

文本检测分析方法、装置及设备

技术领域

本发明涉及图像处理技术领域，特别涉及一种文本检测分析方法、装置及设备。

发明背景

物体检测是一种计算机视觉（CV）和图像处理相关的计算机技术，从数字图像和视频中检测出特定类别（例如人类、建筑和汽车等）的语义对象实例。物体检测在人脸检测和行人检测领域发展比较完善。物体检测在计算机视觉领域有着大量的应用场景，包括图像检索和视频监控。

模板分析是一种光学字符识别相关的计算机技术，从数字文档图像中将对应的文字区域映射到语义对象实例，从而将数字文档图像等非结构化的数据转化成与语义对象实例相关的结构化数据。模板分析将文档中各个文字区域映射到语义对象实例从而将文档图像内容结构化。现有的方法中文字区域的匹配都是基于传统的图像特征匹配方法，图像的亮度、对比度、分辨率等会严重影响文字区域匹配的准确度；现有方法的文字区域的匹配是通过固定的相对位置去确定文字区域，但是手写文字、针式打印文字等会出现不可预知的位置偏移，文字区域的定位往往会出错。

发明内容

为了解决现有技术的问题，本发明实施例提供了一种文本检测分析方法、装置及设备。所述技术方案如下：

第一方面，本发明的实施例提供了一种文本检测分析方法，包括：利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息；根据标注模板信息和带类别的文字区域信息进行模板匹配，生成结构化信息数据。

在本发明某些实施例中，第一方面的方法还包括：通过对模板图像进行模板标注，生成标注模板信息，其中，模板图像与待检测图像的母版相同，待检测图像的文字区域包括第一锚点文字区域和第一非锚点文字区域，模板图像的文字区域包括第二锚点文字区域和第二非锚点文字区域。

在本发明某些实施例中，通过对模板图像进行模板标注，生成标注模板信息，包括：标注模板图像上第二锚点文字区域和第二非锚点文字区域的大小及相对位

置，以及实体与第二锚点文字区域、第二非锚点文字区域的映射关系，生成标注模板信息。

在本发明某些实施例中，利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息，包括：利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，以获取待检测图像上第一锚点文字区域和第一非锚点文字区域的文字区域信息。

在本发明某些实施例中，根据标注模板信息和带类别的文字区域信息进行模板匹配，生成结构化信息数据，包括：通过匹配第二锚点文字区域和第一锚点文字区域、以及第二非锚点文字区域和第一非锚点文字区域，确定待检测图像中第一锚点文字区域和第一非锚点文字区域的位置；根据实体与第二锚点文字区域、第二非锚点文字区域的映射关系，确定实体与第一锚点文字区域、第一非锚点文字区域的映射关系，并基于第一锚点文字区域和第一非锚点文字区域的位置去除待检测图像的文字区域中包括非文本区域和无关文本在内的冗余部分，以生成结构化信息。

在本发明某些实施例中，第一方面的方法还包括：对待检测图像进行图像预处理，图像预处理包括图像矫正和/或缩放到预设尺寸。

在本发明某些实施例中，第一方面的方法还包括：训练预设深度神经网络检测模型。

在本发明某些实施例中，训练预设深度神经网络检测模型，包括：使用样本生成工具生成样本；使用样本训练预设深度神经网络检测模型。

在本发明某些实施例中，待检测图像为新样本，其中，第一方面的方法还包括：利用新样本对预设深度神经网络检测模型进行微调。

第二方面，本发明的实施例提供了一种文本检测分析装置，包括：文字区域检测模块，用于利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息；匹配模块，用于根据标注模板信息和带类别的文字区域信息进行模板匹配，生成结构化信息数据。

在本发明某些实施例中，第二方面的装置还包括：标注模块，用于通过对模板图像进行模板标注，生成标注模板信息，其中，模板图像与待检测图像的母版相同，待检测图像的文字区域包括第一锚点文字区域和第一非锚点文字区域，模板图像的文字区域包括第二锚点文字区域和第二非锚点文字区域。

在本发明某些实施例中，标注模块用于：标注模板图像上第二锚点文字区域和第二非锚点文字区域的大小及相对位置，以及实体与第二锚点文字区域、第二非锚点文字区域的映射关系，生成标注模板信息。

在本发明某些实施例中，文字区域检测模块用于：利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，以获取待检测图像上第一锚点

文字区域和第一非锚点文字区域的文字区域信息。

在本发明某些实施例中，匹配模块用于：通过匹配第二锚点文字区域和第一锚点文字区域、以及第二非锚点文字区域和第一非锚点文字区域，确定待检测图像中第一锚点文字区域和第一非锚点文字区域的位置；根据实体与第二锚点文字区域、第二非锚点文字区域的映射关系，确定实体与第一锚点文字区域、第一非锚点文字区域的映射关系，并基于第一锚点文字区域和第一非锚点文字区域的位置去除待检测图像的文字区域中包括非文本区域和无关文本在内的冗余部分，以生成结构化信息。

在本发明某些实施例中，第二方面的装置还包括：图像预处理模块，用于对待检测图像进行图像预处理，图像预处理包括图像矫正和/或缩放到预设尺寸。

在本发明某些实施例中，第二方面的装置还包括：模型训练模块，用于训练预设深度神经网络检测模型。

在本发明某些实施例中，模型训练模块用于：使用样本生成工具生成样本；使用样本训练预设深度神经网络检测模型。

在本发明某些实施例中，待检测图像为新样本，其中，第二方面的装置还包括：微调模块，用于利用新样本对预设深度神经网络检测模型进行微调。

第三方面，本发明的实施例提供了一种文本检测分析设备，包括：处理器；存储器，用于存储有处理器的可执行指令；其中，处理器配置为经由可执行指令来执行第一方面所述的文本检测分析方法。

第四方面，本发明的实施例提供了一种计算机可读存储介质，存储介质存储有计算机程序，计算机程序用于执行第一方面所述的文本检测分析方法。

本发明实施例提供了一种文本检测分析方法、装置及设备，通过利用预设深度神经网络检测模型进行图像文字区域的检测、分类，然后结合标注模板信息与检测分类得到的带类别的文字区域信息，进行模板匹配，生成结构化信息数据，能够对文档图像中各种字段，实现快速、准确的检测分析，具有实时、准确、通用、鲁棒、可扩展的特点。

附图简要说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1是本发明实施例1提供的文本检测分析方法的流程示意图；

图2是一优选实施方式的模板匹配过程示意图；

图3是本发明实施例2提供的文本检测分析方法的流程示意图；

图4是本发明实施例3提供的文本检测分析装置的结构示意图；

图5是本发明实施例提供的文本检测分析设备的组成示意图；

图6至图9是本发明实施例提供的文本检测分析方法、装置及设备应用实例的票据处理效果图。

实施本发明的方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。需要说明的是，在本发明的描述中，“多个”的含义是两个以上，除非另有明确具体的限定。

下面结合具体实施例及附图，对本发明实施例提供的文本检测分析方法、装置及设备作详细说明。

实施例1

图1是本发明实施例1提供的文本检测分析方法的流程示意图。如图1所示，本发明实施例提供的文本检测分析方法，包括以下内容。

101: 利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息。

神经网络(Neural Network)是一种人为设计的网络结构，其本质是多层感知机(Multi-layer Perceptron)。感知机由若干神经元(Neuron)构成，每个神经元从外部或其它节点收到输入信号，并通过激活函数得到输出信号，就像大脑中神经元的信号传递。神经元按层相连，形成网络结构。与神经细胞不同，人工神经元的信号能够反向传播，这种反馈机制让感知机具备学习功能。除了学习功能，多层感知机可以表示非线性映射，因此神经网络能够帮助人们解决一些相对复杂的问题，如模式识别、自动控制、决策评估、预测等。

卷积神经网络(Convolutional Neural Network)是一种前馈神经网络，它的神经元可以响应一部分覆盖范围内的周围单元，形成感受野(Receptive Field)，对图像处理有较好表现。同时，通过权值共享，卷积神经网络相比传统的全连接神经网络节省空间。用卷积神经网络学习出的特征，具有一定的位置不变性与尺度不变性，用卷积神经网络训练的光学字符识别(Optical Character Recognition, OCR)模型，对字体、字号的变化具有鲁棒性。

深度学习的检测算法主要包含三个部分：检测窗口的选择，特征的提取和分

类器的选择。深度学习早期，大都使用滑动窗口的方式进行窗口提取，近期的检测算法例如 Faster R-CNN 则直接利用区域建议网络（Region Proposal Networks, RPN）网络来计算候选框。RPN 以一张任意大小的图片为输入，输出一批矩形区域，每个区域对应一个目标分数和位置信息。借助于卷积神经网络强大的非线性建模能力，能够提取出契合各种不同子任务的特征。深度学习的检测算法也打破了传统滑框的框架，使用回归的方法解决检测问题。

在本实施例中，预设深度神经网络检测模型可以采用 Faster R-CNN 网络模型、Mask R-CNN 网络模型或者其他任何可能的深度神经网络模型，本发明实施例不对其加以特别限制。

具体地，待检测图像可以是发票、收据、票据等具有文本信息的图像，该图像可以通过拍摄或扫描纸质的文件获得的，也可以直接是电子版的。

例如，待检测图像是发票，该发票上标有一些项目类别，如“发票代码”、“发票号码”、“付款方名称”、“收款方名称”、“开票日期”等。该发票未被填写时的版本，可以称为该发票的母版，母版上的文本信息可以通过电子设备输入的；当发票被填写后（可以是电子填写或手写）获得的图像，可以称为待检测图像。

待检测图像的文字区域可以包括母版上原有的文本信息，以及用户填写的文本信息。

采用深度神经网络检测模型能够精准找出不同尺度(字号)的文字区域，并对文字内容进行初步分类，如金额、项目名、词条、票据编号等类别，从而可以准确地检测出文档中的文字。

在一实施例中，文字区域信息的类别可以分为：母版上原有的文本信息、用户填写的文本信息；进一步地，文字区域信息的类别还可以分为：文本信息为日期、汉字、英文等等。利用预设深度神经网络检测模型检测待检测图像的文字区域，可以获得带类别的文字区域信息。具体地，带类别的文字区域信息可以包括文本的类别、位置信息等。

预设深度神经网络检测模型可以针对一种母版的待检测图像进行检测和分类，也可以针对多种不同母版的待检测图像分别进行检测和分类，这取决于训练时的样本类别。

102：根据标注模板信息和带类别的文字区域信息进行模板匹配，生成结构化信息数据，其中，标注模板信息包括用户需要获得的实体的信息，实体为用户需要从待检测图像上获得的信息。

标注模板信息可以是对待检测图像或待检测图像的母版进行标注后得到的。实体的信息可以包括实体的类别、位置信息等。例如，用户想要获得的实体的类别是“付款方名称”，标注模板信息包括该实体的类别以及该实体（付款方的具体名称）在待检测图像中的位置信息，该位置信息可以是绝对位置信息，也可以是

相对于其他文本信息的相对位置信息。

当位置信息是相对于其他文本信息的相对位置信息时，通过将标注模板信息与带类别的文字区域信息进行模板匹配，可以准确地获得用户所需要的实体的位置，即将待检测图像中对应的文字区域映射到语义对象实例，实现待检测图像内容的结构化。

这里，用户需要获得的实体可以是一个或多个。对实体具体内容的识别过程可以在标注模板信息和带类别的文字区域信息的匹配完成后执行，也可以在 101 中执行。这里涉及的识别过程可以采用 tesseract、CRNN 之类的识别模型，本发明实施例不对其加以特别限制。

本发明实施例提供的文本检测分析方法，通过利用预设深度神经网络检测模型进行图像文字区域的检测、分类，然后结合标注模板信息与检测分类得到的带类别的文字区域信息，进行模板匹配，生成结构化信息数据，能够对文档图像中各种字段，实现快速、准确的检测分析，具有实时、准确、通用、鲁棒、可扩展的特点。

根据本发明一实施例，图 1 的方法还包括：通过对模板图像进行模板标注，生成标注模板信息，其中，模板图像与待检测图像的母版相同，待检测图像的文字区域包括第一锚点文字区域和第一非锚点文字区域，模板图像的文字区域包括第二锚点文字区域和第二非锚点文字区域。

具体地，锚点文字区域可以是母版的文本信息对应的区域；非锚点文字区域可以是除了母版的文本信息之外的其他文本信息对应的区域，例如，需要用户填写的文本区域。

锚点文字区域可以包括一个或多个锚点，每个锚点可以是一个字段，多个锚点可以分布在不同的位置。锚点是每个模板图像中最为稳定的特征，内容和相对位置都不会发生变化，例如表格中的表头，通过设置多个锚点可以完全确定当前模板图像中各个区域的相对位置关系。

模板图像可以是空白的母版，也可以是待检测图像，或者是与待检测图像为同一母版的其他图像（可以为另一用户填写）。当模板图像为空白的母版时，第二非锚点文字区域可以是用户还未填写的空白区域。

进一步地，通过对模板图像进行模板标注，生成标注模板信息，包括：标注模板图像上第二锚点文字区域和第二非锚点文字区域的大小及相对位置，以及实体与第二锚点文字区域、第二非锚点文字区域的映射关系，生成标注模板信息。

具体地，该模板标注过程用于标注所有需要识别字段的位置和类别，类别包括：是否为锚点，以及文本信息是否为日期、汉字、英文等等。生成的标注模板信息用于后续的模板匹配。通过模板标注可以确定第二锚点和第二非锚点之间的相对位置关系，以及每个第二锚点和第二非锚点对应的文字区域的大小。第一锚

点对应的文字区域的大小，可以指第一锚点对应的文字区域的高度和宽度，即第一锚点对应的文字区域的形状可以是矩形，当然也可以是圆形、椭圆或其他形状，本发明实施例对此不做具体限定。

进一步地，当第二锚点和第二非锚点分别为多个时，通过模板标注还可以确定各个第二锚点之间的相对位置关系，以及各个第二非锚点之间的相对位置关系。

与预设深度神经网络检测模型类似，这里进行标注的模板图像并不限于一种或两种，根据实际需要，可以预先设置多种模板图像，进而获得多种标注模板信息。

在该实施例中，101 包括：利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，以获取待检测图像上第一锚点文字区域和第一非锚点文字区域的文字区域信息。

具体地，带类别的文字区域信息可以包括第一锚点文字区域和第一非锚点文字区域的文字区域信息。

进一步地，102 包括：通过匹配第二锚点文字区域和第一锚点文字区域、以及第二非锚点文字区域和第一非锚点文字区域，确定待检测图像中第一锚点文字区域和第一非锚点文字区域的位置；根据实体与第二锚点文字区域、第二非锚点文字区域的映射关系，确定实体与第一锚点文字区域、第一非锚点文字区域的映射关系，并基于第一锚点文字区域和第一非锚点文字区域的位置去除待检测图像的文字区域中包括非文本区域和无关文本在内的冗余部分，以生成结构化信息。

具体地，利用预设深度神经网络检测模型可以确定待检测图像上各个第一锚点和第一非锚点的绝对位置。但是，在实际检测过程中，由于图像本身存在倾斜、或者个别字段不清楚，会导致一些第一锚点、第一非锚点未被检测出来或检测出的位置偏移实际位置。在本实施例中，通过将标注模板信息标注与带类别的文字区域信息结合，即绝对位置结合相对位置，可以调整定位偏移的第一锚点、第一非锚点的位置、或定位出未被检测出的第一锚点、第一非锚点的位置，进而可以准确的定位想要的文字区域位置和内容信息。

此外，还可以通过文字区域的大小匹配来确定第一锚点和第一非锚点的位置。尤其当待检测图像是基于手写发票获得的图像时，不同的用户填写信息时，信息的填写位置会有不同的偏差。此时可以通过第二非锚点对应的文字区域的大小与第一非锚点对应的文字区域的大小进行匹配，当匹配度满足第一阈值时，该第一非锚点被认为匹配成功。

基于各个锚点、非锚点之间的相对位置关系，通过匹配成功的锚点、非锚点的位置可以确定未匹配成功的其他锚点、非锚点的位置，从而可以提高匹配结果的准确度，提高识别率。

在本实施例中，对于位置和内容信息稳定的锚点能够实现准确的匹配，提升

匹配步骤的准确率。此外，通过排除非文本区域与无关文本，更利于抓住关键信息，进一步提升检测结果的准确率。

根据本发明一实施例，图1的方法还包括：对待检测图像进行图像预处理，图像预处理包括图像矫正和/或缩放到预设尺寸。

具体地，由于实际检测的图像可能存在倾斜或大小不一致的问题，这在检测时，会影响对锚点和非锚点的位置进行定位的精确度。因此，在本实施例中，可以在101之前，对待检测图像进行矫正，或缩放。将待检测图像缩放至预设深度神经网络检测模型训练时的样本的尺寸时，可以提高匹配过程的准确率。在本实施例中，模版图像的尺寸可以与预设深度神经网络检测模型训练时的样本的尺寸一致。当然，也可以根据实际情况进行图像预处理操作的相应设置。

图2是一优选实施方式的模板匹配过程示意图，如图2所示，上述102可进一步包括以下内容。

1021：结合标注模板信息与带类别的文字区域信息进行模板匹配。

具体地，将标注模板信息与带类别的文字区域信息进行匹配，以确定第一锚点文字区域、第一非锚点文字区域的位置，即对多个第一锚点、第一非锚点进行定位。在本实施例中，该匹配过程可以弥补预设深度神经网络检测模型未检测出的第一锚点、第一非锚点，或调整定位结果不准的第一锚点、第一非锚点的位置。

1022：通过实体与文字区域的映射，去除冗余部分，生成结构化信息数据。

通过第一锚点、第一非锚点将带类别的文字区域映射到对应的实体，去除包括非文本区域和无关文本在内的冗余部分，进而确定与实体对应的锚点和非锚点。通过识别可以确定非锚点处的文本信息，并根据识别出的文本信息与锚点的相对位置，可以确定该文本信息的类型（及锚点的文本信息，如姓名，性别，年龄等等）。结合该文本信息以及该文本信息的类型可以生成结构化信息数据。可以将结构化信息数据输出，以便进行相应应用。

根据本发明一实施例，图1的方法还包括：训练预设深度神经网络检测模型。

具体地，训练预设深度神经网络检测模型的样本可以是提前标记好锚点的位置。即在训练中，会对样本中的文字区域分类（包括但不限于进行锚点类和非锚点类的分类），然后训练检测模型。

在一实施例中，训练预设深度神经网络检测模型，包括：使用样本生成工具生成样本；使用样本训练预设深度神经网络检测模型。在该实施例中，通过样本生成工具可以快速地获得大量的样本。

根据本发明一实施例，待检测图像为新样本，其中，图1的方法还包括：利用新样本对预设深度神经网络检测模型进行微调。

具体地，在检测应用中形成数据回流，可以获得更多的新样本。利用新样本可以对预设深度神经网络检测模型进行微调（Fine-tuning）。这样，在训练初期，

可以不需要大量标注样本来训练检测模型，形成数据回流之后模型的识别率可以逐步提高。

实施例 2

图 3 是本发明实施例 2 提供的文本检测分析方法的流程示意图,如图 3 所示,本发明实施例提供的文本检测分析方法包括以下内容。

201: 训练预设深度神经网络检测模型。

具体地,该训练过程可以参见图 1 中的描述。

在检测应用中可以形成数据回流,获得更多新样本,并利用新样本对预设深度神经网络检测模型进行微调(Fine-tuning)。

在该训练或微调过程中,会对样本中的文字行分类(包括但不限于进行锚点类和非锚点类的分类),然后训练检测模型。

值得注意的是,201 除了上述实现方式之外,还可以通过其他方式实现,本发明实施例对具体的方式不加以限定。

202: 标注模板图像上第二锚点文字区域和第二非锚点文字区域的大小及相对位置,以及实体与第二锚点文字区域、第二非锚点文字区域的映射关系,生成标注模板信息。

这一过程用于标注模板图像上所有需要识别字段的位置和类别,类别可以包括:是否为锚点,以及文本信息是否为日期、汉字、英文等等。生成的标注模板信息用于后续的模板匹配。需要注意的是,这里进行标注的模板图像并不限于一种或两种,根据实际需要,可以预先设置多种模板图像,进而获得多种标注模板信息。

值得注意的是,202 除了上述实现方式之外,还可以通过其他方式实现,本发明实施例对具体的方式不加以限定。

203: 对待检测图像进行图像预处理,图像预处理包括图像矫正和/或缩放到预设尺寸。这里图像预处理的内容包括但不限于图像矫正和/或缩放到预设尺寸,可以根据实际情况进行图像预处理操作的相应设置。

值得注意的是,203 除了上述实现方式之外,还可以通过其他方式实现,本发明实施例对具体的方式不加以限定。

204: 利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类,以获取待检测图像上第一锚点文字区域和第一非锚点文字区域的文字区域信息。

预设深度神经网络检测模型可以采用 Faster R-CNN 网络模型、Mask R-CNN 网络模型或者现有技术中其他任何可能的深度神经网络模型,本发明实施例不对其加以特别限制。

采用深度神经网络物体检测框架能够精准找出不同尺度(字号)的文字区域,并对文字内容进行初步分类,如金额、项目名、词条、票据编号等类别。

值得注意的是,204 除了上述实现方式之外,还可以通过其他方式实现,本发明实施例对具体的方式不加以限定。

202 可以在 201 之前执行、或与 201 同时执行,本发明实施例对此不做限制。

205:将标注模板信息与带类别的文字区域信息进行模板匹配,定位第一锚点、第一非锚点,并通过第一锚点、第一非锚点将带类别的文字区域映射到对应的实体,去除包括非文本区域和无关文本在内的冗余部分,生成结构化信息数据。

锚点定位成功了,相当于确定了整个版面各个字段的相对位置,然后结合之前的锚点、非锚点文字区域信息和版面的相对位置就可以准确的定位想要的字段对位置和内容信息,因此稳定的锚点能够实现准确的匹配,提升匹配步骤的准确率。

通过识别可以确定非锚点处的文本信息,并根据识别出的文本信息与锚点的相对位置,可以确定该文本信息的类型(及锚点的文本信息,如姓名,性别,年龄等等)。结合该文本信息以及该文本信息的类型可以生成结构化信息数据。可以将结构化信息数据输出,以便进行相应应用。

值得注意的是,205 除了上述实现方式之外,还可以通过其他方式实现,本发明实施例对具体的方式不加以限定。

需要说明的是,本发明实施例提供的文本检测分析方法中具体步骤的操作顺序,并不限定于上述方法中的步骤顺序,在不脱离本发明发明构思的情况下,可以根据实际情况进行相应变换设计。

实施例 3

图 4 是本发明实施例 3 提供的文本检测分析装置 30 的结构示意图,如图 4 所示,本发明实施例提供的文本检测分析装置 30 包括文字区域检测模块 31 和匹配模块 32。

文字区域检测模块 31 用于利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类,生成带类别的文字区域信息;匹配模块 32 用于根据标注模板信息和带类别的文字区域信息进行模板匹配,生成结构化信息数据,其中,标注模板信息包括用户需要获得的实体的信息,实体为用户需要从待检测图像上获得的信息。

具体地,最终可将结构化信息输出,可提供统一接口方便其他外部模块调用。

本发明实施例提供的文本检测分析装置,通过利用预设深度神经网络检测模型进行图像文字区域的检测、分类,然后结合标注模板信息与检测分类得到的带类别的文字区域信息,进行模板匹配,生成结构化信息数据,能够对文档图像中

各种字段，实现快速、准确的检测分析，具有实时、准确、通用、鲁棒、可扩展的特点。

根据本发明一实施例，装置 30 还包括：标注模块 33，用于通过对模板图像进行模板标注，生成标注模板信息，其中，模板图像与待检测图像的母版相同，待检测图像的文字区域包括第一锚点文字区域和第一非锚点文字区域，模板图像的文字区域包括第二锚点文字区域和第二非锚点文字区域。

根据本发明一实施例，标注模块 33 用于：标注模板图像上第二锚点文字区域和第二非锚点文字区域的大小及相对位置，以及实体与第二锚点文字区域、第二非锚点文字区域的映射关系，生成标注模板信息。

根据本发明一实施例，文字区域检测模块 31 用于：利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，以获取待检测图像上第一锚点文字区域和第一非锚点文字区域的文字区域信息。

根据本发明一实施例，匹配模块 32 用于：通过匹配第二锚点文字区域和第一锚点文字区域、以及第二非锚点文字区域和第一非锚点文字区域，确定待检测图像中第一锚点文字区域和第一非锚点文字区域的位置；根据实体与第二锚点文字区域、第二非锚点文字区域的映射关系，确定实体与第一锚点文字区域、第一非锚点文字区域的映射关系，并基于第一锚点文字区域和第一非锚点文字区域的位置去除待检测图像的文字区域中包括非文本区域和无关文本在内的冗余部分，以生成结构化信息。

根据本发明一实施例，装置 30 还包括：图像预处理模块 34，用于对待检测图像进行图像预处理，图像预处理包括图像矫正和/或缩放到预设尺寸。

根据本发明一实施例，装置 30 还包括：模型训练模块 35，用于训练预设深度神经网络检测模型。

根据本发明一实施例，模型训练模块 35 用于：使用样本生成工具生成样本；使用样本训练预设深度神经网络检测模型。

根据本发明一实施例，待检测图像为新样本，其中，装置 30 还包括：微调模块 36，用于利用新样本对预设深度神经网络检测模型进行微调。

应当理解，上述实施例中的文字区域检测模块 31、匹配模块 32、标注模块 33、图像预处理模块 34、模型训练模块 35 以及微调模块 36 的操作和功能可以参考上述图 1、图 2 以及图 3 提供的文本检测分析方法中的描述，为避免重复，在此不再赘述。

实施例 4

图 5 是本发明实施例提供的文本检测分析设备 40 的组成示意图，如图 5 所示，本发明实施例提供的文本检测分析设备 40 包括处理器 41 和存储器 42。

具体地，存储器 42 用于存储有处理器 41 的可执行指令，处理器 41 配置为经由可执行指令来执行实施例 1 和实施例 2 所述的文本检测分析方法的步骤，具体实施过程详见实施例 1、实施例 2 相应内容，在此不再赘述。

应用实例

系统构建方面分为硬件需求与软件需求。

硬件方面需要服务器一台、计算设备一台、图像输入设备一台。

软件方面需要在服务器上架设票据识别的 web 服务、用于终端显示的页面、存储图像及检测结果的数据库。

流程方面分为模型生成流程及文档检测流程。

模型生成方面，首先通过样本生成模块生成训练样本，如果有足够人力，可采用人工标记真实样本。训练样本足够后(通常为 10 万数量级)，在服务器上通过深度神经网络训练票据识别模型，并存储在服务器上，架设 web 服务。

检测方面，首先用图像输入设备对文档进行扫描。得到文档图像后，从终端打开识别服务的页面，将图像上传至服务器。服务器接收到图像后，通过图像预处理模块、文本检测模块以及模板匹配模块获得结构化检测结果，之后利用识别模块对相应区域的具体文本内容进行识别，再根据具体情况对识别结果进行后处理，最终将识别结果传回终端，展现给用户。

下面具体的以一种医疗票据的识别为例。假设我们需要识别票据上的医院名称、医院类型、医保类型、对应项目以及金额。医院名称、医院类型、医保类型、对应项目以及金额，是需要获得的实体。

第一步，获得如图 6 所示的原始票据图像。

第二步，如图 7 所示，对票据图像进行版面分析。

对票据图像进行版面分析，即进行模板标注，生成标注模板信息。这里，待检测图像和模板图像为同一张图像。具体地，对于同一母版的多张票据的识别，可以只进行一次模板标注。

第三步，如图 8 所示，找出不同类别的文字区域。

具体地，可以利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息，以区分不同类别的文字区域。

第四步，如图 9 所示，根据版面分析与文字区域分类，提取出需要识别的区域。

具体地，根据标注模板信息和带类别的文字区域信息进行模板匹配，提取出需要识别的区域。

在本实施例中，根据实体与文字区域的映射关系，需要去除包括非文本区域和无关文本在内的冗余部分。例如，如图 8 所示，箭头 A 所指的文字区域为无关

文本，箭头 B 所指的文字区域为非文本区域，A 和 B 处的文字区域是需要去除的冗余部分。

第五步，用相应的识别模型识别上述区域。

识别结果：

“上海第一妇婴保健院”

“三级甲等专科医院”

“医保自费”

“治疗费(自费)”

“门诊诊疗费”

“6.00”

“800”

上述的识别结果可以是通过 tesseract、CRNN 等识别模型获得的，识别结果也有可能与实际结果存在一些偏差，因此可以通过后处理过程对识别结果进行校正。

第六步，对识别结果进行后处理。

后处理结果为：

“上海市第一妇婴保健院”（查询医院名表匹配）

“三级甲等专科医院”

“医保自费”

“治疗费(自费)”

“门诊诊疗费”

“6.00”

“8.00”（根据金额规则匹配）

在这里，可以通过查询医院名表对识别结果中的医院名“上海第一妇婴保健院”进行校正，校正结果为“上海市第一妇婴保健院”，即识别结果中的医院名少了一个“市”字。类似地，可以对识别结果中的每个信息分别进行校正。例如，识别结果中门诊诊疗费是“800”，但是经过金额规则与之匹配后，将其校正为“8.00”。

第七步，格式化识别结果，生成结构化信息数据。

```
{  
  医院名:上海市第一妇婴保健院  
  医院类型: 三级甲等专科医院  
  医保类型: 医保自费  
  费用清单:  
  {
```

治疗费(自费): 6.00

门诊诊疗费: 8.00

}

}

需要说明的是：上述实施例提供的文本检测分析装置、设备在进行文本检测分析时，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置、设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的文本检测分析装置、设备与文本检测分析方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过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。

本申请实施例中是参照根据本申请实施例中实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

尽管已描述了本申请实施例中的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例作出另外的变更和修改。所以，所附权利要求要求意欲解释为包括优选实施例以及落入本申请实施例中范围的所有变更和修改。

显然，本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样，倘若本发明的这些修改和变型属于本发明权利要求及其等

同技术的范围之内，则本发明也意图包含这些改动和变型在内。

以上所述仅为本发明的较佳实施例，并不用以限制本发明，凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权利要求书

1、一种文本检测分析方法，其特征在于，包括：

利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息；

根据标注模板信息和所述带类别的文字区域信息进行模板匹配，生成结构化信息数据。

2、根据权利要求1所述的方法，其特征在于，还包括：

通过对模板图像进行模板标注，生成所述标注模板信息，其中，所述模板图像与所述待检测图像的母版相同，所述待检测图像的文字区域包括第一锚点文字区域和第一非锚点文字区域，所述模板图像的文字区域包括第二锚点文字区域和第二非锚点文字区域。

3、根据权利要求2所述的方法，其特征在于，所述通过对模板图像进行模板标注，生成所述标注模板信息，包括：

标注所述模板图像上所述第二锚点文字区域和所述第二非锚点文字区域的大小及相对位置，以及实体与所述第二锚点文字区域、所述第二非锚点文字区域的映射关系，生成所述标注模板信息。

4、根据权利要求2或3所述的方法，其特征在于，所述利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类，生成带类别的文字区域信息，包括：

利用所述预设深度神经网络检测模型对所述待检测图像的文字区域进行检测和分类，以获取所述待检测图像上所述第一锚点文字区域和所述第一非锚点文字区域的文字区域信息。

5、根据权利要求2至4中任一项所述的方法，其特征在于，所述根据标注模板信息和所述带类别的文字区域信息进行模板匹配，生成结构化信息数据，包括：

通过匹配所述第二锚点文字区域和所述第一锚点文字区域、以及所述第二非锚点文字区域和所述第一非锚点文字区域，确定所述待检测图像中所述第一锚点文字区域和所述第一非锚点文字区域的位置；

根据实体与所述第二锚点文字区域、所述第二非锚点文字区域的映射关系，确定所述实体与所述第一锚点文字区域、所述第一非锚点文字区域的映射关系，并基于所述第一锚点文字区域和所述第一非锚点文字区域的位置去除所述待检测图像的文字区域中包括非文本区域和无关文本在内的冗余部分，以生成所述结构化信息。

6、根据权利要求1至5任一项所述的方法，其特征在于，还包括：

对所述待检测图像进行图像预处理,所述图像预处理包括图像矫正和/或缩放到预设尺寸。

7、根据权利要求1至6任一项所述的方法,其特征在于,还包括:

训练所述预设深度神经网络检测模型。

8、根据权利要求7所述的方法,其特征在于,所述训练所述预设深度神经网络检测模型,包括:

使用样本生成工具生成样本;

使用所述样本训练所述预设深度神经网络检测模型。

9、根据权利要求7或8所述的方法,其特征在于,

所述待检测图像为新样本,其中,所述方法还包括;

利用所述新样本对所述预设深度神经网络检测模型进行微调。

10、一种文本检测分析装置,其特征在于,包括:

文字区域检测模块,用于利用预设深度神经网络检测模型对待检测图像的文字区域进行检测和分类,生成带类别的文字区域信息;

匹配模块,用于根据标注模板信息和所述带类别的文字区域信息进行模板匹配,生成结构化信息数据。

11、根据权利要求10所述的装置,其特征在于,还包括:

标注模块,用于通过对模板图像进行模板标注,生成所述标注模板信息,其中,所述模板图像与所述待检测图像的母版相同,所述待检测图像的文字区域包括第一锚点文字区域和第一非锚点文字区域,所述模板图像的文字区域包括第二锚点文字区域和第二非锚点文字区域。

12、根据权利要求11所述的装置,其特征在于,所述标注模块用于:

标注所述模板图像上所述第二锚点文字区域和所述第二非锚点文字区域的大小及相对位置,以及实体与所述第二锚点文字区域、所述第二非锚点文字区域的映射关系,生成所述标注模板信息。

13、根据权利要求11或12所述的装置,其特征在于,所述文字区域检测模块用于:

利用所述预设深度神经网络检测模型对所述待检测图像的文字区域进行检测和分类,以获取所述待检测图像上所述第一锚点文字区域和所述第一非锚点文字区域的文字区域信息。

14、根据权利要求11至13中任一项所述的装置,其特征在于,所述匹配模块用于:

通过匹配所述第二锚点文字区域和所述第一锚点文字区域、以及所述第二非锚点文字区域和所述第一非锚点文字区域,确定所述待检测图像中所述第一锚点文字区域和所述第一非锚点文字区域的位置;根据实体与所述第二锚点文字区域、

所述第二非锚点文字区域的映射关系，确定所述实体与所述第一锚点文字区域、所述第一非锚点文字区域的映射关系，并基于所述第一锚点文字区域和所述第一非锚点文字区域的位置去除所述待检测图像的文字区域中包括非文本区域和无关文本在内的冗余部分，以生成所述结构化信息。

15、根据权利要求 10 至 14 任一项所述的装置，其特征在于，还包括：

图像预处理模块，用于对所述待检测图像进行图像预处理，所述图像预处理包括图像矫正和/或缩放到预设尺寸。

16、根据权利要求 10 至 15 任一项所述的装置，其特征在于，还包括：

模型训练模块，用于训练所述预设深度神经网络检测模型。

17、根据权利要求 16 所述的装置，其特征在于，所述模型训练模块用于：

使用样本生成工具生成样本；

使用所述样本训练所述预设深度神经网络检测模型。

18、根据权利要求 16 或 17 所述的装置，其特征在于，所述待检测图像为新样本，其中，所述装置还包括：

微调模块，用于利用所述新样本对所述预设深度神经网络检测模型进行微调。

19、一种文本检测分析设备，其特征在于，包括：

处理器；

存储器，用于存储有所述处理器的可执行指令；

其中，所述处理器配置为经由所述可执行指令来执行权利要求 1 至 9 中任一项所述的文本检测分析方法。

20、一种计算机可读存储介质，所述存储介质存储有计算机程序，所述计算机程序用于执行上述权利要求 1 至 9 中任一项所述的文本检测分析方法。