别的准确率,对印章的背景文本干扰的鲁棒性较好。

[0090] 本发明的文字识别模型利用真实复杂背景图进行训练,所述真实复杂背景图包括公司名称和印章类型字符串数据,作为印章文本内容,以及复杂背景,其中,复杂背景指的是印章图片背景中包括的底色或其他干扰文字等,例如,图2b中的“2019年4月2日”,作为印章图片的背景文字,其不属于印章本身包括的内容,而属于文档的内容。在训练文字识别模型时,将印章类型字符串放置于复杂背景中,并对印章文字做形变以及位置变换处理,位置变换包括旋转、上下左右移动等,使得训练样本具有多样性。

[0091] 在一个实施例中,印章提取模块102利用YOLOv3检测模型对所述文档图片进行印章检测定位。具体地,通过下述方式实现:训练YOLOv3检测模型;利用训练得到的YOLOv3检测模型获取印章在文档图片中的位置坐标。

[0092] 其中,YOLOv3检测模型是基于YOLOv2模型改进的目标检测算法模型,包括多个卷积层,通过所述卷积层获取不同尺度下的特征图,每个特征图均包含了预测目标区域(印章图片)的中心点坐标、目标区域的尺寸以及分类,根据不同尺度下的特征图得到印章在文档图片中的位置坐标。采用这种多尺度特征预测的方法使得预测结果更加精确。

[0093] 其中,训练YOLOv3检测模型的训练样本为带有印章的文本图片,并标注有文本图片中的印章位置坐标。利用这批数据训练模型可以实现印章的精准定位。

[0094] 对于已经检测到的印章,现有的文字识别只能处理横排文字的识别,而印章中可能同时存在曲形文本和矩形文本,因此,需要将印章图片中的所有文本区域检测出来,根据需要分别进行提取。由于印章中各个文字之间的间距比较小,在此通过印章的定位检测得到印章图片之后,将印章图片裁剪下来(参照图2b所示),然后,将裁剪得到的印章图片放大,可以拉开各个文字之间的间隔,方便通过检测算法检测出不同的文本区域。

[0095] 在一个实施例中,检测分割模块103采取PSENet(Shape Robust Text Detection with Progressive Scale Expansion Network)文字检测网络对所述印章图片进行文本检测,检测出印章图片中的各个文本区域,包括曲形文本区域和/或矩形文本区域。对于矩形文本区域,可以直接输入文字识别模型中进行文字识别,而对于曲形文本区域,则需要将曲形文本区域转化为矩形文本区域,再输入文字识别模型中进行处理。

[0096] 进一步地,检测分割模块103对印章图片进行文本检测的步骤包括:

[0097] 将印章图片输入PSENet文字检测网络,获取与输入的印章图片对应的低维特征图,输入图片维度为 $[B, 3, H, W]$ ,其中H表示图片高度,W表示图片宽度;

[0098] 对输入的印章图片进行下采样处理,得到高维特征图;

[0099] 对所述高维特征图进行上采样处理,并与所述低维特征图进行特征融合,得到与输入的印章图片相同尺寸的输出图片,维度为 $[B, C, H, W]$ ,其中C为设置的核个数,核根据从小到大表示为 $S_1 \dots S_n$ ;

[0100] 使用广度优先搜索算法(Breadth First Search,BFS)搜索所述输出图片,从 $S_1$ 开始,根据 $S_2$ 加入更多像素来扩展区域,直到 $S_n$ 搜索结束,获取文本连通域,得到印章图片中的文本区域。

[0101] 由于印章中可能存在有曲形文本区域和矩形文本区域,对于曲形文本区域,在印章文字识别时,最关键的就是将曲形文字转为横排直的文本。通过检测分割可以得到一块曲形文本区域,其边缘线的点坐标,记为 $\{p_i\}_{i=1}^N$ ,其中, $p_i$ 表示第i个点的坐标 $(x_i, y_i)$ ,这N个点包围的区域记为S。

[0102] 在一个实施例中,文本转化模块104通过下述方式将曲形文本区域由曲线形转变为直线形,得到直线形文本图片,具体地,包括:

[0103] 步骤S41,假设曲形文本区域S为圆形区域的一部分,获取所述圆形区域的圆心坐标和圆半径;具体地,根据印章的特点,印章曲形文本,主要是圆形和椭圆形。因此,可以假设曲形文本区域S是一个圆形区域的一部分。求解圆形区域的圆心和半径,可以转化为求解下述优化问题:

$$[0104] \quad \min_{r,c} \sum_{i=1}^N (\|p_i - c\|^2 - r^2)^2$$

[0105] 其中,r为圆半径,c为圆心坐标, $p_i$ 表示边缘线上第i个点的点坐标,N为曲形文本区域S的边缘线上点的数量;

[0106] 步骤S42,根据所述圆心坐标和圆半径估算曲形文本区域对应的圆弧区域,得到所述圆弧区域起点及终点所对应的弧度,并获取所述圆弧区域的最小半径和最大半径;具体地,采取试探法,通过边缘线上每个点 $\{p_i\}_{i=1}^N(x_i, y_i)$ 与圆心 $c = (c_0, c_1)$ 的连线,可以求得

相应的弧度 $\{\theta_i\}_{i=1}^N$ ,即

$$[0107] \quad \theta_i = \arctan \frac{c_1 - y_i}{c_0 - x_i}$$

[0108] 其中,  $(x_i, y_i)$  为边缘线上点坐标,  $(c_0, c_1)$  为圆心坐标。

[0109] 现在需要找到文本区域的起始点和终点所对应的弧度值 $\alpha$ 。显然就是找到一个直线, 其与圆心连线对应的弧度值不在 $\{p_i\}_{i=1}^N$ 与圆心连线的弧度值集合中, 也即现需要找到一个弧度 $\alpha$ , 使得 $\min_{i \in \{1, 2, \dots, N\}} |\theta_i - \alpha| > 0.05$ 。具体地, 通过采样的方法, 从 $[0, 2\pi]$ 区间中每隔0.01取样, 判断是否满足公式 $\min_{i \in \{1, 2, \dots, N\}} |\theta_i - \alpha| > 0.05$ , 如果满足, 则将弧度值 $\alpha$ 作为曲形文本区域起点及终点所对应的弧度。

[0110] 参照图2c所示, 曲形文本区域是由两个同心圆组成的, 分别通过下述公式估计最小半径 $r$ 和最大半径 $R$ :

$$[0111] \quad \begin{cases} r = \min_{i=1, 2, \dots, N} \|p_i - c\| \\ R = \max_{i=1, 2, \dots, N} \|p_i - c\| \end{cases}$$

[0112] 其中,  $p_i$ 为边缘线上的第 $i$ 个点坐标,  $c$ 为圆心坐标,  $r$ 为最小半径,  $R$ 为最大半径。

[0113] 为了保证这对同心圆能包括所有的文本, 可以将 $R$ 适当的增大以及 $r$ 适当缩小一点, 具体调整量根据实际情况确定。

[0114] 步骤S43, 根据所述圆心坐标、最小半径、最大半径和所述圆弧区域起点及终点所对应的弧度, 将直线形文本图片中的坐标对应到印章图片中的坐标, 从而将曲形文本区域映射到矩形区域, 得到直线形文本图片。具体地, 假设曲形文本区域展开为矩形区域, 高为 $R-r$ , 长为 $2\pi R$ 。通过下式将直线形文本图片中的坐标对应到印章图片中的坐标,

$$[0115] \quad \begin{cases} x' = c_0 + (r + y) \cos(\alpha - \frac{x}{R}) \\ y' = c_1 - (r + y) \sin(\alpha - \frac{x}{R}) \end{cases}$$

[0116] 其中,  $(c_0, c_1)$ 表示圆心坐标,  $r$ 表示圆弧区域的最小半径,  $R$ 表示圆弧区域的最大半径,  $\alpha$ 表示圆弧区域起点及终点所对应的弧度,  $(x, y)$ 表示直线形文本图片中的坐标,  $(x', y')$ 表示坐标 $(x, y)$ 对应到印章图片中的坐标。

[0117] 由于印章图片中分割得到的既可能是曲形文本区域, 也可能是矩形文本区域。优选地, 将曲形文本区域由曲线形转变为直线形之前, 还包括: 判断文本区域是否为曲形(包括圆形或椭圆形), 若文本区域的边缘线的点坐标包围区域的面积与所述包围区域的最小外接矩形面积的比值小于预设阈值, 则所述文本区域为曲形。具体地, 利用OPENCV算法, 可以计算 $\{p_i\}_{i=1}^N$ 包围区域的面积 $A_1$ , 同时计算 $\{p_i\}_{i=1}^N$ 包围区域的最小外接矩形面积 $A_2$ , 如果 $\frac{A_1}{A_2} < \sigma$ , 则判断出 $\{p_i\}_{i=1}^N$ 包围的区域是曲形区域, 其中,  $\sigma$ 为设定阈值, 其值在 $[0, 1]$ 区间内, 优选为 $\sigma = 0.7$ 。若其包围的区域面积 $A_1$ 与最小外接矩形面积 $A_2$ 越接近, 比值越接近于1, 则文本区域为矩形。

[0118] 如图4所示, 是本发明实现印章文字检测识别方法的电子设备的结构示意图。

[0119] 所述电子设备1可以包括处理器10、存储器11和总线,还可以包括存储在所述存储器11中并可在所述处理器10上运行的计算机程序,如印章文字检测识别程序12。

[0120] 其中,所述存储器11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。所述存储器11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。所述存储器11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,所述存储器11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。所述存储器11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,例如印章文字检测识别程序的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0121] 所述处理器10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。所述处理器10是所述电子设备的控制核心(Control Unit),利用各种接口和线路连接整个电子设备的各个部件,通过运行或执行存储在所述存储器11内的程序或者模块(例如印章文字检测识别程序等),以及调用存储在所述存储器11内的数据,以执行电子设备1的各种功能和处理数据。

[0122] 所述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。所述总线被设置为实现所述存储器11以及至少一个处理器10等之间的连接通信。

[0123] 图4仅示出了具有部件的电子设备的,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。

[0124] 例如,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10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。

[0125] 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。

[0126] 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。

[0127] 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。

[0128] 所述电子设备1中的所述存储器11存储的印章文字检测识别程序12是多个指令的组合,在所述处理器10中运行时,可以实现:

[0129] 获取待处理的文档图片;

[0130] 对所述文档图片进行印章检测定位,并根据检测定位结果提取印章图片,其中,所述印章图片为包括印章的最小矩形图片;

[0131] 对所述印章图片进行文本检测,并分割得到印章中的曲形文本区域;

[0132] 将曲形文本区域由曲线形转变为直线形,得到直线形文本图片;

[0133] 将所述直线形文本图片输入文字识别模型,得到印章中的文字信息;

[0134] 其中,所述文字识别模型采用SAR网络进行文字识别,所述SAR网络包括ResNet模块,用于提取文字特征,并获取特征向量;基于LSTM编码器-解码器的框架,所述框架包括LSTM编码器和解码器;注意力模块,用于向解码器施加注意力机制;

[0135] 通过所述ResNet模块获取特征向量,将所述特征向量输入所述LSTM编码器中获得隐藏状态向量;将隐藏状态向量输入施加了注意力机制的解码器中,获得印章中的文字信息。

[0136] 具体地,所述处理器10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。

[0137] 进一步地,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,可以存储在一个计算机可读取存储介质中。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)。

[0138] 在本发明所提供的几个实施例中,应该理解到,所揭露的设备,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。

[0139] 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。

[0140] 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。

[0141] 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。

[0142] 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。

[0143] 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。

[0144] 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。

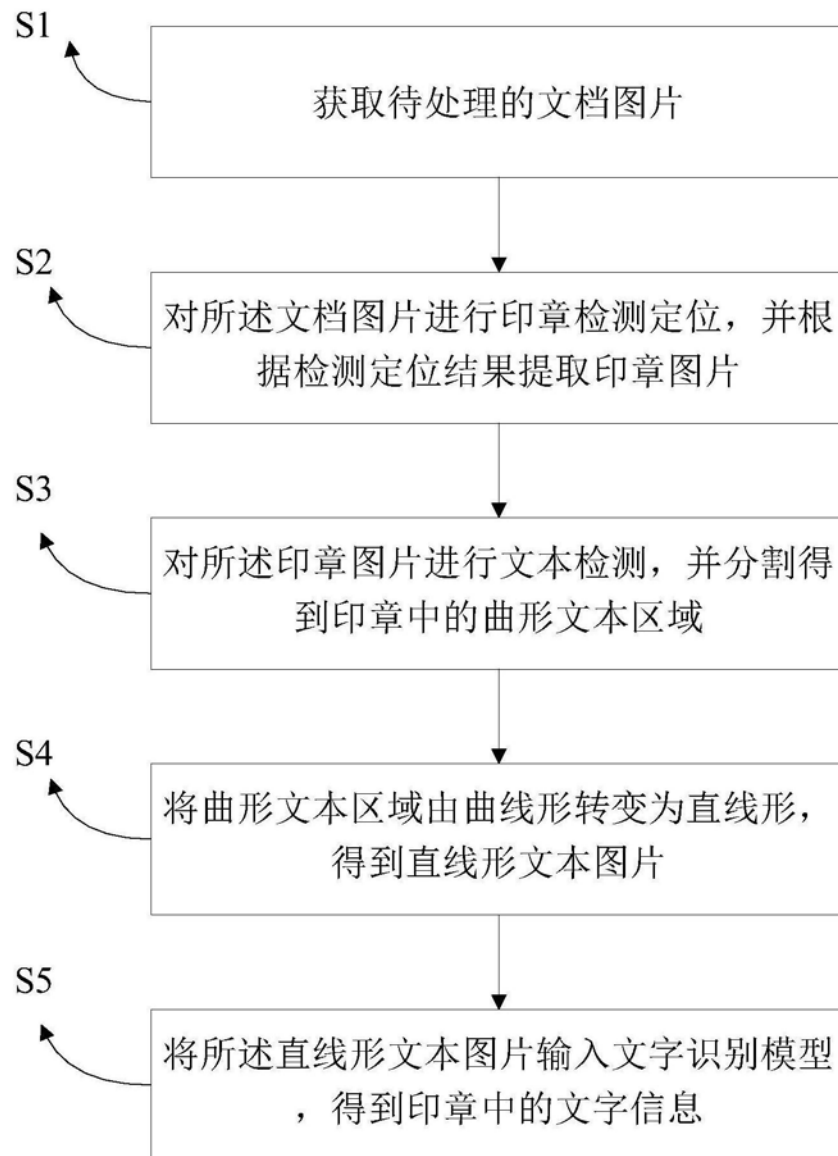


图1





图2a



图2b



图2c



图2d



图3



图4