



(12) 发明专利申请

(10) 申请公布号 CN 113065400 A

(43) 申请公布日 2021.07.02

(21) 申请号 202110242359.X

(51) Int.Cl.

(22) 申请日 2021.03.04

G06K 9/00 (2006.01)

G06K 9/32 (2006.01)

(71) 申请人 国网河北省电力有限公司

G06N 3/04 (2006.01)

地址 050022 河北省石家庄市富强大街32号

申请人 国网河北省电力有限公司雄安新区供电公司

(72) 发明人 刘义江 姜琳琳 李云超 辛锐

陈曦 侯栋梁 魏明磊 杨青

池建昆 范辉 陈蕾 阎鹏飞

吴彦巧 姜敬 檀小亚 师孜晗

(74) 专利代理机构 石家庄新世纪专利商标事务
所有限公司 13100

代理人 呼春辉

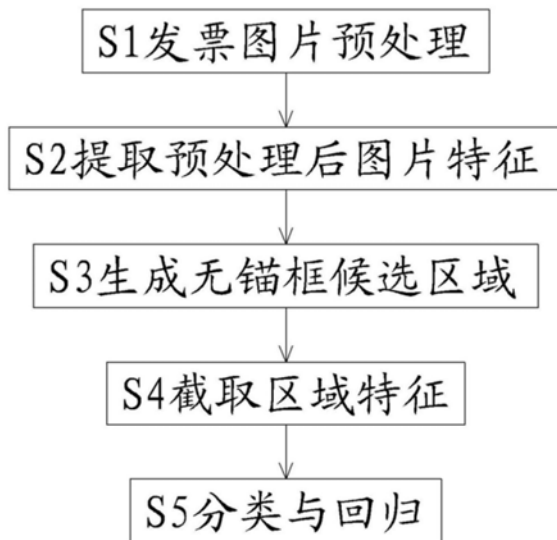
权利要求书2页 说明书8页 附图2页

(54) 发明名称

基于无锚框两阶段网络检测发票印章的方法及装置

(57) 摘要

本发明公开了基于无锚框两阶段网络检测发票印章的方法及装置,涉及票据文本检测技术领域;方法包括S1发票图片预处理,处理器对发票图片图像预处理并获得统一尺寸的预处理图片;S2提取预处理后图片特征,处理器将预处理图片输入至特征提取卷积神经网络并获得特征图;S3生成无锚框候选区域,处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域;装置包括发票图片预处理模块、提取预处理后图片特征模块和生成无锚框候选区域模块;其通过步骤S1至步骤S4等,实现发票印章检测第一阶段中锚框冗余度较小,提升了检测发票印章的工作效率。



1. 一种基于无锚框两阶段网络检测发票印章的方法,其特征在于:包括如下步骤,S1发票图片预处理,处理器从存储器获取发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片;S2提取预处理后图片特征,处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图;S3生成无锚框候选区域,处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。

2. 根据权利要求1所述的基于无锚框两阶段网络检测发票印章的方法,其特征在于:步骤S1具体包括如下步骤,S101旋转处理,处理器对预处理图片做随机旋转处理,以50%的概率进行水平旋转并获得旋转图片;S102归一化处理,处理器将旋转图片进行归一化处理并获得归一化图片;S103统一图片,处理器将上述归一化图片进行填充并获得统一尺寸的预处理图片;在步骤S2中,所述特征图为最后得到的大小为 $C \times H \times W$ 的特征向量矩阵F,其中,C为图像的通道,H为图像的高度,W为图像的宽度。

3. 根据权利要求1所述的基于无锚框两阶段网络检测发票印章的方法,其特征在于:在步骤S3之后还包括如下步骤,S4截取区域特征,处理器通过无锚框候选区域对特征图进行截取并获得区域特征图;S5分类与回归,处理器基于 $K \times K \times C$ 的区域特征图进行分类和回归处理。

4. 根据权利要求3所述的基于无锚框两阶段网络检测发票印章的方法,其特征在于:在步骤S4中,基于任一候选框沿特征图的高度和宽度方向都平均切成K份,获得 $K \times K$ 个方格,对每一个方格进行最大池化,获得 $K \times K \times C$ 的区域特征图, $K=5, C=512$;在步骤S5中,对于每一个区域特征图,分别经过分类分支和回归分支,每个分支都是四层 3×3 的卷积层,分类分支最后输出的特征图形状为 $H \times W \times N$,回归分支最后输出的特征图形状为 $H \times W \times 4$,其中N为待分类的类别数目,4为回归得到的到四个边的距离。

5. 一种基于无锚框两阶段网络检测发票印章的装置,其特征在于:包括发票图片预处理模块、提取预处理后图片特征模块和生成无锚框候选区域模块,发票图片预处理模块为程序模块,用于处理器从存储器获取发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片;提取预处理后图片特征模块为程序模块,用于处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图;生成无锚框候选区域模块为程序模块,用于处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。

6. 根据权利要求5所述的基于无锚框两阶段网络检测发票印章的装置,其特征在于:所述发票图片预处理模块,还用于处理器对预处理图片做随机旋转处理,以50%的概率进行水平旋转并获得旋转图片,处理器将旋转图片进行归一化处理并获得归一化图片,处理器将上述归一化图片进行填充并获得统一尺寸的预处理图片;在提取预处理后图片特征模块中,所述特征图为最后得到的大小为 $C \times H \times W$ 的特征向量矩阵F,其中,C为图像的通道,H为图像的高度,W为图像的宽度。

7. 根据权利要求5所述的基于无锚框两阶段网络检测发票印章的装置,其特征在於:还包括截取区域特征模块和分类与回归模块,截取区域特征模块为程序模块,用于处理器通过无锚框候选区域对特征图进行截取并获得区域特征图;分类与回归模块为程序模块,用于处理器基于 $K*K*C$ 的区域特征图进行分类和回归处理。

8. 根据权利要求7所述的基于无锚框两阶段网络检测发票印章的装置,其特征在於:在截取区域特征模块中,基于任一候选框沿特征图的高度和宽度方向都平均切成 K 份,获得 $K*K$ 个方格,对每一个方格进行最大池化,获得 $K*K*C$ 的区域特征图, $K=5,C=512$;在分类与回归模块中,对于每一个区域特征图,分别经过分类分支和回归分支,每个分支都是四层 $3*3$ 的卷积层,分类分支最后输出的特征图形状为 $H*W*N$,回归分支最后输出的特征图形状为 $H*W*4$,其中 N 为待分类的类别数目, 4 为回归得到的到四个边的距离。

9. 一种基于无锚框两阶段网络检测发票印章的装置,其特征在於:包括存储器、处理器以及存储在存储器中并可在处理器上运行的权利要求5~权利要求8中的程序模块,所述处理器执行程序模块时实现权利要求1~权利要求4中任意一项基于无锚框两阶段网络检测发票印章方法的步骤。

10. 一种基于无锚框两阶段网络检测发票印章的装置,其特征在於:为计算机可读存储介质,所述计算机可读存储介质存储有权利要求5~权利要求8中的程序模块,所述程序模块被处理器执行时实现权利要求1~权利要求4中任意一项基于无锚框两阶段网络检测发票印章方法的步骤。

基于无锚框两阶段网络检测发票印章的方法及装置

技术领域

[0001] 本发明涉及票据文本检测技术领域,尤其涉及一种基于无锚框两阶段网络检测发票印章的方法及装置。

背景技术

[0002] 发票是企业费用报销中的一个重要组成部分,包含了发票名称、开票日期、开票金额、印章等多项报销所必需信息,其中对于印章的检测识别目前主要由人工对比,其缺点是人为因素多、准确性差、工作效率低、非常的耗时耗力,如果能够将深度学习技术要用到发票印章上,实现信息的自动提取,将会极大的节省人力资源成本。

[0003] 发票印章信息自动提取流程包含了候选区域生成、区域坐标调整和内容识别两个阶段。作为整个流程的基础步骤,第一个环节候选区域生成面临着更多的问题。现有的基于深度学习的方法主要分为基于锚框的方法以及基于无锚框的方法。基于锚框的方法预先在图像的特征图上生成密集固定大小和尺寸比的先验锚框,然后基于这些锚框进行后续的优化。这种方法通常是两阶段的,第一阶段通过区域生成网络将先验框调整生成候选框,第二阶段对这些候选框内的特征进行进一步内容分析与判断。但是在利用锚框时需要设置超参数,并且会产生大量冗余先验框,增加了问题的复杂程度。基于无锚框的方法直接对特征图上的每个像素点建模到四个边的距离以及类别,这种方法简单快速,然而准确度不如拥有第二阶段微调的两阶段方法。在发票印章的检测场景下,漏检、边界不正确会极大的影响后续其他处理。

[0004] 现有技术问题及思考:

[0005] 如何解决发票印章检测第一阶段中锚框冗余的技术问题。

发明内容

[0006] 本发明所要解决的技术问题是提供一种基于无锚框两阶段网络检测发票印章的方法及装置,其通过步骤S1至步骤S4等,实现发票印章检测第一阶段中锚框冗余度较小,提升了检测发票印章的工作效率。

[0007] 为解决上述技术问题,本发明所采取的技术方案是:一种基于无锚框两阶段网络检测发票印章的方法包括如下步骤,S1发票图片预处理,处理器从存储器获取发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片;S2提取预处理后图片特征,处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图;S3生成无锚框候选区域,处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。

[0008] 进一步的技术方案在于:步骤S1具体包括如下步骤,S101旋转处理,处理器对预处理图片做随机旋转处理,以50%的概率进行水平旋转并获得旋转图片;S102归一化处理,处

处理器将旋转图片进行归一化处理并获得归一化图片;S103统一图片,处理器将上述归一化图片进行填充并获得统一尺寸的预处理图片;在步骤S2中,所述特征图为最后得到的大小为 $C \times H \times W$ 的特征向量矩阵 F ,其中, C 为图像的通道, H 为图像的高度, W 为图像的宽度。

[0009] 进一步的技术方案在于:在步骤S3之后还包括如下步骤,S4截取区域特征,处理器通过无锚框候选区域对特征图进行截取并获得区域特征图;S5分类与回归,处理器基于 $K \times K \times C$ 的区域特征图进行分类和回归处理。

[0010] 进一步的技术方案在于:在步骤S4中,基于任一候选框沿特征图的高度和宽度方向都平均切成 K 份,获得 $K \times K$ 个方格,对每一个方格进行最大池化,获得 $K \times K \times C$ 的区域特征图, $K=5, C=512$;在步骤S5中,对于每一个区域特征图,分别经过分类分支和回归分支,每个分支都是四层 3×3 的卷积层,分类分支最后输出的特征图形状为 $H \times W \times N$,回归分支最后输出的特征图形状为 $H \times W \times 4$,其中 N 为待分类的类别数目,4为回归得到的到四个边的距离。

[0011] 一种基于无锚框两阶段网络检测发票印章的装置包括发票图片预处理模块、提取预处理后图片特征模块和生成无锚框候选区域模块,发票图片预处理模块为程序模块,用于处理器从存储器获取发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片;提取预处理后图片特征模块为程序模块,用于处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图;生成无锚框候选区域模块为程序模块,用于处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。

[0012] 进一步的技术方案在于:所述发票图片预处理模块,还用于处理器对预处理图片做随机旋转处理,以50%的概率进行水平旋转并获得旋转图片,处理器将旋转图片进行归一化处理并获得归一化图片,处理器将上述归一化图片进行填充并获得统一尺寸的预处理图片;在提取预处理后图片特征模块中,所述特征图为最后得到的大小为 $C \times H \times W$ 的特征向量矩阵 F ,其中, C 为图像的通道, H 为图像的高度, W 为图像的宽度。

[0013] 进一步的技术方案在于:还包括截取区域特征模块和分类与回归模块,截取区域特征模块为程序模块,用于处理器通过无锚框候选区域对特征图进行截取并获得区域特征图;分类与回归模块为程序模块,用于处理器基于 $K \times K \times C$ 的区域特征图进行分类和回归处理。

[0014] 进一步的技术方案在于:在截取区域特征模块中,基于任一候选框沿特征图的高度和宽度方向都平均切成 K 份,获得 $K \times K$ 个方格,对每一个方格进行最大池化,获得 $K \times K \times C$ 的区域特征图, $K=5, C=512$;在分类与回归模块中,对于每一个区域特征图,分别经过分类分支和回归分支,每个分支都是四层 3×3 的卷积层,分类分支最后输出的特征图形状为 $H \times W \times N$,回归分支最后输出的特征图形状为 $H \times W \times 4$,其中 N 为待分类的类别数目,4为回归得到的到四个边的距离。

[0015] 一种基于无锚框两阶段网络检测发票印章的装置包括存储器、处理器以及存储在存储器中并可在处理器上运行的上述程序模块,所述处理器执行程序模块时实现上述基于无锚框两阶段网络检测发票印章方法的步骤。

[0016] 一种基于无锚框两阶段网络检测发票印章的装置为计算机可读存储介质,所述计

计算机可读存储介质存储有上述程序模块,所述程序模块被处理器执行时实现上述基于无锚框两阶段网络检测发票印章方法的步骤。

[0017] 采用上述技术方案所产生的有益效果在于:

[0018] 一种基于无锚框两阶段网络检测发票印章的方法包括如下步骤,S1发票图片预处理,处理器从存储器获取发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片;S2提取预处理后图片特征,处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图;S3生成无锚框候选区域,处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。该技术方案,其通过步骤S1至步骤S4等,实现发票印章检测第一阶段中锚框冗余度较小,提升了检测发票印章的工作效率。

[0019] 一种基于无锚框两阶段网络检测发票印章的装置包括发票图片预处理模块、提取预处理后图片特征模块和生成无锚框候选区域模块,发票图片预处理模块为程序模块,用于处理器从存储器获取发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片;提取预处理后图片特征模块为程序模块,用于处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图;生成无锚框候选区域模块为程序模块,用于处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。该技术方案,其通过发票图片预处理模块、提取预处理后图片特征模块和生成无锚框候选区域模块等,实现发票印章检测第一阶段中锚框冗余度较小,提升了检测发票印章的工作效率。

[0020] 一种基于无锚框两阶段网络检测发票印章的装置包括存储器、处理器以及存储在存储器中并可在处理器上运行的上述程序模块,所述处理器执行程序模块时实现上述基于无锚框两阶段网络检测发票印章方法的步骤。该技术方案,其通过该装置,实现发票印章检测第一阶段中锚框冗余度较小,提升了检测发票印章的工作效率。

[0021] 一种基于无锚框两阶段网络检测发票印章的装置为计算机可读存储介质,所述计算机可读存储介质存储有上述程序模块,所述程序模块被处理器执行时实现上述基于无锚框两阶段网络检测发票印章方法的步骤。该技术方案,其通过该计算机可读存储介质,实现发票印章检测第一阶段中锚框冗余度较小,提升了检测发票印章的工作效率。

[0022] 详见具体实施方式部分描述。

附图说明

[0023] 图1是本发明实施例1的流程图;

[0024] 图2是本发明实施例1的数据流图。

具体实施方式

[0025] 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清楚、完

整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。以下对至少一个示例性实施例的描述实际上仅仅是说明性的,决不作为对本申请及其应用或使用的任何限制。基于本申请中的实施例,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,都属于本申请保护的范围。

[0026] 在下面的描述中阐述了很多具体细节以便于充分理解本申请,但是本申请还可以采用其他不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本申请内涵的情况下做类似推广,因此本申请不受下面公开的具体实施例的限制。

[0027] 实施例1:

[0028] 如图1所示,本发明公开了一种基于无锚框两阶段网络检测发票印章的方法包括如下步骤:

[0029] S1发票图片预处理

[0030] 通过扫描设备或者拍照设备获取发票图片并发送至处理器,处理器接收到发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片。

[0031] S101旋转处理

[0032] 处理器对预处理图片做随机旋转处理,以50%的概率进行水平旋转并获得旋转图片。

[0033] S102归一化处理

[0034] 处理器将旋转图片进行归一化处理并获得归一化图片。

[0035] S103统一图片

[0036] 处理器将上述归一化图片进行填充并获得统一尺寸的预处理图片。

[0037] S2提取预处理后图片特征

[0038] 处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图,即最后得到的大小为 $C \times H \times W$ 的特征向量矩阵 F ,其中, C 为图像的通道, H 为图像的高度, W 为图像的宽度。

[0039] S3生成无锚框候选区域

[0040] 处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。

[0041] S4截取区域特征

[0042] 处理器通过无锚框候选区域对特征图进行截取并获得区域特征图。具体为,基于任一候选框沿特征图的高度和宽度方向都平均切成 K 份,获得 $K \times K$ 个方格,对每一个方格进行最大池化,获得 $K \times K \times C$ 的区域特征图, $K=5, C=512$ 。

[0043] S5分类与回归

[0044] 处理器基于 $K \times K \times C$ 的区域特征图进行分类和回归处理。对于每一个区域特征图,分别经过分类分支和回归分支,每个分支都是四层 3×3 的卷积层,分类分支最后输出的特征图形状为 $H \times W \times N$,回归分支最后输出的特征图形状为 $H \times W \times 4$,其中 N 为待分类的类别数目,4为回归得到的到四个边的距离。

[0045] 实施例2:

[0046] 本发明公开了一种基于无锚框两阶段网络检测发票印章的装置包括发票图片预处理模块、提取预处理后图片特征模块、生成无锚框候选区域模块、截取区域特征模块和分类与回归模块,所述发票图片预处理模块包括发票图片预处理模块、旋转处理为程序模块和归一化处理模块,以上模块均为程序模块。

[0047] 发票图片预处理模块,用于通过扫描设备或者拍照设备获取发票图片并发送至处理器,处理器接收到发票图片,对发票图片图像预处理并获得统一尺寸的预处理图片。

[0048] 旋转处理为程序模块,用于处理器对预处理图片做随机旋转处理,以50%的概率进行水平旋转并获得旋转图片。

[0049] 归一化处理模块,用于处理器将旋转图片进行归一化处理并获得归一化图片。

[0050] 统一图片模块,处理器将上述归一化图片进行填充并获得统一尺寸的预处理图片。

[0051] 提取预处理后图片特征模块,处理器将预处理图片输入至特征提取卷积神经网络并获得特征图,所述特征提取卷积神经网络为基于ResNet-50卷积神经网络去掉其最后的全连接层和池化层而获得的神经网络,所述特征图为经过特征提取卷积神经网络而形成的最后一层的特征图,即最后得到的大小为 $C \times H \times W$ 的特征向量矩阵 F ,其中, C 为图像的通道, H 为图像的高度, W 为图像的宽度。

[0052] 生成无锚框候选区域模块,用于处理器对特征图分别进行类别判断分支和位置回归分支处理并生成无锚框候选区域,所述类别判断分支和位置回归分支处理为分别取两个 3×3 的窗口与特征图进行卷积。

[0053] 截取区域特征模块,用于处理器通过无锚框候选区域对特征图进行截取并获得区域特征图。具体为,基于任一候选框沿特征图的高度和宽度方向都平均切成 K 份,获得 $K \times K$ 个方格,对每一个方格进行最大池化,获得 $K \times K \times C$ 的区域特征图, $K=5, C=512$ 。

[0054] 分类与回归模块,用于处理器基于 $K \times K \times C$ 的区域特征图进行分类和回归处理。对于每一个区域特征图,分别经过分类分支和回归分支,每个分支都是四层 3×3 的卷积层,分类分支最后输出的特征图形状为 $H \times W \times N$,回归分支最后输出的特征图形状为 $H \times W \times 4$,其中 N 为待分类的类别数目,4为回归得到的到四个边的距离。

[0055] 实施例3:

[0056] 本发明公开了一种基于无锚框两阶段网络检测发票印章的装置包括存储器、处理器以及存储在存储器中并可在处理器上运行的实施例2中的程序模块,所述处理器执行程序模块时实现实施例1的步骤。

[0057] 实施例4:

[0058] 本发明公开了一种计算机可读存储介质,所述计算机可读存储介质存储有实施例2中的程序模块,所述程序模块被处理器执行时实现实施例1中的步骤。

[0059] 本申请的技术贡献:

[0060] 为了解决发票印章检测第一阶段中锚框冗余的技术问题,本发明提供了一种基于无锚框候选区域生成的两阶段检测算法,能够有效的检测发票印章的位置与内容。

[0061] 本发明的技术方案主要包括三个部分:

[0062] 第一部分是基于ResNet-50进行图片特征提取模块。

[0063] 第二部分采用无锚框候选区域生成网络。

[0064] 第三部分为常规的第二阶段检测分支,用来进一步调整和识别候选区域的内容。在第一部分中,我们使用ResNet-50作为骨干网络,并去掉了最后的池化层和全连接层,得到输入图片的空间特征。第二部分中,我们将骨干网络提取的特征输入到一个无锚框候选区域生成网络,该部分会对每一个像素点进行判断是否可能含有印章,并且直接回归出一个候选框,然后将那些得分较高的候选框输出给第二阶段进行调整;其中为了衡量不同像素点的重要程度,我们加入了中心损失函数的优化目标,主要为了让在印章中心的像素点能够得到较高的相应结果。在第三部分中,利用产生的候选区域得到对应的特征并且输入到后续的卷积层,最终输入每个区域的类别以及回归调整后的区域坐标作为最后的检测结果。

[0065] 如图1所示,该发票印章检测方法包含以下主要步骤:

[0066] S1发票图片预处理

[0067] 利用扫描设备或者拍照设备将单张发票的图片上传系统并进行图像预处理。由于发票数据量有限,为了让模型见到更多更丰富的数据,使用以下图像预处理和图像增强方法。第一步,对上传的图片做随机旋转处理,以50%的概率进行水平旋转;第二步,为了方便后续神经网络的更好收敛,将所有图像数据进行归一化处理,得到归一化图像;第三步,将上述结果进行填充到指定尺寸,得到固定尺寸的图片,输入神经网络进行后续处理。

[0068] S2提取预处理后图片特征

[0069] 处理后的发票图片先经过ResNet-50卷积神经网络做特征提取。在ResNet-50中我们去掉了最后的全连接层与池化层,只使用前五个stage,其中第二至第五个stage的输出特征图大小依次为输入图片的1/4、1/8、1/16、1/32,不同于通常利用多尺度的方法,往往印章的大小是固定的,因此本方法选择只使用最后一层特征图,即最后得到大小C×H×W的特征向量矩阵F,C、H、W分别表示图像的通道、高度和宽度。之后在该特征图上进行无锚框候选区域生成。

[0070] S3生成无锚框候选区域

[0071] 对得到的特征图进行无锚框候选区域生成。我们分别使用类别判断分支以及位置回归分支对上述特征图进行处理:分别取两个不同的3×3的窗口与特征图进行卷积,即每个点与周围3×3区域做特征提取获得长度为1和长度为4的特征向量,前者表示当前像素点中可能含有印章的概率大小P,后者表示当前像素点所产生的候选框的编码,只有当P大于给定的阈值时,才会将该像素点列入候选项。最后便可以得到大小为形状为(N,L,T,R,B)的候选区域集合,其中N表示候选区域个数,剩下分别表示当前中心像素点到候选框左边界、上边界、右边界以及下边界的距离。除此之外,为了衡量一个候选区域内部不同像素点的重要性,我们使用了中心损失函数来让处于目标中心的点得到更高的响应。该损失函数定义具体如下公式:

$$[0072] \quad \text{centerness} = \sqrt{\frac{\min(L,R)}{\max(L,R)} * \frac{\min(T,B)}{\max(T,B)}} \quad \text{式 1}$$

[0073] S4截取区域特征

[0074] 接下来是利用步骤S3中生成的候选区域框对步骤S2中提取到特征图进行截取。具体来说,对任一候选框,即使该候选框形状不同,沿其高度和宽度方向都平均切成K份,则会

得到 $K \times K$ 个方格,然后每个方格取最大池化可得最后 $K \times K \times C$ 的特征图。

[0075] S5分类与回归

[0076] 最后便是基于上述的 $K \times K \times C$ 的特征图进行第二阶段分类与回归。对于每一个特征图,分别经过分类分支以及回归分支,每个分支都是四层 3×3 卷积,分类分支和回归分支最后输出的特征图形状分别为 $H \times W \times N$ 以及 $H \times W \times 4$,其中 N 为待分类的类别数目,4为回归得到的到四个边的距离。

[0077] 技术方案说明:

[0078] S1发票图片预处理

[0079] 利用扫描设备或者拍照设备将单张发票的图片上传系统并进行图像预处理。本案例采用图像预处理和图像增强方法。第一步,对上传的图片做随机旋转处理,以50%的概率进行水平旋转;第二步,为了方便后续神经网络的更好收敛,将所有图像数据进行归一化处理,得到归一化图像;第三步,将上述结果进行填充到指定尺寸,得到固定尺寸的图片,案例采用固定尺寸为 800×640 ,输入神经网络进行后续处理。

[0080] S2提取预处理后图片特征

[0081] 处理后的发票图片先经过ResNet-50卷积神经网络做特征提取。在ResNet-50中我们去掉了最后的全连接层与池化层,只使用前五个stage,其中第二至第五个stage的输出特征图大小依次为输入图片的 $1/4$ 、 $1/8$ 、 $1/16$ 、 $1/32$,不同于通常利用多尺度的方法,往往印章的大小是固定的,因此本方法选择只使用最后一层特征图,即最后得到大小 $512 \times 20 \times 25$ 像素的特征向量矩阵 F ,512、20、25分别表示图像的通道、高度和宽度。之后在该特征图上进行无锚框候选区域生成。

[0082] S3生成无锚框候选区域

[0083] 对得到的特征图进行无锚框候选区域生成。我们分别使用类别判断分支以及位置回归分支对上述特征图进行处理:分别取两个不同的 3×3 的窗口与特征图进行卷积,即每个点与周围 3×3 区域做特征提取获得长度为1和长度为4的特征向量,前者表示当前像素点中可能含有印章的概率大小 P ,后者表示当前像素点所产生的候选框的编码,只有当 P 大于给定的阈值时,才会将该像素点列入候选项,在案例中通常设定阈值为0.95,即 P 大于0.95才认为此处有印章。最后便可以得到大小为形状为 (N, L, T, R, B) 的候选区域集合,其中 N 表示候选区域个数,剩下分别表示当前中心像素点到候选框左边界、上边界、右边界以及下边界的距离。除此之外,为了衡量一个候选区域内部不同像素点的重要性,我们使用了中心损失函数来让处于目标中心的点得到更高的响应。

[0084] S4截取区域特征

[0085] 接下来是利用上一步生成的候选区域框在特征向量 F 进行截取。具体来说,对任一候选框(即使形状不同),沿其高度和宽度方向都平均切成5份,则会得到 5×5 个方格,然后每个方格取最大池化可得最后 $5 \times 5 \times 512$ 的特征向量 G ,由于 F 在案例中是 $512 \times 20 \times 25$ 像素,宽高都平均切分为 5×5 ,即宽度每四个像素点,高度每5个像素点,共20个像素点,选择其中最大值作为结果,因此 F 会变成 $G_{5 \times 5 \times 512}$ 像素特征向量矩阵。

[0086] S5分类与回归

[0087] 如图2所示,最后便是基于 $G_{5 \times 5 \times 512}$ 像素特征向量矩阵进行第二阶段分类与回归。对于每一个特征图,分别经过分类分支以及回归分支,每个分支都是四层 3×3 卷积,分类

分支和回归分支最后输出的特征图形状分别为 $5 \times 5 \times 2$ 以及 $5 \times 5 \times 4$,其中2为待分类的类别数目,此处仅有是印章和不是印章两种分类可能,4为回归得到的到四个边的距离。在得到分类和回归结果之后,将分类为是印章的区域中心点以及回归得到的距离四个边界的距离既可以得出印章检测的结果。

[0088] 本申请保密运行一段时间后,现场技术人员反馈的有益之处在于:

[0089] 整个系统采用ResNet-50进行特征抽取,然后分为两个阶段:第一阶段利用无锚框方式预测印章的候选区域及背景信息,第二阶段将上述候选区域进行进一步的分类与回归得到最后的印章检测结果。

[0090] 该方法主要针对发票中的印章进行检测,通过将第一阶段的候选区域生成的方式由基于锚框改为基于无锚框,降低了模型的复杂程度,有利于更好更快地实现对发票印章的准确检测,能够有效的解决发票中印章检测问题。

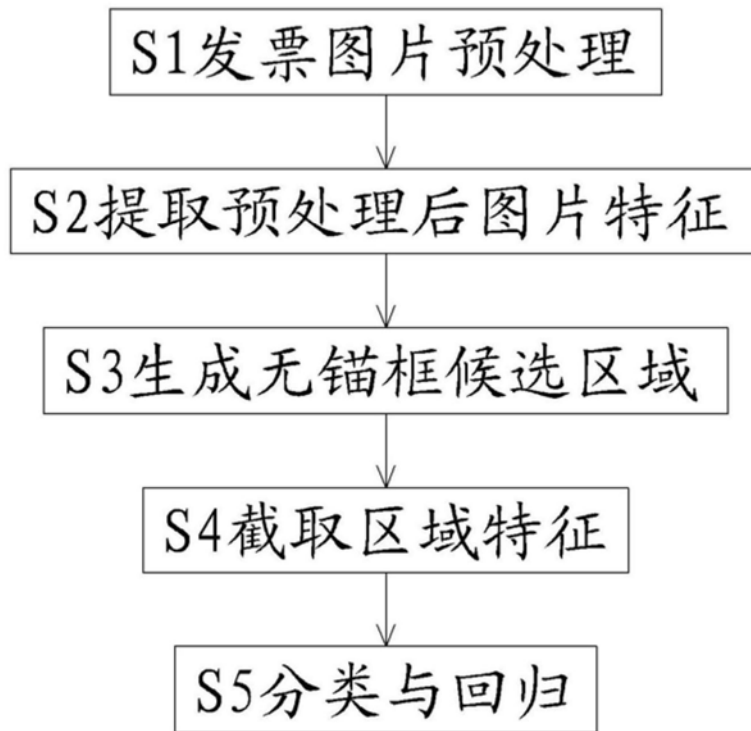


图1

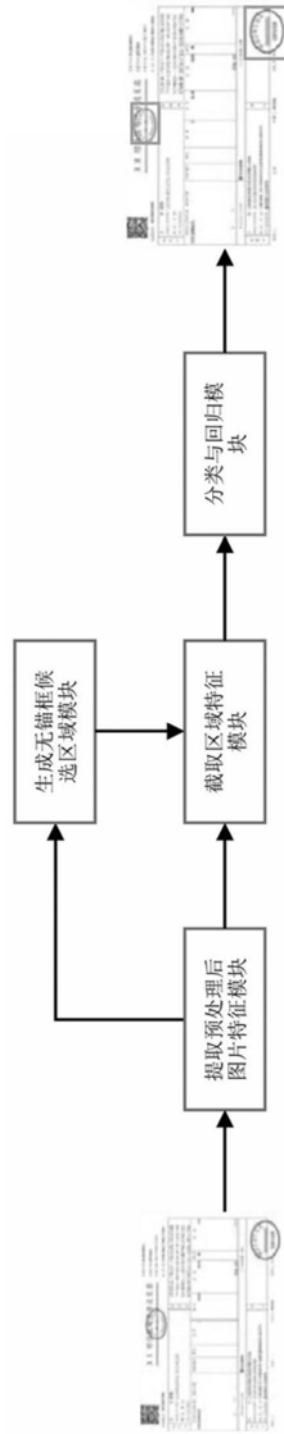


图2