



(12) 发明专利申请

(10) 申请公布号 CN 114038004 A

(43) 申请公布日 2022.02.11

(21) 申请号 202111375715.1 *G06V 30/146* (2022.01)

(22) 申请日 2021.11.19 *G06V 30/19* (2022.01)

(71) 申请人 北京恒安嘉新安全技术有限公司 *G06V 30/148* (2022.01)  
地址 100086 北京市海淀区北三环西路25号27号楼五层5001室 *G06V 30/10* (2022.01)  
*G06V 10/764* (2022.01)  
*G06V 10/82* (2022.01)

(72) 发明人 刘瑞 李玉惠 傅强 蔡琳 *G06K 9/62* (2022.01)  
阿曼太 梁彧 马寒军 田野 *G06N 3/04* (2006.01)  
王杰 杨满智 金红 陈晓光 *G06N 3/08* (2006.01)

(74) 专利代理机构 北京品源专利代理有限公司  
11332  
代理人 马迪

(51) Int. Cl.  
*G06V 30/42* (2022.01)  
*G06V 30/413* (2022.01)  
*G06V 10/25* (2022.01)

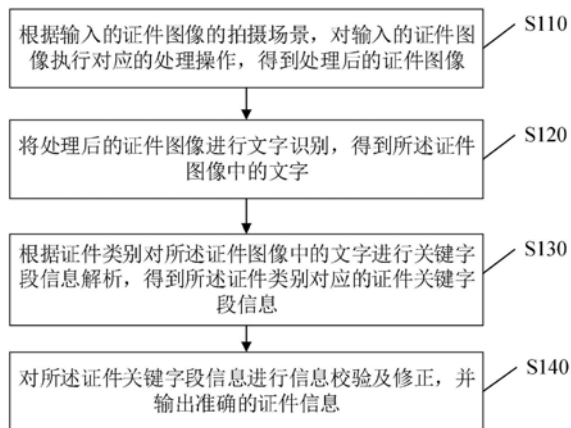
权利要求书2页 说明书12页 附图9页

(54) 发明名称

一种证件信息提取方法、装置、设备及存储介质

(57) 摘要

本发明实施例公开了一种证件信息提取方法、装置、设备及存储介质。该方法包括：根据输入的证件图像的拍摄场景，对输入的证件图像执行对应的处理操作，得到处理后的证件图像；将处理后的证件图像进行文字识别，得到所述证件图像中的文字；根据证件类别对所述证件图像中的文字进行关键字段信息解析，得到所述证件类别对应的证件关键字段信息；对所述证件关键字段信息进行信息校验及修正，并输出准确的证件信息。通过本发明实施例的技术方案，能够在多场景下便捷、快速且识别结果有保障的进行证件信息提取，提升了证件信息提取的工作效率，提高了用户的使用感。



1. 一种证件信息提取方法,其特征在于,包括:

根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像;

将处理后的证件图像进行文字识别,得到所述证件图像中的文字;

根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息;

对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

2. 根据权利要求1所述的方法,其特征在于,根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,包括:

若输入的证件图像的拍摄场景为指定场景,则对输入的证件图像进行文本行检测及倾斜校正操作;

若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作。

3. 根据权利要求1所述的方法,其特征在于,所述根据证件类别对所述证件图像中的文字进行关键字段信息解析,包括:

识别证件图像中包含文字的各个文本框;

将各个所述文本框存入文本框列表中;

从所述文本框列表选择一个文本框,作为目标文本框,将所述目标文本框的坐标作为当前文本行的基准坐标,并遍历所述文本框列表中的剩余文本框;

当存在与目标文本框的Y轴坐标有交叉的文本框时,按X轴坐标的排序添加至当前文本行;

当文本框列表中不存在与目标文本框Y轴坐标有交叉的文本框,返回从所述文本框列表选择一个文本框,作为目标文本框的操作,直至所述文本框列表中不存在文本框;

遍历所有文本行,按照Y轴坐标对所有文本行进行排序,输出排序后的文本行对应的文本行列表;

基于证件类别对应的关键字段与位置信息对所述文本行列表中的文本行进行文本内容解析,得到证件关键字段信息。

4. 根据权利要求2所述的方法,其特征在于,对输入的证件图像进行证件分割以及证件图像预处理操作,包括:

若分割后的证件图像存在倾斜、变形及颠倒的情况,则对分割后的所述证件图像进行证件校正、插值固定尺寸以及证件图像掩膜操作。

5. 根据权利要求2所述的方法,其特征在于,在若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作之前,还包括:

对输入的证件图像进行尺寸缩放以及边框填充,得到符合证件分割模型要求的证件图像;

在若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作之后,还包括:

去除所述证件图像中误检的矩形,获取所述证件图像的轮廓多边形坐标集;

使用聚类算法将所述证件图像的轮廓多边形坐标集进行聚类,得到对应四个角点的角

点集；

对四个角点的所述角点集按照设定规则进行旋转拟合,得到四个角点坐标；

对四个所述角点坐标进行坐标映射,获取四个所述角点坐标在证件原图中对应的四个角坐标,完成证件图像的轮廓检测。

6. 根据权利要求2所述的方法,其特征在于,所述若输入的证件图像的拍摄场景为指定场景,则对输入的证件图像进行文本行检测及倾斜校正操作,包括:

使用深度学习算法对输入的证件图像进行文本检测,获得输入的证件图像中文本框的位置坐标；

遍历输入的证件图像中所有文本框,并获取长度最长的文本框线段,作为目标文本框线段；

计算目标文本框线段的倾斜角度,并判断所述倾斜角度是否小于设定阈值角度；

若否,对所述输入的证件图像进行倾斜校正；

若是,通过文本方向分类模型将所述输入的证件图像中文本的方向分类为0度或者180度,当所述输入的证件图像中文本的方向为180度时,将所述输入的证件图像按照顺时针方向进行180度旋转。

7. 根据权利要求1所述的方法,其特征在于,所述证件关键字段信息包括身份证支持提取的关键字段信息、驾驶证支持提取的关键字段信息以及行驶证支持提取的关键字段信息。

8. 一种证件信息提取装置,其特征在于,包括:

证件区域预处理模块,用于根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像；

文本检测及识别模块,用于将处理后的证件图像进行文字识别,得到所述证件图像中的文字；

关键字段信息提取模块,用于根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息；

信息校验及修正模块,用于对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

9. 一种计算机设备,其特征在于,所述计算机设备包括:

一个或多个处理器；

存储装置,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-7中任一所述的证件信息提取方法。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-7中任一所述的证件信息提取方法。

## 一种证件信息提取方法、装置、设备及存储介质

### 技术领域

[0001] 本发明实施例涉及图像识别技术,尤其涉及一种证件信息提取方法、装置、设备及存储介质。

### 背景技术

[0002] 目前,市面上有多家公司为用户提供了证件信息提取的服务,但大多数厂家对待识别的证件原图做了较为苛刻的限制,如规定了证件的拍摄角度需摆正,证件的正反面需分开单独识别,不同类型的证件需分开单独识别等等。

[0003] 但由于对待识别的证件原图的要求限制,使得用户无法便捷、快速的实现对证件信息的提取,降低了用户的使用感。因此,如何提供一种便捷、快速且识别结果有保障的证件信息提取方法,是目前亟需解决的问题。

### 发明内容

[0004] 本发明实施例提供一种证件信息提取方法、装置、设备及存储介质,可以实现在多场景下便捷、快速且识别结果有保障的进行证件信息提取。

[0005] 第一方面,本发明实施例提供了一种证件信息提取方法,包括:

[0006] 根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像;

[0007] 将处理后的证件图像进行文字识别,得到所述证件图像中的文字;

[0008] 根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息;

[0009] 对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

[0010] 第二方面,本发明实施例还提供了一种证件信息提取装置,该装置包括:

[0011] 证件区域预处理模块,用于根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像;

[0012] 文本检测及识别模块,用于将处理后的证件图像进行文字识别,得到所述证件图像中的文字;

[0013] 关键字段信息提取模块,用于根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息;

[0014] 信息校验及修正模块,用于对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

[0015] 第三方面,本发明实施例还提供了一种计算机设备,所述计算机设备包括:

[0016] 一个或多个处理器;

[0017] 存储装置,用于存储一个或多个程序,

[0018] 当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如本发明任一实施例所述的证件信息提取方法。

[0019] 第四方面,本发明实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如本发明任一实施例所述的证件信息提取方法。

[0020] 本发明实施例的技术方案,通过根据输入的证件图像的拍摄场景将输入的证件图像执行对应的处理,再将处理后的证件图像进行文字识别,并根据证件类别对证件图像中的文字进行关键字段信息解析,得到证件类别对应的证件关键字段信息,最后,对得到的证件关键字段信息进行信息校验及修正,以输出准确的证件信息,解决了现有技术中由于对待识别的证件原图的要求限制,使得用户无法便捷、快速的实现对证件信息的提取的问题,能够实现在多场景下便捷、快速且识别结果有保障的进行证件信息提取,提升了证件信息提取的工作效率,提高了用户的使用感。

## 附图说明

[0021] 图1是本发明实施例提供的一种证件信息提取方法的流程图;

[0022] 图2a是本发明实施例提供的一种证件信息提取方法的流程图;

[0023] 图2b是本发明实施例提供的一种文本行检测及倾斜校正方法的流程示意图;

[0024] 图2c是本发明实施例提供的一种证件校正、插值固定尺寸以及证件图像掩膜方法的流程示意图;

[0025] 图2d是本发明实施例提供的一种实例分割算法的流程示意图;

[0026] 图2e是本发明实施例提供的一种关键字段信息解析方法的流程示意图;

[0027] 图3是本发明实施例提供的一种指定场景下证件信息提取方法的流程示意图;

[0028] 图4是本发明实施例提供的一种自然场景下证件信息提取方法的流程示意图;

[0029] 图5是本发明实施例提供的一种证件信息提取装置的结构示意图;

[0030] 图6是本发明实施例提供的一种计算机设备的结构示意图。

## 具体实施方式

[0031] 下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅仅用于解释本发明,而非对本发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与本发明相关的部分而非全部结构。

[0032] 图1为本发明实施例提供的一种证件信息提取方法的流程图,本实施例可适用于在多种场景下对证件信息进行提取的情况,该方法可以由证件信息提取装置来执行,该装置可采用硬件和/或软件的方式实现,并一般可以集成在计算机设备中。

[0033] 如图1所示,本发明实施例提供的一种证件信息提取方法包括以下具体步骤:

[0034] S110:根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像。

[0035] 其中,输入的证件图像的拍摄场景可以分为利用指定应用程序(Application, App)拍摄输入的证件图像的场景,例如银行身份证上传认证等场景,也可以为在自然场景下拍摄输入的证件图像的场景。值得注意的是,对于指定App拍摄的证件图像,证件图像占拍摄图像的主体部分,且背景较整洁;而自然场景下拍摄的证件图像,证件图像非拍摄图像的主体部分,且背景较杂乱。因此,需要根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,示例性的,去除自然场景下拍摄的证件图像的背景,以使后续操作

可以通过处理后的证件图像对证件信息进行提取。

[0036] 在本发明实施例的一个可选的实施方式中,根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,包括:若输入的证件图像的拍摄场景为指定场景,则对输入的证件图像进行文本行检测及倾斜校正操作;若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作。其中,指定场景可以指利用指定App拍摄输入的证件图像的场景;证件分割可以指将证件图像从输入的证件图像中分割出来的操作;证件图像预处理可以指对证件分割后的证件图像进行倾斜校正或固定尺寸等操作,以使处理后的证件图像满足后续操作的要求。具体的,由于输入的证件图像的拍摄场景为指定场景,可以保证证件图像占输入的证件图像的主体部分,因此,无需进行证件分割获取证件图像的轮廓,但由于拍摄时可能会存在一定的倾斜情况,需对证件图像进行倾斜校正,以使后续证件信息提取可以顺利的进行。相应的,由于输入的证件图像的拍摄场景为自然场景,导致证件图像非输入的证件图像的主体部分,且背景较杂乱,因此,需要对输入的证件图像进行证件分割以及证件图像预处理操作,以保证证件信息提取可以顺利的进行。

[0037] S120:将处理后的证件图像进行文字识别,得到所述证件图像中的文字。

[0038] 其中,文字识别可以指利用相关文字识别算法对处理后的证件图像中的文字进行文字识别,示例性的,可以利用深度学习中的卷积循环神经网络结构(Convolutional Recurrent Neural Network,CRNN)对处理后的证件图像中的文字进行文字识别,CRNN算法可以用于解决基于图像的序列识别问题,特别是场景文字识别问题,CRNN算法最大的特点就是不用先对单个文字进行切割,而是将文本识别转化为时序依赖的序列学习问题,就是基于图像的序列识别。CRNN认为文字识别是对序列的预测方法,所以采用了对序列预测的循环神经网络(Recurrent Neural Network,RNN),通过CNN将图片的特征提取出来后采用RNN对序列进行预测,最后通过一个时序分类(Connectionist Temporal Classification,CTC)的翻译层得到最终结果。

[0039] S130:根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息。

[0040] 其中,证件类别可以指输入的证件图像所属证件的类别,例如,身份证、驾驶证或行驶证。关键字段信息解析可以指根据不同证件类别对证件图像中的关键字段信息进行拆解分析,以获取证件类对应的证件关键字段信息。可选的,所述证件关键字段信息包括身份证支持提取的关键字段信息、驾驶证支持提取的关键字段信息以及行驶证支持提取的关键字段信息。在本发明实施例中,身份证支持提取的关键字段信息可以包含姓名、性别、民族、出生年月、住址、身份证号码、签发机关及有效期等信息;驾驶证支持提取的关键字段信息可以包含证件号、姓名、性别、国籍、出生年月、初次领证日期、有效期、驾照类型、准驾车型及档案编号等信息;行驶证支持提取的关键字段信息可以包含号牌号码、车辆类型、所有人、使用性质、品牌型号、车辆识别代号(Vehicle Identification Number,VIN)、发动机号码、注册日期及发证日期等信息。

[0041] S140:对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

[0042] 其中,信息校验可以指按照证件信息的编码规则对解析出的部分证件关键字段信息进行校验,通常信息校验的对象为证件关键字段信息中的数字关键字段信息或者可以依

据相应编码规则可以推导的关键字段信息,如身份证号码或驾驶证档案编号等。

[0043] 示例性的,若证件类别为身份证,身份证号码由17位本体码和1位校验码组成,其编码规则可以如下:前6位是地址码,表示登记户口时所在地的行政区划代码;7到14位是出生年月日,采用YYYYMMDD格式,如,出生年月日为2021年11月1日,则7到14位为20211101;15到17位是顺序码,表示在同一地址码所标识的区域范围内,对同年、同月、同日出生的人编订的顺序号,顺序码的奇数分配给男性,偶数分配给女性,即第17位奇数表示男性,偶数表示女性;第18位是校验码,采用ISO7064:1983,MOD11-2校验字符系统。若证件类别为驾驶证,驾驶证中的证件号即为身份证号码,编码规则同身份证号码,其档案编号编码规则可以如下:档案编号通常是12位数字连成的串码,前2位为该驾驶证的省份代码;第3-4位为该驾驶证的区域代码;后8位为该驾驶证的顺序编号,由各地车管所自行划分号段,其中省份代码和区域代码均存在一定的范围,可据此进行校验。若证件类别为行驶证,VIN编码规则可以如下:行驶证中的VIN由17个英文或数字组成,为了避免与数字的1或0混淆,英文字母“I”、“O”及“Q”均不使用,同时VIN每一位数字或字母均对应一位数字及加权值,VIN从第一位开始,数字的对应值 $\times$ 该位的加权值,计算全部17位的乘积值相加除以11,所得的余数,即为第九位校验值。

[0044] 其中,修正可以指按照证件信息的编码规则对出现问题的证件关键字段信息进行相应的修正,以对身份证的证件关键字段信息进行相应的修正为例,可以根据身份证号的编码规则从前几位中计算出最后一位,以此对最后一位的数字进行修正;也可以根据身份证号的最后一位判断出具体性别,若性别关键字段中的关键字段信息为与男相近的词,则可以此对其进行修正。

[0045] 由此,根据编码规则可以对关键字段信息的唯一标识信息进行校验及修改,提升关键字段信息输出的准确率。

[0046] 本发明实施例的技术方案,通过根据输入的证件图像的拍摄场景将输入的证件图像执行对应的处理,再将处理后的证件图像进行文字识别,并根据证件类别对证件图像中的文字进行关键字段信息解析,得到证件类别对应的证件关键字段信息,最后,对得到的证件关键字段信息进行信息校验及修正,以输出准确的证件信息,解决了现有技术中由于对待识别的证件原图的要求限制,使得用户无法便捷、快速的实现对证件信息的提取的问题,能够实现在多场景下便捷、快速且识别结果有保障的进行证件信息提取,提升了证件信息提取的工作效率,提高了用户的使用感。

[0047] 图2a为本发明实施例提供的一种证件信息提取方法的流程图。在本实施例中,可选的,对输入的证件图像进行证件分割以及证件图像预处理操作,包括:若分割后的证件图像存在倾斜、变形及颠倒的情况,则对分割后的所述证件图像进行证件校正、插值固定尺寸以及证件图像掩膜操作。

[0048] 如图2a所示,本发明实施例提供的一种证件信息提取方法包括以下具体步骤:

[0049] S210:若输入的证件图像的拍摄场景为指定场景,则对输入的证件图像进行文本检测及倾斜校正操作。

[0050] 对于要识别的证件图像,由于拍摄的设备、场景以及拍摄人等因素,使得输入的证件图像可能是360度内任意角度的图像。由于倾斜的证件图像不仅影响文本检测的效果,也会影响后续的证件关键字段信息提取,因此很有必要对输入的证件图像进行倾斜校正。在

一个可选的实施方式中,所述若输入的证件图像的拍摄场景为指定场景,则对输入的证件图像进行文本行检测及倾斜校正操作,包括:使用深度学习算法对输入的证件图像进行文本检测,获得输入的证件图像中文本框的位置坐标;遍历输入的证件图像中所有文本框,并获取长度最长的文本框线段,作为目标文本框线段;计算目标文本框线段的倾斜角度,并判断所述倾斜角度是否小于设定阈值角度;若否,对所述输入的证件图像进行倾斜校正;若是,通过文本方向分类模型将所述输入的证件图像中文本的方向分类为0度或者180度,当所述输入的证件图像中文本的方向为180度时,将所述输入的证件图像按照顺时针方向进行180度旋转。其中,可以使用深度学习中的可微二值化(Differentiable Binarization, DB)文本检测算法对输入的证件图像进行文本检测,DB文本行检测算法提出了一种可微分的二值化操作,通过将二值化操作插入到分割网络中进行组合优化,从而实现阈值在热力图各处的自适应,相比于传统文本检测方法,可以对自然场景中出现的各种文本行起到一个很好的检测效果。设定阈值角度可以指预先设定的用于对倾斜角度进行评估的数值,在本发明实施例中可将设定阈值角度设为10度。倾斜校正可以指按照倾斜角度与设定阈值角度之间的差值,对输入的证件图像进行角度校正。

[0051] 具体的,如图2b所示为本发明实施例提供的一种文本行检测及倾斜校正方法的流程图示意图。在图2b中,step1:使用DB文本检测算法对输入的证件图像进行文本行检测,获取文本框的位置坐标bbox;step2:遍历当前输入的证件图像中所有的文本框,按照公式  $line_{max} = \max(line_{11}, line_{12}, line_{21}, line_{22}, \dots, line_{n1}, line_{n2})$  获取长度最长的文本框线段  $line_{max}$ ,在本发明实施例中,  $line_{11}$  表示第一个文本框的左上角坐标与右上角坐标确定的线段,规定每个文本框的左上角坐标为  $(x_1, y_1)$ ,顺时针旋转,右上角坐标为  $(x_2, y_2)$ ,右下角坐标为  $(x_3, y_3)$ ,左下角坐标为  $(x_4, y_4)$ ,则根据直线上两点的距离公式可得  $line_{11} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ ,  $line_{21} = \sqrt{(x_1 - x_4)^2 + (y_1 - y_4)^2}$ ,其中,  $line_{n1}, line_{n2}$  则表示第n个文本框的线段长度;step3:获取  $line_{max}$  后,将其带入如下所示的公式中计算倾斜角度angle:

$$angle = \arctan2[line_{max}(y_2) - line_{max}(y_1), line_{max}(x_2) - line_{max}(x_1)] \frac{180}{\pi};$$

[0052] step4:判断当前angle是否大于设定的阈值角度angle\_t(如10度),若angle小于等于angle\_t,说明当前输入的证件图像倾斜程度符合需求,进入step7;若angle大于angle\_t,说明当前输入的证件图像倾斜程度不符合需求,需按照当前angle进行倾斜校正后返回step2;step5:得到校正为0度或180度,上下偏差10度的证件图像;step6:使用基于深度学习技术的文本方向分类模型,判断证件图像中文本方向是0度还是180度,若文字方向为0度进入输出校正后的证件图像;若文字方向为180度,则将输入的证件图像按照顺时针方向进行180度旋转。在本发明实施例中,arctan2为计算两个点之间方位角的函数,可以如下公式所示进行表达:



$$[0053] \quad \arctan2(y,x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & y \geq 0, x < 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & y < 0, x < 0 \\ +\frac{\pi}{2} & y > 0, x = 0 \\ -\frac{\pi}{2} & y < 0, x = 0 \\ 0 & y = 0, x = 0 \end{cases}$$

[0054] 由此,通过利用目标文本框线段的倾斜角度完成对输入的证件图像的倾斜校正,相比于使用传统图像处理算法首先检测证件的最小外接矩形,再根据外接矩形的倾斜角度来校正的方法,鲁棒性更好;并且,相比于额外使用一个新模型来检测证件轮廓再进行校正的方法,本发明实施例所使用的模型为后续操作必需的文本行检测模型,可以在不增加额外算法开销的基础上完成倾斜校正。

[0055] S220:若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割,若分割后的证件图像存在倾斜、变形及颠倒的情况,则对分割后的所述证件图像进行证件校正、插值固定尺寸以及证件图像掩膜操作。

[0056] 其中,证件分割可以采用实例分割算法进行实现,相比于目标检测算法,采用实例分割算法可直接获取证件图像精确的轮廓,有效提升后续信息提取的准确度。

[0057] 如图2c所示为本发明实施例提供的一种证件校正、插值固定尺寸以及证件图像掩膜方法的流程示意图。其中,证件校正用于处理证件图像存在倾斜及变形情况,若证件图像不存在倾斜及变形则无需进行此部分处理,在本发明实施例中,可以利用证件分割输出的4个证件角坐标,获取证件图像的倾斜角度或变形情况,然后使用仿射变换算法对证件图像进行矫正。插值固定尺寸可以将证件图像缩放或扩大至一个固定的尺寸,值得注意的是,将证件图像长度插值至640时,证件图像整体识别率最高。以身份证为例,标准二代身份证长宽比例为1.58:1,身份证图像需插值至640\*410大小;以驾驶证和行驶证为例,两个证件的长宽比例均为1.47:1,驾驶证及行驶证图像需插值至640\*435大小。证件图像掩膜可以指根据空白模板图片关键字段信息坐标位置对校正后的证件图像进行掩膜,由此可以去除大量无需识别的区域的干扰,在提升效率的同时也提升了后续文本检测与文字识别的准确率。

[0058] 由于实例分割算法输出结果是一个证件区域的掩膜,但后续的图像预处理中需要算法输出证件图像的坐标,因此需要对实例分割算法输出的掩模结果添加后处理逻辑,以满足图像预处理操作的要求。在本发明一个可选的实施方式中,在若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作之前,还包括:对输入的证件图像进行尺寸缩放以及边框填充,得到符合证件分割模型要求的证件图像;在若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作之后,还包括:去除所述证件图像中误检的矩形,获取所述证件图像

的轮廓多边形坐标集;使用聚类算法将所述证件图像的轮廓多边形坐标集进行聚类,得到对应四个角点的角点集;对四个角点的所述角点集按照设定规则进行旋转拟合,得到四个角点坐标;对四个所述角点坐标进行坐标映射,获取四个所述角点坐标在证件原图中对应的四个角坐标,完成证件图像的轮廓检测。其中,按照设定规则进行旋转拟合可以指按照设定的角度将角点集进行旋转,示例性的,若实例分割后得到的结果为45度倾斜,则可以将设定的角度定15度,对角点集进行旋转,之后再角点集中的坐标进行拟合,解决了若证件图像经实例分割后得到的结果为倾斜的,则角点聚合的效果将会不佳的问题。

[0059] 具体的,如图2d所示为本发明实施例提供的一种实例分割算法的流程示意图。在图2d中,Step1:将输入的证件图像尺寸缩放为一个固定的尺寸以满足实例分割算法的模型需求;Step2:在图像周围填充一圈边框,防止证件图像在输入的证件图像的边缘时,导致分割不准确的现象发生;Step3:将处理后的输入的证件图像输入至训练好的实例分割算法中,完成证件图像的分割;Step4:对分割后的结果进行分层,使证件图像与背景分开,示例性的,可以将证件图像的像素点置为0,把非证件区域的像素点置为1;Step5:对获取的证件图像进行实例面积校验,示例性的,可以通过计算轮廓面积在矩形中的占比,过滤占比面积比较小的误检矩形;Step6:提取证件图像的轮廓坐标,组成证件图像的轮廓多边形坐标集poly;Step7:使用聚类算法如,K-means,将poly的所有坐标点聚为4类,分别对应4个角点;Step8:对四个角点的角点集按照设定规则进行旋转展平;Step9:对四个角点的角点集分别进行角点坐标拟合,得到得到四个角点坐标;Step10:对四个角点坐标进行坐标映射,获取四角点坐标在证件原图中对应的四个角坐标,完成证件图像的轮廓检测。

[0060] 由此,通过为实例分割算法输出的掩模结果添加后处理逻辑,可以满足后续的图片预处理中的需求,使证件信息提取的结果更加准确、可靠。

[0061] S230:将处理后的证件图像进行文字识别,得到所述证件图像中的文字。

[0062] S240:根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息。

[0063] 如图2e所示为文发明实施例提供的一种关键字段信息解析方法的流程示意图。在本发明实施例中,可选的,所述根据证件类别对所述证件图像中的文字进行关键字段信息解析,包括:识别证件图像中包含文字的各个文本框;将各个所述文本框存入文本框列表中;从所述文本框列表选择一个文本框,作为目标文本框,将所述目标文本框的坐标作为当前文本行的基准坐标,并遍历所述文本框列表中的剩余文本框;当存在与目标文本框的Y轴坐标有交叉的文本框时,按X轴坐标的排序添加至当前文本行;当文本框列表中不存在与目标文本框Y轴坐标有交叉的文本框,返回从所述文本框列表选择一个文本框,作为目标文本框的操作,直至所述文本框列表中不存在文本框;遍历所有文本行,按照Y轴坐标对所有文本行进行排序,输出排序后的文本行对应的文本行列表;基于证件类别对应的关键字段与位置信息对所述文本行列表中的文本行进行文本内容解析,得到证件关键字段信息。其中,将各个所述文本框存入文本框列表中可以指将DB文本检测算法输出的文本框位置坐标bbox以数组的形式存放于文本框列表list\_all中,示例性的,ist\_all=[bbox<sub>1</sub>,bbox<sub>2</sub>,...,bbox<sub>n</sub>],其中,n表示当前证件图像共有n个文本框,每个文本框bbox的形式可以如下所示:bbox<sub>i</sub>=[(x<sub>1</sub>,y<sub>1</sub>),(x<sub>2</sub>,y<sub>2</sub>),(x<sub>3</sub>,y<sub>3</sub>),(x<sub>4</sub>,y<sub>4</sub>)],1≤i≤n。与目标文本框的Y轴坐标有交叉的文本框可以指与目标文本框在同一文本行中的文本框;按X轴坐标的排序添加至当

前文本行可以指按X轴坐标从小到大的书序将文本框添加至当前文本行;按照Y轴坐标对所有文本行进行排序可以指按照Y轴坐标从大到小的顺序对所有文本行进行排序。证件类别对应的关键字段可以指不同证件类别中的各关键字段,例如,身份证中的姓名、年龄或者民族等。基于证件类别对应的关键字段与位置信息对所述文本行列表中的文本行进行文本内容解析可以指,根据证件图像中已知的关键字段与位置信息解析出对应证件类别的证件关键字段信息位置,以此得到证件关键字段信息,示例性的,已知身份证中第二个文本行中包含的关键字段为性别及民族,则可知性别与民族之间的位置为具体的性别关键字段信息,民族后面的位置为具体的民族关键字段信息,由此可获得证件关键字段信息。

[0064] 由此,通过将证件图像中的文本行进行重新排序,可以使得关键字段的位置信息更加准确,保证了基于证件类别对应的关键字段与位置信息对文本行列表中的文本行进行文本内容解析,得到证件关键字段信息的准确性。

[0065] S250:对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

[0066] 本发明实施例的技术方案,在输入的证件图像的拍摄场景为指定场景时,对输入的证件图像进行文本行检测及倾斜校正操作,在输入的证件图像的拍摄场景为自然场景时,对输入的证件图像进行证件分割,并当分割后的证件图像存在倾斜、变形及颠倒的情况时,对分割后的所述证件图像进行证件校正、插值固定尺寸以及证件图像掩膜操作,再将处理后的证件图像进行文字识别,并根据证件类别对证件图像中的文字进行关键字段信息解析,得到证件类别对应的证件关键字段信息,最后,对得到的证件关键字段信息进行信息校验及修正,以输出准确的证件信息,解决了现有技术中由于对待识别的证件原图的要求限制,使得用户无法便捷、快速的实现对证件信息的提取的问题,能够实现在多场景下便捷、快速且识别结果有保障的进行证件信息提取,提升了证件信息提取的工作效率,提高了用户的使用感。

[0067] 在上述实施例的基础上,本发明实施例还包括了训练数据收集、训练数据制作、训练数据标注、模型搭建、模型训练、模型测试以及模型输出转换的操作。由于本发明实施例的三种证件图像均属于敏感数据,目前尚未有相关的公开数据集,因此,所需训练数据需自行收集制作,本发明实施例的技术方案在数据收集过程中通过多来源数据的整合,保证数据完整、多态。首先,将获取的少量的身份证、驾驶证与行驶证图片保存到本地计算机中,挑选其中证件图片分辨率在355\*288以上,质量符合训练要求的数据。然后,收集三种证件图片原件,通过使用不同的终端摄影设备(不同收集、摄像头、相机等),在不同条件(不同背景、光照、距离、角度等)采集三种证件的真实数据。之后,对所有原始数据进行人工标注,其中,标注分为身份证人像面与国徽面、驾驶证主页及副页、行驶证主页及副页6类,每个类别对应一个标签,分别用idcard\_0、idcard\_1、dlcard\_0、dlcard\_1、vlcard\_0、vlcard\_1表示,具体的,可以采用标注软件对以上6类证件数据的轮廓进行多边形标注,经人工标注完成后会生成与标注图片名称对应的标注文件。但由于目前训练样本存在的数据量不足并且各个种类的标注多边形样本不均衡,本发明实施例通过CutMix的方式对人工标注的证件数据进行裁剪,以随机将证件粘贴在其他不包含证件的背景图片中的方式,获得指定数量的证件数据,可以同时解决训练不足及训练样本种类不均衡的问题。最后,将准备好的图片数据和标注文件输入卷积神经网络中,同时对整个卷积神经网络模型的参数进行初始化操作。其中,卷积神经网络的结构是一种层级结构,是由一系列卷积层、激活层、池化层以及归一化

层排列组合而成,最后连接到全连接层和损失层。损失层用于计算预测值与真实值之间的差值,为使差值最小化,将通过卷积神经网络的反向传播算法更新整个卷积神经网络模型的参数。卷积神经网络前向传播和反向传播如此反复迭代N次就可以得到最优的参数,最终得到卷积神经网络算法模型。由此,获得训练好的卷积神经网络算法模型,使得输入的证件图像的拍摄场景为自然场景时,可以将处理后的输入的证件图像输入至训练好的实例分割算法中,完成证件图像的分割。

[0068] 图3为本发明实施例提供的一种指定场景下证件信息提取方法的流程示意图。在图3中,输入的证件图像的拍摄场景为指定场景,首先,对输入的证件图像进行文本行检测及倾斜校正操作;之后,对处理好的证件图像进行文本行定位(即文本行检测)及文字识别,以得到含有文本的文本框;然后,根据证件分类,即身份证、驾驶证及行驶证,对识别出的证件图像中的文字进行关键字段信息解析及信息校验,若信息校验不符合相应的编码规则,则按照对应的编码规则对关键字段信息进行信息修正,之后返回执行关键字段信息解析的步骤,直到信息校验的结果准确,对证件信息进行提取,将结构化识别结果进行输出。由此,能够实现在指定场景下便捷、快速且识别结果有保障的进行证件信息提取,完成对关键字段信息的提取,输出结构化的识别结果,极大地节省了人工手动输入的时间成本,提升了证件信息提取的工作效率,提高了用户的使用感。

[0069] 图4为本发明实施例提供的一种自然场景下证件信息提取方法的流程示意图。在图4中,输入的证件图像的拍摄场景为自然场景,首先,将输入的证件图像输入至证件分割网络中进行证件分割,其中,在证件分割网络中可以判断出是否存在证件图像以及证件图像的类别,示例性的,若证件分割网络输出了证件图像的坐标即证明有证件,若输出了0或者空版本则证明不存在证件图像。当存在证件图像时,对分割后的证件图像依次进行证件图像预处理、文本定位(即文本行检测)及文字识别操作,并对文字识别的结果进行关键字段信息解析及信息校验,若信息校验不符合相应的编码规则,则按照对应的编码规则对关键字段信息进行信息修正,之后返回执行关键字段信息解析的步骤,直到信息校验的结果准确,对证件信息进行提取,将结构化识别结果进行输出。当不存在证件图像时,即输入的证件图像的类别可能不属于预先定义的证件类别,或输入的证件图像质量较差,无法被证件分割网络检测到,若对该证件图像进行证件信息提取将会导致提取出的证件信息准确率较低,无实用价值,因此,直接输出一个空的识别结果,即对此种情况的证件图像暂时不进行识别。由此,能够实现在自然场景下便捷、快速且识别结果有保障的进行证件信息提取,完成对关键字段信息的提取,输出结构化的识别结果,极大地节省了人工手动输入的时间成本,提升了证件信息提取的工作效率,提高了用户的使用感。

[0070] 图5为本发明实施例提供的一种证件信息提取装置的结构示意图,该装置可以执行上述各实施例中涉及到的证件信息提取方法。该装置可采用软件和/或硬件的方式实现,如图5所示,所述证件信息提取装置具体包括:证件区域预处理模块510、文本检测及识别模块520、关键字段信息提取模块530、信息校验及修正模块540。

[0071] 其中,证件区域预处理模块510,用于根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像;

[0072] 文本检测及识别模块520,用于将处理后的证件图像进行文字识别,得到所述证件图像中的文字;

[0073] 关键字段信息提取模块530,用于根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息;

[0074] 信息校验及修正模块540,用于对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

[0075] 本发明实施例的技术方案,通过根据输入的证件图像的拍摄场景将输入的证件图像执行对应的处理,再将处理后的证件图像进行文字识别,并根据证件类别对证件图像中的文字进行关键字段信息解析,得到证件类别对应的证件关键字段信息,最后,对得到的证件关键字段信息进行信息校验及修正,以输出准确的证件信息,解决了现有技术中由于对待识别的证件原图的要求限制,使得用户无法便捷、快速的实现对证件信息的提取的问题,能够实现在多场景下便捷、快速且识别结果有保障的进行证件信息提取,提升了证件信息提取的工作效率,提高了用户的使用感。

[0076] 可选的,证件区域预处理模块510可以包括第一证件区域预处理单元及第二证件区域预处理单元;

[0077] 其中,第一证件区域预处理单元,用于若输入的证件图像的拍摄场景为指定场景,则对输入的证件图像进行文本行检测及倾斜校正操作;

[0078] 第二证件区域预处理单元,用于若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作。

[0079] 可选的,关键字段信息提取模块530具体可以用于,识别证件图像中包含文字的各个文本框;将各个所述文本框存入文本框列表中;从所述文本框列表选择一个文本框,作为目标文本框,将所述目标文本框的坐标作为当前文本行的基准坐标,并遍历所述文本框列表中的剩余文本框;当存在与目标文本框的Y轴坐标有交叉的文本框时,按X轴坐标的排序添加至当前文本行;当文本框列表中不存在与目标文本框Y轴坐标有交叉的文本框,返回从所述文本框列表选择一个文本框,作为目标文本框的操作,直至所述文本框列表中不存在文本框;遍历所有文本行,按照Y轴坐标对所有文本行进行排序,输出排序后的文本行对应的文本行列表;基于证件类别对应的关键字段与位置信息对所述文本行列表中的文本行进行文本内容解析,得到证件关键字段信息。

[0080] 可选的,第二证件区域预处理单元具体可以用于,若分割后的证件图像存在倾斜、变形及颠倒的情况,则对分割后的所述证件图像进行证件校正、插值固定尺寸以及证件图像掩膜操作。

[0081] 可选的,证件信息提取装置还包括前处理模块,用于在若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作之前,对输入的证件图像进行尺寸缩放以及边框填充,得到符合证件分割模型要求的证件图像;

[0082] 相应的,证件信息提取装置还包括后处理模块,用于在若输入的证件图像的拍摄场景为自然场景,则对输入的证件图像进行证件分割以及证件图像预处理操作之后,去除所述证件图像中误检的矩形,获取所述证件图像的轮廓多边形坐标集;使用聚类算法将所述证件图像的轮廓多边形坐标集进行聚类,得到对应四个角点的角点集;对四个角点的所述角点集按照设定规则进行旋转拟合,得到四个角点坐标;对四个所述角点坐标进行坐标映射,获取四个所述角点坐标在证件原图中对应的四个角坐标,完成证件图像的轮廓检测。

[0083] 可选的,第一证件区域预处理单元具体可以用于,使用深度学习算法对输入的证

件图像进行文本检测,获得输入的证件图像中文本框的位置坐标;遍历输入的证件图像中所有文本框,并获取长度最长的文本框线段,作为目标文本框线段;计算目标文本框线段的倾斜角度,并判断所述倾斜角度是否小于设定阈值角度;若否,对所述输入的证件图像进行倾斜校正;若是,通过文本方向分类模型将所述输入的证件图像中文本的方向分类为0度或者180度,当所述输入的证件图像中文本的方向为180度时,将所述输入的证件图像按照顺时针方向进行180度旋转。

[0084] 可选的,所述证件关键字段信息包括身份证支持提取的关键字段信息、驾驶证支持提取的关键字段信息以及行驶证支持提取的关键字段信息。

[0085] 本发明实施例所提供的证件信息提取装置可执行本发明任意实施例所提供的证件信息提取方法,具备执行方法相应的功能模块和有益效果。

[0086] 图6为本发明实施例提供的一种计算机设备的结构示意图,如图6所示,该计算机设备包括处理器610、存储器620、输入装置630和输出装置640;计算机设备中处理器610的数量可以是一个或多个,图6中以一个处理器610为例;计算机设备中的处理器610、存储器620、输入装置630和输出装置640可以通过总线或其他方式连接,图6中以通过总线连接为例。

[0087] 存储器620作为一种计算机可读存储介质,可用于存储软件程序、计算机可执行程序以及模块,如本发明实施例中的证件信息提取方法对应的程序指令/模块(例如,证件信息提取装置中的证件区域预处理模块510、文本检测及识别模块520、关键字段信息提取模块530和信息校验及修正模块540)。处理器610通过运行存储在存储器620中的软件程序、指令以及模块,从而执行计算机设备的各种功能应用以及数据处理,即实现上述的证件信息提取方法。

[0088] 存储器620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;存储数据区可存储根据终端的使用所创建的数据等。此外,存储器620可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实例中,存储器620可进一步包括相对于处理器610远程设置的存储器,这些远程存储器可以通过网络连接至计算机设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0089] 输入装置630可用于接收输入的数字或字符信息,以及产生与计算机设备的用户设置以及功能控制有关的键信号输入。输出装置640可包括显示屏等显示设备。

[0090] 本发明实施例还提供一种包含计算机可执行指令的存储介质,所述计算机可执行指令在由计算机处理器执行时用于执行一种证件信息提取方法,该方法包括:

[0091] 根据输入的证件图像的拍摄场景,对输入的证件图像执行对应的处理操作,得到处理后的证件图像;

[0092] 将处理后的证件图像进行文字识别,得到所述证件图像中的文字;

[0093] 根据证件类别对所述证件图像中的文字进行关键字段信息解析,得到所述证件类别对应的证件关键字段信息;

[0094] 对所述证件关键字段信息进行信息校验及修正,并输出准确的证件信息。

[0095] 当然,本发明实施例所提供的一种包含计算机可执行指令的存储介质,其计算机可执行指令不限于如上所述的方法操作,还可以执行本发明任意实施例所提供的证件信息

提取方法中的相关操作。

[0096] 通过以上关于实施方式的描述,所属领域的技术人员可以清楚地了解到,本发明可借助软件及必需的通用硬件来实现,当然也可以通过硬件实现,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在计算机可读存储介质中,如计算机的软盘、只读存储器(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory, RAM)、闪存(FLASH)、硬盘或光盘等,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,服务器,或者网络设备等)执行本发明各个实施例所述的方法。

[0097] 值得注意的是,上述证件信息提取装置的实施例中,所包括的各个单元和模块只是按照功能逻辑进行划分的,但并不局限于上述的划分,只要能够实现相应的功能即可;另外,各功能单元的具体名称也只是为了便于相互区分,并不用于限制本发明的保护范围。

[0098] 注意,上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解,本发明不限于这里所述的特定实施例,对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此,虽然通过以上实施例对本发明进行了较为详细的说明,但是本发明不仅仅限于以上实施例,在不脱离本发明构思的情况下,还可以包括更多其他等效实施例,而本发明的范围由所附的权利要求范围决定。

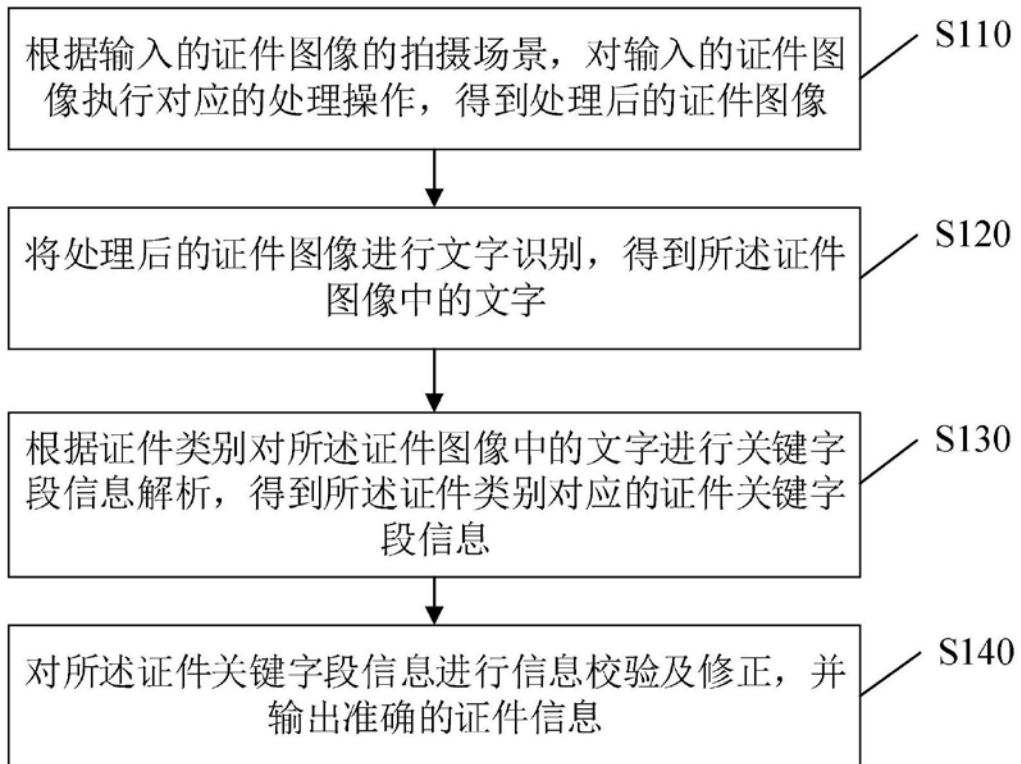


图1



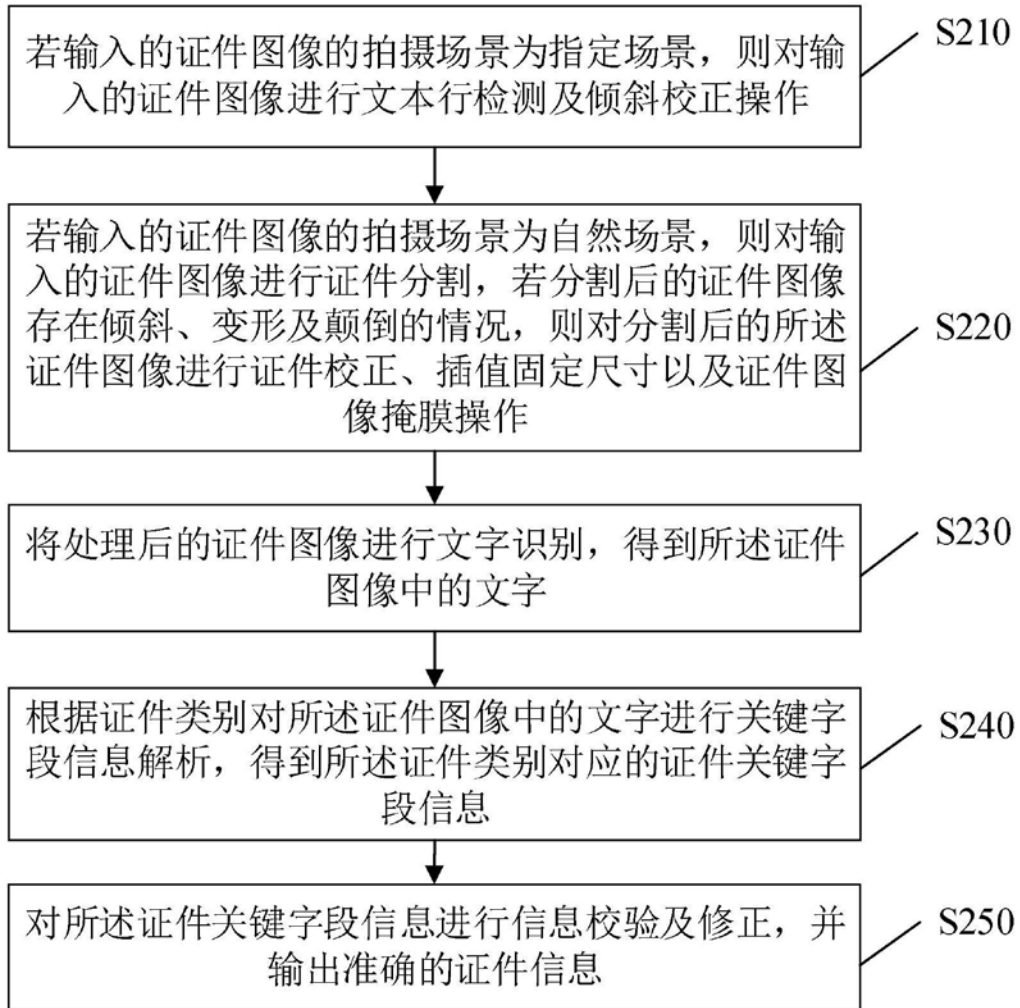


图2a

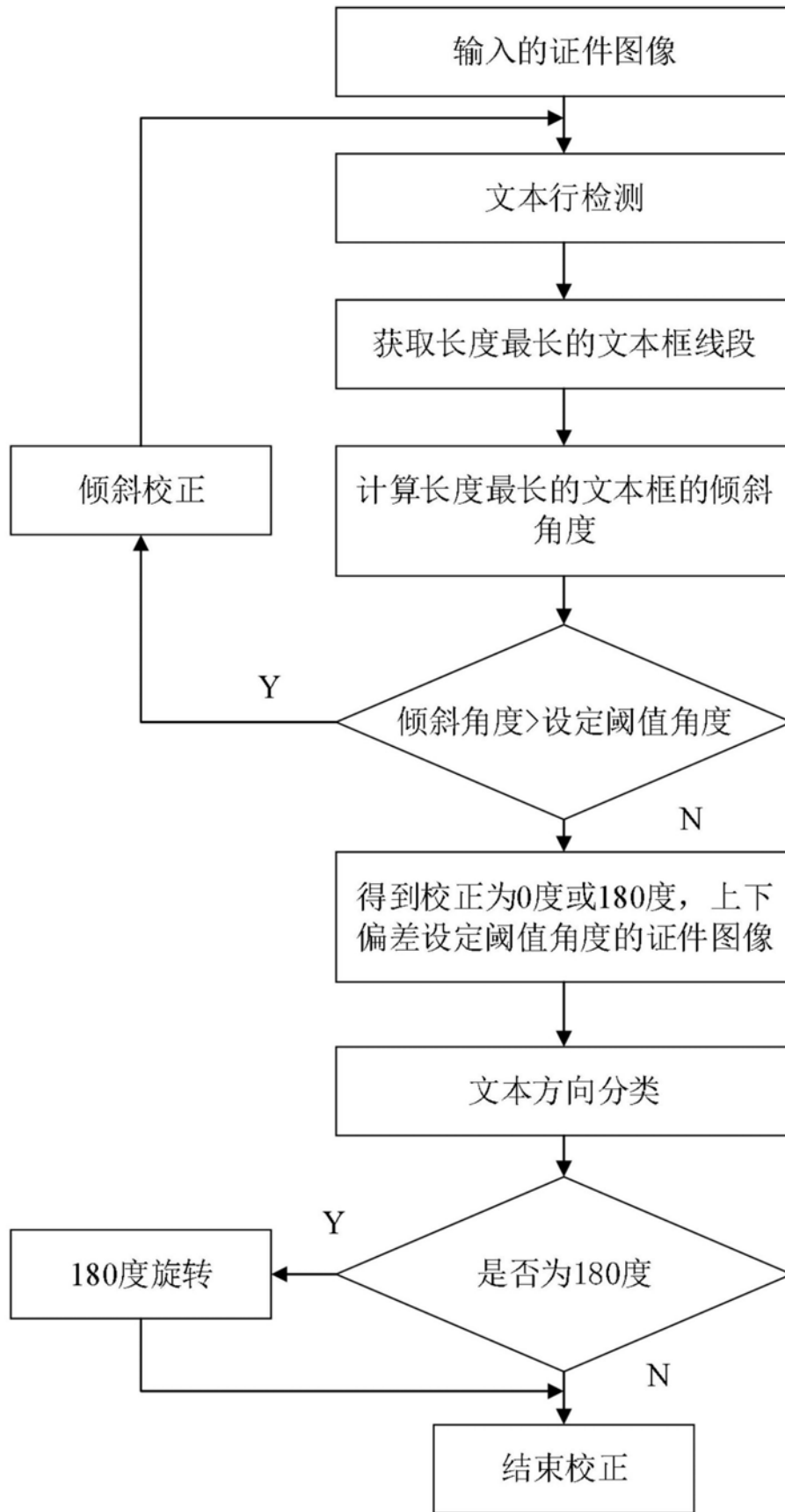


图2b

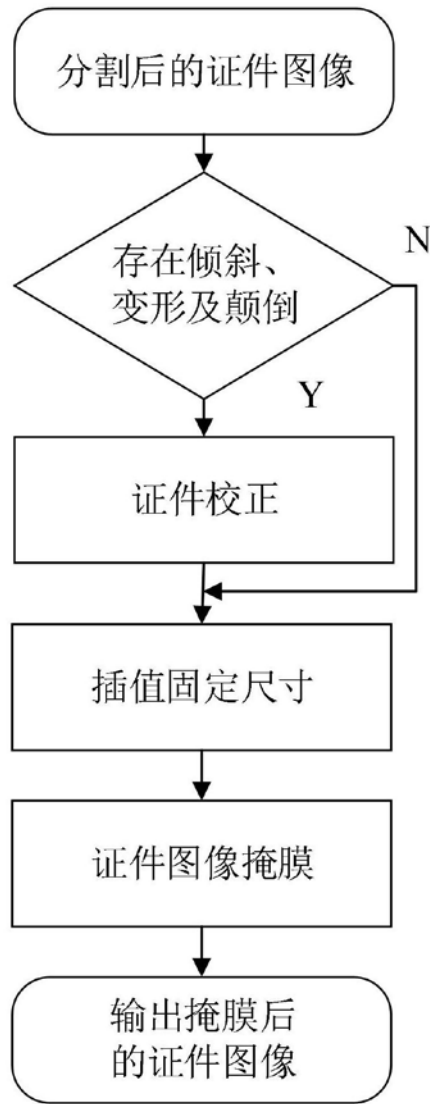


图2c

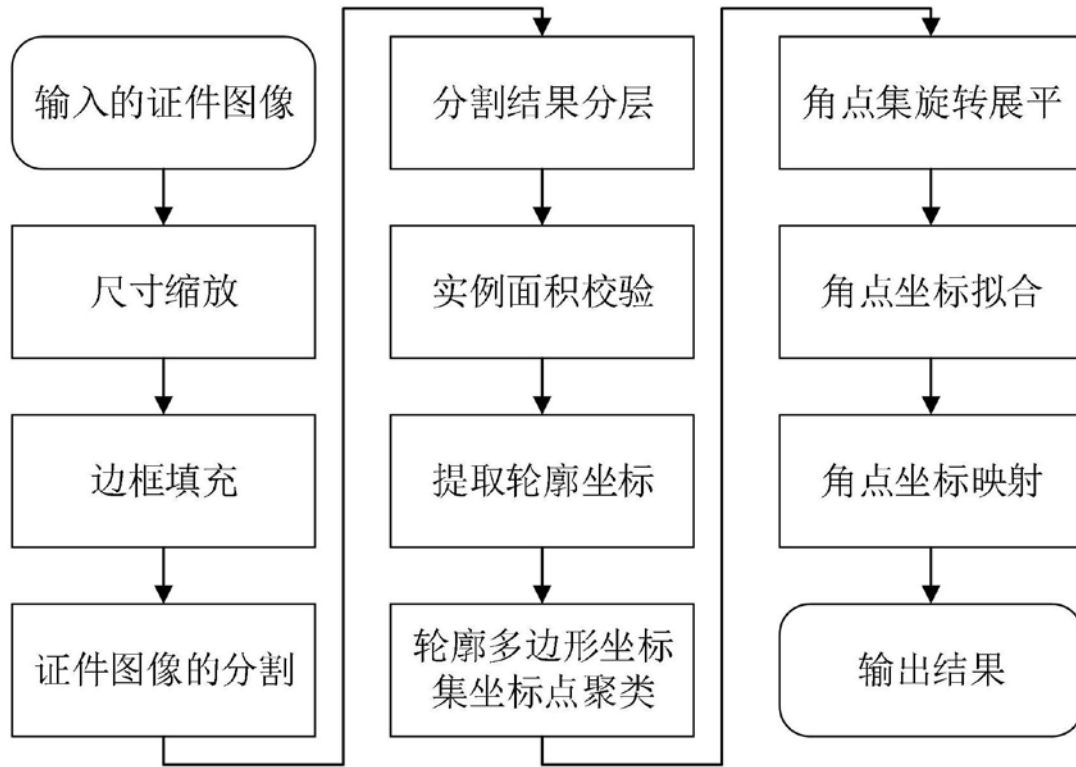


图2d

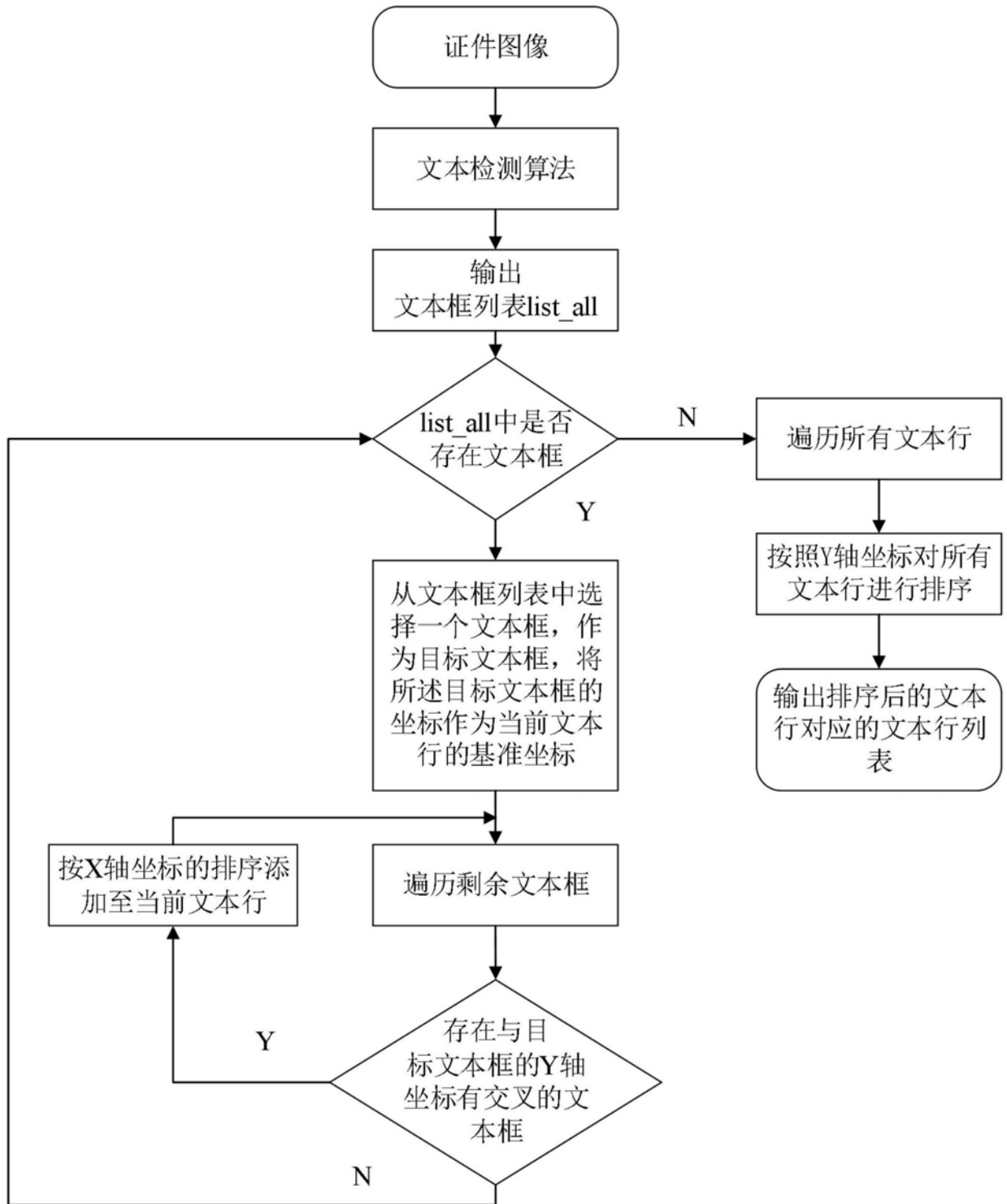


图2e

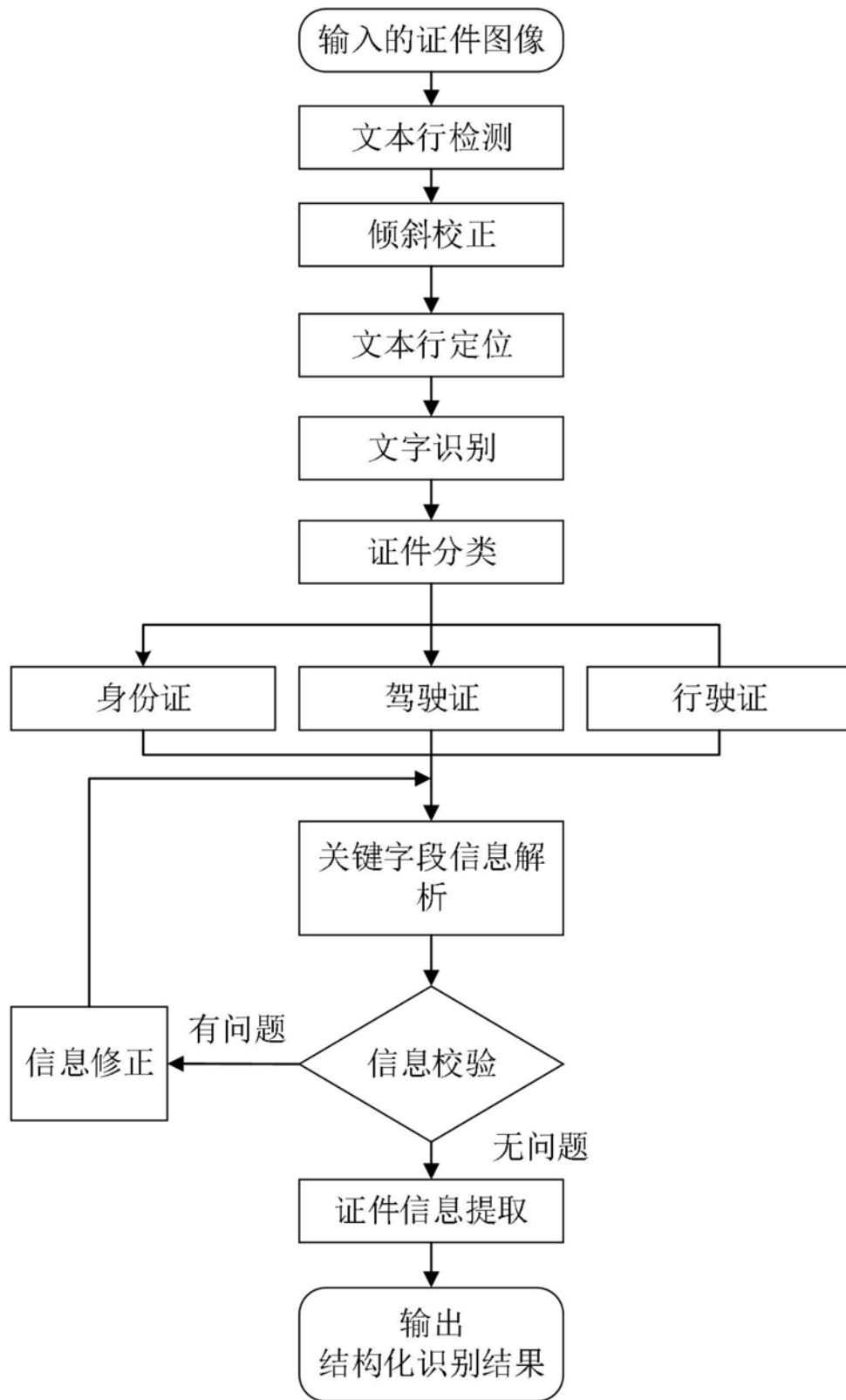


图3

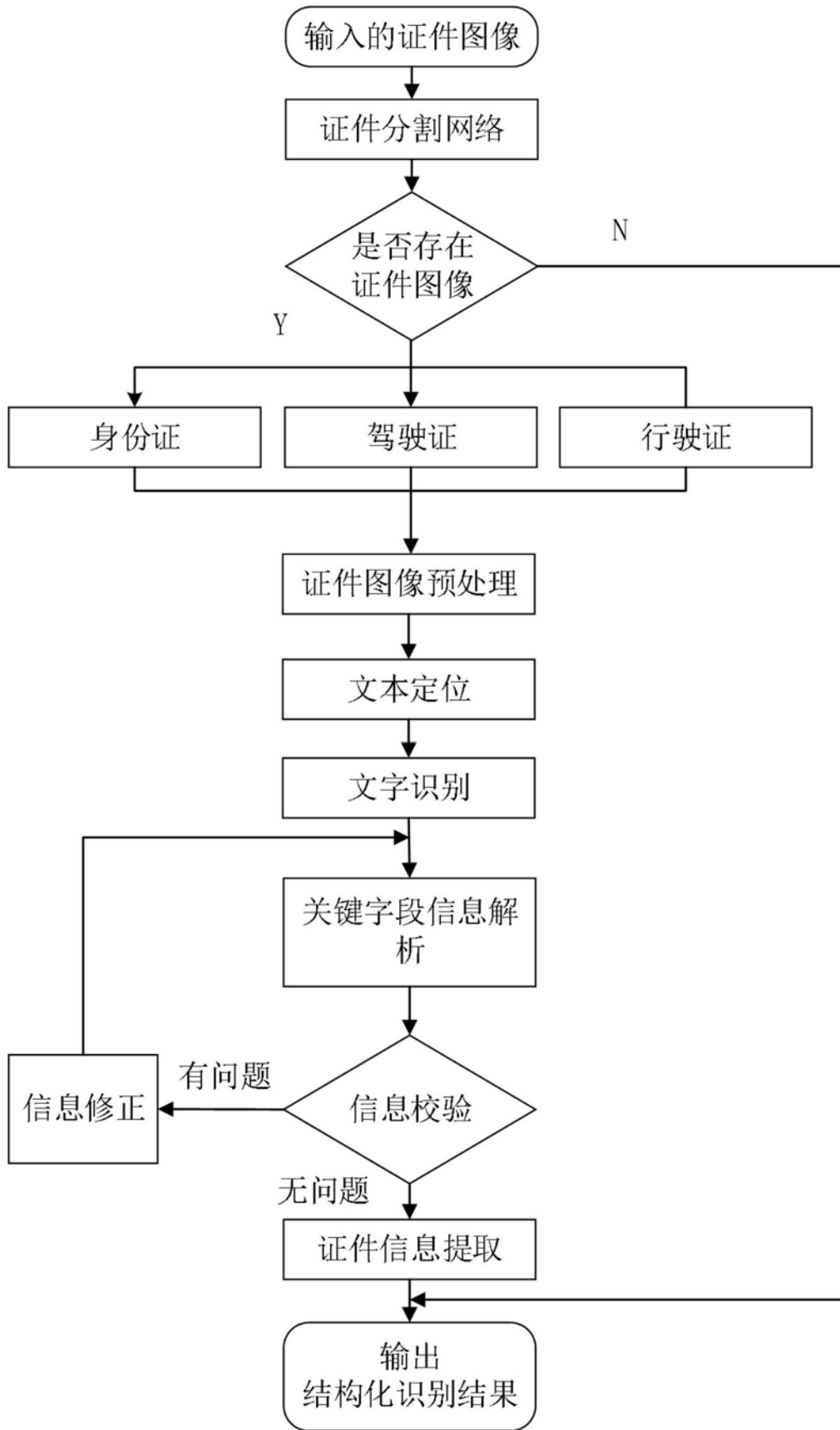


图4

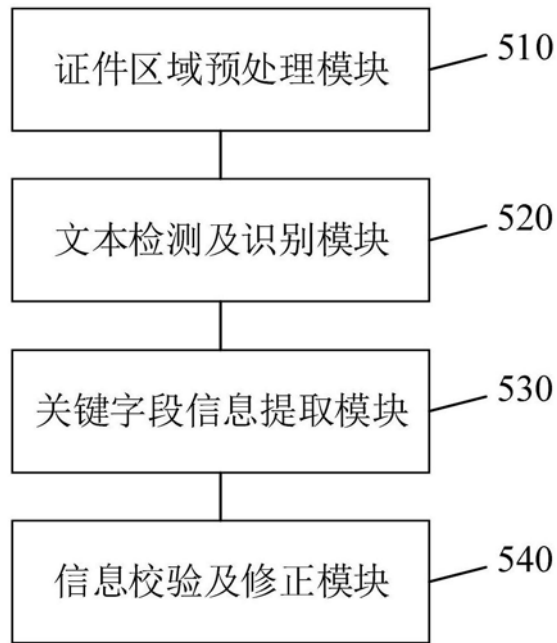


图5

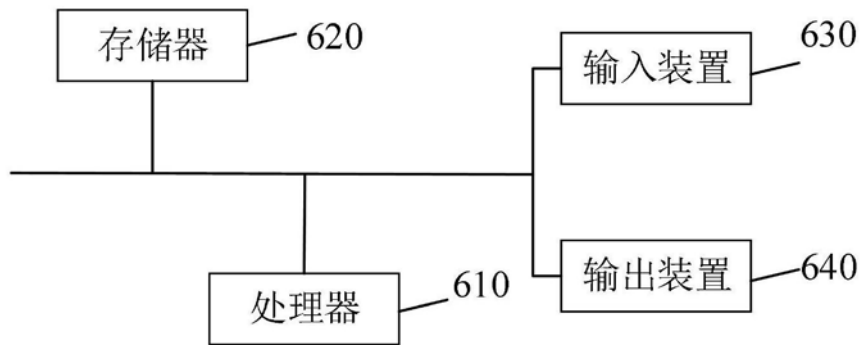


图6