

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 8 月 6 日 (06.08.2020)



(10) 国际公布号
WO 2020/155763 A1

(51) 国际专利分类号:

G06K 9/00 (2006.01)

(21) 国际申请号:

PCT/CN2019/117914

(22) 国际申请日:

2019 年 11 月 13 日 (13.11.2019)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201910078744.8 2019 年 1 月 28 日 (28.01.2019) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国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, Guangdong 518000 (CN)。

(72) 发明人: 许洋(XU, Yang); 中国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, Guangdong 518000 (CN)。

刘鹏(LIU, Peng); 中国广东省深圳市福田区福田街道福安

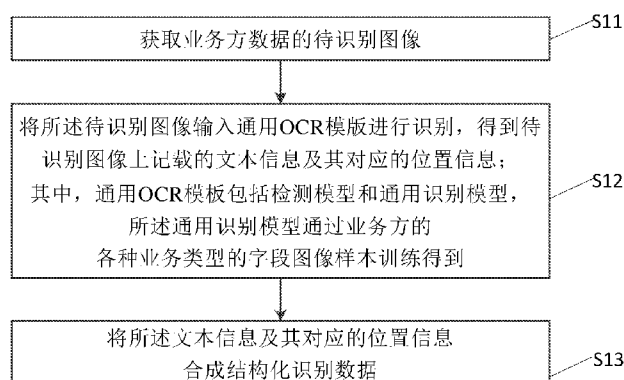
社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。王健宗(WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 广州三环专利商标代理有限公司 (SCIHEAD IP LAW FIRM); 中国广东省广州市越秀区先烈中路 80 号汇华商贸大厦 1508 室, Guangdong 510070 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: OCR RECOGNITION METHOD AND ELECTRONIC DEVICE THEREOF

(54) 发明名称: OCR 识别方法及其电子设备



S11 Obtain an image to be recognized of service party data

S12 Input the image to be recognized into a universal OCR template for recognition to obtain text information recorded in the image to be recognized and position information corresponding to the text information, wherein the universal OCR template comprises a detection model and a universal recognition model, and the universal recognition model is trained by means of field image samples of each service type of the service party

S13 Synthesize the text information and the position information corresponding thereto into structured recognition data

图 1

(57) Abstract: An OCR recognition method and OCR recognition electronic device. The method comprises the steps of: obtaining an image to be recognized of service party data (S11); inputting the image to be recognized into a universal OCR template for recognition to obtain text information recorded in the image to be recognized and position information corresponding to the text information, wherein the universal OCR template comprises a detection model and a universal recognition model (S12); and synthesizing the text information and the position information corresponding thereto into structured recognition data (S13). According to the described method, the image of an object to be recognized (such as a contract, an invoice, a bill and a certificate) can be efficiently and rapidly recognized by means of a universal OCR template, thus generating structured recognition data and completing recognition from an optical character to text information. The universal OCR template employed in the method can be trained in a short amount of time, has high adaptability, can adapt to various objects to be recognized, has high recognition accuracy, and has an overall highly efficient recognition process.

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种OCR识别方法和OCR识别电子设备。所述方法包括步骤: 获取业务方数据的待识别图像(S11); 将待识别图像输入通用OCR模版识别, 得到待识别图像记载的文本信息及其对应的位置信息; 其中, 通用OCR模版包括检测模型和通用识别模型(S12); 将文本信息及其对应位置信息合成结构化识别数据(S13)。所述方法能高效快速通过通用OCR模板对待识别对象(例如合同、发票、票据、证件等)的图像进行识别, 生成结构化识别数据, 完成光学字符到文本信息的识别, 其采用的通用OCR模板的训练时间短, 适应性强, 能适应多种不同的待识别对象, 识别准确率高, 识别过程整体效率高。

OCR 识别方法及其电子设备

本申请要求于 2019 年 01 月 28 日提交中国专利局、申请号为 201910078744.8、申请名称为“OCR 识别方法及其电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及图像识别领域，更具体地，涉及一种 OCR 识别方法及其电子设备。

背景技术

OCR (Optical Character Recognition, 光学字符识别) 主要通过对载体上显示的光学字符进行识别,生成文本输出。以纸质文件的 OCR 识别为例,通过采集纸质文件上的印刷体得到的光学字符,对其进行识别,即可得到文本信息等数据。

现有技术中的 OCR 识别方法往往依赖于识别的对象的特点,进行个性化的模版定制,例如针对票据、报纸、教材等识别的对象,甚至针对不同字号、字体的光学字符识别,都需要重新定制相应的光学字符识别模板,才能采用特定的光学字符识别模板的进行识别。

现有技术的 OCR 识别方法中,定制光学字符识别模板的训练数据量很高,训练时间长,定制识别模板的效率低,很难转移到其他识别对象中应用,定制光学字符识别模板容易受字符变化等因素的影响,OCR 识别方法应用的定制光学字符识别模板对对象的依赖性强,影响了 OCR 识别效率。

发明内容

鉴于上述问题,本申请提出了一种 OCR 识别方法及其电子设备,其能够解决训练时间长,定制识别模板的效率低,难以转移到其他识别对象中应用,定制光学字符识别模板容易受字符变化等因素的影响,定制光学字符识别模板对对象的依赖性强,影响了 OCR 识别效率的至少一种技术缺陷。

本申请提供一种 OCR 识别方法,包括:

获取业务方数据的待识别图像;

将所述待识别图像输入通用 OCR 模版进行识别,得到待识别图像上记载的文本信息及其对应的位置信息;其中,所述通用 OCR 模板包括检测模型和通用识别模型,所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到;

将所述文本信息及其对应的位置信息合成结构化识别数据。

本申请还提供一种电子设备,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为执行上述任一实施例的所述 OCR 识别方法的步骤。

本申请还提供一种非临时性计算机可读存储介质,当所述存储介质中的指令由移动终

端的处理器执行时，使得移动终端能够执行上述任一实施例的所述 OCR 识别方法。

本申请还提供了一种 OCR 识别装置，该 OCR 识别装置包括用于执行本申请的 OCR 识别方法的单元。

本申请还提供了一种计算机非易失性可读存储介质，所述计算机非易失性可读存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时使所述处理器执行本申请的 OCR 识别方法。

相对于现有技术，本申请提供的方案——OCR 识别方法及其电子设备，通过获取业务方数据的待识别图像；将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息；其中，所述通用 OCR 模板包括检测模型和通用识别模型，所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到；将所述文本信息及其对应的位置信息合成结构化识别数据的技术方案，能够高效快速地通过通用 OCR 模板对待识别对象（例如是合同、发票、票据、证件等对象）的图像进行识别，生成结构化识别数据，完成光学字符到文本信息之间的识别。本申请中采用的通用 OCR 模板的训练时间短，适应性强，能够适应多种不同的待识别对象，识别准确率高，整体的效率高。

本申请的这些方面或其他方面在以下实施例的描述中会更加简明易懂。

附图说明

为了更清楚地说明本申请实施例技术方案，下面将对实施例描述中所需要使用的附图进行说明。

图 1 示出本申请实施例中 OCR 识别方法的方法流程图；

图 2 示出本申请实施例中所述 OCR 识别方法识别的发票样本的示意图；

图 3 示出本申请中根据业务类型训练通用识别模型的方法流程示意图；

图 4 示出本申请实施例中构建通用识别模型的方法流程示意图；

图 5 示出本申请根据预标注字段子图像训练检测模型的方法流程示意图；

图 6 示出本申请根据行高信息、长度信息生成检测模型的方法流程示意图

图 7 示出本申请中根据识别准确率调整模型参数的方法流程示意图；

图 8 示出本申请验证结构化识别数据是否符合验证条件的流程示意图；

图 9 示出本申请实施例中 OCR 识别方法识别的合同样本的示意图；

图 10 示出本申请实施例提供的终端相关的部分结构的框图。

具体实施方式

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描述。

在本申请的说明书和权利要求书及上述附图中的描述的一些流程中，包含了按照特定顺序出现的多个操作，但是应该清楚了解，这些操作可以不按照其在本文中出现的顺序来执行或并行执行，操作的序号如 S1、S21 等，仅仅是用于区分各个不同的操作，序号本身不代表任何的执行顺序。另外，这些流程可以包括更多或更少的操作，并且这些操作可

以按顺序执行或并行执行。需要说明的是，本文中的“第一”、“第二”等描述，是用于区分不同的消息、设备、模块等，不代表先后顺序，也不限定“第一”和“第二”是不同的类型。

请参考图 1，图 1 示出图 1 示出本申请实施例中 OCR 识别方法的方法流程图。其中，
5 OCR 是指光学字符识别，所述 OCR 识别方法包括：

步骤 S11：获取业务方数据的待识别图像。

本申请中的 OCR 识别方法可以应用到 OCR 模版开发的过程中，用于开发通用型 OCR 识别模版。在上述过程中，业务方是指需要 OCR 模版的一方。待识别图像是指 OCR 识别方法的待识别对象通过拍照、扫描等方式获得的图像信息。通过 OCR 识别方法将待识别图
10 像上以光学字符记载的文字信息转化为文本信息输出。

步骤 S12：将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息；其中，所述通用 OCR 模版包括检测模型和通用识别模型，所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到；

在上述识别的过程中，将待识别图像输入通用 OCR 模版进行识别，通用 OCR 模版中
15 包括检测模型和通用识别模型。其中，检测模型识别对应文本信息的位置并将待识别图像对应位置截取后，转至通用识别模型进行文本识别。

步骤 S13：将所述文本信息及其对应的位置信息合成结构化识别数据。

在上述过程中，检测模型识别对应文本信息的位置时，对应匹配该文本信息的位置的结构化信息。结构化信息可以是文本信息的类别、分类、特点等等信息，在一些场景中结
20 构化信息可以是身份证号、邮编、卡号、识别码等等用于提示对应文本内容种类的信息。相应地，通用识别模型识别得到文本信息，根据前述结构化信息与文本信息相结合，生成结构化数据。

为了更好地展示本申请的技术方案，在下文将以一个具体的场景和识别对象结合解释本方案。请参考图 2，图 2 中展示了 OCR 识别方法识别的发票样本的待识别图像。应用本
25 申请的方法，对发票样本进行识别：

首先，从业务方处获取业务方数据中发票样本的待识别图像，如图 2。

其后，将发票样本的待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息。

其中的过程包括：

30 通用 OCR 模版的检测模型识别得到发票样本中的“纳税识别号”所在的区域 A，就是对应的位置信息。检测模型将“纳税识别号”所在区域 A 对应的图像（可以称为“待识别子图像”）截取并发送至通用识别模型；

通用 OCR 模版的通用识别模型识别得到发票样本中纳税识别号 A 所在区域对应的图像，通过光学字符与文字之间的映射关系的识别得到文本信息“12345”。通用 OCR 模板
35 还可以将文本信息和根据位置信息匹配得到的结构化信息输出得到结构化数据。例如在本例中，通用 OCR 模板可以根据在发票样本中的区域 A 的位置信息，匹配得到相应的“纳

税识别号”的结构化信息，将得到的文本信息“12345”和结构化信息“纳税人识别号”合成结构化识别数据“纳税人识别号：12345”输出。

上述 OCR 识别方法，相对于采用传统的 OCR 识别模型的现有技术，现有技术中需要大量的数据用于训练用于定位和文本识别的模型，而且每次更换一种识别对象都需要重新训练用于定位和文本识别的模型。此时需要的用于训练的数据量很大和训练时间很长，严重制约了 OCR 识别的效率。本申请的技术方案由于采用通用 OCR 模板可以对大多数光学字符实现光学信息到文本信息之间的转化，不需要针对每一种待识别对象都进行训练。因此在建立通用 OCR 模板时，可以继续沿用已经训练的通用识别模型，而不需要单独对通用识别模型进行额外的训练，节省了训练时间，降低了训练数据量的要求，更快速形成 OCR 识别的模板，最终提升 OCR 识别的整体的效率。

请参考图 3，为了更好实现 OCR 识别，提升识别效果，本实施例还提供一种技术方案，用于训练所述通用识别模型，所述 OCR 识别方法还包括：

步骤 S31：从业务方数据中确定业务方处理的各个业务类型。

在上述过程中，对来自业务方的业务方数据确定业务类型的分类。业务类型主要是指识别对象的类型或者与业务方业务相关的类型，例如是发票识别业务、证件识别业务、课本识别业务、包装识别业务或者说明书识别业务。

步骤 S32：根据各个业务类型分别获取对应的样本。

根据不同的业务类型获取对应业务类型的样本，每一种业务类型都提供相应的样本数量。为了更好地训练通用识别模型，可以根据业务类型的比例相应调整样本中各个业务类型对应样本的占比。

例如，当业务方的发票识别事项达到 50%，合同识别事项达到 30%，报销单识别事项达到 20%。此时，样本中的待识别图像可以来源于 50%的发票样本，30%的合同样本，20%的报销单样本。

步骤 S33：利用所述样本训练得到通用识别模型。

上述的各个业务类型的样本用于训练通用识别模型，不同业务类型的样本可以令训练得到的通用识别模型对不同业务类型的待识别对象的光学字符进行有效的识别。

为了进一步节省训练的时间，本实施例还提供一种技术方案，在训练通用识别模型之前还可以获取在其他业务类型中应用的识别模型，并采用不同业务类型的样本对该识别模型进行适应性训练，得到通用识别模型。当然，在一些场景下该识别模型本身就具有较高的识别率时，可以测试对不同业务类型的样本的识别率，当识别率达到识别阈值时，可以采用该识别模型作为通用识别模型。

请参考图 4，为了更进一步训练更好的通用识别模型，提升通用识别模型的识别准确性，本申请实施例中还提供一种 OCR 识别方法，其中利用所述样本训练得到通用识别模型的步骤 S33，包括：

步骤 S41：提取所述样本中训练图像所记载文本信息的文字特征信息。

在上述过程中，样本的训练图像记载的文本信息提取的文字特征信息。文字特征信息

指的是可以反映文本信息的载体——字体本身的特征信息。由于在同一训练图像中可能存在多段文本信息，如果这些文本信息的字体都相同，也就是具有同样的文字特征信息时，可以同时提取。当同一训练图像中存在多段文本信息且字体不相同，此时需要截取或者标注特定的文本信息。根据不同字体的特点，将文本信息的字体的特点除去，仅保留字体本身用于表示外形的特征信息，也就是文字特征信息。

请参考图 2 中，当发票样本作为训练图像用于构建通用识别模型时，区域 A 的“纳税识别号”的相关内容可以是以楷体四号字体并以光学字符显示，区域 B 的“单位名称”可以是以黑体小四号字体并以光学字符显示时，此时需要提取文本信息对应的光学字符，滤去与黑体、楷体相关的字体特征，将字号根据光学字符占据的面积大小按照比例缩小或放大至合适的比例。进一步，在同一展示比例和滤去字体特征后，得到文字以光学字符排布的特征就是文字特征信息。

同样的字体在行书、楷书、黑体等字体的展示下，光学字符都会有区别。即使是相同的文本内容以同样的字体展示，在不同的字号、光学采集条件、环境条件下都会有所区别。因此，本申请的实施例还提供一种方案，通过提取文本信息中每一个字的主体结构作为字体特征信息。提取文本信息中每一个字的主体结构时，将滤去一些对识别确认文字作用较低的光学字符，例如是笔锋、笔画收尾、笔画粗细等特征。

步骤 S42：获取文字特征信息对应的训练文本信息，分析文字特征信息与训练文本信息之间的对应关系，得到映射信息。

根据文字特征信息意义对应的文本信息，根据两者之间的关系，得到文字特征信息与训练文本信息之间的映射关系。

步骤 S43：根据所述映射信息构建通用识别模型。

根据反映字体特征信息与文本信息之间的映射关系的映射信息构建得到通用识别模型。

通过提取上述反映文字主体特征的文字特征信息构建的通用识别模型能够有效识别不同字体、不同字号的文本信息。

请参考图 5，为了得到检测字段位置效果更优的检测模型，提升检测模型的识别准确性，本申请实施例中还提供一种 OCR 识别方法。其中，将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息的步骤 S12，包括：

步骤 S51：获取预标注字段子图像位置的训练图像。

以图 2 中的发票样本作为训练图像时，预先在发票样本上标注字段子图像的位置，例如是图 2 中的“纳税识别号”区域 A 和“单位名称”区域 B。

步骤 S52：提取所述文本信息的位置特征信息，根据所述位置特征信息构建所述检测模型。

根据上述区域 A、区域 B 以及其他区域在图像上的相对距离、边距等等因素提取得到上述“纳税识别号”和“单位名称”对应内容的文本信息所在的位置特征信息。

由于不同识别对象中文本信息对应的位置特征都不尽相同，检测模型需要根据不同的待识别对象进行训练或者构建。检测模型用于识别待识别对象中文本信息所在的位置和截

取对应位置的图像。

待识别对象中的文本信息的长度往往都不一致，例如图 2 中的区域 B 的“单位名称”的文本信息的长度可以不同，此时将会影响检测模型截取待识别图像的区域或长度。请参考图 2 中的发票样本中，记载“纳税识别号”的区域 A 与记载发票“票头”的区域 C 的文本信息的长度不一致。

为此请参考图 6，本申请的实施例还提供一种技术方案，用于训练变长的识别范围的检测模型以便于识别不同面积和不同形状的显示区域，所述 OCR 识别方法中的提取所述文本信息的位置特征信息，根据所述位置特征信息构建所述检测模型步骤 S52 包括：

步骤 S61：根据文本信息的行高信息对用于训练检测模型的训练图像进行分割，得到训练子图像。

在上述过程中，行高信息可以通过输入获取的方式得到。更进一步，为了提升效率，可以通过光学字符的排布确定行间距，再根据行间距确定行高的技术方案，提取样本中的行高信息。为了更好地解释本实施例，当图 2 中的发票样本作为训练检测模型的样本时，本申请中的技术方案可以是对各个文本信息展示区域的光学字符进行区域化识别。例如是对“纳税识别号”的区域 A 中的光学字符通过提取边缘线，将光学字符边缘线向外拓展设定的边距，就可以得到相应区域 A 中文本信息的行高信息。

根据上述方法得到的行高信息对训练图像（也就是图 2 中的发票样本）进行分割将若干展示区域分割为多个训练子图像，以区域 A 为例，区域 A 中有光学字符的图像分割为若干小的片段，即生成区域 A 的训练子图像。

步骤 S62：将所述训练子图像输入全连接网络模型，通过识别字符数据库中的字符，计算得到在训练子图像的置信度。

得到区域 A 对应的训练子图像后，将上述区域 A 的多个训练子图像输入全连接网络模型输出一维的向量。根据这些一维的向量，在通过识别字符数据库中的字符的匹配，计算对应训练子图像的置信度。置信度表明训练子图像中的光学字符与识别字符数据库中的字符之间的匹配可能性，一定概率范围内可以认为对应训练子图像存在识别字符数据库中的相应字符。其中，全连接网络模型可以通过神经卷积网络算法构建。

步骤 S63：根据训练子图像的置信度生成文本信息的长度信息。

在上述过程中，根据多个训练子图像的置信度可以确定文本信息的长度信息。换言之，在行高信息确定的情况下，根据多个训练子图像的置信度所指示多个训练子图像中存在具有识别字符数据库中的字符特性的训练子图像，得到区域 A 的长度信息。简单来说，通过识别字符数据库中的字符匹配得到区域 A 中多个训练子图像是否具有可识别字符的信息。步骤 S63 通过训练子图像置信度得到的长度信息，可以将图 2 发票样本中“货物信息”区域 E 和“税额”区域 F 及两者之间存在可识别光学字符的“货物数量”和“货物价格”的区域区分开。

步骤 S64：将所述文本信息的行高信息和长度信息生成文本信息的位置特征信息。

综合文本信息的行高信息和长度信息就能够确定文本信息的位置特征信息，文本信息

的位置特征信息能够指示区域 A 位置和范围。

步骤 S65：根据所述位置特征信息构建所述检测模型。

根据发票样本和需要提取文字信息的区域 A 的位置特征信息以及两者之间的映射关系，构建检测模型。当然，在上述过程中还需要通过与发票样本同类的训练图像进一步训练，直到检测模型的检测准确率达到预设的要求。

在一些情况下，例如是“脚注”的区域 D，由于区域 D 比较靠近样本的边缘区域，此时可以只对区域 D 上方进行光学字符边缘线的识别，可以降低运算量，快捷地确定相应区域。

在一些情况下，例如是区域 D 的文本信息比较紧凑，文本信息内部没有较大的空隙，相应的光学字符也会比较紧凑，此时也可以通过区域 D 两端的光学字符边缘线确定区域 D 的范围，此时可以直接得到区域 D 的位置特征信息。但是，如果出现图 2 的发票样本中的记载“货物信息”区域 E 至记载“税额”的区域 F 之间比较紧凑的情形时，直接提取光学字符边缘线可能会将区域 E 和区域 F 以及两者之间的展示区域都并为一个区域。因此，本实施例中通过上述步骤 S61 至步骤 S65 的方案可以克服这个问题。

为了实现更好的识别准确率和识别精度，本实施例中提供一种准确性评估和对应调整模型参数的技术方案。请参考图 7，OCR 识别方法中，所述将所述文本信息及其对应的位置信息合成结构化识别数据的步骤 S13 之后，还包括：

步骤 S71：对所述结构化识别数据进行准确性评估，得到识别准确率。

根据上述情形，计算结构化识别数据的识别准确率。识别准确率可以根据多张待识别图像输出的结构化识别数据进行评估。

步骤 S72：根据识别准确率调整通用 OCR 模版的模型参数，生成调整后的通用 OCR 模版。

根据上述识别准确率调整通用 OCR 模版的模型参数。其中可以根据识别准确率调整通用识别模型的相关识别参数，此时可以采用神经卷积网络算法，在识别的过程中进一步优化通用 OCR 模版中的通用识别模型。根据参数优化后的通用识别模型与检测模型生成新的通用 OCR 模版。后续的 OCR 识别采用新的通用 OCR 模版。

为了验证所述 OCR 识别方法结构化识别数据的识别准确率和识别精度，本实施例中提供一种验证的技术方案。请参考图 8，OCR 识别方法中，所述将所述文本信息及其对应的位置信息合成结构化识别数据的步骤 S13 之后，还包括：

步骤 S81：验证所述结构化识别数据是否符合验证条件。

步骤 S82：若否，将所述结构化识别数据中不符合验证条件的文本信息对应的待识别图像，输入调整后的通用 OCR 模版中进行重新识别。

步骤 S83：若是，输出所述结构化识别数据。

通过验证公式或者联合验证的方式，验证所述结构化识别数据是否符合验证的条件。在得到结构化识别数据后，以结构化识别数据“纳税识别号：12345”为例。可以通过人工识别或者通过对应结构化识别数据的校验公式进行校验。

再举一个例子,当识别得到的结构化识别数据是“身份证号:4401*11999*****2459”(为了规避隐私权风险,某些位置的数据使用“*”遮蔽,而此处在实际识别场景中存在相应数字)时,可以通过身份证号的校验公式得到上述结构化识别数据是否识别准确。校验的内容包括结构化识别数据的数字位数、结构、末尾校验码等等。

5 此外,结构化识别数据还可以相应结合其他类型的结构化识别数据进行联合校验,例如结构化识别数据中“身份证号:4401*11999*****2459”与“户籍所在地:广东省广州市天河区...”,此时可以判断结构化识别数据中身份证号的前四位识别是准确的,可以执行步骤 S83。

10 如果结构化识别数据不符合验证条件那么执行步骤 S82,将待识别图像输入上述经过调整的通用 OCR 模版中进行二次识别。更进一步,还可以通过继续检验的方式,直到识别准确率达到要求,否则继续调整相应的模型参数。

当通过通用 OCR 模版得到多段所述文本信息时,为了得到结构化和可读性更强的结构化识别数据,本实施例提供的 OCR 识别方法中,将所述待识别图像输入通用 OCR 模版进行识别,得到待识别图像上记载的文本信息及其对应的位置信息的步骤 S12 之后,还包括:
15 根据检测模型识别得到多段文本信息在待识别图像上的相对位置,将多段文本信息依序拼合。

在上述过程中,当待识别图像通过通用 OCR 模版识别得到多段文本信息时,根据文本信息出现的具体位置将多段文本信息依序拼合。以图 2 中的发票样本作为示范性举例,当“货物信息”的区域 E 中存在多段货物信息时,例如是某用户购买了水果时开具的发票,
20 购买了苹果、香蕉和雪梨。此时用户开具的纸质发票就是待识别对象,对用户购买水果开具的发票图像输入通用 OCR 模板,在区域 E 得到三段对应的文本信息——“苹果、香蕉、雪梨”。此时根据上述三段对应文本信息的相对位置拼合,在本例中分三行将文本信息根据相对位置的结构化信息以结构化的形式展示。例如,根据识别得到的文本信息和对应结构化信息,拼合成发票样式进行展示。除此之外,根据图 2 中识别的若干项目也可以作为
25 多段文字信息,通用 OCR 模板根据其相对位置在结构化的发票模板中进行拼接和展示。本实施例还提供一种方案,当通用 OCR 模板识别到无法相应以文字展示的信息时,例如是图 2 的发票样本中的“签章”区域 H 中记载某单位的公章信息,上述的通用 OCR 模板可以采集区域 H 的图像信息,并将其矢量化,生成签章矢量图。在拼接和展示时,将签章矢量图拼接到上述发票模板中。在另一种常用的场景中,当通用 OCR 模板识别到无法相应以文字
30 展示的信息时,例如是图 2 的发票样本中的“密码区域”区域 G 中记载的密码信息,该密码以二维码形式展示。上述的通用 OCR 模板可以采集区域 G 的图像信息,并通过识别二维码的方式得到相应的密码信息。在拼接和展示时,将密码信息以明文或者反向生成条码的方式,将密码信息拼接到上述发票模板中。

35 为了调整生成根据多段文本信息在待识别图像上的相对位置依照顺序拼合时的位置偏差,本实施例在上述方案的基础上还提供一种 OCR 识别方法,其中在根据检测模型识别得到多段文本信息在待识别图像上的相对位置,将多段文本信息依序拼合的步骤之后,还包

括：根据多段文本信息在图像信息上的相对位置，调整通用 OCR 模板的检测模型中的定位间距参数。

仍以上述例子作为解释的基础，根据多段文本信息“苹果、香蕉、雪梨”在图像信息的相对位置，重新确定多段文本信息之间的行距、同行文字之间的间距信息。根据间距信息调整通用 OCR 模板中检测模型的定位间隔参数。其中，定位间隔参数可以用于检测模型在识别待识别图像时，定位文本信息中每一个字之间的间隔以及多段文本信息之间的行距，便于检测模型截取相应的图像信息。在上述过程中，间距信息可以通过识别的文本信息在显示字号、相同行距、单字间隔相同的情况下，比较识别的相应字段的显示区域，进而得到间隔信息。

本实施例还相应提供一种电子设备，包括：
处理器；

用于存储处理器可执行指令的存储器；

其中，所述处理器被配置为执行上述任一项实施例所述 OCR 识别方法的步骤。

本实施例除了上述用于示例性识别图 2 中发票样本的用途，还可以识别合同、票据、证件等待识别对象。为了更好解释本申请的技术方案，本实施例现结合图 9 的合同样本的待识别图像进行进一步解释。

请结合图 1 和图 9，上述的 OCR 识别方法包括：

步骤 S11：获取业务方数据的待识别图像。

获取业务方数据中合同样本的待识别图像，待识别图像的获取方式可以通过扫描或者拍照等方式获得待识别图像。

步骤 S12：将所述待识别图像输入通用 OCR 模板进行识别，得到待识别图像上记载的文本信息及其对应的位置信息；其中，所述通用 OCR 模板包括检测模型和通用识别模型，所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到。

在上述过程中，将合同样本的待识别图像输入通用 OCR 模板进行识别，得到合同样本的待识别图像上记载的文本信息及其对应的位置信息。其中，识别合同样本的通用 OCR 模板中的检测模型需要利用业务方提供的业务方数据进行训练，业务方数据中包括与合同样本同类型的训练图像作为训练的对象。利用与合同样本同类型的训练图像作为训练的对象，训练得到的检测模型能够根据合同样本中文本信息的位置截取相应的待识别子图像，以供通用识别模型进行光学字符到文本信息的识别。本申请中，针对不同的待识别对象，只需要重新训练检测模型，而通用识别模型可以不需要重复训练。例如，在识别合同样本前，只需要相应训练检测模型，而通用识别模型可以采用识别发票样本时通用 OCR 模板中的通用识别模型。

在上述过程中，通用 OCR 模板通过识别合同样本中的“合同名称”区域 I、“当事人信息”区域 J、“合同正文”区域 K、“签章信息”区域 L、“落款与日期”区域 M 等区域的待识别子图像，得到相应的文本信息。其中，包括两个过程：首先，通用 OCR 模板中的检测模型将通过预先训练的映射关系，检测得到“合同名称”区域 I、“当事人信息”区

域 J、“合同正文”区域 K、“签章信息”区域 L、“落款与日期”区域 M 的相对位置的位置信息,并截取相应区域的待识别子图像;其后,通用 OCR 模板中的通用识别模型识别上述“合同名称”区域 I、“当事人信息”区域 J、“合同正文”区域 K、“签章信息”区域 L、“落款与日期”区域 M 的待识别子图像中的光学字符,并根据光学字符和文字之间的映射关系,最终识别得到待识别子图像对应的文本信息。通用 OCR 模板中还可以根据每个区域的相对位置匹配结构化信息,这些结构化信息可以是“合同名称”、“当事人信息”、“合同正文”、“签章信息”、“落款与日期”等与识别区域相对应的信息。

步骤 S13:将所述文本信息及其对应的位置信息合成结构化识别数据。

根据上述通用 OCR 模板识别得到的文本信息及其对应的位置信息生成结构化识别数据。其中,检测模型得到待识别的合同样本中的各个展示区域的相对位置生成与图 9 中合同样本对应的合同模板。此时,将文本信息根据位置信息中对应的位置写入合同模板中,生成结构化识别数据。

此外,还可以通过通用 OCR 模板匹配得到的结构化信息,结合识别得到的文本信息及其对应位置信息生成结构化识别数据。

本实施例还提供一种电子设备,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为执行上述任一实施例的所述 OCR 识别方法的步骤。

本申请的实施例提供的电子设备,如图 10 所示,为了便于说明,仅示出了与本申请实施例相关的部分,具体技术细节未揭示的,请参照本申请实施例方法部分。该终端可以为包括手机、平板电脑、PDA(Personal Digital Assistant,个人数字助理)、POS(Point of Sales,销售终端)、车载电脑等任意终端设备,以终端为手机为例:

图 10 示出的是与本申请实施例提供的终端相关的手机的部分结构的框图。参考图 10,手机包括:射频(Radio Frequency, RF)电路 1010、存储器 1020、输入单元 1030、显示单元 1040、传感器 1010、音频电路 1060、无线保真(wireless fidelity, WiFi)模块 1070、处理器 1080、以及电源 1090 等部件。本领域技术人员可以理解,图 10 中示出的手机结构并不构成对手机的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

下面结合图 10 对手机的各个构成部件进行具体的介绍:

RF 电路 1010 可用于收发信息或通话过程中,信号的接收和发送,特别地,将基站的下行信息接收后,给处理器 1080 处理;另外,将设计上行的数据发送给基站。通常,RF 电路 1010 包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器(Low Noise Amplifier, LNA)、双工器等。此外,RF 电路 1010 还可以通过无线通信与网络和其他设备通信。上述无线通信可以使用任一通信标准或协议,包括但不限于全球移动通讯系统(Global System of Mobile communication, GSM)、通用分组无线服务(General Packet Radio Service, GPRS)、码分多址(Code Division Multiple Access, CDMA)、宽带码分

多址(Wideband Code Division Multiple Access, WCDMA)、长期演进(Long Term Evolution, LTE)、电子邮件、短消息服务(Short Messaging Service, SMS)等。

存储器 1020 可用于存储软件程序以及模块,处理器 1080 通过运行存储在存储器 1020 的软件程序以及模块,从而执行手机的各种功能应用以及数据处理。存储器 1020 可主要包
5 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手机的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器 1020 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、或其他易失性固态存储器件。

10 输入单元 1030 可用于接收输入的数字或字符信息,以及产生与手机的用户设置以及功能控制有关的键信号输入。具体地,输入单元 1030 可包括触控面板 1031 以及其他输入设备 1032。触控面板 1031,也称为触摸屏,可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在触控面板 1031 上或在触控面板 1031 附近的操
15 作),并根据预先设定的程式驱动相应的连接装置。可选的,触控面板 1031 可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测用户的触摸方位,并检测触摸操作带来的信号,将信号传送给触摸控制器;触摸控制器从触摸检测装置上接收触摸信息,并将它转换成触点坐标,再送给处理器 1080,并能接收处理器 1080 发来的命令并加以执行。此外,可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触控面板 1031。除了触控面板 1031,输入单元 1030 还可以包括其他输入设备 1032。具体地,其他输入设备
20 1032 可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆等中的一种或多种。

显示单元 1040 可用于显示由用户输入的信息或提供给用户的信息以及手机的各种菜单。显示单元 1040 可包括显示面板 1041,可选的,可以采用液晶显示器(Liquid Crystal Display, LCD)、有机发光二极管(Organic Light-Emitting Diode, OLED)等形式来配置显
25 示面板 1041。进一步的,触控面板 1031 可覆盖显示面板 1041,当触控面板 1031 检测到在其上或附近的触摸操作后,传送给处理器 1080 以确定触摸事件的类型,随后处理器 1080 根据触摸事件的类型在显示面板 1041 上提供相应的视觉输出。虽然在图 10 中,触控面板 1031 与显示面板 1041 是作为两个独立的部件来实现手机的输入和输入功能,但是在某些实施例
30 中,可以将触控面板 1031 与显示面板 1041 集成而实现手机的输入和输出功能。

手机还可包括至少一种传感器 1050,比如光传感器、运动传感器以及其他传感器。具体地,光传感器可包括环境光传感器及接近传感器,其中,环境光传感器可根据环境光线的明暗来调节显示面板 1041 的亮度,接近传感器可在手机移动到耳边时,关闭显示面板 1041 和/或背光。作为运动传感器的一种,加速计传感器可检测各个方向上(一般为三轴)加速度的大小,静止时可检测出重力的大小及方向,可用于识别手机姿态的应用(比如横
35 竖屏切换、相关游戏、磁力计姿态校准)、振动识别相关功能(比如计步器、敲击)等;至于手机还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器,在此

不再赘述。

音频电路 1060、扬声器 1061，传声器 1062 可提供用户与手机之间的音频接口。音频电路 1060 可将接收到的音频数据转换后的电信号，传输到扬声器 1061，由扬声器 1061 转换为声音信号输出；另一方面，传声器 1062 将收集的声音信号转换为电信号，由音频电路 1060 接收后转换为音频数据，再将音频数据输出处理器 1080 处理后，经 RF 电路 1010 以发送给比如另一手机，或者将音频数据输出至存储器 1020 以便进一步处理。

WiFi 属于短距离无线传输技术，手机通过 WiFi 模块 1070 可以帮助用户收发电子邮件、浏览网页和访问流式媒体等，它为用户提供了无线的宽带互联网访问。虽然图 10 示出了 WiFi 模块 1070，但是可以理解的是，其并不属于手机的必须构成，完全可以根据需要在不改变发明的本质的范围内而省略。

处理器 1080 是手机的控制中心，利用各种接口和线路连接整个手机的各个部分，通过运行或执行存储在存储器 1020 内的软件程序和/或模块，以及调用存储在存储器 1020 内的数据，执行手机的各种功能和处理数据，从而对手机进行整体监控。可选的，处理器 1080 可包括一个或多个处理单元；优选的，处理器 1080 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器 1080 中。

手机还包括给各个部件供电的电源 1090（比如电池），优选的，电源可以通过电源管理系统与处理器 1080 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。

尽管未示出，手机还可以包括摄像头、蓝牙模块等，在此不再赘述。

在本申请实施例中，该终端所包括的处理器 1080 还具有以下功能：

获取业务方数据的待识别图像；

将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息；其中，所述通用 OCR 模板包括检测模型和通用识别模型，所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到；

将所述文本信息及其对应的位置信息合成结构化识别数据。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请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（电子设备）实施例仅仅是示意性的，例如，所述单元的划分，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，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元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单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元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，也可以采用软件功能单元的形式实现。

5 本实施例还提供一种非临时性计算机可读存储介质，当所述存储介质中的指令由移动终端的处理器执行时，使得移动终端能够执行上述任一实施例的所述 OCR 识别方法。

10 在本申请的 OCR 识别方法及其电子设备，通过获取业务方数据的待识别图像；将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息；其中，所述通用 OCR 模板包括检测模型和通用识别模型，所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到；将所述文本信息及其对应的位置信息合成结构化识别数据的技术方案，能够高效快速地通过通用 OCR 模板对待识别对象（例如是合同、发票、票据、证件等对象）的图像进行识别，生成结构化识别数据，完成光学字符到文本信息之间的识别。本申请中采用的通用 OCR 模板的训练时间短，适应性强，能够适应多种不同的待识别对象，识别准确率高，整体的效率高。

15 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：只读存储器（ROM，Read Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁盘或光盘等。

20 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分步骤是可以通程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

 以上对本申请所提供的一种电子设备进行了详细介绍，对于本领域的一般技术人员，依据本申请实施例的思想，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申请的限制。

权利要求

1.一种 OCR 识别方法，OCR 是指光学字符识别，其特征在于，包括：
获取业务方数据的待识别图像；

5 将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息；其中，所述通用 OCR 模板包括检测模型和通用识别模型，所述检测模型用于识别对应文本信息的位置并将所述待识别图像对应位置截取后，转至所述通用识别模型进行文本识别，所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到；

将所述文本信息及其对应的位置信息合成结构化识别数据。

10 2.根据权利要求 1 所述的 OCR 识别方法，其特征在于，还包括：
从业务方数据中确定业务方处理的各个业务类型；
根据各个业务类型分别获取对应的样本；
利用所述样本训练得到通用识别模型。

15 3.根据权利要求 2 所述的 OCR 识别方法，其特征在于，所述利用所述样本训练得到通用识别模型的步骤，包括：

提取所述样本中训练图像所记载文本信息的文字特征信息；

获取文字特征信息对应的训练文本信息，分析文字特征信息与训练文本信息之间的对应关系，得到映射信息；

根据所述映射信息构建通用识别模型。

20 4.根据权利要求 1 所述的 OCR 识别方法，其特征在于，所述将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息的步骤之前，还包括：

获取预标注字段子图像位置的训练图像；

提取所述文本信息的位置特征信息，根据所述位置特征信息构建所述检测模型。

25 5.根据权利要求 4 所述的 OCR 识别方法，其特征在于，所述提取所述文本信息的位置特征信息，根据所述位置特征信息构建所述检测模型的步骤，包括：

根据文本信息的行高信息对用于训练检测模型的训练图像进行分割，得到训练子图像；

将所述训练子图像输入全连接网络模型，通过识别字符数据库中的字符，计算得到在训练子图像的置信度；

30 根据训练子图像的置信度生成文本信息的长度信息；

将所述文本信息的行高信息和长度信息生成文本信息的位置特征信息；

根据所述位置特征信息构建所述检测模型。

6.根据权利要求 1 所述的 OCR 识别方法，其特征在于，所述将所述文本信息及其对应的位置信息合成结构化识别数据的步骤之后，还包括：

35 对所述结构化识别数据进行准确性评估，得到识别准确率；

根据识别准确率调整通用 OCR 模版的模型参数，生成调整后的通用 OCR 模版。

7.根据权利要求 6 所述的 OCR 识别方法,其特征在于,所述将所述文本信息及其对应的位置信息合成结构化识别数据的步骤之后,还包括:

验证所述结构化识别数据是否符合验证条件;

若是,输出所述结构化识别数据;

5 若否,将所述结构化识别数据中不符合验证条件的文本信息对应的待识别图像,输入调整后的通用 OCR 模版中进行重新识别。

8.根据权利要求 1 所述的 OCR 识别方法,其特征在于,将所述待识别图像输入通用 OCR 模版进行识别,得到待识别图像上记载的文本信息及其对应的位置信息的步骤之后,当通过通用 OCR 模版得到多段所述文本信息时,还包括:根据检测模型识别得到多段文本
10 信息在待识别图像上的相对位置,将多段文本信息依序拼合。

9.根据权利要求 8 所述的 OCR 识别方法,其特征在于,根据检测模型识别得到多段文本信息在待识别图像上的相对位置,将多段文本信息依序拼合的步骤之后,还包括:根据多段文本信息在图像信息上的相对位置,调整通用 OCR 模版的检测模型中的定位间距参数。

10.本申请还提供一种电子设备,包括:

15 处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为执行以下步骤:

获取业务方数据的待识别图像;

将所述待识别图像输入通用 OCR 模版进行识别,得到待识别图像上记载的文本信息及其对应的位置信息;其中,所述通用 OCR 模板包括检测模型和通用识别模型,所述检测模型用于识别对应文本信息的位置并将所述待识别图像对应位置截取后,转至所述通用识别模型进行文本识别,所述通用识别模型通过业务方的各种业务类型的字段图像样本训练得到;
20 将所述文本信息及其对应的位置信息合成结构化识别数据。

11.根据权利要求 10 所述的电子设备,其特征在于,所述处理器还用于执行以下步骤:

25 从业务方数据中确定业务方处理的各个业务类型;

根据各个业务类型分别获取对应的样本;

利用所述样本训练得到通用识别模型。

12.根据权利要求 11 所述的电子设备,其特征在于,所述处理器在执行所述利用所述样
30 本训练得到通用识别模型时,具体执行以下步骤:

提取所述样本中训练图像所记载文本信息的文字特征信息;

获取文字特征信息对应的训练文本信息,分析文字特征信息与训练文本信息之间的对
应关系,得到映射信息;

根据所述映射信息构建通用识别模型。

13.根据权利要求 10 所述的电子设备,其特征在于,所述处理器在执行所述将所述待
35 识别图像输入通用 OCR 模版进行识别,得到待识别图像上记载的文本信息及其对应的位置

信息的步骤之前，还执行以下步骤：

获取预标注字段子图像位置的训练图像；

提取所述文本信息的位置特征信息，根据所述位置特征信息构建所述检测模型。

5 14.根据权利要求 13 所述的电子设备，其特征在于，所述处理器在执行所述提取所述文本信息的位置特征信息，根据所述位置特征信息构建所述检测模型的步骤时，具体执行以下步骤：

根据文本信息的行高信息对用于训练检测模型的训练图像进行分割，得到训练子图像；

将所述训练子图像输入全连接网络模型，通过识别字符数据库中的字符，计算得到在训练子图像的置信度；

10 根据训练子图像的置信度生成文本信息的长度信息；

将所述文本信息的行高信息和长度信息生成文本信息的位置特征信息；

根据所述位置特征信息构建所述检测模型。

15.根据权利要求 10 所述的电子设备，其特征在于，所述处理器在执行所述将所述文本信息及其对应的位置信息合成结构化识别数据的步骤之后，还执行以下步骤：

15 对所述结构化识别数据进行准确性评估，得到识别准确率；

根据识别准确率调整通用 OCR 模版的模型参数，生成调整后的通用 OCR 模版。

16.根据权利要求 15 所述的电子设备，其特征在于，所述处理器在执行所述将所述文本信息及其对应的位置信息合成结构化识别数据的步骤之后，还执行以下步骤：

验证所述结构化识别数据是否符合验证条件；

20 若是，输出所述结构化识别数据；

若否，将所述结构化识别数据中不符合验证条件的文本信息对应的待识别图像，输入调整后的通用 OCR 模版中进行重新识别。

17.根据权利要求 10 所述的电子设备，其特征在于，所述处理器在执行将所述待识别图像输入通用 OCR 模版进行识别，得到待识别图像上记载的文本信息及其对应的位置信息的步骤之后，当通过通用 OCR 模版得到多段所述文本信息时，还执行以下步骤：根据检测模型识别得到多段文本信息在待识别图像上的相对位置，将多段文本信息依序拼合。

18.根据权利要求 17 所述的电子设备，其特征在于，所述处理器在执行根据检测模型识别得到多段文本信息在待识别图像上的相对位置，将多段文本信息依序拼合之后，还执行以下步骤：根据多段文本信息在图像信息上的相对位置，调整通用 OCR 模版的检测模型中的定位间距参数。

19、一种光学字符识别 OCR 识别装置，其特征在于，包括用于执行如权利要求 1-9 任一项权利要求所述的方法的单元。

20、一种计算机非易失性可读存储介质，其特征在于，所述计算机非易失性可读存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时使所述处理器执行如权利要求 1-9 任一项所述的方法。

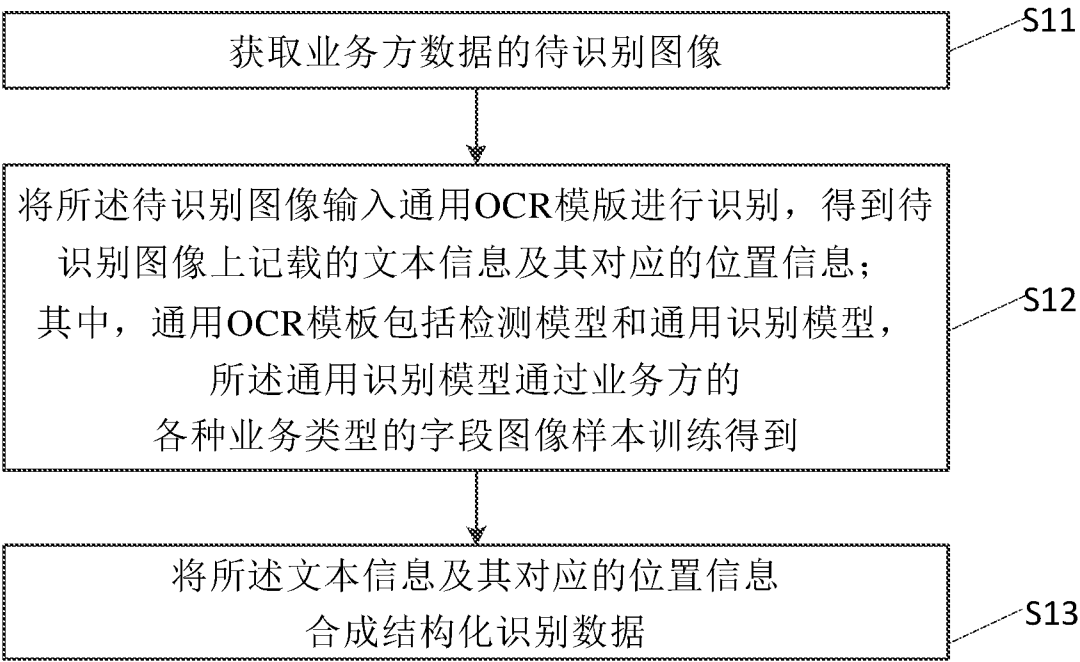


图 1

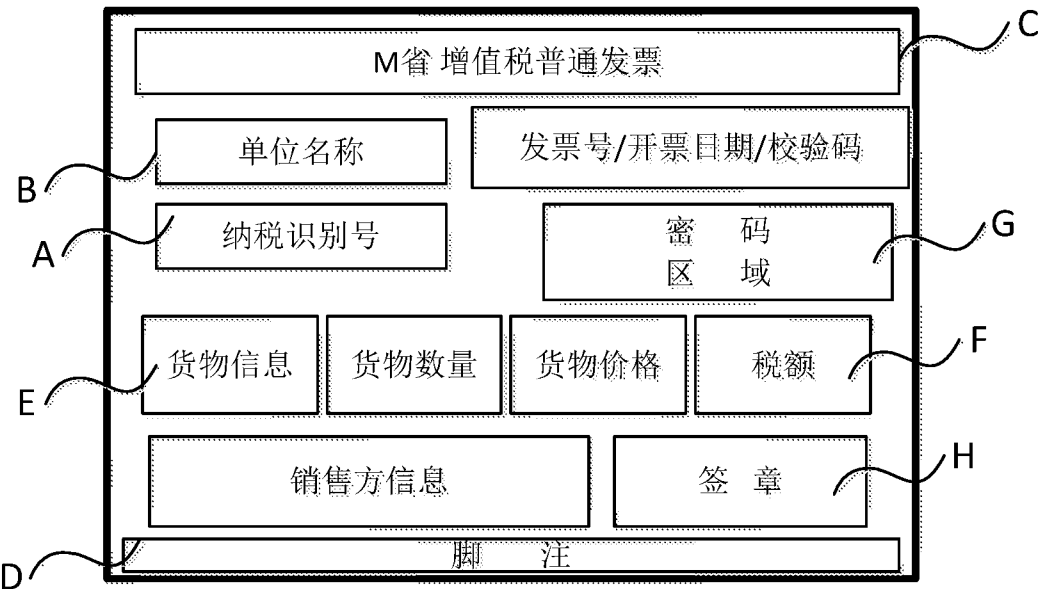


图 2

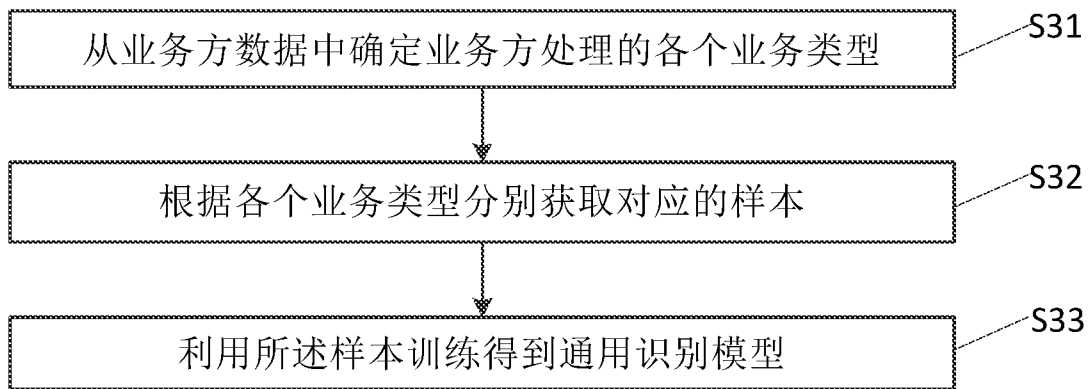


图 3

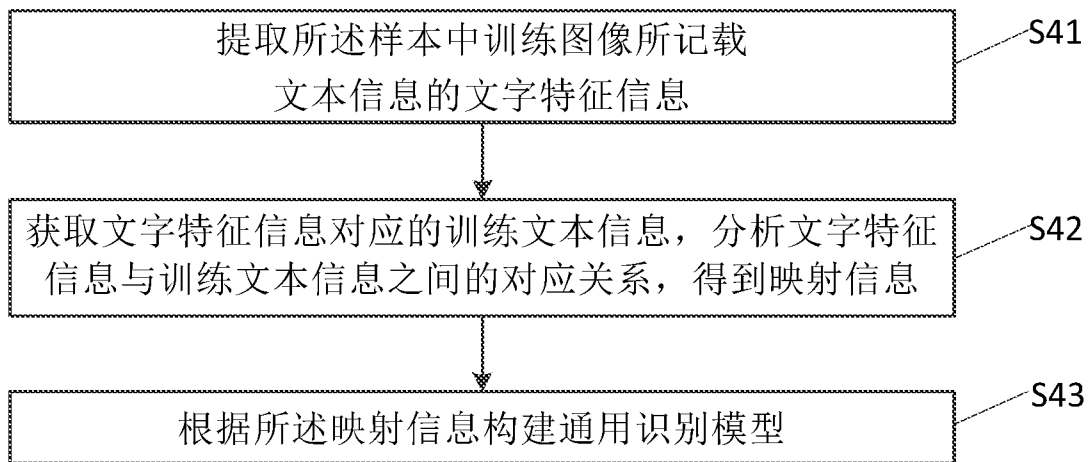


图 4

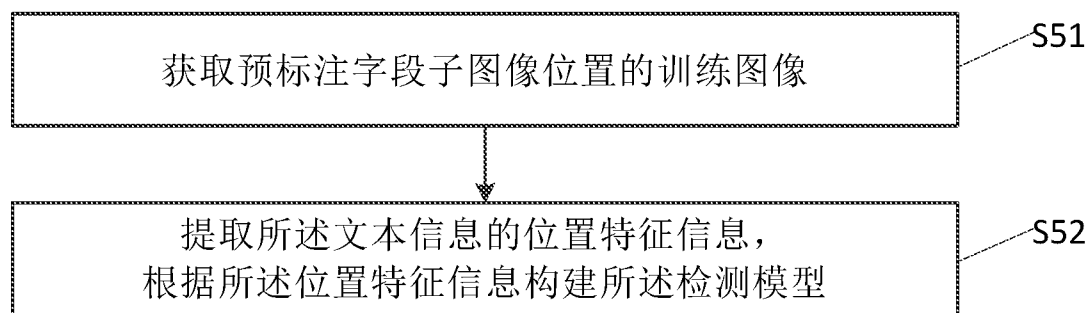


图 5

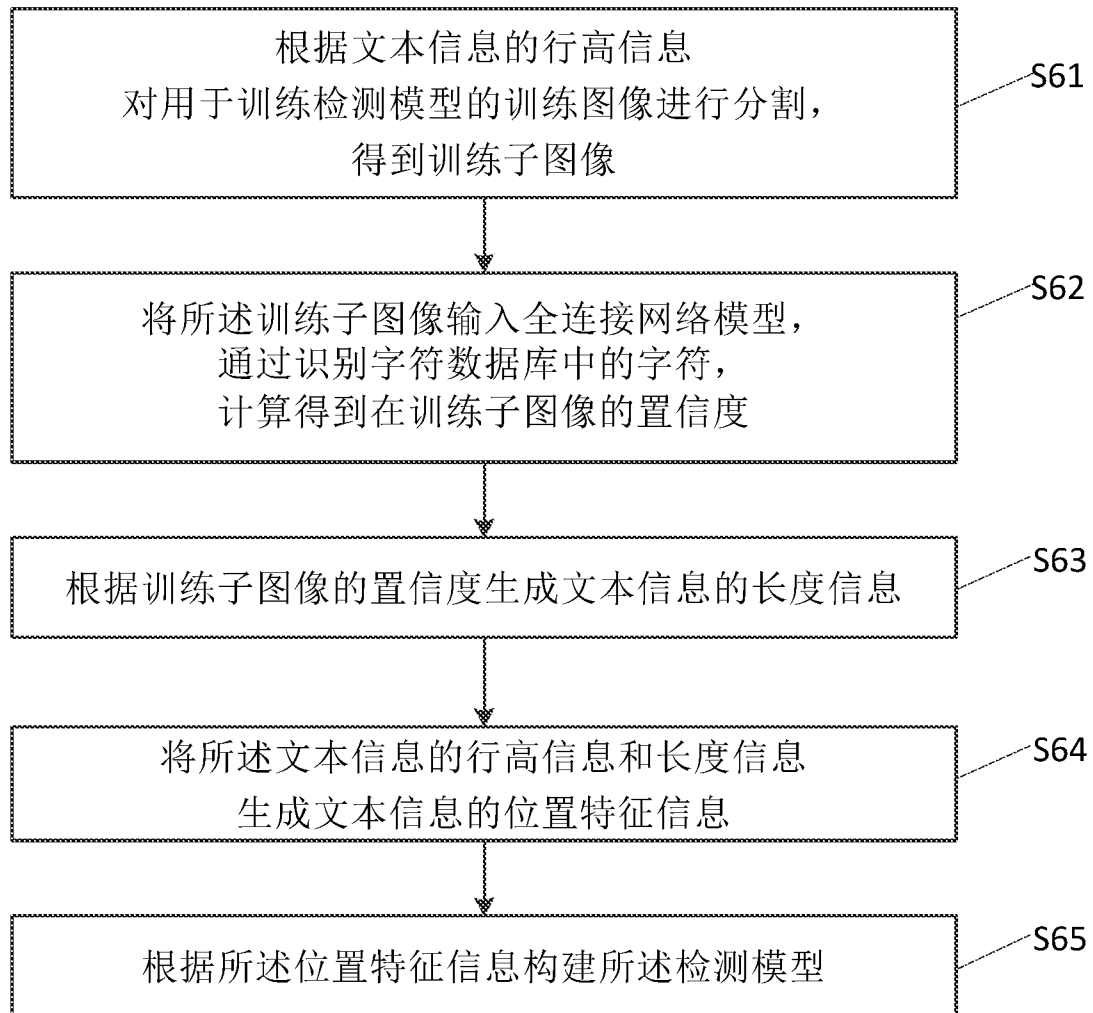


图 6

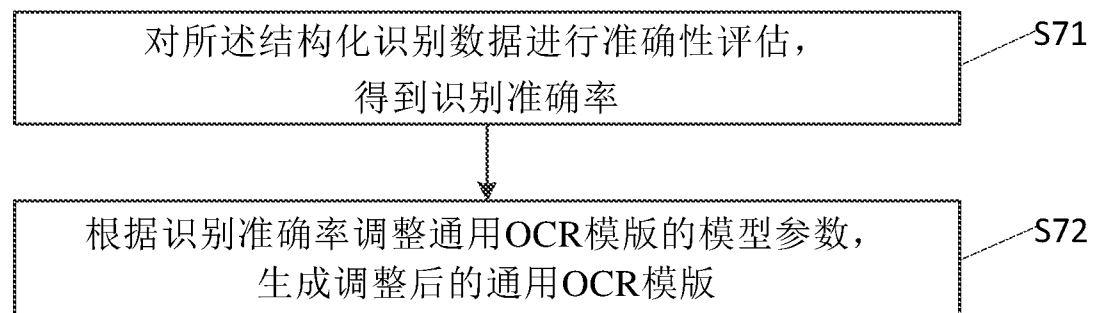


图 7

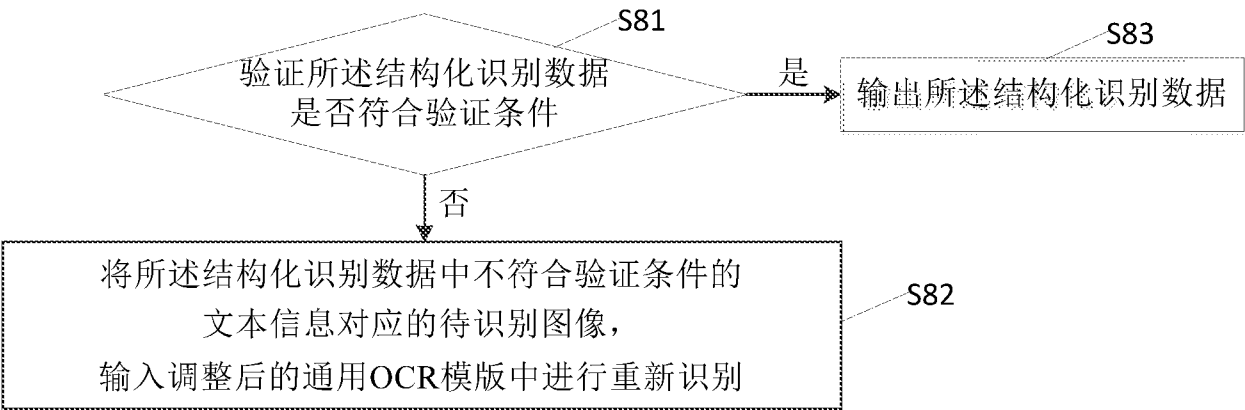


图 8

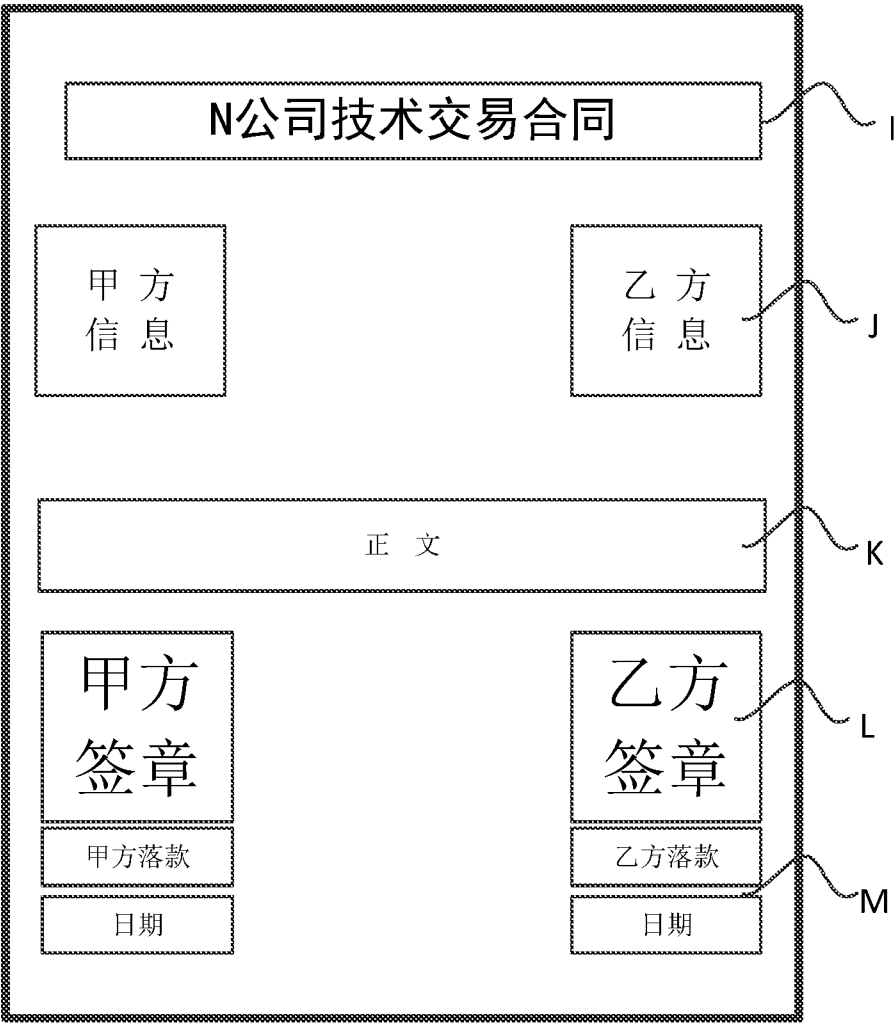


图 9

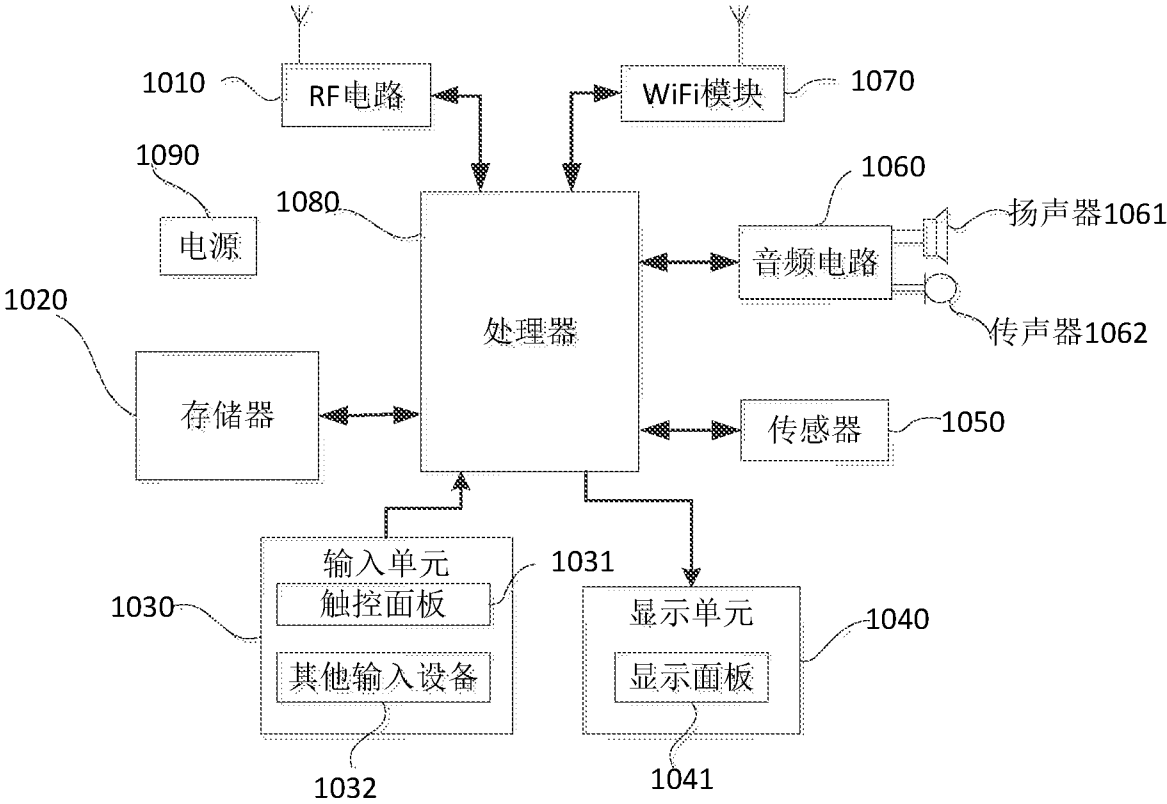


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/117914

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC, IEEE: 区域, 位置, 文字识别, 字符识别, 文本识别, 图像, 模型, 样本, character, recogni+, OCR, location, region, position, image, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109919014 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 21 June 2019 (2019-06-21) claims 1-10, and description, paragraphs [0167]-[0170]	1-20
X	CN 109034159 A (BEIJING SINOVOICE TECHNOLOGY CO., LTD.) 18 December 2018 (2018-12-18) description, paragraphs [0044]-[0086] and [0131]-[0138], and figure 1	1-4, 6-13, 15-20
X	CN 105512657 A (BEIJING KUANGSHI TECHNOLOGY CO., LTD. et al.) 20 April 2016 (2016-04-20) description, paragraphs [0021]-[0029]	1-4, 6-13, 15-20
A	CN 108009546 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. et al.) 08 May 2018 (2018-05-08) entire document	1-20
A	US 2018150956 A1 (INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE) 31 May 2018 (2018-05-31) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

14 January 2020

Date of mailing of the international search report

23 January 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/117914

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109919014	A	21 June 2019	None			
CN	109034159	A	18 December 2018	None			
CN	105512657	A	20 April 2016	None			
CN	108009546	A	08 May 2018	None			
US	2018150956	A1	31 May 2018	TW	201820203	A	01 June 2018
				CN	108108732	A	01 June 2018

国际检索报告

国际申请号

PCT/CN2019/117914

A. 主题的分类 G06K 9/00 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																				
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNPAT, CNKI, WPI, EPDOC, IEEE: 区域, 位置, 文字识别, 字符识别, 文本识别, 图像, 模型, 样本, character, recogni+, OCR, location, region, position, image, model																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 109919014 A (平安科技深圳有限公司) 2019年 6月 21日 (2019 - 06 - 21) 权利要求1-10, 说明书第[0167]-[0170]段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 109034159 A (北京捷通华声科技股份有限公司) 2018年 12月 18日 (2018 - 12 - 18) 说明书第[0044]-[0086]、[0131]-[0138]段, 图1</td> <td>1-4, 6-13, 15-20</td> </tr> <tr> <td>X</td> <td>CN 105512657 A (北京旷视科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第[0021]-[0029]段</td> <td>1-4, 6-13, 15-20</td> </tr> <tr> <td>A</td> <td>CN 108009546 A (北京京东尚科信息技术有限公司 等) 2018年 5月 8日 (2018 - 05 - 08) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2018150956 A1 (INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE) 2018年 5月 31日 (2018 - 05 - 31) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 109919014 A (平安科技深圳有限公司) 2019年 6月 21日 (2019 - 06 - 21) 权利要求1-10, 说明书第[0167]-[0170]段	1-20	X	CN 109034159 A (北京捷通华声科技股份有限公司) 2018年 12月 18日 (2018 - 12 - 18) 说明书第[0044]-[0086]、[0131]-[0138]段, 图1	1-4, 6-13, 15-20	X	CN 105512657 A (北京旷视科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第[0021]-[0029]段	1-4, 6-13, 15-20	A	CN 108009546 A (北京京东尚科信息技术有限公司 等) 2018年 5月 8日 (2018 - 05 - 08) 全文	1-20	A	US 2018150956 A1 (INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE) 2018年 5月 31日 (2018 - 05 - 31) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 109919014 A (平安科技深圳有限公司) 2019年 6月 21日 (2019 - 06 - 21) 权利要求1-10, 说明书第[0167]-[0170]段	1-20																		
X	CN 109034159 A (北京捷通华声科技股份有限公司) 2018年 12月 18日 (2018 - 12 - 18) 说明书第[0044]-[0086]、[0131]-[0138]段, 图1	1-4, 6-13, 15-20																		
X	CN 105512657 A (北京旷视科技有限公司 等) 2016年 4月 20日 (2016 - 04 - 20) 说明书第[0021]-[0029]段	1-4, 6-13, 15-20																		
A	CN 108009546 A (北京京东尚科信息技术有限公司 等) 2018年 5月 8日 (2018 - 05 - 08) 全文	1-20																		
A	US 2018150956 A1 (INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE) 2018年 5月 31日 (2018 - 05 - 31) 全文	1-20																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																			
国际检索实际完成的日期 2020年 1月 14日		国际检索报告邮寄日期 2020年 1月 23日																		
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 孙国辉 电话号码 86-(10)-53961538																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/117914

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109919014	A	2019年 6月 21日	无	
CN	109034159	A	2018年 12月 18日	无	
CN	105512657	A	2016年 4月 20日	无	
CN	108009546	A	2018年 5月 8日	无	
US	2018150956	A1	2018年 5月 31日	TW 201820203 A	2018年 6月 1日
				CN 108108732 A	2018年 6月 1日