



(12)发明专利申请

(10)申请公布号 CN 107545223 A

(43)申请公布日 2018.01.05

(21)申请号 201610506416.X

(22)申请日 2016.06.29

(71)申请人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市福田区振兴路
赛格科技园2栋东403室

申请人 北京大学

(72)发明人 孙广宇

(74)专利代理机构 北京派特恩知识产权代理有
限公司 11270

代理人 李梅香 张颖玲

(51)Int.Cl.

G06K 9/00(2006.01)

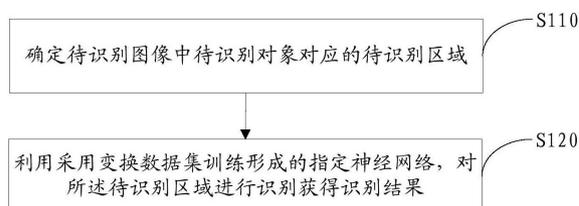
权利要求书3页 说明书16页 附图5页

(54)发明名称

图像识别方法及电子设备

(57)摘要

本发明实施例公开了一种图像识别方法及装置,所述图像识别方法包括:确定待识别图像中待识别对象对应的待识别区域;利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。



1. 一种图像识别方法,其特征在于,包括:
确定待识别图像中待识别对象对应的待识别区域;
利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;
其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。
2. 根据权利要求1所述的方法,其特征在于,
所述方法还包括:
在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。
3. 根据权利要求2所述的方法,其特征在于,
所述进行神经网络训练,获得所述指定神经网络,包括:
利用数学变换对图形对象进行所述预设变换,获得所述包含有预设变换的图形对象的训练图像;
利用所述包含有预设变换的图形对象的训练图像及预设结果,对神经网络进行训练获得所述网络参数。
4. 根据权利要求1、2或3所述的方法,其特征在于,
所述确定待识别图像中待识别对象对应的待识别区域,包括:
定位所述待识别图像的待识别对象;
确定所述待识别对象的边界参数;其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域;
所述利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果,包括:
根据所述边界参数从所述待识别图像中提取出所述待识别区域;
根据所述待识别区域中图形识别出待识别对象,获得所述识别结果。
5. 根据权利要求4所述的方法,其特征在于,
所述待识别对象包括待识别字符;
所述定位所述待识别图像的待识别对象,包括:
检测出所述待识别图像中边缘点;所述边缘点包括第一边缘点p;
确定所述p的第一梯度方向;
以所述p为起点,沿所述第一梯度方向搜索所述边缘点中的第二边缘点q;其中,所述q的梯度方向为第二梯度方向;所述第一梯度方向和所述第二梯度方向满足第一预设条件;
若存在至少一个所述q,则确定所述p和所述q属于同一待识别字符;
根据所述p和所述q的距离,确定所述待识别字符的笔画宽度;
将所有笔画宽度满足第二预设条件的边缘点合并,组成所述待识别字符。
6. 根据权利要求5所述的方法,其特征在于,
所述第一梯度方向和所述第二梯度方向满足第一预设条件,包括:
所述第一梯度方向 d_p 和所述第二梯度方向 d_q 满足如下函数关系:

$$dp = -dq \mp \theta ;$$

其中,所述 θ 为预设值。

7. 根据权利要求4所述的方法,其特征在于,

所述边界参数包括:所述待识别对象的外接边框的顶点坐标参数。

8. 一种图像识别装置,其特征在于,包括:

确定单元,用于确定待识别图像中待识别对象对应的待识别区域;

识别单元,用于利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

9. 根据权利要求8所述的装置,其特征在于,

所述装置还包括:

训练单元,用于在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。

10. 根据权利要求9所述的装置,其特征在于,

所述训练单元,具体用于利用数学变换对图形对象进行所述预设变换,获得所述包含有预设变换的图形对象的训练图像;利用所述包含有预设变换的图形对象的训练图像及预设结果,对神经网络进行训练获得所述网络参数。

11. 根据权利要求9、10或11所述的装置,其特征在于,

所述确定单元,具体用于定位所述待识别图像的待识别对象;确定所述待识别对象的边界参数;其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域;

所述识别单元,具体用于根据所述边界参数从所述待识别图像中提取出所述待识别区域;根据所述待识别区域中图形识别出待识别对象,获得所述识别结果。

12. 根据权利要求11所述的装置,其特征在于,

所述待识别对象包括待识别字符;

所述确定单元,具体用于检测出所述待识别图像中边缘点;所述边缘点包括第一边缘点 p ;确定所述 p 的第一梯度方向;

以所述 p 为起点,沿所述第一梯度方向搜索所述边缘点中的第二边缘点 q ;其中,所述 q 的梯度方向为第二梯度方向;所述第一梯度方向和所述第二梯度方向满足第一预设条件;若存在至少一个所述 q ,则确定所述 p 和所述 q 属于同一待识别字符;根据所述 p 和所述 q 的距离,确定所述待识别字符的笔画宽度;将所有笔画宽度满足第二预设条件的边缘点合并,组成所述待识别字符。

13. 根据权利要求12所述的装置,其特征在于,

所述第一梯度方向和所述第二梯度方向满足第一预设条件,包括:

所述第一梯度方向 dp 和所述第二梯度方向 dq 满足如下函数关系:

$$dp = -dq \mp \theta ;$$

其中,所述 θ 为预设值。

14.根据权利要求11所述的装置,其特征在于,

所述边界参数包括:所述待识别对象的外接边框的顶点坐标参数。

图像识别方法及电子设备

技术领域

[0001] 本发明涉及信息技术领域,尤其涉及一种图像识别方法及电子设备。

背景技术

[0002] 随着信息技术发展,出现了图像识别技术,利用图像识别技术识别出图像中的信息内容,例如,提取图像中的字符。

[0003] 在现有技术中对图像中字符的识别,可包括以下步骤:

[0004] 对图像进行预处理,提取出只包括字符的图像部分;

[0005] 对图像进行特征提取和参数估计;

[0006] 根据参数进行仿射变换或投影变换等数学变化,进行图像中字符的形态校正;

[0007] 最后识别校正之后的字符,完成对图像信息的提取。

[0008] 这种识别方法存在以下问题:

[0009] 一:对图像进行数学变换,再提取特征,可能会出现很多误差,从而导致识别精度比较低,且大量的数学变换会导致很多复杂的中间运算。

[0010] 二:这种识别技术通常都无法识别孤立的字符,能够识别的范围有限。

发明内容

[0011] 有鉴于此,本发明实施例期望提供一种图像识别方法及电子设备,能够提升识别效率或提升准确率。

[0012] 为达到上述目的,本发明的技术方案是这样实现的:

[0013] 本发明实施例第一方面提供一种图像识别方法,包括:

[0014] 确定待识别图像中待识别对象对应的待识别区域;

[0015] 利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0016] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0017] 基于上述方案,所述方法还包括:

[0018] 在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。

[0019] 基于上述方案,所述进行神经网络训练,获得所述指定神经网络,包括:

[0020] 利用数学变换对图形对象进行所述预设变换,获得所述包含有预设变换的图形对象的训练图像;

[0021] 利用所述包含有预设变换的图形对象的训练图像及预设结果,对神经网络进行训练获得所述网络参数。

[0022] 基于上述方案,所述确定待识别图像中待识别对象对应的待识别区域,包括:

- [0023] 定位所述待识别图像的待识别对象；
- [0024] 确定所述待识别对象的边界参数；其中，所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域；
- [0025] 所述利用采用变换数据集训练形成的指定神经网络，对所述待识别区域进行识别获得识别结果，包括：
- [0026] 根据所述边界参数从所述待识别图像中提取出所述待识别区域；
- [0027] 根据所述待识别区域中图形识别出待识别对象，获得所述识别结果。
- [0028] 基于上述方案，所述待识别对象包括待识别字符；
- [0029] 所述定位所述待识别图像的待识别对象，包括：
- [0030] 检测出所述待识别图像中边缘点；所述边缘点包括第一边缘点p；
- [0031] 确定所述p的第一梯度方向；
- [0032] 以所述p为起点，沿所述第一梯度方向搜索所述边缘点中的第二边缘点q；其中，所述q的梯度方向为第二梯度方向；所述第一梯度方向和所述第二梯度方向满足第一预设条件；
- [0033] 若存在至少一个所述q，则确定所述p和所述q属于同一待识别字符；
- [0034] 根据所述p和所述q的距离，确定所述待识别字符的笔画宽度；
- [0035] 将所有笔画宽度满足第二预设条件的边缘点合并，组成所述待识别字符。
- [0036] 基于上述方案，所述第一梯度方向和所述第二梯度方向满足第一预设条件，包括：
- [0037] 所述第一梯度方向 dp 和所述第二梯度方向 dq 满足如下函数关系：
- [0038] $dp = -dq \mp \theta$ ；
- [0039] 其中，所述 θ 为预设值。
- [0040] 基于上述方案，所述边界参数包括：所述待识别对象的外接边框的顶点坐标参数。
- [0041] 本发明实施例第二方面提供一种图像识别装置，包括：
- [0042] 确定单元，用于确定待识别图像中待识别对象对应的待识别区域；
- [0043] 识别单元，用于利用采用变换数据集训练形成的指定神经网络，对所述待识别区域进行识别获得识别结果；
- [0044] 其中，所述变换数据集包括：包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。
- [0045] 基于上述方案，所述装置还包括：
- [0046] 训练单元，用于在对所述待识别区域进行识别获得识别结果之前，利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果，进行神经网络训练，获得所述指定神经网络。
- [0047] 基于上述方案，所述训练单元，具体用于利用数学变换对图形对象进行所述预设变换，获得所述包含有预设变换的图形对象的训练图像；利用所述包含有预设变换的图形对象的训练图像及预设结果，对神经网络进行训练获得所述网络参数。
- [0048] 基于上述方案，所述确定单元，具体用于定位所述待识别图像的待识别对象；确定所述待识别对象的边界参数；其中，所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域；
- [0049] 所述识别单元，具体用于根据所述边界参数从所述待识别图像中提取出所述待识

别区域;根据所述待识别区域中图形识别出待识别对象,获得所述识别结果。

[0050] 基于上述方案,所述待识别对象包括待识别字符;

[0051] 所述确定单元,具体用于检测出所述待识别图像中边缘点;所述边缘点包括第一边缘点p;确定所述p的第一梯度方向;

[0052] 以所述p为起点,沿所述第一梯度方向搜索所述边缘点中的第二边缘点q;其中,所述q的梯度方向为第二梯度方向;所述第一梯度方向和所述第二梯度方向满足第一预设条件;若存在至少一个所述q,则确定所述p和所述q属于同一待识别字符;根据所述p和所述q的距离,确定所述待识别字符的笔画宽度;将所有笔画宽度满足第二预设条件的边缘点合并,组成所述待识别字符。

[0053] 基于上述方案,所述第一梯度方向和所述第二梯度方向满足第一预设条件,包括:

[0054] 所述第一梯度方向 d_p 和所述第二梯度方向 d_q 满足如下函数关系:

[0055] $d_p = -d_q \mp \theta$;

[0056] 其中,所述 θ 为预设值。

[0057] 基于上述方案,所述边界参数包括:所述待识别对象的外接边框的顶点坐标参数。

[0058] 本发明实施例提供的图像识别方法及电子设备,采用是利用包含有预设变换图形对象的训练图像训练得到的指定神经网络,来识别待识别对象对应的待识别区域;显然本实施例中采用的指定神经网络是直接可以识别包含有预设变换的待识别对象的,故当确定出待识别区域之后,可以直接输入到指定神经网络中进行识别,获得识别结果。在获得识别结果的过程中,无需进行数学变换,故节省了数学变换所需的时间,提升了识别效率;因为减少数学变换的步骤,减少了数学变换引入的误差,从而提升了识别精确度。

附图说明

[0059] 图1为本发明实施例提供的第一种图像识别方法的流程示意图;

[0060] 图2为本发明实施例提供的一个待识别图像的显示示意图;

[0061] 图3为本发明实施例提供的另一个待识别图像的显示示意图;

[0062] 图4为本发明实施例提供的第二种图像识别方法的流程示意图;

[0063] 图5为本发明实施例提供的一种神经网络的训练流程示意图;

[0064] 图6为本发明实施例提供的又一个待识别图像的显示示意图;

[0065] 图7a至7c为本发明实施例提供的确定待识别字符的示意图;

[0066] 图8a至图8b为本发明实施例提供的笔画宽度示意图;

[0067] 图9为本发明实施例提供的电子设备的结构示意图;

[0068] 图10为本发明实施例提供的第三种图像识别方法的流程示意图。

具体实施方式

[0069] 以下结合说明书附图及具体实施例对本发明的技术方案做进一步的详细阐述。

[0070] 实施例一:

[0071] 如图1所示,本实施例提供一种图像识别方法,包括:

[0072] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0073] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进

行识别获得识别结果；

[0074] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0075] 本实施例所述的图形识别方法可应用于包括神经网络的识别装置中。

[0076] 在步骤S110中将确定出待识别图像中待识别对象对应的待识别区域。在本实施例中所述待识别图像为包括需要进行图像识别的各种图像。所述待识别对象可包括各种类型的待识别对象,例如,待识别对象中展示的人物对象、景物对象、动物对象和/或字符等。所述待识别区域为所述待识别图像中显示所述待识别对象的图像区域。如图2表示的即为本实施例提供的一个待识别图像。图2中笔记本上书写的“A”即为本实施例所述的待识别对象;所述虚线框框住的区域即可为本实施例中待识别对象对应的待识别区域。

[0077] 在步骤S120中将利用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果。所述变换数据集为所述指定神经网络的训练集,不同一般的数据集,该变换数据集中包括的训练图像至少部分是包含有预设变换的图形对象的。在本实施例中所述指定神经网络进行识别。这样,若所述待识别区域包括预设变换的图形,则指定神经网络就能够直接识别出所述待识别对象,且具有识别结果精确的特点。所述预设变换可包括倾斜、不同视角的透视、弯曲、旋转以及图形的各种滤镜效果。例如,图形挤压效果;总之这里的预设变换可包括各种图形不同的呈现形式。所述预设变换的训练图像可为包括图形不同的呈现形式的图像。所述图像内容与识别结果对应。例如,字母“A”是一个经过预设角度旋转的字符,旋转后的字母“A”对应的图像作为所述训练图像;所述字母“A”这一图像内容即为所述预设结果。所述预设结果为所述训练图像的图像内容,与指定神经网络输出的识别结果对应。

[0078] 例如,当前获得一个神经网络B,旋转后的字母“A”及“A”这一个预设结果输入神经网络B,对神经网络B进行训练;训练后的神经网络,识别旋转后的字母“A”,输出的识别结果将为“A”。故所述训练图像对应于待识别区域;所述预设结果对应于所述识别结果。显然图2中字母“A”相对于图2所示的图像是倾斜的,若利用本实施例提供的指定神经网络就可以直接识别,无需利用数据变换进行倾斜校正处理。

[0079] 如图3所示,在待识别图像中包括弯曲排列的文字行“我爱北京”,在现有技术中在识别该文字行之前,需要通过数学变化,将该弯曲排列的文字行,校正成直线排列的文字行之后,再由学习机等具有识别能力的装置进行识别。显然这样会引入因数学变化自身的误差、数学变化识别所消耗的时间和系统资源,这样就会导致识别结果精确度低、识别效率低等问题。

[0080] 若利用本实施例提供的图像识别方法,由于所述神经网络本身在训练时,就可能采用弯曲排列文字行作为训练集进行训练的,故可以直接将弯曲排列的文字行作为所述指定神经网络的输入,对弯曲排列的文字行进行识别,指定神经网络对倾斜的文字识别后,输出识别结果。显然这中间省略了对待识别对象的数学变化处理,减少了数学变换引入的误差,同时节省了数学变化所需的时间,故相对现有技术提升了识别精确度及识别效率。

[0081] 实施例二:

[0082] 如图1所示,本实施例提供一种图像识别方法,包括:

[0083] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0084] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0085] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0086] 如图4所示,所述方法还包括:

[0087] 步骤S101:在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。

[0088] 在本实施例中,在执行所述步骤S120之前,需要对神经网络进行训练得到所述指定神经网络。

[0089] 在本实施例中所述指定神经网络的训练可包括:

[0090] 确定训练图像和预设结果;

[0091] 将所述训练图像和所述预设结果,输入网络参数未定的神经网络,得到该神经网络的网络参数;

[0092] 利用校验集,校验训练得到的网络参数获得校验结果;

[0093] 若校验结果表明该神经网络的识别精确度满足预设要求,则可以停止该神经网络的训练;

[0094] 若校验结果表明该神经网络的识别精确度不满足所述预设要求,则调整训练集和神经网络的结构参数的至少其中之一,继续训练直至训练得到的神经网络的识别精确度满足所述预设要求。

[0095] 所述神经网络的参数包括:神经网络的层数,每一层的节点个数;所述网络参数可包括每个节点的运算符和/或权值等至少其中之一。

[0096] 在本实施例中所述校验集也要求包括:包含有所述预设变换的校验图像和所述预设变换的校验图像对应的校验结果。所述校验图像与所述训练图像和待识别区域对应;所述校验结果与所述预设结果及识别结果对应。

[0097] 在本实施例中所述预设变换的训练图像可为:从网络等各种图像平台搜集的包含有所述预设变换的图像和人工输入的预设结果等;或从训练数据库中抽取的包含有所述预设变化的图像和对应预设结果的数据集。

[0098] 总之,本实施例在前一实施例的基础上,增设了利用包含有预设变换的训练图像和预设结果,训练神经网络获得所述指定神经网络的步骤。

[0099] 实施例三:

[0100] 如图1所示,本实施例提供一种图像识别方法,包括:

[0101] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0102] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0103] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0104] 如图5所示,所述方法还包括:

[0105] 步骤S101:在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变

换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。

[0106] 所述步骤S101中,进行神经网络训练,获得所述指定神经网络,包括:

[0107] 利用数学变换对图形对象进行所述预设变换,获得所述包含有预设变换的图形对象的训练图像;

[0108] 利用所述包含有预设变换的图形对象的训练图像及预设结果,对神经网络进行训练获得所述网络参数。

[0109] 所述利用数学变换对图形对象进行所述预设变换,可包括对未包含有预设变换的图形对象的图形进行预设变换。所述未包含有预设变换的图形对象的图像可包括:仅包括正视视角的图像、仅包含有非透视图形对象的图像、未包括倾斜图形对象的图像、未包含有旋转图形对象的图像。所述预设结果为这些训练图像对应的图像内容,与识别结果对应。

[0110] 在本实施例中,利用数学变换将这些未包含有预设变换图形对象的图像进行所述预设变化,可例如,将正视视角的图像中的正视视角图形对象,通过3维建模等变换成包括有侧视视角的图形对象、顶视视角的图形对象或顶视视角的图形对象的图像,这样就简便的形成了所述包含有预设变换图形对象的训练图像。再比如,将非倾斜的图形对象,转换成倾斜图形对象;将未旋转的图形对象旋转得到旋转不同角度后的图形对象。在本实施例中在进行神经网络的训练时,利用数学变换获得包含有预设变换图形对象的训练图像,一方面避免了需要费力搜寻包含有预设变换图形对象的训练图像导致的训练难度大的问题,另一方面,利用数据变换得到的训练图像进行神经网络训练,再通过校验集的校验,可以保证神经网络的识别精确度,从而能够使数学变化的误差通过神经网络的处理,减少数学变换对精确度的不利影响。

[0111] 当然值得注意的是:在进行所述指定神经网络的训练过程中,作为训练图像可为一开始就包含有预设变换的图形对象的原始图像对象,不局限于本实施例中利用数学变换形成的包含有预设变换的图形对象的训练图像。

[0112] 图6所示的为本实施例提供的训练神经网络的示意图。

[0113] 在图6中获得了包括字符为变换的“中国人”的图像,通过数学变换处理,使得形成了一个包括变换的“中国人”的图片。显然从图5中该图片的字符发生了扭曲。将发生了扭曲的字符所在的图片输入到待训练的神经网络,通过训练之后将得到上述指定神经网络。在图5显示的神经网络有n层节点,显然第1层节点和第n层节点都仅有一个节点,第2层节点包括3个节点。第3层节点包括5个节点,第n-1层节点包括3个节点。这些两层之间的节点之间的运算符号和权重等网络参数,可以通过训练确定。当然一个结构的神经网络经过训练之后无法达到较高的识别准确率,可以通过改变神经网络的结构,再次经过如图5所示的训练就能够获得所述指定神经网络。改变神经网络的结构,可以包括改变神经网络的层数和每一层的节点数等。通常所述n为不小于3的整数。

[0114] 实施例四:

[0115] 如图1所示,本实施例提供一种图像识别方法,包括:

[0116] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0117] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0118] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0119] 所述待识别对象包括待识别字符。在本实施例中所述待识别对象包括待识别字符。这里的字符可各种语言的字符,例如,英语中的英文字母,还包括中文中的汉字,包括韩文中的韩文字符等各种类型的字符。

[0120] 所述步骤S110可包括:

[0121] 定位所述待识别图像的待识别对象;

[0122] 确定所述待识别对象的边界参数;

[0123] 其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。

[0124] 在本实施例中定位所述待识别对象,可包括:图形对象的外轮廓的处理,初步确定待识别对象的大致位置。例如,所述待识别对象为待识别字符,则定位待识别字符可包括:提取出待识别图像中的字符区域;这里的字符区域相当于显示有字符的区域。确定字符的边界参数;这里的边界参数为显示每一个字符的边界的位置参数或坐标参数,这样就可以实现各个待识别字符之间的分割。这里的位置参数或坐标参数可包括所述待识别图像像素的坐标参数或位置参数。在本实施例中所述边界参数可为待识别对象的外轮廓的参数。

[0125] 例如,如图6所示中的字符串“Vision”在本实施例中步骤S110中首先,定位出待识别图像中显示该字符串显示的字符区域;然后分割每一个字符,得到每一个字符的边界参数。例如,字母“V”的边界参数,可以用于定位所述字母“V”,这样指定神经网络在识别的过程中,就不会将字母“V”和字母“i”视为一个字母来识别,这样一方面能够实现单字母的识别,另一方面通过确定边界参数,确定待识别区域,不再采用滤波等去噪处理方法对字符区域进行去噪处理,可以减少去噪处理损失字符的特征导致的降低识别精确度的现象。

[0126] 在本实施例中所述边界参数和所述待识别图像将作为所述指定神经网络的共同输入,所述指定神经网络就可以根据所述边界参数直接提取出待识别对象对应的图像区域,可以不用利用去噪处理,去除待识别图像中待识别对象以外的背景信息等干扰,从而可以避免去噪处理去除了待识别对象的部分特征导致的识别精确度降低现象。

[0127] 例如,在现有技术中,当提取出字符之后,会对待识别图像进行去噪处理,在去噪处理的过程中就可能将字母“i”的点视为噪声去掉,从而导致字母“i”被识别为数字“1”等其他信息,而本实施例中通过边界参数的提取,不再进行对待识别图像进行去噪处理,直接将边界参数作为神经网络的一个输出参数,方便神经网络从待识别图像中提取待识别字符对应的待识别区域,一方面可以避免去噪导致的识别精确度降低的现象,另一方面减少因去噪处理等操作导致的识别效率低等现象。

[0128] 实施例五:

[0129] 如图1所示,本实施例提供一种图像识别方法,包括:

[0130] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0131] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0132] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

- [0133] 所述步骤S110可包括：
- [0134] 定位所述待识别图像的待识别对象；
- [0135] 确定所述待识别对象的边界参数；
- [0136] 其中，所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。
- [0137] 所述步骤S120可包括：
- [0138] 根据所述边界参数从所述待识别图像中提取出所述待识别区域；
- [0139] 根据所述待识别区域中图形识别出待识别对象，获得所述识别结果。
- [0140] 所述待识别对象可包括待识别字符。在本实施例中所述待识别对象包括待识别字符。这里的字符可各种语言的字符，例如，英语中的英文字母，还包括中文中的汉字，包括韩文中的韩文字符等各种类型的字符。但是值得注意的所述待识别对象不局限于待识别字符，还可是待识别的图标、待识别的建筑物或待识别的人物等。
- [0141] 在本实施例中所述边界参数为所述指定神经网络，会根据边界参数将所述待识别区域从待识别图像中直接提取出来，不再进行噪音过滤等处理，一方面减少去噪处理导致的待识别区域的图形特征的丢失，另一方面能够减少运算量，提升识别速率。
- [0142] 实施例六：
- [0143] 如图1所示，本实施例提供一种图像识别方法，包括：
- [0144] 步骤S110：确定待识别图像中待识别对象对应的待识别区域；
- [0145] 步骤S120：利用采用变换数据集训练形成的指定神经网络，对所述待识别区域进行识别获得识别结果；
- [0146] 其中，所述变换数据集包括：包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。
- [0147] 所述步骤S110可包括：
- [0148] 定位所述待识别图像的待识别对象；
- [0149] 确定所述待识别对象的边界参数；
- [0150] 其中，所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。
- [0151] 作为本实施例的进一步改进，所述待识别对象包括待识别字符，则所述定位所述待识别图像的待识别对象，包括：
- [0152] 检测出所述待识别图像中边缘点；所述边缘点包括第一边缘点p；
- [0153] 确定所述p的第一梯度方向；
- [0154] 以所述p为起点，沿所述第一梯度方向搜索所述边缘点中的第二边缘点q；其中，所述q的梯度方向为第二梯度方向；所述第一梯度方向和所述第二梯度方向满足第一预设条件；
- [0155] 若存在至少一个所述q，则确定所述p和所述q属于同一待识别字符；
- [0156] 根据所述p和所述q的距离，确定所述待识别字符的笔画宽度；
- [0157] 将所有笔画宽度满足第二预设条件的边缘点合并，组成所述待识别字符。
- [0158] 在本实施例中所述第一边缘点和所述第二边缘点都为求解出的边缘点的一个。在本实施例中检测边缘点，可以采用现有技术各种边缘检测方法来进行检测。例如，可以利

用canny边缘检测方法检测所述边缘点。

[0159] 例如,可以利用如下函数关系检测边缘点:

[0160] `cvCanny(grayImage,edgeImage,threshold_low,threshold_high,M)`;其中,所述`cvCanny`表示的canny边缘检测的检测算法。所述`grayImage`为所述待识别图像对应的灰度图;所述`edgeImage`为所述待识别图像对应的边缘图像;所述`threshold_low`为检测边缘点的低阈值;所述`threshold_high`为检测边缘点的高阈值。所述`M`为利用`cvCanny`算法中用于求取边缘的卷积矩阵的矩阵的规模。

[0161] 如图7a所示的图像,利用所述canny边缘检测方法,检测例如图7b所示的边缘点。

[0162] 接下来利用预设函数关系,求解出每一个边缘点的梯度方向。

[0163] 例如,`cvSobel(gaussianImage,gradientX,1,0,CV_SCHARR)`;其中,所述`cvSobel`为梯度求解函数;所述`gaussianImage`为所述待识别图像的高斯图像;所述`gradientX`表示的为求解的X轴方向的梯度值;函数中的“1,0”为卷积方向;所述`CV_SCHARR`为卷积矩阵。再例如,`cvSobel(gaussianImage,gradientY,0,1,CV_SCHARR)`;其中,所述`cvSobel`为梯度求解函数;所述`gaussianImage`为所述待识别图像的高斯图像;所述`gradientY`表示的为求解的Y轴方向的梯度值;函数中的“0,1”为卷积方向;所述`CV_SCHARR`为卷积矩阵。在本实施例中所述X轴方向和Y轴方向为所述待识别图像中相互垂直的两个方向;X轴方向的梯度值和Y轴方向的梯度值,为同一个边缘点在不同方向上的梯度值。

[0164] 所述第一边缘点p的第一梯度方向可利用如下函数关系计算;

[0165] $dp = \sqrt{\text{gradientX}[p]^2 + \text{gradientY}[p]^2}$,其中,`gradientX[p]`表示p的X轴方向上的梯度值;所述`gradientY[p]`表示的p的Y轴方向上的梯度值。所述`sqrt`表示开放。所述`dp`表示的p的第一梯度方向。

[0166] 如图7b所示,所述p的第一梯度方向大致垂直字符的笔画延伸的方向。在图7b中利用箭头表示所述p的第一梯度反向,字母W表示的为确定的笔画宽度。

[0167] 接下来以p点为起点,沿所述第一梯度方向搜索另一个边缘点。若搜索到另一个边缘点,该边缘点可称为不同于所述p的第二边缘点q。在本实施例中,接下来判断q的第二梯度方向,是否与p的第一梯度方向满足第一预设条件。这里的第二梯度方向的计算与所述第一梯度方向的计算相同,在此就不重复了。

[0168] 若所述第一梯度方向和第二梯度反向满足所述第一预设条件,就可认为是所述p和q属于同一个字符。通常位于两个字符中的两个边缘点,一方面由于梯度方向之间不满足所述第一预设条件被过滤掉了,且由于q是以所述p为起点,沿第一梯度方向搜索的边缘点,再次减少了不同字符的边缘点被归属到一个字符中的记录。在具体实现的过程中,若未发现所述p对应的第二边缘点q,则放弃沿第一梯度方向搜索边缘点。图7c所示的通过逐像素搜索,搜索到了一个p对应的q;显然p和q之间的距离大致等于笔画宽度。在本实施例中,并将根据p和q的距离确定所述待识别字符的笔画宽度。在本实施例具体可以利用所述p和q之间的欧式距离作为所述待识别字符的笔画宽度。

[0169] 图8a为字符中为转弯处的笔画宽度 w_1 ,图8b为字符中转弯处的笔画宽度 w_2 ,显然通过图8a和图8b比对可知, w_2 大于 w_1 。通常一个字符的笔画宽度是一致,但是可能出现如图8a和8b所示的情况,在笔画转弯的地方可能会出现字符宽度比笔画宽度的现象。为了处理这种现象,可以实现确定一个笔画宽度浮动值,该浮动值可为预设装置内的预设值,可以

根据笔画宽度与浮动值之间的关系,可以确定出界定宽度。这里的界定宽度用于确定待识别图像中两个边缘点是否属于同一字符。这里的界定宽度可为所述笔画宽度加上或减去所述浮动值。在本实施例中一个边缘点,对应于一个像素。若一个像素对应的笔画宽度都未超过所述界定宽度,则将这些相邻像素或相距在指定范围值内的像素进行合并,得到所述待识别字符。

[0170] 在本实施例中的边缘点合并可以仅包括边缘点,也可以是边缘点对应的像素的合并及包括边缘点围成的轮廓内的像素的合并。

[0171] 所述笔画宽度满足所述第二预设条件还可包括:

[0172] 根据所述笔画宽度动态确定出字符中转弯或拐角处的笔画宽度,得到所述界定宽度;在进行边缘点合并时,就笔画宽度小于所述节点界定宽度的边缘点进行合并,得到所述待识别字符。若进行边缘点的合并,得到的所述待识别字符的轮廓;在进行合并时,还合并所述边缘点形成的轮廓内的像素,则确定的识别字符不仅包括轮廓和轮廓内的填充。

[0173] 采用本实施例所述的图像识别方法,能够快速地进行字符的定位和分割,以方便后续指定神经网络快速识别出待识别字符,且具有识别精确度高及识别效率高的特点。

[0174] 实施例七:

[0175] 如图1所示,本实施例提供一种图像识别方法,包括:

[0176] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0177] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0178] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0179] 所述步骤S110可包括:

[0180] 定位所述待识别图像的待识别对象;

[0181] 确定所述待识别对象的边界参数;

[0182] 其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。

[0183] 作为本实施例的进一步改进,所述待识别对象包括待识别字符,则所述定位所述待识别图像的待识别对象,包括:

[0184] 检测出所述待识别图像中边缘点;所述边缘点包括第一边缘点p;

[0185] 确定所述p的第一梯度方向;

[0186] 以所述p为起点,沿所述第一梯度方向搜索所述边缘点中的第二边缘点q;其中,所述q的梯度方向为第二梯度方向;所述第一梯度方向和所述第二梯度方向满足第一预设条件;

[0187] 若存在至少一个所述q,则确定所述p和所述q属于同一待识别字符;

[0188] 根据所述p和所述q的距离,确定所述待识别字符的笔画宽度;

[0189] 将所有笔画宽度满足第二预设条件的边缘点合并,组成所述待识别字符。

[0190] 所述第一梯度方向和所述第二梯度方向满足第一预设条件,包括:

[0191] 所述第一梯度方向 dp 和所述第二梯度方向 dq 满足如下函数关系:

[0192] $dp = -dq \mp \theta$;

[0193] 其中,所述 θ 为预设值。

[0194] 在本实施例所述 θ 可为30度、40度或45度等值,该 θ 可为经验值或仿真值。一个笔画两边对应位置的边缘点将满足上述函数关系,这样就可以根据该函数关系简便的确定出所述p是否有对应的q,若有对应的q则执行笔画宽度确定及识别字符的定位等处理。若p没有对应的q则进入下一个边缘点的处理,直至将所有边缘点或所有边缘点的1/2的对应的q的搜索。

[0195] 值得注意的是:本实施例的dp即可为前一实施例中的dp;本实施例中的dq即可为前一实施例中的dq。

[0196] 实施例八:

[0197] 如图1所示,本实施例提供一种图像识别方法,包括:

[0198] 步骤S110:确定待识别图像中待识别对象对应的待识别区域;

[0199] 步骤S120:利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0200] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0201] 所述步骤S110可包括:

[0202] 定位所述待识别图像的待识别对象;

[0203] 确定所述待识别对象的边界参数;

[0204] 其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。

[0205] 所述边界参数包括:所述待识别对象的外接边框的顶点坐标参数。

[0206] 本实施例所述的待识别对象可包括待识别字符。

[0207] 如图8所示,在本实施例中所述边界参数可为所述待识别对象的外接边框的顶点坐标参数。在具体的实现过程中,所述边界参数还可包括所述待识别对象的外接圆的圆心和半径参数等。当然,所述边界参数还可直接包括边缘点的坐标参数。但是在本实施例中一方面减少输入到所述指定神经网络的参数值,且外接边框的顶点坐标参数足以定位所述待识别对象所在图像区域;且采用外接边框的顶点参数相对于圆心和半径参数,更加直观;指定神经网络可以在不做任何计算的情况下,直接根据所述外接边框的顶点坐标参数,定位并提取出待识别对象所在的图像区域,即提取出所述待识别区域,减少运算量,加速识别效率。

[0208] 实施例九:

[0209] 如图9所示,本实施例提供一种图像识别装置,包括:

[0210] 确定单元110,用于确定待识别图像中待识别对象对应的待识别区域;

[0211] 识别单元120,用于利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0212] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0213] 本实施例中所述图像识别装置可为应用于各种电子设备中的信息处理装置。所述确定单元110可对应处理器或处理电路。所述处理器可包括中央处理器、微处理器、数字信

号处理器、应用处理器或可编程阵列等处理结构。所述处理电路可包括专用集成电路等。所述处理器或处理电路通过信息处理定位出所述待识别区域。

[0214] 所述识别单元120同样可对应于处理器或处理电路,但是该处理器或处理电路可为运行神经网络的处理器或处理电路。该处理器或处理电路的类型可以与所述确定单元110对应的处理器或处理电路类似。且所述确定单元110和识别单元120可对应于相同的处理器或处理电路,也可以对应不同的处理器或处理电路。

[0215] 所述处理器或处理电路,可以通过执行预定代码,实现所述确定单元110和所述识别单元120的功能。

[0216] 本实施例中提供的图像识别装置,对图像进行识别时,采用的利用包括预设变换训练图像及预设结果的训练集得到的指定神经网络,该指定神经网络能够识别出图像中各种变换之后的图形对象,故本实施例提供的图像识别装置,在确定出待识别区域之后,将直接利用指定神经网络进行待识别区域中待识别对象的识别,不管当前待识别区域中的识别对象是否出现倾斜、弯曲或透视等各种变换。本实施例提供的图像识别装置,同样具有识别效率高及识别精确度高的特点。

[0217] 实施例十:

[0218] 如图9所示,本实施例提供一种图像识别装置,包括:

[0219] 确定单元110,用于确定待识别图像中待识别对象对应的待识别区域;

[0220] 识别单元120,用于利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0221] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0222] 所述装置还包括:

[0223] 训练单元,用于在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。

[0224] 在本实施例中所述训练单元同样可对应于前述的处理器或处理电路。所述训练单元将训练集输入待训练的神经网络,经过训练之后可以得到本实施例提供的指定神经网络。值得注意的是本实施例提供的图像识别装置的训练单元,在训练神经网络时,采用的是包含有预设变换的训练图像和与这些训练图像对应的预设结果,显然这样训练出的指定神经网络是能够直接用于识别包含有预设变换的待识别对象,就减少利用数学变换对待识别区域的变换处理,减少数学变换所需的时间,故提升了识别效率,减少数学变化引入的误差,故能够提升识别精确度。

[0225] 在本实施例中所述训练单元,训练神经网络,可包括确定神经网络的结构参数和网络参数。这里结构参数可包括神经网络的层数、每一层的节点数;所述网络参数可包括每一个节点的符号和权值等参数,当所述结构参数和网络参数确定之后,则得到了能够识别出预设变换的待识别对象的指定神经网络。

[0226] 实施例十一:

[0227] 如图9所示,本实施例提供一种图像识别装置,包括:

[0228] 确定单元110,用于确定待识别图像中待识别对象对应的待识别区域;

[0229] 识别单元120,用于利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0230] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0231] 所述装置还包括:

[0232] 训练单元,用于在对所述待识别区域进行识别获得识别结果之前,利用包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果,进行神经网络训练,获得所述指定神经网络。

[0233] 所述训练单元,具体用于利用数学变换对图形对象进行所述预设变换,获得所述包含有预设变换的图形对象的训练图像;利用所述包含有预设变换的图形对象的训练图像及预设结果,对神经网络进行训练获得所述网络参数。

[0234] 在本实施例中所述训练单元在训练神经网络时,为了降低到处寻找包括预设变换的训练图像,在本实施例中可以利用数学变换直接对为变换的训练图像进行预设变换,从而获得包含有预设变换的图像。当然值得注意的是,在具体的实现过程中,训练单元可以直接利用原本就包含有预设变换的图形对象的原始图像作为训练图像进行神经网络的训练。

[0235] 实施例十二:

[0236] 如图9所示,本实施例提供一种图像识别装置,包括:

[0237] 确定单元110,用于确定待识别图像中待识别对象对应的待识别区域;

[0238] 识别单元120,用于利用采用变换数据集训练形成的指定神经网络,对所述待识别区域进行识别获得识别结果;

[0239] 其中,所述变换数据集包括:包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0240] 所述确定单元110,具体用于定位所述待识别图像的待识别对象;确定所述待识别对象的边界参数;

[0241] 其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。

[0242] 在本实施例中所述待识别对象包括待识别字符,则训练形成所述指定神经网络的变换数据集至少包括:包含有预设变换的文字图形的训练图像和预设结果。

[0243] 在本实施例中所述确定单元110,将通过图像处理技术,提取出待识别字符的待识别区域中的待识别字符,并界定待识别字符的边界参数,这里的边界参数可为该待识别字符在待识别图像中的位置参数或坐标参数,能够用于定位所述待识别字符。

[0244] 在具体的实现过程中,所述确定单元110可采用斯托克城宽度变换(Stroke Width Transform, SWT)算法,定位待识别字符,并提取出边界参数。进一步地,所述识别单元120,具体用于根据所述边界参数从所述待识别图像中提取出所述待识别区域;根据所述待识别区域中图形识别出待识别对象,获得所述识别结果。

[0245] 在本实施例中在提取出边界参数之后,所述识别单元120可直接根据所述边界参数定位出所述待识别字符在所述待识别图像的位置,不用利用去噪处理等去除待识别图像中所述待识别对象以外的信息,避免去噪处理过程中导致的字符特征的损失,避免字符特征损失导致的识别精确低的现象。

[0246] 实施例十三：

[0247] 如图9所示，本实施例提供一种图像识别装置，包括：

[0248] 确定单元110，用于确定待识别图像中待识别对象对应的待识别区域；

[0249] 识别单元120，用于利用采用变换数据集训练形成的指定神经网络，对所述待识别区域进行识别获得识别结果；

[0250] 其中，所述变换数据集包括：包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0251] 所述确定单元110，具体用于定位所述待识别图像的待识别对象；确定所述待识别对象的边界参数；

[0252] 其中，所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。

[0253] 所述待识别对象包括待识别字符。

[0254] 所述确定单元110，具体用于检测出所述待识别图像中边缘点；所述边缘点包括第一边缘点p；确定所述p的第一梯度方向；

[0255] 以所述p为起点，沿所述第一梯度方向搜索所述边缘点中的第二边缘点q；其中，所述q的梯度方向为第二梯度方向；所述第一梯度方向和所述第二梯度方向满足第一预设条件；若存在至少一个所述q，则确定所述p和所述q属于同一待识别字符；根据所述p和所述q的距离，确定所述待识别字符的笔画宽度；将所有笔画宽度满足第二预设条件的边缘点合并，组成所述待识别字符。

[0256] 在本实施例中所述确定单元110将用于进行边缘检测确定出边缘点，并通过计算边缘点的梯度方向，即两个边缘点之间的梯度方向的比较，确定出哪些边缘点是属于同一个字符的，将属于同一个字符的边缘点进行合并，形成所述待识别字符。

[0257] 本实施例所述的确定单元110可对应于处理器或处理电路，一方面可以进行像素的搜索，同时还可以进行相关的运算，进而能够简便的提取出待识别字符。

[0258] 作为本实施例的进一步改进，所述第一梯度方向和所述第二梯度方向满足第一预设条件，包括：所述第一梯度方向 d_p 和所述第二梯度方向 d_q 满足如下函数关系： $d_p = -d_q \mp \theta$ ；其中，所述 θ 为预设值。在本实施例中所述 θ 可为30度或40度或50度，可选的取值范围为15度至75度。本实施例在前一实施例的基础上，具有提取待识别字符简便及精确的特点。

[0259] 值得注意的本实施例是在前述任意一个装置实施例的基础上的进一步改进，在不冲突的前提下，能够与前述任意一个装置的技术方案结合，形成新的装置的实施例。例如，本实施例中所述图像识别装置也包括进行神经网络训练，获得所述神经网络的训练单元。

[0260] 实施例十四：

[0261] 如图9所示，本实施例提供一种图像识别装置，包括：

[0262] 确定单元110，用于确定待识别图像中待识别对象对应的待识别区域；

[0263] 识别单元120，用于利用采用变换数据集训练形成的指定神经网络，对所述待识别区域进行识别获得识别结果；

[0264] 其中，所述变换数据集包括：包含有预设变换的图形对象的训练图像和所述训练图像对应的预设结果。

[0265] 所述确定单元110,具体用于定位所述待识别图像的待识别对象;确定所述待识别对象的边界参数;

[0266] 其中,所述边界参数能够用于所述指定神经网络定位所述待识别对象对应的所述待识别区域。

[0267] 所述边界参数包括:所述待识别对象的外接边框的顶点坐标参数。

[0268] 所述边界参数可为各种能够定位出所述待识别对象的位置参数或坐标参数等,在本实施例中优选所述边界参数为待识别对象的外接边框的顶点坐标参数,这样确定单元110确定出外边框之后,仅需将外边框的顶点坐标参数发送给指定神经网络,指定神经网络接收的参数量,能够根据顶点坐标参数简便的定位出所述待识别对象在待识别图像中的位置,从而能够简便的提取出待识别区域,进行图像识别,一方面不用进行去噪处理,故因避免去噪导致对象特征丢失的识别精确低的现象,另一方面能够根据顶点坐标参数提取出待识别区域,以减少待识别图像中其他图形的干扰,提升了识别精确度。

[0269] 以下结合上述任意实施例提供一个具体示例:

[0270] 示例一:

[0271] 如图10所示,本示例提供一种图像识别方法,包括:

[0272] 步骤S1:获取待识别图像;

[0273] 步骤S2:利用canny检测待识别图像中各个图形对象的边缘点;

[0274] 步骤S3:利用SWT算法提取确定图形对象的笔画宽度,将认为属于通过一待识别字符的像素进行标记,被标记的像素就表示是属于同一待识别字符的。

[0275] 步骤S4:提取每一个待识别字符的外边框。

[0276] 步骤S5:将所述外边框和待识别图像输入指定神经网络,进行识别。

[0277] 在本示例中,所述指定神经网络可为卷积神经网络。在利用神经网络进行图像识别之前,需要进行神经网络的训练,神经网络的训练可包括如下步骤:

[0278] 第一步:数据集扩充,利用图形文本数据集ICDAR'11Graphic Text Dataset作为初始数据集;由于需要训练得到的指定神经网络能够处理倾斜、旋转、透视变换等字符,对训练集进行扩充。具体方法就是对原训练集中的图片进行相应的变换,使得原来的训练图像中的图形对象出现倾斜、弯曲、选择或透视等变换。

[0279] 第二步:网络训练。基于Caffe的深度神经网络的设计、配置,网络参数的训练在一块图形处理器(Graphics Processing Unit,GPU)卡上完成,训练数据集为扩充后的训练集提供。网络的结构会根据最终识别效果和网络运行速度进行调整。其中,所述Caffe为当前一种深度学习算法。

[0280] 在步骤S5中识别字符时,可利用Fast R-CNN可以同时完成待识别字符检测,也可已由SWT算法提取出字符,然后经过由扩充后的训练集训练出来的神经网络,将每个字符进行分割并分别识别出来。具体的识别过程可以会利用训练好的网络参数作为卷积操作的运算核,卷积操作可以用来提取原始输入图片的一些特征。与此同时,对原始的待识别图像的应用字符分割算法来划定哪几个区域中有待识别字符并在特征图中标记对应位置,最后由训练后的神经网络完成字符识别。

[0281] 在本申请所提供的几个实施例中,应该理解到,所揭露的设备和方法,可以通过其它的方式实现。以上所描述的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,如:多个单元或组件可以结合,或可以集成到另一个系统,或一些特征可以忽略,或不执行。另外,所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以通过一些接口,设备或单元的间接耦合或通信连接,可以是电性的、机械的或其它形式的。

[0282] 上述作为分离部件说明的单元可以是、或也可以不是物理上分开的,作为单元显示的部件可以是、或也可以不是物理单元,即可以位于一个地方,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全部单元来实现本实施例方案的目的。

[0283] 另外,在本发明各实施例中的各功能单元可以全部集成在一个处理模块中,也可以是各单元分别单独作为一个单元,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单元的形式实现。

[0284] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于一计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:移动存储设备、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0285] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

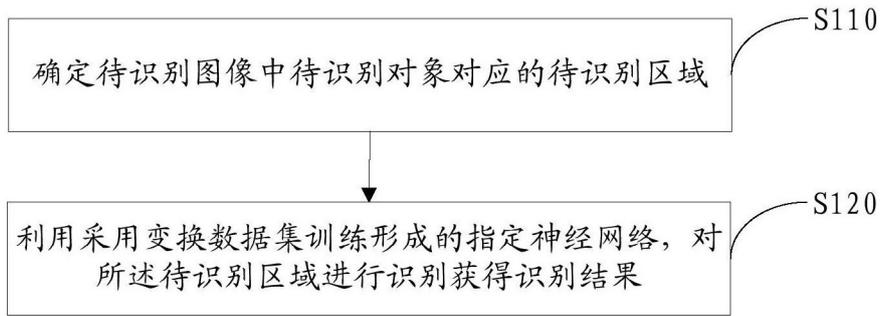


图1



图2

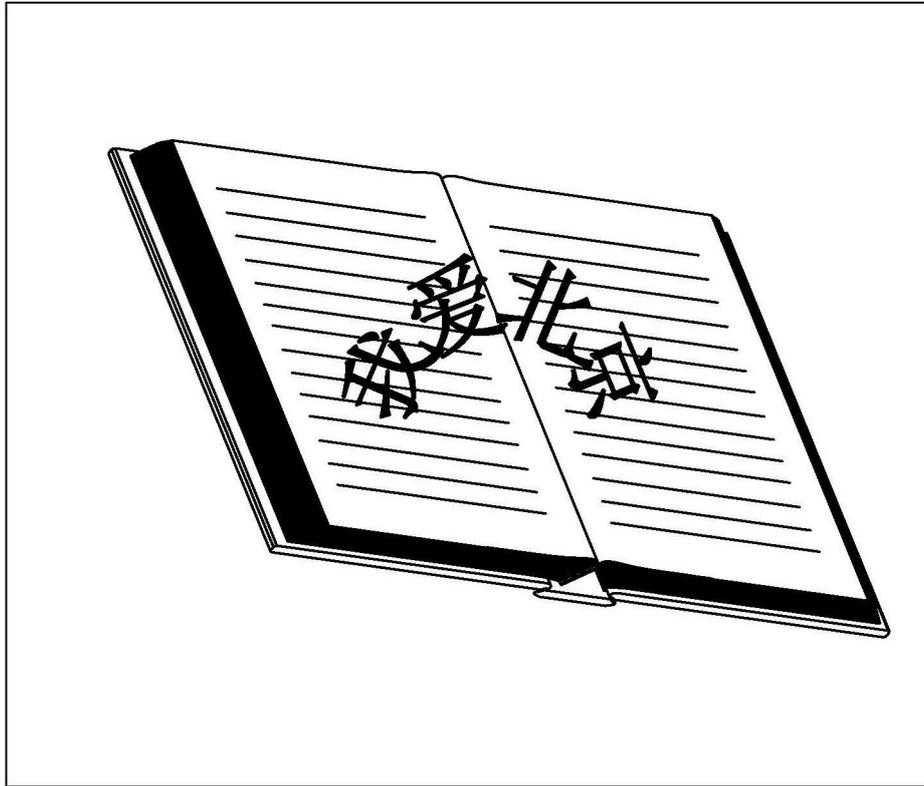


图3

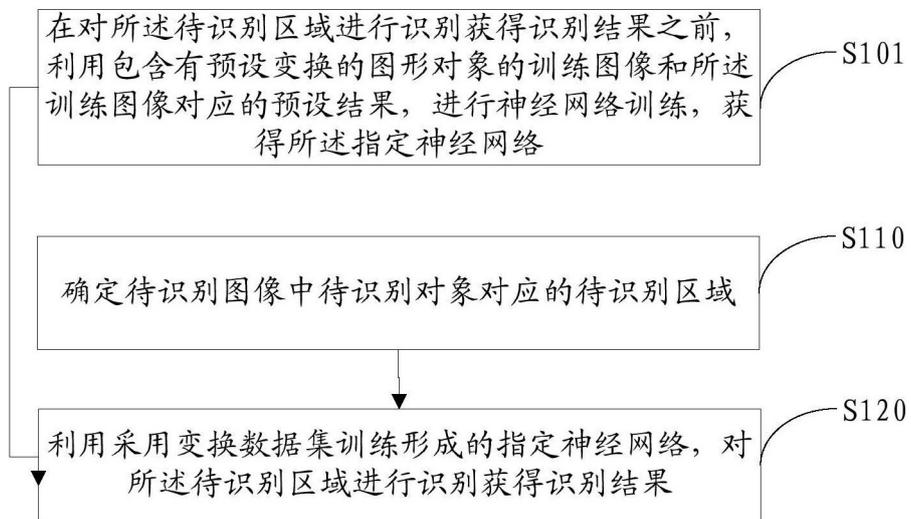


图4

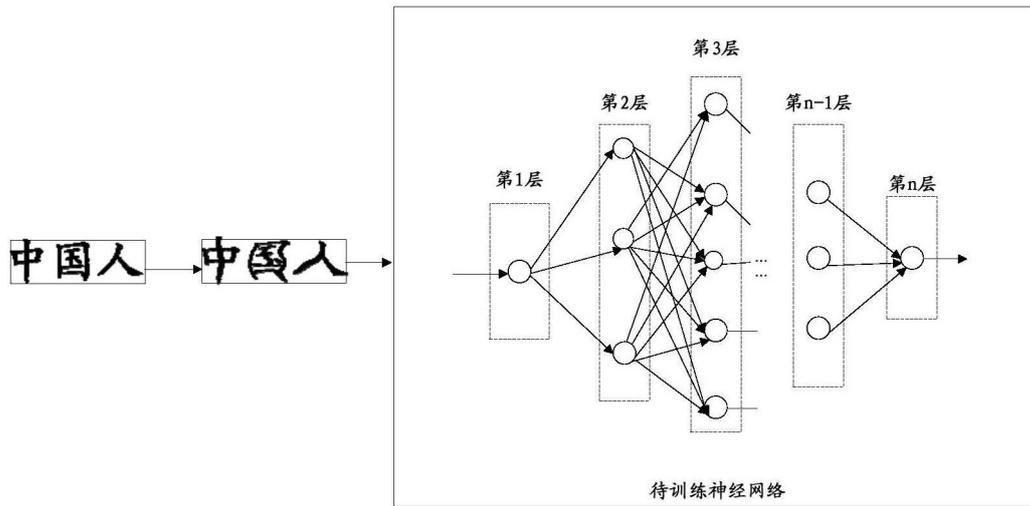


图5

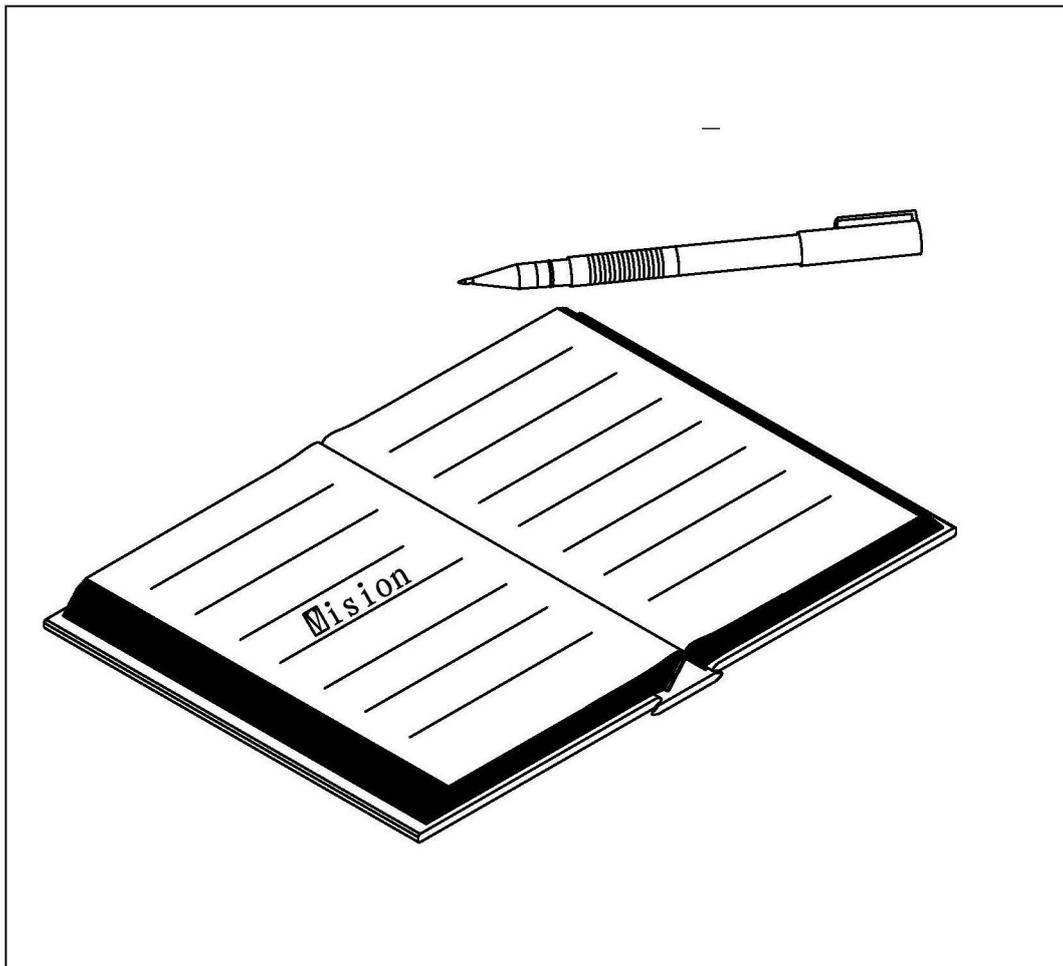


图6

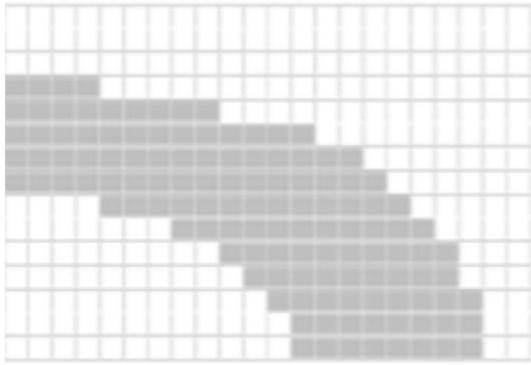


图7a

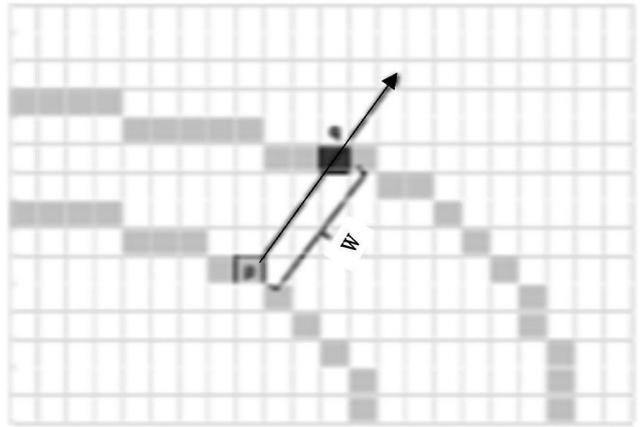


图7b

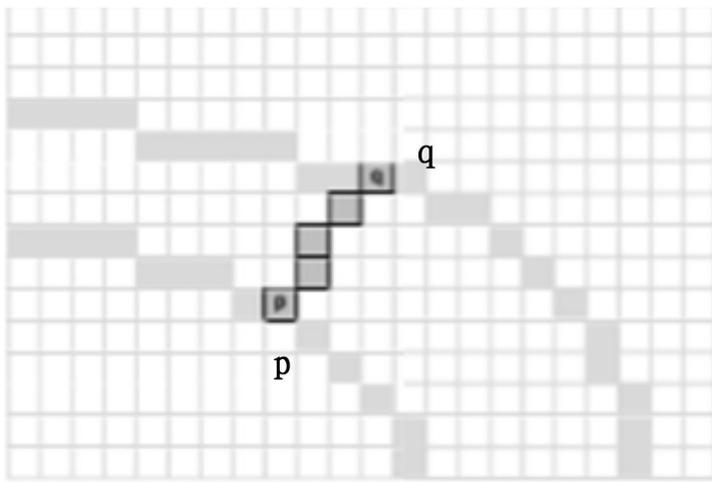


图7c

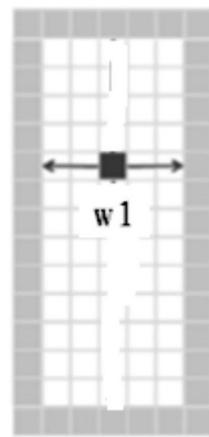


图8a

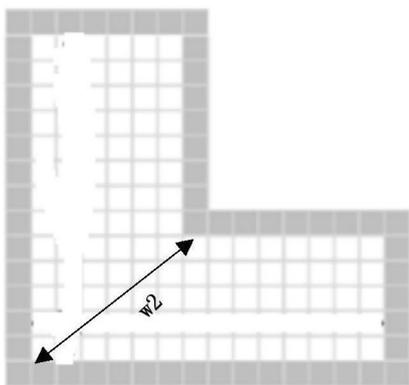


图8b



图9

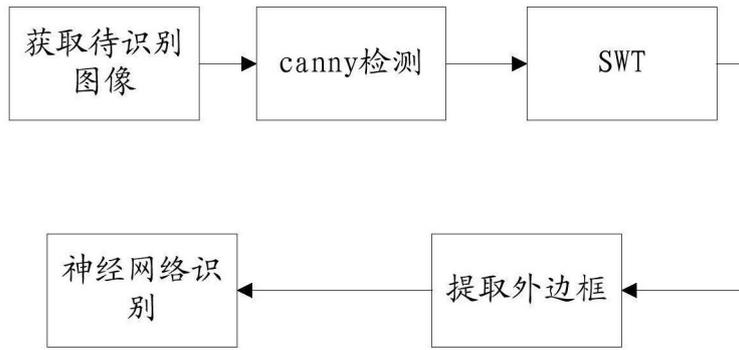


图10