

导读:随着智能时代的来临,深度学习和图像识别相关技术得到了越来越广泛的关注与应用。深度学习和图像识别已成为推动工业 4.0 和双碳目标的重要实施路径。目前,深度学习已经广泛应用于语音识别、文本分析、知识图谱、图像识别、视频分析等诸多领域。为了促进深度学习和图像识别的核心技术交流,《电子技术应用》特推出“深度学习和图像识别”主题专栏。本专栏针对深度神经网络模型理论研究及应用、深度强化学习理论研究、图像增强/特征提取/对齐/检索方法研究、数字图像相关方法研究、图形/文字识别相关研究、图像识别在工程/医学/生物等应用研究、图像识别模型轻量化进行研究和讨论。期待对深度学习和图像识别的研究和应用提供有益的借鉴。

特约主编:



朱小燕,清华大学计算机系教授、博士生导师,清华大学智能技术与系统国家重点实验室主任,北京市计算机学会副理事长,《电子技术应用》编委会委员。入选“人工智能全球女性榜单”和“2019 年度中国人工智能 100 人榜单”。从事人工智能领域相关工作 30 余年,主要研究领域为智能信息处理,包括人工神经网络、机器学习、自然语言处理、智能信息获取、智能问答系统等。获得国家发明专利十余项,在国际刊物和会议上发表论文 100 余篇。近年研究工作集中于人机对话/交互的研究,所研究的原型系统应用于多家知名企业。华来知识首席科学家,华来知识创始人之一。



利节,重庆科技学院副教授、硕士生导师,重庆市人工智能学会智慧交通专业委员会委员,西藏自治区高原复杂环境信息传输与处理工程研究中心特聘研究员,IIC-SPI 国际会议程序委员会主席。在 IEEE Transactions on Industrial Informatics、IEEE Transactions on Industrial Electronics 等期刊上发表学术论文 40 余篇;出版著作《人工智能》广受业内好评。研究方向包括机器学习、图像识别、视频分析、智能芯片、边缘加速和片内计算。



王东,英国爱丁堡大学博士,清华大学副研究员、语音语言技术中心副主任,曾任 IBM 中国高级软件工程师,英国爱丁堡大学 Marie Curie 研究员,法国 EURECOM 博士后研究员,美国 Nuance 公司高级研究科学家。研究方向包括机器学习、语音信号处理、自然语言理解等,在领域主流期刊和会议发表论文 150 余篇,著有《机器学习导论》《人工智能》等著作。

基于漫水填充算法的中文印章识别方法*

张 祥¹, 秦 毅¹, 董志诚², 黄琦麟¹, 利 节¹

(1.重庆科技学院 智能技术与工程学院, 重庆 401331; 2.西藏大学 信息科学技术学院, 西藏 拉萨 850000)

摘 要: 印章识别是智能化办公中必不可少的一环, 现阶段的印章识别方法是将扫描后的电子文档直接输入神经网络模型中进行识别, 面临着无法准确定位印章位置、弯曲文本识别准确率低等问题。针对上述问题, 提出了一种高效的中文印章文本识别方法, 该方法使用漫水填充算法处理灰度图像进行印章图像特征增强, 保证了中文印章检测精度, 同时引入了极坐标转换操作以确保文本特征完整性。为了评估本文方法的有效性, 在现有的文本识别网络模型进行了多组对比实验。经实验表明, 已有的网络模型融合该方法所提取的文本特征均表现出优异的识别结果。

关键词: 漫水填充; 特征增强; 印章特征; 印章文本识别

中图分类号: TP183

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.223130

中文引用格式: 张祥, 秦毅, 董志诚, 等. 基于漫水填充算法的中文印章识别方法[J]. 电子技术应用, 2022, 48(11): 2-6, 12.

英文引用格式: Zhang Xiang, Qin Yi, Dong Zhicheng, et al. Chinese seal recognition method based on flood filling algorithm[J]. Application of Electronic Technique, 2022, 48(11): 2-6, 12.

Chinese seal recognition method based on flood filling algorithm

Zhang Xiang¹, Qin Yi¹, Dong Zhicheng², Huang Qilin¹, Li Jie¹

(1.School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing 401331, China;

2.School of Information Science and Technology, Tibet University, Lhasa 850000, China)

Abstract: Seal recognition is an indispensable part of the intelligent office. The current stage of seal recognition method is to directly input the scanned electronic documents into the neural network model for identification, facing the problems of unable to accurately locate the position of the seal, low accuracy of bending text recognition. Aiming at the above problems, this paper proposes an efficient stamp text recognition method, which uses the diffuse water filling algorithm to process the grayscale image for seal image feature enhancement, which ensures the accuracy of Chinese seal detection, and introduces the polar coordinate conversion operation to ensure the integrity of text features. In order to evaluate the effectiveness of the proposed method, multiple sets of comparative experiments were carried out in the existing and other network models. Experimental results show that the existing network model fused the text features extracted by the method shows excellent recognition results.

Key words: flooded fill; feature enhancements; seal image excision; seal text recognition

0 引言

印章作为一种具有法律效力的工具, 广泛应用于政府文件、法律文件等各领域发布的文件中。印章可能存在模糊、褶皱等特征, 这对基于深度学习的印章识别方法具有极大困难, 该方法需要构建网络结构、制作印章数据集以及训练网络模型等多个步骤。此方法存在一定的局限性, 受网络模型的不可解释性以及参数权重多等因素的影响, 该方法检测效果不好, 时间周期长, 鲁棒性差, 所以不适合对印章的检测。

随着计算机视觉的发展, 印章识别技术越来越受到人们重视。姚敏^[1]等人利用 SIFT 描述子对尺度、旋转、亮

度具有不变性的特点, 采用 SIFT 算法实现印章图像的粗匹配, 再通过 RANSAC 算法根据一组包含异常数据的数据集计算数学模型参数, 从而达到印章的精确定位。马丽霞^[2]等人通过对已切除印章图像进行去噪、二值化、边缘提取等图像预处理操作, 对污染而残缺的部分进行了修复, 实现了去除具有噪声、划痕、孔洞的印章图像背景。肖进胜^[3]等人通过对印章图像的颜色空间的转换, 利用带角度信息的联结文本提议网络结合贝塞尔文本区域实现了对印章区域的准确检测。赵勇涛^[4]等人对采集的印章图像构建 HIS 彩色模型, 通过提取红色分量的方法将印章分离出来, 然后进行二值化、填充、去噪得

* 基金项目: 重庆市自然科学基金面上项目(cstc2020jcyj-msxmX0683); 重庆市教委科学技术研究项目(KJQN201801504); 四川省科技计划项目(2022YFG0043, 2022YFG0046)

到预处理图像。该方法的优点是速度快,能够实时处理请求。蔡亮^[5]等人提出了基于外轮廓骨架线套位法,通过获取不同印泥的形状特征,实现了不同类型印章的快速检测方法。

针对倾斜文本的检测问题,SegLink 算法^[6]将文本拆分为两个可局部检测的元素,增加了角度的检测,对具有倾斜方向的文本行检测表现出优秀的效果。但是该网络无法检测间隔大、弯曲文本的图像。

针对弯曲文本的检测问题,Zhang^[7]等人通过文本组件建议网络与深度关系推理图网络共享卷积特征,骨干网络采用 VGG-16 网络,将文本实例划分为一系列矩形组件,用于估计组件的高度、宽度等几何属性,该方法对任意形状文本检测具有良好的性能。Zhu^[8]等人提出了一种全新的傅里叶轮廓嵌入方法,并构造了一个具有主干特征金字塔和后处理的傅里叶变换的卷积神经网络,FCENet 可以以端到端方式进行优化,并且不需要任何复杂的后处理,经实验证明该网络对场景文本的轮廓具有优秀的准确性和鲁棒性。Zhou^[9]等人通过单一的卷积神经网络,直接预测完整图像中任意方向的四边形文本,消除了不必要的中间步骤,比如文本区域的划分和形成等,并且该网络是全卷积神经网络,可以对输出的单词或者文本的每一个像素进行预测。Li^[10]等人提出了一种渐近尺度扩展网络,一种基于图像分割的文本检测模型,它可以实现对每个文本实例进行预测,通过具有最小尺度的内核扩展到具有最大和完整形状的文本实例。由于内核之间存在很大的几何边距,因此该方法对任意形状具有良好的鲁棒性。

针对文本识别问题,Shi^[11]等人基于图像序列解决了场景文本识别的问题,提出了一种新的神经网络结构,该网络可以进行端到端训练,能够自然处理任意长度的文本序列,无需进行字符分割或水平尺度归一化处理。该方法集成了卷积神经网络和递归神经网络的优点来完成文本识别问题。此外,该方法的主干网络放弃了传统神经网络中的全连接层,成为了一个更加紧凑和高效的模型。并且该方法在光学音乐识别数据集上的效果同样优于其他网络。

目前,无论是通过传统图像或者基于深度学习的印章识别算法很多,但是各种算法都存在相应的局限性。比如基于传统图像通过字符切割对文本进行识别,但是未考虑到文本弯曲程度大无法准确切割的问题。基于深度学习的网络模型需要大量数据集进行训练,而印章与其他目标对象不同,它是个人、公司的重要物品,因此无法公开印章数据集,这给基于卷积神经网络的印章文本检测与识别带来了困难。为了解决上述问题,本文提出了一种印章特征增强结构(Seal Feature Enhancement Structure, SFES),该结构用于提取印章文本,并有利于在电子文档场景下对中文印章进行识别。

1 相关技术

1.1 滤波算法

电子文档图像大多数是将纸质文档通过扫描仪器进行扫描而产生的,扫描过程中会出现杂质以及纸质文档氧化等情况,导致扫描后的电子文档图像会出现黑点、褶皱等噪声。常见的图像滤波算法包括均值滤波算法、双边滤波算法、高斯滤波算法等。

均值滤波是获取窗口函数覆盖的区域中的所有像素值的平均值,并代替窗口函数原先的所有像素值。该滤波算法计算速度快,但是在滤波的同时会丢失图像的细节。中值滤波算法与上述算法的区别是通过获取区域中像素值的中位数来代替原先的所有像素值,该算法对处理“胡椒盐”类噪声的效果好。

高斯滤波算法与均值滤波算法相似,返回滤波器窗口中的像素均值,可以对图像平滑化。不同点在于窗口模板系数,均值滤波算法的模板系数均为 1,高斯滤波算法的系数与模板中心的距离成反比,服从二维正态分布的系数分配模式。式(1)为二维高斯分布公式。

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

其中 x, y 为像素点坐标,在电子图像中可以当作整数来处理, σ 为标准差。可以通过该公式计算高斯核,再与图像的像素值相乘叠加得到最终结果。

1.2 图像饱和度

图像饱和度是色彩中彩色成分和消色成分的占比,该比例决定多通道图像中色彩的鲜艳程度。在多数电子文档图像中的印章均为红色,并且印章与周围像素值具有明显的差值。

饱和度调整算法是将红(R)、绿(G)和蓝(B)组成的颜色空间转换为色调(H)、饱和度(S)、明亮度(V)颜色空间,用来对饱和度 S 进行上下限控制,此处无需考虑色相 H 差异,当增量小于 0 时, R 可通过线性减弱方式调整;当增量取 -1 时,可直接采用线性方式调整。同时将饱和度调整过的 RGB 进行校验,最终将 RGB 的图像输出即可。

2 印章识别方法

2.1 整体架构

SFES(Seal Feature Enhancement Structure)的架构如图 1 所示,该结构总体分为模块一和模块二。为了提高识别速度,模块一通过漫水填充算法对灰度图进行处理,而不是直接对原图进行处理,从而将印章检测问题转换为印章分割问题,为了保证轮廓检测的准确性,再对图像进行二值化处理以达到印章检测的目的。但是印章被切割后仍然无法识别印章中的文本。模块二根据模块一的输出进行极坐标转换,将印章中的弯曲文本转换为矩形文本,方便后续进行识别。

2.2 灰度处理

彩色图像是通过 (R, G, B) 三通道来表征一幅图像,

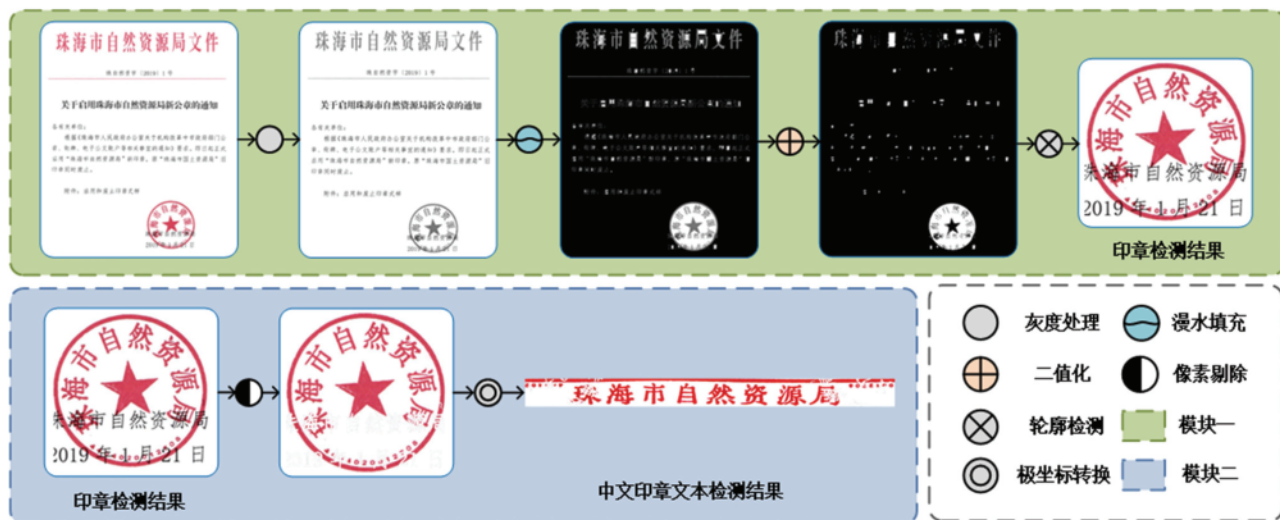


图1 SFES 架构图

在图像处理过程中,算法更加关注图像梯度信息,而不是彩色信息。三通道图像的参数量会导致算法运行缓慢等消极影响,为保证图像梯度信息的同时,也减少了参数量,印章识别过程采用单通道的灰度图。灰度图是只包含亮度信息,通常划分0至255个级别,其中0代表全黑,255代表全白。彩色图转换为灰度图公式如式(2)所示。

$$\text{Gray} = (0.114 \quad 0.587 \quad 0.299) \begin{pmatrix} B \\ G \\ R \end{pmatrix} \quad (2)$$

该式通过 1×3 矩阵与 3×1 矩阵相乘得到灰度值。

2.3 漫水填充算法

漫水填充算法是图像预处理中常用的区域填充算法,该算法根据像素灰度值之间的差异寻找相同区域实现图像分割。图2为漫水填充原理图。

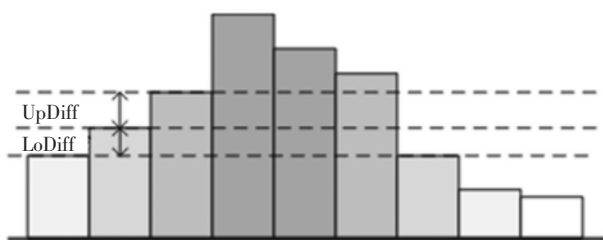


图2 漫水填充原理图

该算法在处理过程中随机选定初始种子点作为起始点,向周围像素点扩散,将相似的像素值容纳在内,得到一组无交集的连通区域,并将其填充新的像素值。同时需要设置掩码图像,要求比原图片宽高各大两个像素以防止边缘被填充。具体算法步骤如下:

(1)在原图中随机标记种子像素点,标记为 (x, y) 。

(2)搜索步骤(1)中的种子点的邻域点,如果该点未被遍历过且满足填充条件,则填充该位置。

(3)检测已填充位置,继续执行步骤(2)。

(4)当全部像素点被遍历,则填充完成。

如图1中所示为电子文本图像进行漫水填充算法后的结果,印章像素值和周围像素值差值更大,更方便寻找轮廓。

2.4 二值化

二值化是将图像中像素点的灰度值设置为0或者255。通过阈值法实现图像二值化,将图像所有像素点的像素值与某个临界灰度值 x 做对比,如果大于 x 则设置为灰度最大值,反之则设置为灰度最小值。阈值的选取方式可以选择基于平均值的方式、基于直方图双峰的方式、基于最大类间方差的方式。

基于平均值的方式的阈值取为图像本身像素值的平均值。如果背景和物体的灰度值差异较大,则可以基于直方图双峰的方式进行二值化。但是上述阈值的选取方式都具有本身的局限性,且直方图包含低谷和不规则抖动,导致寻找准确的极值点十分困难。故采用基于最大类间方差的方式进行图像二值化。方差可以评估灰度分布的均匀性,目标物体和背景之间的类间方差越大,说明两部分的差距越大,当部分目标物体和背景之间的像素混淆时会引起类间差距减小,因此采用基于类间方差方式的阈值法进行二值化会更加准确。

2.5 轮廓检测

轮廓检测是在电子文本图像中,忽略图像背景以及纹理信息等噪声的干扰,实现目标轮廓提取的过程。轮廓检测算法采用连通域思想进行轮廓,存在两个像素 a_1 和 a_2 ,如果 a_1 在 a_2 的四邻域中,称两个像素为四连通;如果 a_1 在 a_2 的八邻域中,称两个像素为八连通,如图3所示,其中 P 代表当前像素。

假设输入图像为 $F=f(i, j)$,其中 i, j 为像素坐标, $f(i, j)$ 为在该点的像素值。确定边界开始点 i, j 进行边界追

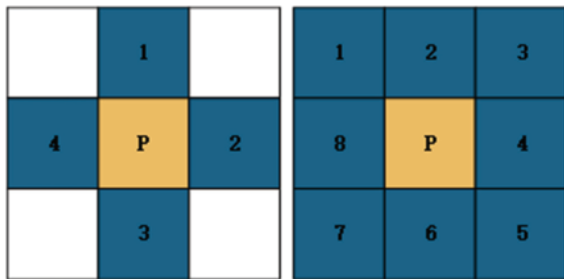


图3 四连通和八连通

溯,按顺时针方向查找该点的四连通是否存在非0像素点,对像素值为1的 (i,j) 点若与其四连通也存在某点的值是1,则两点划为同一物体,图4所示为四连通运算结果。

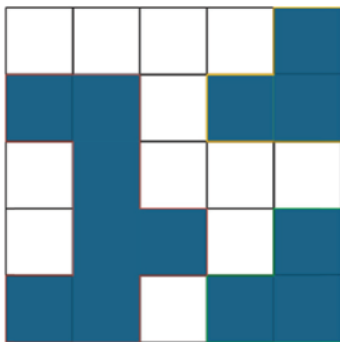


图4 四连通结果图

2.6 极坐标转换

如图5所示,可通过极坐标转换将坐标轴中的圆形转换为矩形。

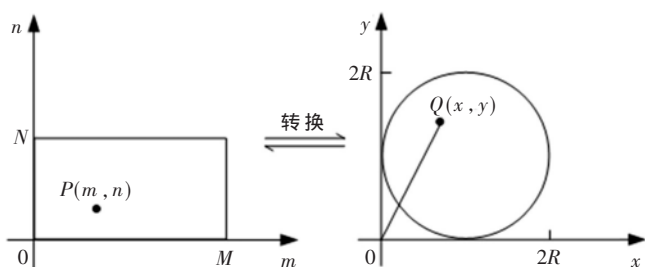


图5 极坐标转换

假设圆形图片中的某个坐标点 $Q(x,y)$,矩形中的某个坐标点 $P(m,n)$,圆形半径为 R ,矩形长宽比为 $N:M$,则圆形上任意点坐标 (x,y) 和矩形内任意点坐标的转换规则如式(3)所示。

$$\begin{cases} x = \frac{2R-1}{2} + n \frac{1}{r} \cos(mt) \\ y = \frac{2R-1}{2} - n \frac{1}{r} \sin(mt) \end{cases} \quad (3)$$

其中缩放因子 $r=N/R$,角度因子 $t=2\pi/M$ 。

2.7 像素剔除

为了保证最终的文本识别率,通过判断像素值把印章检测结果中的黑色像素值剔除。具体算法步骤如下:

(1)读取印章文本图像,并获取图像的 weight、high 属性。

(2)遍历文本图像像素,判断当前像素点是否满足剔除条件。

(3)如果满足条件则剔除像素,并赋予新的像素值;如果不满足,则继续判断下一个像素点,直到完成遍历所有像素点。

3 实验结果

3.1 数据集

CTW(Chinese Text in the Wild)^[12]是一个自然场景下的中文数据集,包含30 000多张街景图片,并标记了3 850万个汉字。该数据集自然场景种类多样,包括平面文本、凸起文本、遮挡文本等场景。标记格式不仅包含了字符以及边界框,还具备环境复杂性,比如外观、风格等。本节所提及的文本检测网络的训练均在该数据集上进行。

为了验证 SFES 对中文印章识别的有效性,本实验制作了100张包含中文印章文本图像的电子文档数据。

3.2 结果与分析

如图6所示,设直接输入图像为方法一,输入图像+模块一为方法二,输入图像+模块一+模块二为方法三。

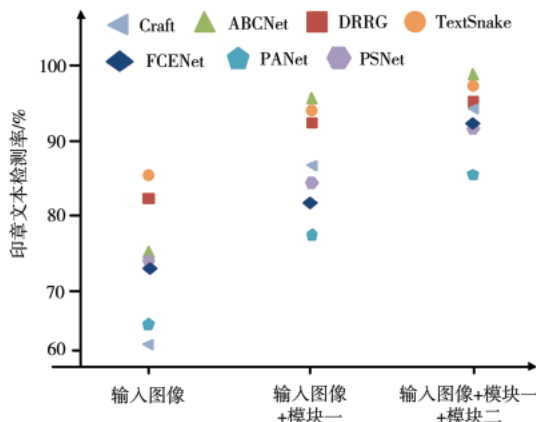


图6 文本检测模型数据对比

文本检测网络模型采用 Craft^[13]、ABCNet^[14]、DRRG^[7]、TextSnake^[15]、FCENet^[8]、PANet^[16]、PSNet^[10]网络。实验数据表明,Craft 模型和 ABCNet 模型通过方法二有显著的检测率上升的趋势。DRRG、TextSnake、FCENet、PANet、PSNet 网络模型通过方法二和方法三均呈现上升的趋势。

将上述实验的数据结果分析可得,在基础的文本检测网络中的原始图像输入通过方法二、方法三更改后,印章文本的检测都有明显的提升,这是因为 SFES 填补了网络模型对中文弯曲文本检测的不足,即增强了原始图像的文本特征,又为后续的文本识别准备了充足的工作。

图7所示为文本检测模型+模块一的效果图,不难

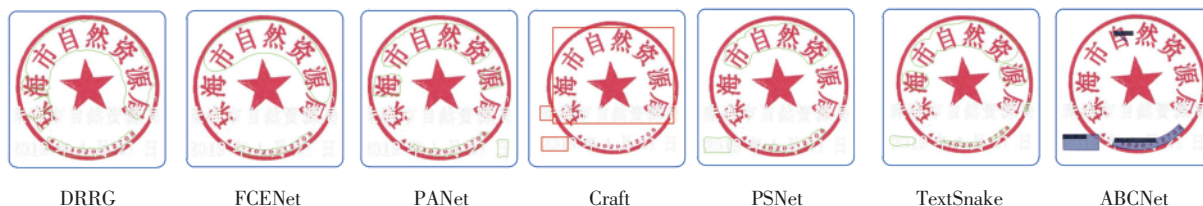


图7 文本检测模型通过模块一的数据

发现,文本检测网络对弯曲文本的检测效果不佳,每个网络的检测结果都存在相应的问题。具体表现为:FCENet、PSNet以及TextSnake网络无法检测到“珠海”“局”三个字;PANet网络无法检测“珠”“局”两个字;Craft网络无法准确定位每个汉字的位置;ABCNet无法识别到汉字字符;DRRG网络检测“珠”字出现了定位不准确的问题。

通过模块一的检测结果直接输入到文本识别网络模型是无法进行准确识别汉字的,所以需要通过模块二进行极坐标转换,将弯曲文本转换为矩形文本。如图8所示,可以看到极坐标转换后的结果极佳,只有PANet网络出现字符漏检情况。

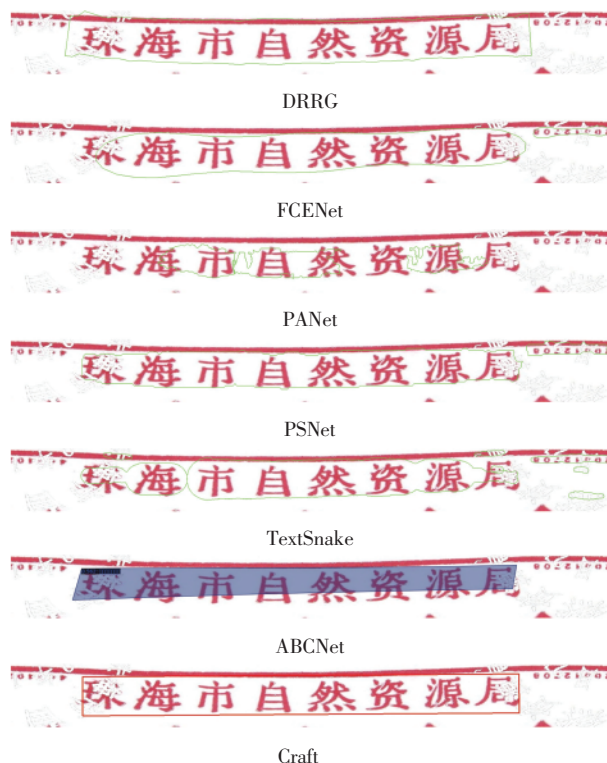


图8 文本检测模型通过模块一和模块二的数据

与文本检测网络的任务不同,文本识别网络更加专注于印章文本的内容而不是印章文本的位置。因此设文本识别网络+模块一为方法一,文本识别网络+模块一+模块二为方法二,图9所示为文本识别网络融合SFES所产生的实验结果。文本识别网络采用PaddleOCR、SAR^[17]、

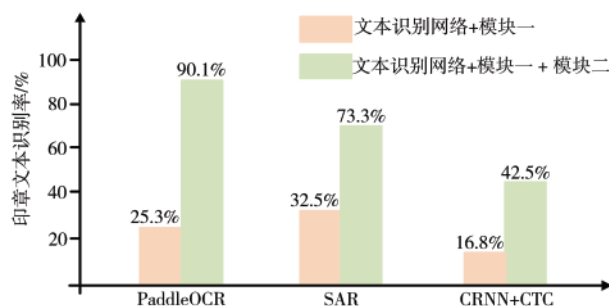


图9 文本识别网络通过模块一和模块二的数据对比

CRNN^[11]。根据实验结果可以观察到,方法一所展示的上述文本识别网络针对弯曲文本的识别均未达到预期的效果,识别率均在40%以下。方法二的中文印章文本识别率得到了一定幅度的提升,证明了SFES在中文印章文本识别的有效性。

4 结论

本文提出了一种基于漫水填充算法的中文印章文本识别方法,该方法通过印章特征增强结构完成印章图像与扫描电子文档的分割,以及将弯曲文本转换为矩形文本。实验表明,文本检测网络与识别网络融合SFES后检测率和识别率都得到了明显提升。尽管该方法在实验中表现出优异的性能,但是该方法在鲁棒性上表现不佳,后续可通过傅里叶变换等方式进行更加深入的研究。

参考文献

- [1] 姚敏,牟雪儿,陈鹏,等.图像中印章的检测定位与识别研究[J].信息技术与信息化,2018(12):148-150.
- [2] 马丽霞,朱秋萍.图像处理技术在印章识别预处理中的应用[J].武汉大学学报(信息科学版),2004(8):691-693.
- [3] 肖进胜,赵陶,熊闻心,等.角度优化网络的印章文字检测与识别算法[J].电子与信息学报,2021,43(11):3327-3334.
- [4] 赵永涛,李志敏,王洪建,等.印章识别中的图像预处理研究[J].仪器仪表学报,2004(S2):401-403,410.
- [5] 蔡亮,梅丽.基于楔环结构的彩色印章检测和对齐方法[J].浙江大学学报(工学版),2006(10):1696-1700.
- [6] SHI B, BAI X, BELONGIE S. Detecting oriented text in natural images by linking segments[C]//IEEE Computer Society. IEEE Computer Society, 2017.
- [7] ZHANG S X, ZHU X, HOU J B, et al. Deep relational rea-

(下转第12页)

- [4] CRAIG P S. Echinococcosis working group in China epidemiology of human alveolar echinococcosis in China[J]. Parasitology International, 2006, 55: S221-S225.
- [5] 木拉提·哈密提, 员伟康. 基于综合特征的新疆地方性肝包虫病图像特征的提取与分析[J]. 科技通报, 2015, 31(5): 58-62.
- [6] STOJKOVIC M, ZWAHLEN M, TEGGI A, et al. Treatment response of cystic echinococcosis to benzimidazoles: a systematic review[J]. PLoS Neglect. Trop. Dis., 2009, 3(9): e524.
- [7] 杜燕, 李书兵. 超声对肝包虫病的筛查及诊断价值[J]. 临床超声医学杂志, 2019, 21(12): 957-958.
- [8] 李宏霄, 李姝, 石霞飞, 等. 基于 BiT 的早期胃癌内镜图像识别[J]. 激光与光电子学进展, 2022, 59(6): 354-362.
- [9] 张雪芹, 李天任. 基于 Cycle-GAN 和改进 DPN 网络的乳腺癌病理图像分类[J]. 浙江大学学报(工学版), 2022, 56(4): 727-735.
- [10] 茹仙古丽·艾尔西丁. 基于深度学习的肝囊型包虫病 CT 图像分类方法研究[D]. 乌鲁木齐: 新疆医科大学, 2020.
- [11] WU M, YAN C B, WANGS X R, et al. Automatic classification of hepatic cystic echinococcosis using ultrasound images and deep learning[J]. J. Ultrasound Med., 2021, 41: 163-174.
- [12] CHENG J, WANG H, LI R, et al. A two-stage multiresolution neural network for automatic diagnosis of hepatic echinococcosis from ultrasound images: a multicenter study[J]. Med. Phys., 2022, 49(5): 3199-3212.
- [13] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Neural Information Processing Systems, 2017.
- [14] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[J]. arXiv preprint arXiv: 2010.11929, 2020.
- [15] LIU P, SONG Y, CHAI M, et al. Swin-UNet++: a nested swin transformer architecture for location identification and morphology segmentation of dimples on 2.25Cr1Mo0.25V[J]. Fractured Surface. Materials(Basel). 2021, 14(24): 7504.
- [16] CHEN Y, ZHANG P, KONG T, et al. Scale-aware automatic augmentations for object detection with dynamic training[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2022.
- [17] 肝两型包虫病诊断与治疗专家共识(2019 版)[J]. 中华消化外科杂志, 2019(8): 711-721.
- [18] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical

(下转第 18 页)

(上接第 6 页)

- soning graph network for arbitrary shape text detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2020.
- [8] ZHU Y, CHEN J, LIANG L, et al. Fourier contour embedding for arbitrary-shaped text detection[C]//CVPR 2021, 2021.
- [9] ZHOU X, YAO C, WEN H, et al. EAST: an efficient and accurate scene text detector[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2017.
- [10] LI X, WANG W, HOU W, et al. Shape robust text detection with progressive scale expansion network[J]. arXiv: 1806-02539, 2018.
- [11] SHI B, XIANG B, CONG Y. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(11): 2298-2304.
- [12] YUAN T L, ZHU Z, XU K, et al. A large chinese text dataset in the wild[J]. Journal of Computer Science and Technology, 2019, 34(3): 509-521.
- [13] BAEK Y, LEE B, HAN D, et al. Character region awareness for text detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2019.
- [14] LIU Y, CHEN H, SHEN C, et al. ABCNet: real-time scene text spotting with adaptive bezier-curve network[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2020.
- [15] Long Shangbang, Ruan Jiaqiang, Zhang Wenjie, et al. TextSnake: a flexible representation for detecting text of arbitrary shapes[C]//ECCV, 2018.
- [16] WANG W, XIE E, SONG X, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[C]//2019 IEEE/CVF International Conference on Computer Vision(ICCV). IEEE, 2019.
- [17] LI H, WANG P, SHEN C, et al. Show, attend and read: a simple and strong baseline for irregular text recognition[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 8610-8617.

(收稿日期: 2022-06-29)

作者简介:

张祥(1997-), 男, 硕士研究生, 主要研究方向: 图像分类、文本识别。

秦毅(1986-), 男, 博士, 讲师, 主要研究方向: 计算机视觉、目标检测。

利节(1985-), 通信作者, 女, 博士, 副教授, 主要研究方向: 对抗生成网络、工业缺陷检测, E-mail: jiel@cqust.edu.cn。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所