

## 角度优化网络的印章文字检测与识别算法

肖进胜<sup>①</sup> 赵陶<sup>①</sup> 熊闻心<sup>①</sup> 杨天<sup>①</sup> 姚渭箐<sup>\*②</sup>

<sup>①</sup>(武汉大学电子信息学院 武汉 430072)

<sup>②</sup>(国网湖北省电力有限公司信息通信公司 武汉 430077)

**摘要:** 利用光学字符识别方法对印章文字进行检测与识别,能够加快各类合同的分类处理速度与鉴别效率。该文针对圆形印章文字呈环形排列的特点,利用极坐标展开对印章文字进行预处理,克服了印章文字方向不统一的问题。对于展开后上下起伏的文本区域,利用带角度信息的联结文本提议网(CTPN)对印章文字区域进行检测,并使用贝塞尔拟合文本区域,实现了对印章区域的准确检测。最后利用注意力转移机制和该文匹配算法对检测的文字区域进行识别,输出印章文字内容。运用该算法对输出印章文字内容自制的中文印章数据集进行实验,印章内容的文字检测F值可以达到84.73%,文字识别召回率达到84.4%,表明该算法可以有效地检测识别印章内容,对文档的分类与鉴别研究具有重要的意义。

**关键词:** 图像处理; 印章识别; 循环神经网络; 极坐标转换

中图分类号: TN911.73; TP391.1

文献标识码: A

文章编号: 1009-5896(2021)11-3327-08

DOI: 10.11999/JEIT201008

## Seal Text Detection and Recognition Algorithm with Angle Optimization Network

XIAO Jinsheng<sup>①</sup> ZHAO Tao<sup>①</sup> XIONG Wenxin<sup>①</sup>

YANG Tian<sup>①</sup> YAO Weiqing<sup>②</sup>

<sup>①</sup>(School of Electronic Information, Wuhan University, Wuhan 430072, China)

<sup>②</sup>(State Grid Hubei Information & Telecommunication Company Limited, Wuhan 430077, China)

**Abstract:** Using the methods of Optical Character Recognition (OCR) to detect and recognize the seal characters can speed up the classification speed and identification efficiency of all kinds of contracts. According to the characteristics of the cycle seal characters arranged in a ring, polar coordinate conversion is used to preprocess the seal characters, which overcomes the problem that the direction of the seal characters is not uniform. The Connectionist Text Proposal Network (CTPN) with angle information is used to detect the undulating text area, and the Bezier curve is used to achieve the accurate detection of the seal area. Finally, a method combined with the attention mechanism and the matching algorithm is used to recognize the detected text area and the seal text content is obtained. Using this algorithm to test the self-made Chinese seal data set, the F-measure of the seal content can reach 84.73%, and the recall rate of the character recognition is 84.4%, which shows that this algorithm can detect and recognize the seal content effectively, and has an important meaning for the research of document classification and identification.

**Key words:** Image processing; Seal recognition; Recurrent neural networks; Polar coordinate conversion

### 1 引言

在当代社会,各级国家机关、社会团体、企事

业单位都使用印章作为一种重要的标记信物。随着图像处理技术发展,光学字符识别(Optical Character Recognition, OCR)等技术也取得了巨大的发展,同时随着电子化与无纸化办公的进一步普及,印章电子化也成为一种趋势。如果能够使用图像处理与光学字符识别的相关技术对中文印章内的文字进行检测与识别,使用机器快速地对印章的内容进行检测与分类,可以有效地提高文档分类的效率与准确率,节省人力,具有巨大的应用价值。

收稿日期: 2020-11-30; 改回日期: 2021-03-26; 网络出版: 2021-04-15

\*通信作者: 姚渭箐 ywq1005@whu.edu.cn

基金项目: 国家自然科学基金(61471272), 国网湖北省电力有限公司2019年科技项目(52153318004G)

Foundation Items: The National Natural Science Foundation of China (61471272), The Technology Project of State Grid Hubei Electric Power Co., Ltd. (52153318004G)

印章内容检测与识别借鉴自然场景文字识别的相关方法,一般分为文字区域检测与文字内容识别。Lin等人<sup>[1]</sup>对近几年场景文字检测与识别的各类方法进行了分析。联结文本提议网络(Connectionist Text Proposal Network, CTPN)<sup>[2]</sup>是文本检测的常用网络,奠定文本检测算法的基础,但只能检测水平文本行。SegLink算法<sup>[3]</sup>将每个字符切割为更易检测的细粒度文字块,然后将各个小文字块连接成单词,便于识别长度变化范围很大、带方向的单词和文本行。Zhou等人<sup>[4]</sup>提出一种高效和高精度的场景文本检测算法(an Efficient and Accurate Scene-Text detector, EAST),该模型使用全卷积网络(Fully Convolutional Networks, FCN)<sup>[5]</sup>生成多尺度融合的特征图,然后在此基础上进行像素级的文本块预测,可以实现多尺度的自然文本文字框检测。目前时序分类机制(Connectionist Temporal Classification, CTC)<sup>[6]</sup>和注意力机制(Seq2Seq)<sup>[7]</sup>广泛应用于文字识别领域。CTC算法用双向长短时记忆网络(Bi-directional Long Short-Term Memory, BLSTM)<sup>[8]</sup>对字符图像进行特征提取,用CTC损失函数对特征图进行编码与去冗。基于注意力机制的文字识别方法利用中间向量完成文字的编码与解码输出,能够提高检测效率和效果<sup>[9]</sup>。

方形印章中的文字呈水平分布,这类文本检测已经十分成熟,而中文常见印章内容都呈圆环排列,其识别工作仍具有很大的挑战。本文深入研究国内外印章内容识别技术的发展状况,提出了角度优化网络的印章内容识别算法。利用极坐标对印章文字进行展开,克服圆形印章文字方向不统一的问题,同时为了防止展开过程中印章文字发生形变,在展开的同时进行双线性插值。针对印章文字展开后依旧上下起伏和倾斜等情况,将旋转区域提议网络(Rotation Region Proposal Networks, RRPN)<sup>[10]</sup>运用至文本检测之中,使得文本提议框融入角度信息,并利用贝塞尔曲线对文本进行拟合。印章识别部分将卷积循环神经网络(Convolutional Recurrent Neural Network, CRNN)文字识别网络中的CTC转录层替换为注意力机制Seq2Seq,并利用文本匹配算法对识别结果进行修正。本文算法在印章预处理、文字区域提取和文字

内容识别等各个方面有一定的优势,具有一定的参考价值 and 实用性。

## 2 印章文字检测与识别

算法的预处理部分的主要任务是对印章进行提取与展开,对于呈环形排列的中文印章,利用极坐标展开将环形文字展开为矩形文字,在转换的同时也融合了双线性插值,克服了环形文字转换之后产生的形变问题。文字检测与识别部分提取印章文字区域和识别具体内容。将展开的印章图像使用包含RRPN的CTPN对文字区域进行提取,然后利用贝塞尔曲线对文本框进行拟合,最后将提取的文字区域输入至使用注意力转移机制的CRNN网络进行文字识别,并进行文本匹配输出印章内部的文字。整个架构如图1所示。

### 2.1 印章预处理

预处理的主要任务是对合同内的印章进行提取与展开,将印章图像由RGB颜色空间转换为CIE LAB颜色空间,由A, B通道确定印章所在位置,调节L通道阈值克服印章颜色亮度带来的影响。在CIE LAB颜色空间中对印章区域进行提取,可以有效地将印章文字与背景文字分离。同时利用极坐标转换,可以将需要检测的印章环形文字展开为横向排列的文字,克服了印章内部文字方向不统一的问题。在进行极坐标转换的过程中同时结合双线性插值,使得展开后每个像素点的值不只是简单地取决于原图中的单一像素点,而是由原图中的多个像素点共同来决定。印章提取和极坐标转换的示意分别如图2、图3所示。

如图3所示,圆图中点 $p(i, j)$ 在方图中对应的坐标为 $q(x, y)$ ,方图中的每一列对应圆图中的一个圆环,圆图的半径为 $R$ ,方图的长宽为 $N \times M$ , $M$ , $N$ 与实际的印章大小有关,一般展开后方图的每一行对应于印章的圆周,每一列对应于圆周上的每一条半径。定义缩放因子 $\lambda_t = N/R$ ,角度因子为 $\lambda_r = 2\pi/M$ ,圆图上图像坐标系 $(i, j)$ 和矩形坐标系 $(x, y)$ 的变换关系为

$$\left. \begin{aligned} i &= \frac{2R-1}{2} + y \cdot \lambda_r \cdot \cos(x \cdot \lambda_t) \\ j &= \frac{2R-1}{2} + y \cdot \lambda_r \cdot \sin(x \cdot \lambda_t) \end{aligned} \right\} \quad (1)$$

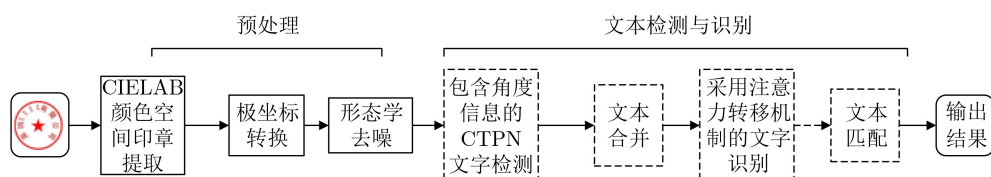


图1 印章识别架构图



图 2 印章提取效果

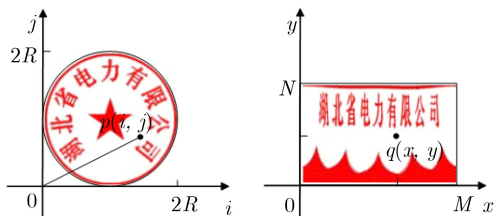


图 3 极坐标转换示意图

极坐标转换为直角坐标图像的过程中，采用了双线性插值，对图像进行拉伸。极坐标展开的效果如图4所示。



图 4 印章展开效果图

### 2.2 包含角度信息的CTPN印章文字检测

展开后的印章文字会呈波浪起伏排列，如果只检测文本提议框的中心坐标与矩形宽高4个参数，会包含较多的背景信息。本文为了减少印章文字提取过程中的信息干扰，将CTPN网络中的文本候选区域提取网络(Region Proposal Networks, RPN)替换为包含角度信息的旋转候选区域网络(Rotation Region Proposal Networks, RRRPN)<sup>[10,11]</sup>。RRPN在目标候选网络中加入了旋转因子，能够生成任意方向的候选区域<sup>[10]</sup>，使得网络更加适合弯曲的文本。RRPN网络中的每一个文本提议框可以表示为 $\{x, y, w, h, \theta\}$ 。 $(x, y)$ 确定框的中心位置， $(w, h)$

确定框的宽度与高度， $\theta$ 为框的垂边与水平轴的夹角。RRPN中文本提议框的宽度只有8, 16, 32 3类，高度值根据文本宽度决定，其比例为1:2, 1:5, 1:8。文本框的偏转角度信息只包含5个方向，分别是 $\left\{ \frac{4\pi}{9}, \frac{17\pi}{36}, \frac{\pi}{2}, \frac{19\pi}{36}, \frac{5\pi}{9} \right\}$ ，RRPN网络如图5所示。

通过RRPN网络，会产生大量的文本提议框，然后利用文本分类网络对这些文本提议框进行分类。RRPN网络损失函数由两部分组成

$$L(p, l, v^*, v) = L_{cls}(p, l) + \lambda L_{reg}(v^*, v) \quad (2)$$

其中， $l$ 为文本提议框的分类标签值， $p$ 是文本区域概率值。 $v = \{x, y, h, w, \theta\}$ 为文本提议框的检测输出位置信息与角度信息， $L_{cls}$ 代表文本分类损失函数， $L_{reg}$ 为位置回归损失函数，计算公式为

$$\left. \begin{aligned} L_{cls}(p, l) &= -\ln(pl) \\ L_{reg}(v^*, v) &= \sum_i \text{smooth}_{L1}(v_i^* - v_i) \end{aligned} \right\} \quad (3)$$

文本提议框分类标签值 $l = 1, 0$ 分别代表该提议框为正样本、负样本， $p$ 是通过Softmax函数计算的文本概率值。 $i$ 为细粒度文本提议框的索引值。 $v^*$ 为真值中标注的文本框位置标签值。 $L_{cls}$ 采用对数似然概率损失函数， $L_{reg}$ 采用L1损失函数。回归与分类损失函数具有不同的变化范围。为克服不同尺寸文本提议框对损失函数的影响，需对这些坐标信息进行归一化

$$\left. \begin{aligned} v_x &= \frac{x - x_a}{w_a}, v_y = \frac{y - y_a}{h_a} \\ v_h &= \ln \frac{h}{h_a}, v_w = \ln \frac{w}{w_a} \end{aligned} \right\} \quad (4)$$

其中， $\{v_x, v_y, v_h, v_w\}$ 代表归一化后的位置比例值， $\{x_a, y_a, h_a, w_a\}$ 为文本提议框的位置信息值，真值的坐标位置信息 $v^*$ 也同样需要进行相同的归一化操作，操作符 $a \otimes b = a - b + k\pi$ ，通过 $k$ 值的变化可以保证夹角 $\theta$ 在 $\left( \frac{4\pi}{9}, \frac{5\pi}{9} \right)$ 范围。网络训练时通过反

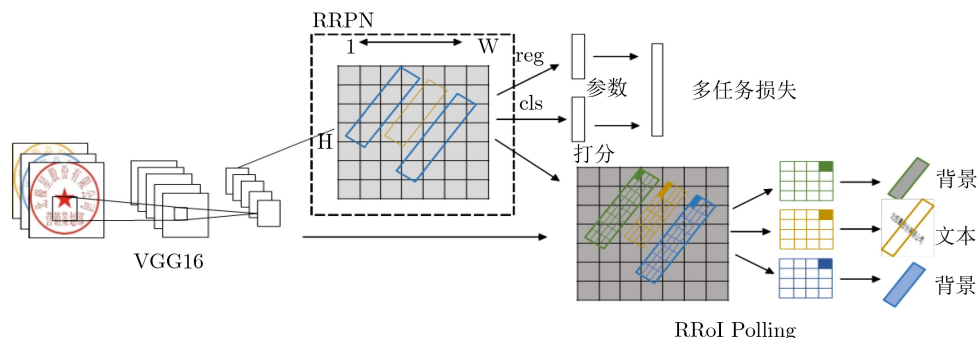


图 5 RRRPN网络

向传播算法使得 $v$ 与 $v^*$ 的差距最小,分类网络的训练是通过RRPN所产生的正负文本提议框进行的。

### 2.3 文本合并

改进后的CTPN网络的文本提议框带有一定的角度,其文本覆盖范围也包含水平方向上的变化,为了克服角度带来的影响,在合并的过程中考虑交并比的同时也应该保证角度信息。带角度信息的文本合并过程包含两类:(1)对于存在交并比(Intersection Over Union, IOU)大于0.7的文本提议框,选择最大IOU的文本提议框进行合并;(2)若文本提议框的IOU都在[0.3, 0.7]范围内,选择角度差值最小的文本提议框,并且该角度值必须小于 $\pi/12$ 。文本框合并过程如图6所示。

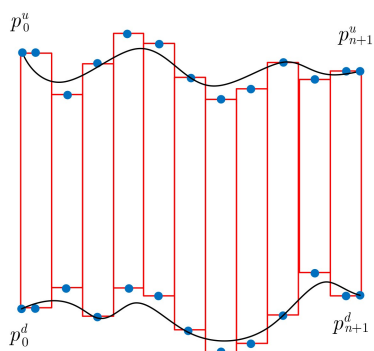


图6 贝塞尔曲线拟合文本线

假设当前文本检测框由 $n$ 个文本提议框构成,选取第1个文本区域的左上角顶点为上边界点 $p_0^u$ ,每个文本区域的上边界框的中点为当前文本区域的上边界点 $p_i^u, i \in (1, n)$ ,第 $n$ 个文本区域的右上角顶点为上边界点 $p_{n+1}^u$ 。同理选取第1个文本区域的左下角顶点为上边界点 $p_0^d$ ,每个文本区域的上边界框的中点为当前文本区域的上边界点 $p_i^d, i \in (1, n)$ ,第 $n$ 个文本区域的右上角顶点为上边界点 $p_{n+1}^d$ ,对于所有的上边界与下边界点采用3阶贝塞尔曲线进行拟合,其拟合的公式为

$$B(t) = (1-t)^3 p_i + 3(1-t)^2 t p_{i+1} + 3(1-t)t^2 p_{i+2} + t^3 p_{i+3}, t \in (0, 1] \quad (5)$$

为了验证RRPN和贝塞尔曲线在印章文字检测中的有效性,设计了关于RRPN和贝塞尔曲线的消融实验,图7是实验效果对比图。

由图7对比可以发现,加入贝塞尔曲线相对于加入RRPN能够使得文本检测更精准,而通过对比图7(c)和图7(d),可以发现RRPN的加入为之后贝塞尔曲线拟合文本区域提供了更贴合文本的文本提议框。

### 2.4 采用注意力机制的文本识别算法

本文采用CRNN网络对印章区域进行特征的提

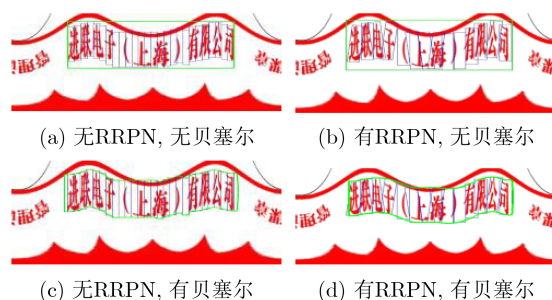


图7 RRPN与贝塞尔曲线消融实验

取,然后利用Seq2Seq对提取的特征进行解码,检测输出印章的文字内容。文本识别网络如图8所示。

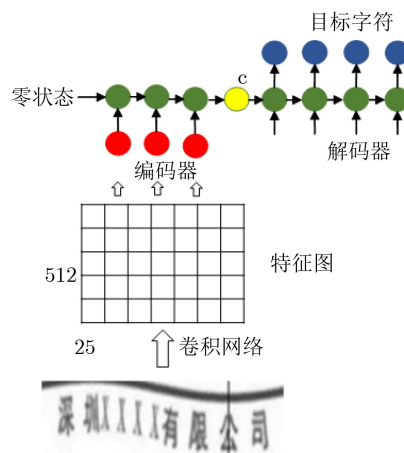


图8 文本识别网络

将提取的印章图片输入网络,得到尺寸为 $\{1 \times 25 \times 512\}$ 的特征图 $x_t, t \in [1, n]$ 。特征图 $x$ 作为Seq2Seq语义转录层的输入,通过编码器(Enc)把输出的特征图转换为一个语义特征向量 $C$ ,然后通过解码器(Dec)将语义特征向量转换为输出字符。加入Seq2Seq转录层之后的网络计算过程如下:

编码器输入由当前时刻的特征向量 $x_t$ 与上一时刻编码器输出 $h_{t-1}$ 共同决定,当前时刻编码器的输出为 $h_t$ ,计算公式为

$$h_t = f(h_{t-1}, x_t) \quad (6)$$

其中, $f(\cdot)$ 为特征图提取器,然后通过编码器将特征图中所有的向量进行汇集,融合为一个语义特征向量 $C$ ,融合的过程表示为enc,计算公式为

$$C = \text{enc}(h_1, h_2, \dots, h_t) \quad (7)$$

然后利用解码器对该语义特征向量 $C$ 进行解码,当前时刻的输出结果 $y_t$ 由该语义特征向量 $C$ 和之前时刻的所有输出共同表示,解码的过程为

$$y_t = \prod_{t=1}^T \text{dec}(y_t | y_1, y_2, \dots, y_{t-1}, C) \quad (8)$$

在上述的Seq2Seq模型中所有的输入都先转化为中间语义向量 $C$ , 然后再进行解码, 在进行编码的过程中其实每个字符的权重都是不相同的。模型能够选择性的关注输入序列有用的部分, 即注意力机制。在Seq2Seq中注意力机制表现为给每个特征图中的序列学习一个独立的权重 $a_{ij}$ 。即

$$h'_i = a_{ij} f(h'_{j-1}, x_j) \quad (9)$$

$a_{ij}$ 为编码器 $j$ 时刻的输入对 $i$ 时刻编码器输出的影响权重,  $a_{ij}$ 由式(10)和式(11)的迭代方程计算

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^T \exp(e_{ik})} \quad (10)$$

$$e_{ij} = \mathbf{W}^T \tanh(\mathbf{W}h_i + \mathbf{V}y_j + b) \quad (11)$$

其中,  $\mathbf{W}$ 为编码器网络参数,  $\mathbf{V}$ 为解码器网络参数,  $b$ 为网络偏置量,  $T$ 为输入时长。  $h_i$ 为解码器输入,  $y_j$ 为解码器的加权输出。加入关系权重后的编码层输出为 $h'_i$ 。语义特征向量 $C$ 的计算为

$$C = \text{enc}(h'_1, h'_2, \dots, h'_t) \quad (12)$$

网络可以学习不同长度的序列, 通过交叉熵损失函数进行联合训练, 训练过程中的损失函数为

$$L_{\text{Seq2Seq}} = - \sum_{t=1}^n \ln P(y_t^* | x, y_{t-1}, \dots, y_1) \quad (13)$$

其中,  $y_t^*$ 代表输出标签序列, 解码器在检测到开始符时进行解码, 检测到终止符时完成字符串的预测。

### 2.5 文本匹配

实际的印章内容检测应用过程中, 我们可以获得目标印章的先验信息, 只需保证被检测印章内容能够与预设的印章内容相近即可。加入文本匹配, 文本自动调整成与其相近的先验印章内容, 能使整个文本识别的结果更准确。文本相似度算法基于字面含义匹配算法, 常见的方法是将两个字符串通过哈希映射后, 再比较每个字符串哈希值的差异。在印章的文字信息识别过程中必须严格按照字符的字面信息进行匹配, 所以本文采用基于字面匹配的文本相似度计算方法—Simhash算法<sup>[12]</sup>。其处理步骤如下:

(1)分词: 计算每个字符在文本内容中的权重分量, 构造字符权重数组;

(2)计算字符哈希值: 利用哈希函数计算每一个字符权重数组元素的哈希值, 该哈希值由二进制表示;

(3)加权: 根据权重数组中每一个元素对应的权重计算其加权值;

(4)纵向相加: 将第3步得到的加权值纵向相加以得到加权和;

(5)降维: 将纵向相加后的加权和归一化, 大于0取1, 反之则取为0。得到最终字符串的指纹签名;

(6)相似度比较: 将上述生成的字符串指纹签名与预存字符串指纹签名进行比较, 依据距离值选出最接近的印章内容。

## 3 实验结果与分析

目前国内并没有公开的中文印章数据集, 因此, 本文采集了大量的纸质合同建立了一个中文印章数据集, 其中该数据集是本文作者建立的, 还未公开。该数据包含真实的印章360张, 电子印章340张。数据集中的每一份样本都统一缩放至 $360 \times 360$ , 并以24位深度bmp格式存储。实验阶段中, 将该数据集分成两组, 第1组300张用于训练, 第2组60张用于测试。

### 3.1 评价指标

文字检测常使用DetEval<sup>[13]</sup>评价方式计算精确率 $P$ 、召回率 $R$ 与 $F$ 值。DetEval评价方法基于真实框 $G$ 的面积与实际检测框 $D$ 的面积关系。计算方法为

$$\left. \begin{aligned} P(G_i, D_j) &= \frac{\text{Area}(G_i \cap D_j)}{\text{Area}(D_j)} \\ R(G_i, D_j) &= \frac{\text{Area}(G_i \cap D_j)}{\text{Area}(G_i)} \\ F(G_i, D_j) &= \frac{2 \cdot P(G_i, D_j) \cdot R(G_i, D_j)}{P(G_i, D_j) + R(G_i, D_j)} \end{aligned} \right\} \quad (14)$$

其中,  $G_i$ 代表实际标注的样本文字区域框,  $i = 1, 2, \dots, |G|$ 。  $D_j$ 代表网络输出的文本检测框,  $j = 1, 2, \dots, |D|$ 。只有当任意检测框与实际框的精确率与召回率大于该阈值时才认为检测框检测正确。

而对于文字识别常用的评价指标为召回率, 只有该识别字符与样本标签中的字符对应, 才认为该字符被正确检出。召回率 $r$ 计算公式为

$$r = \frac{\text{正确识别字符个数}}{\text{识别字符总数}} \quad (15)$$

### 3.2 印章内容检测效果分析

针对印章文字的检测, 我们使用相同的训练集和测试集在多种文字检测算法进行了比较, 由于本文的文字检测算法是针对印章文字的特点提出的, 所以在文字检测效果上有较大的进步, 对于印章文字区域的拟合效果最佳, 详细的实验结果如表1所示。

各类对比方法的实际效果对比如图9所示。

基于图像灰度分布的MSER算法进行文字提

表1 印章文字检测效果对比(%)

方法	召回率	准确率	F值
MSER	36.27	44.36	39.69
Faster-RCNN	36.55	40.23	37.77
CTPN	51.56	74.22	60.82
EAST	77.33	85.27	81.06
PSENET	74.24	96.07	83.76
本文方法	82.17	87.47	84.73

取时会包含大量的背景信息。而Faster-RCNN<sup>[14]</sup>算法的anchor并不适合文字细长的特点,容易使得各个文字区域割裂,并且MSER与Faster-RCNN并不包含角度信息,所以对于印章内的倒立文字无法去除。将Faster-RCNN中的文本提议

框改进后产生的CTPN网络,将文字检测分割为对细粒度的文本提议框的检测,同样不包含角度信息,造成文字检测的召回率较低。EAST方法包含角度信息,能够检测倾斜的文字,但是该方法对文字角度没有限制,容易产生一对多的情况,也容易将背景信息容纳入检测结果中。PSENET利用宽度优先搜索算法对文字区域进行扩展,可以检测任意形状的文字,因而具有很高的检测精度,然而具有较低的召回率。本文方法采用RRPN网络,对每个文本提议框融入了角度信息,并且限定了角度的范围,使得该算法能够适合印章文字呈波浪形排列的特点。同时通过贝塞尔拟合对印章的文本框线进行拟合,能够有效去除背景信息,减少印章文字检测中的干扰信息。

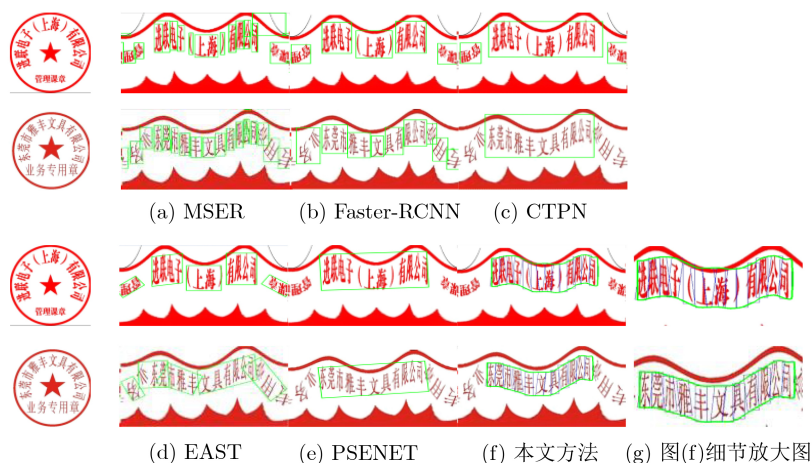


图9 文字提取效果对比

### 3.3 印章内容识别效果分析

本文所有对比方法均在本文自制的印章数据集上进行训练测试,数据集中的文字图片都经过了文字检测算法的正确检测,去除了非识别步骤带来的影响。

将CRNN中的CTC转录层替换为Seq2Seq注意力机制后,文字识别的召回率和效率都有一定的提升。使用CTC转录层的CRNN文字识别网络的召回率为71.3%,而使用Seq2Seq转录模块的文本识别网络的召回率为74.8%,并且平均识别时间几乎相同。使用Seq2Seq转录模块的文本识别网络在采用文本匹配模块后召回率由74.8%提升为84.4%,相比较CRNN识别效果提升13.1%。而本文采用Seq2Seq转录模块和Simhash文本匹配后的文字识别准确度和效率都有较大的提升。详细的实验数据对比如表2所示。

针对文本匹配模块可能会造成的文字识别缺失或者识别错误的问题,本文通过预设的字符对识

表2 印章文字识别效果对比

方法	印章数据集	
	召回率(%)	耗时(s)
CNN+RNN+CTC	71.3	0.722
CNN+RNN+Seq2Seq	74.8	0.612
CNN+RNN+CTC+Simhash	82.3	0.833
CNN+RNN+Seq2Seq+Simhash (本文)	84.4	0.640

别结果进行校正可以输出正确的结果。同时也对Simhash, BM25<sup>[15]</sup>和TF-IDF等多种文本匹配方法在相同测试集上做了比较,文本匹配算法严格依照文字的字面含义进行匹配,并且通过调整各类算法在进行文本匹配时的阈值,当采用Simhash文本匹配阈值为20时,其文字识别的召回率84.4%为最高。文本匹配对比数据如图10所示。

从表2和图10中可以看出本文方法的优势,而印章检测中所出现的错误样例大部分是由于印章在合同上的印迹颜色较浅,字迹模糊;印章识别中所

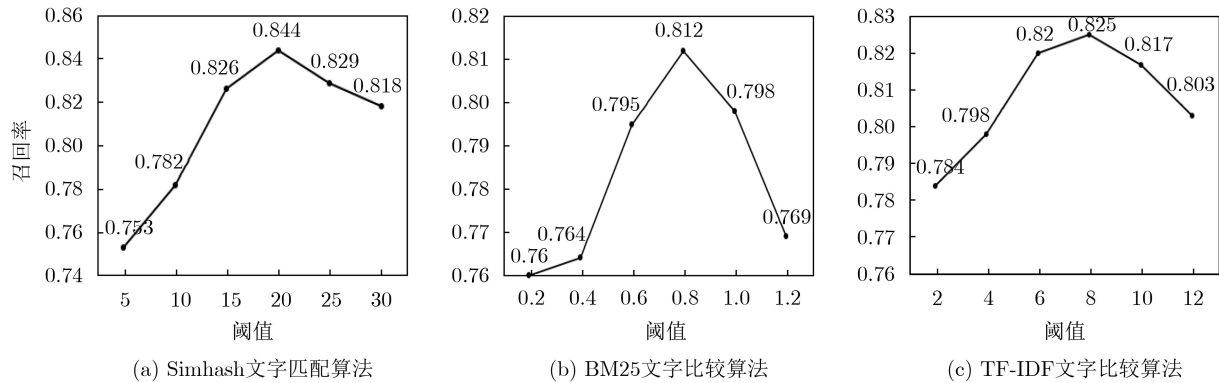


图10 文本匹配对比

出现的错误样例大部分是由于背景文字对印章的内容造成了遮盖。

#### 4 结论

根据合同内部中文印章的文字分布特点,对印章的提取、预处理、检测与识别各部分的关键技术进行了研究,结合印章文字的特点对各个部分做出了针对性的改进,本文方法的印章文字召回率为82.17%,检测准确率为87.47%,F值为84.73%,印章文字识别的召回率为84.4%。在实验平台上的平均检测耗时为640 ms。说明了本文方法在具有较好检测效果的同时也具有较快的检测速度。

#### 参考文献

- [1] LIN Han, YANG Peng, and ZHANG Fanlong. Review of scene text detection and recognition[J]. *Archives of Computational Methods in Engineering*, 2020, 27(2): 433–454. doi: [10.1007/s11831-019-09315-1](https://doi.org/10.1007/s11831-019-09315-1).
- [2] TIAN Zhi, HUANG Weilin, HE Tong, *et al.* Detecting text in natural image with connectionist text proposal network[C]. The 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016: 56–72.
- [3] SHI Baoguang, BAI Xiang, and BELONGIE S. Detecting oriented text in natural images by linking segments[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 3482–3490.
- [4] ZHOU Xinyu, YAO Cong, WEN He, *et al.* EAST: An efficient and accurate scene text detector[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 2642–2651.
- [5] LONG J, SHELHAMER E, DARRELL T, *et al.* Fully convolutional networks for semantic segmentation[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 3431–3440.
- [6] SHI Baoguang, BAI Xiang, and YAO Cong. 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 and Machine Intelligence*, 2017, 39(11): 2298–2304. doi: [10.1109/TPAMI.2016.2646371](https://doi.org/10.1109/TPAMI.2016.2646371).
- [7] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[C]. The 31st International Conference on Neural Information Processing Systems, Long Beach, USA, 2017: 5998–6008.
- [8] STAUFEMEYER R C and MORRIS E R. Understanding LSTM—a tutorial into long short-term memory recurrent neural networks[J]. arXiv: 1909.09586, 2019.
- [9] WANG Qingqing, HUANG Ye, JIA Wenjing, *et al.* FACLSTM: ConvLSTM with focused attention for scene text recognition[J]. *Science China Information Sciences*, 2020, 63(2): 120103. doi: [10.1007/s11432-019-2713-1](https://doi.org/10.1007/s11432-019-2713-1).
- [10] MA Jianqi, SHAO Weiyuan, YE Hao, *et al.* Arbitrary-oriented scene text detection via rotation proposals[J]. *IEEE Transactions on Multimedia*, 2018, 20(11): 3111–3122. doi: [10.1109/TMM.2018.2818020](https://doi.org/10.1109/TMM.2018.2818020).
- [11] 李彩林, 张清华, 陈文贺, 等. 基于深度学习的绝缘子定向识别算法[J]. *电子与信息学报*, 2020, 42(4): 1033–1040. doi: [10.11999/JEIT190350](https://doi.org/10.11999/JEIT190350).
- [12] LI Cailin, ZHANG Qinghua, CHEN Wenhe, *et al.* Insulator orientation detection based on deep learning[J]. *Journal of Electronics & Information Technology*, 2020, 42(4): 1033–1040. doi: [10.11999/JEIT190350](https://doi.org/10.11999/JEIT190350).
- [12] 杨旻, 杨书略, 柯闽. 加密云数据下基于Simhash的模糊排序搜索方案[J]. *计算机学报*, 2017, 40(2): 431–444. doi: [10.11897/SP.J.1016.2017.00431](https://doi.org/10.11897/SP.J.1016.2017.00431).
- [12] YANG Yang, YANG Shulue, and KE Min. Ranked fuzzy keyword search based on Simhash over encrypted cloud data[J]. *Chinese Journal of Computers*, 2017, 40(2): 431–444. doi: [10.11897/SP.J.1016.2017.00431](https://doi.org/10.11897/SP.J.1016.2017.00431).
- [13] KARATZAS D, SHAFAIT F, UCHIDA S, *et al.* ICDAR 2013 robust reading competition[C]. The 12th International Conference on Document Analysis and Recognition,

- Washington, USA, 2013: 1484–1493.
- [14] REN Shaoqing, HE Kaiming, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[C]. The 28th International Conference on Neural Information Processing Systems, Montreal, Canada, 2015: 91–99.
- [15] 陈乐乐, 黄松, 孙金磊, 等. 基于BM25算法的问题报告质量检测方法[J]. 清华大学学报: 自然科学版, 2020, 60(10): 829–836.
- CHEN Lele, HUANG Song, SUN Jinlei, *et al.* Bug report quality detection based on the BM25 algorithm[J]. *Journal of Tsinghua University: Science and Technology*, 2020, 60(10): 829–836.
- 肖进胜: 男, 1975年生, 博士, 副教授, 硕士生导师, 研究方向为图像与视频处理.
- 赵 陶: 男, 1996年生, 硕士生, 研究方向为图像与视频处理.
- 熊闻心: 女, 1998年生, 硕士生, 研究方向为图像处理.
- 杨 天: 男, 1996年生, 硕士, 研究方向为后台开发.
- 姚渭箐: 女, 1990年生, 博士, 技术工程师, 研究方向为文档信息处理.

责任编辑: 余 蓉