

文章编号: 2096-1618(2021)03-0280-06

复杂背景下银行卡号识别方法研究

谭诗雨¹, 杨玲¹, 师春香², 徐梓欣¹

(1. 成都信息工程大学电子工程学院, 四川 成都 610225; 2. 国家气象信息中心气象数据研究室, 北京, 100081)

摘要:针对复杂背景下的银行卡号提取及识别问题提出一种基于深度学习和传统光学字符识别(optical character recognition, OCR)方法相结合的自动提取并识别卡号的算法。算法采用改进的深度学习文本检测模型对文本内容进行检测,然后利用OCR方法对数字部分图像进行分割,最后通过改进的神经网络识别数字得到连续的银行卡号。实验结果表明,改进了卷积核的神经网络对复杂背景下的银行卡号字符识别效果有显著提升,能较好提取图像中的字符特征信息,在浅色背景数据集下准确率可达到98.87%。该系统能有效识别复杂背景下的银行卡号,平均时效约为6.3 s。

关键词:银行卡号识别;深度学习;OCR;神经网络

中图分类号:TP391.41

文献标志码:A

doi:10.16836/j.cnki.jcuit.2021.03.007

0 引言

随着互联网技术的进步,支付方式已经发生变革,各种金融类和消费类应用软件需绑定银行卡进行支付、资金转账等金融活动。使用过程中,利用采集的银行卡图像进行卡号识别可以避免手动输入卡号的失误及繁琐操作,能给人们提供极大的便利性。

对于银行卡号识别,很多学者采用不同的方法进行了研究。目前,针对自然场景下图像文本的检测有两类传统方法,一类是寻找连通域的分析方法^[1-3],另一类是基于滑窗的方法^[4-5]。近年来,深度学习方法用于文本检测成为主流,Cheng-Yang Fu等^[6]在SSD^[7](single shot multiBox detector)的基础上,利用反卷积层使检测小目标的精度得到了提高,但同时导致了检测速度下降。Zheng Zhang等^[8]提出利用全卷积网络进行文本块检测的方法。He等^[9]利用CNN(convolutional neural networks)和RNN(recurrent neural network)结合的方法进行文本检测及识别。数字分割上,目前广泛使用的是垂直投影法,其难点在于确定字符的间距阈值。在文本识别方面,部分文献^[10-12]采用卷积神经网络结构识别手写字符,刘永雪等^[13]将手写体数字识别方法应用于车牌号识别问题,但应用于不同印刷方式、制作工艺差异大以及背景图案设计多样的复杂背景下银行卡号字符识别的较少。

目前大部分移动端都带有拍照功能,但复杂背景、清晰度不高、图片角度倾斜等问题极大限制了现有传统方法识别卡号的应用场景。基于此,提出深度学习结合传统光学字符识别方法来对复杂背景下的银行卡

号进行自动检测和识别的系统。该系统调整了CTPN(connectionist text proposal network)^[14]使其更适用于银行卡号的文本检测工作;针对不同银行卡号印刷技术提出不同文本切割方法;采用改进的卷积神经网络进行数字识别。

1 数据集

进行深度学习训练需要大量的图像作为训练集,而由于银行卡持有人的隐私保护,无法采集到大量银行卡图像,利用图像增强技术来扩展银行卡号图像数据集。采集的银行卡图像信息来源于网络,约3000张完整的银行卡号图像。由于识别模型需要,对大部分图像进行人工处理标记单个字符。对系统进行测试的数据集均为完整的银行卡号图像。

针对图像采集时的背景图案的复杂程度,将银行卡图像分为浅色背景、深色背景以及强干扰背景3类。根据印刷方式的不同,黑色印刷的银行卡号在深色背景和强干扰背景下识别困难;但凹凸压印的银行卡号图像在3种背景下的识别都较为困难。原因在于凹凸压印的银行卡制作时采用烫金、烫银和仅保留背景图案的工艺,且烫金、烫银字符易受磨损,使得背景信息干扰更强。除此之外,银行卡号识别还和采集图像时的光照条件有较大相关性。

2 原理及方法

2.1 整体模型流程

提出的银行卡号识别模型主要由3部分组成:卡

号检测、字符分割以及卡号字符识别。该方法识别银行卡号的整体流程如图 1 所示。

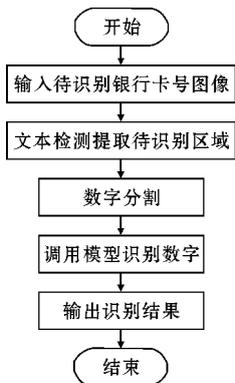


图 1 银行卡号检测及识别整体流程

2.2 文本检测

文本检测的方法较多,其中较有影响力的是 CTPN^[14],是于 ECCV2016 提出的一种文字检测算法。CTPN 结合 LSTM (long short-term memory networks) 和 CNN,可以对复杂场景中横向分布的文字进行有效的检测。CTPN 的核心思想为检测由字符、字符的一部分、多字符组成的序列的文本线。检测文本线又由检测各个单独的、固定宽度的、小的文本段后,经处理把检测到的小文本段拼接起来形成文本线。CTPN 网络结构如图 2 所示。

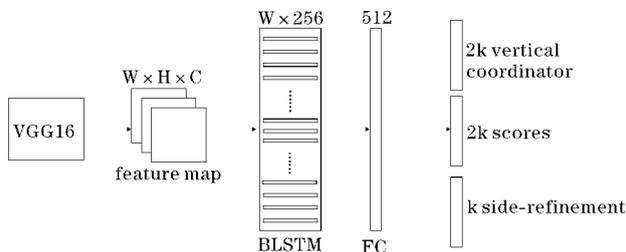


图 2 CTPN 网络结构示意图

检测银行卡号主要是检测横向文本,针对银行卡号的数字序列检测对该模型参数进行调整。分析该方法的特性以及该算法输入的银行卡号图像,锚点宽度保持不变,锚点高度由之前的 11~283 的 10 个值改为限制在 25~60 的 10 个值。将高度范围缩小是因为采集的银行卡号图像中,不同印刷方式和复杂背景影响下的卡号数字的像素点大小在 25~60,避免对不同高度的文本进行检测以抑制背景中有复杂文本干扰。再对文本连接处的参数进行调整,如加大最大水平间隙宽度,利于检测卡号中间的空格以提取完整的银行卡号区域。调整参数后的模型更适用于复杂背景下银行卡号检测任务。最后得到检测文本框的坐标值用于提取目标文本区域。

检测文本部分算法的流程主要有以下几个步骤:

(1)用 VGG16 进行特征的提取,获得的特征当做

特征向量(feature map),大小为 $W \times H \times C$ 。

(2)在上一步获得的特征向量上做窗口为 3×3 的大小滑窗。即各个窗口都可以获得一个单独的特征向量,其大小为 $3 \times 3 \times C$ 。获得的特征向量用来对 10 个锚点(anchor)间的偏移大小进行预测。

(3)把第二步所获得的特征添加至一个双向长短时记忆网络里,能够获得 $256 \times W$ 长度的输出,下一步接一个 512 全连接层。

(4)输出层主要是由 3 个输出构成的。 $2k$ 个纵向坐标(vertical coordinate)、 $2k$ 个分数(score)、 k 个水平偏移量(side-refinement)。分别表示单个锚点所对应的中心位置高度、矩形框的高度、以及各个文本段所对应的 2 个分数和文本 2 个端点的水平方向的平移量。

(5)得到各项输出后,即得到了密集预测的小文本段,采用非极大值抑制算法对多余的文本框进行过滤。

(6)采取基于图的文本行构造算法,把获得的各个单独的文本段拼接成完整的文本行。

2.3 数字分割

字符分割的整体流程图如图 3 所示。首先针对图像倾斜采用霍夫变换进行图像旋转校正,便于后面进行垂直投影分割。完成角度校正以后,需要判断提取出的图像是黑色印刷类还是压印类。将图片从 RGB 形式转变成 HSV 形式后,经过对 V 分量阈值进行设定判定不同的银行卡号印刷类型。再针对不同印刷类型的银行卡号进行不同的数字分割策略。对于黑色印刷类,若为浅色背景则使用大津阈值的方法,若为深色背景则进行掩膜垂直投影分割。针对压印类卡片,选择 YUV 空间变换法来对分割字符的干扰进行剔除。首先对旋转校正后的图像提取 Y、U、V 3 个分量的图像,之后对这 3 个分量分别采用 canny 算子做边缘检测,再对其进行字符分割。在分割过程中会出现个别字符之间的边缘粘连重叠现象,因此在进行分割时,会先提取文本框的图像,然后对提取到的图像进行粗分割,然后在粗分割基础上继续判断分割出的字符区域是否需要第二次的细分割。

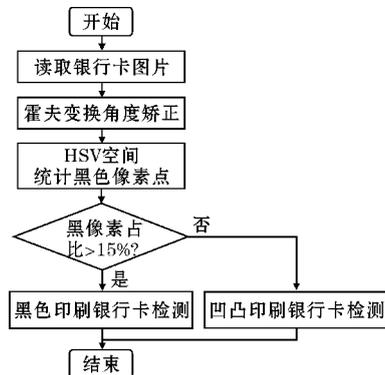


图 3 字符分割整体流程

2.4 数字识别

2.4.1 LeNetCNN 模型

采用 LeNet CNN 模型及在此基础上的改进模型,该网络模型由 LeCun 等^[15]于 1998 年提出。利用传统神经网络识别图像中的数字时,输入图像的像素点个数和输入层的神经元个数相等,需要训练大量参数。传统神经网络随着层数加深,训练的时间会呈指数增长,训练效果不佳。和传统的神经网络比较,CNN 模型除了输入层和输出层,还包含了卷积层、池化层以及全连接层。在数字识别问题中,核心任务即为提取数字图像特征,而在银行卡背景信息多样且干扰层次丰富的情况下,捕捉字符的图像特征较为困难。而 CNN 模型的优点在于能自动处理和提取图像中的高维特征,并且 CNN 的卷积核共享,可以有效地处理高维数据,快速训练,能很好地训练层数较多的网络,从而提高识别准确率。

LeNet CNN 模型的原理如图 4 所示,采用两个卷积层和两个池化层以及两个全连接层以及一个输出层。卷积操作可以增强原图像特征及降低噪声,这是特征提取的核心步骤。池化层采用平均池化,能对提取的特征图像进行抽样,降低数据量以及留存有用信息,能有效防止网络训练参数过多及模型的过拟合。经过两层“卷积层+池化层”的组合处理,提取到的特征信道增多,宽度逐渐降低。然后,将特征图像的像素点全部展开构成全连接层,全连接层和输出层之间也采用了全连接。最后,输出数字的识别分类结果。

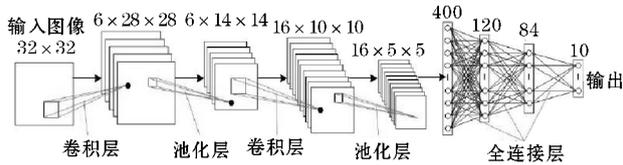


图 4 LeNet CNN 模型结构

对采集的字符样本进行分析后,将数据集分为 3 种类型:第一类是浅色背景下黑色印刷的银行卡号以及具有烫金、烫银工艺的凹凸压印银行卡号图像;第二类是深色背景干扰下的多种印刷方式的银行卡号图像;第三类是包含强干扰背景的银行卡号图像。第一类数字识别任务与 LeNet CNN 模型处理的原任务具有较高相似性;第二类背景下的数字在识别需要针对深色背景下字符提取特征,由于深色背景的干扰,字符与背景之间的对比度较小,整体亮度较低,较难捕捉字符的轮廓信息;第三类数据大多是凹凸压印的银行卡号图像,由于强干扰背景存在,在图像采集时受光照影响可能出现边缘模糊和边缘丢失以及产生多条无效边缘的情况。

2.4.2 模型改进

针对上述亟需解决的复杂问题,在 LeNet CNN 模型基础上采用扩展卷积核的方式进行改进。普通卷积神经网络对卷积核进行随机初始化,然后对误差函数后向传播,采用随机梯度下降法对权重进行迭代更替直到收敛,采用随机初始化可能会使网络陷入局部最优。有学者针对卷积核随机初始化提出改进方法,将随机初始化卷积核改进为利用 PCA (principal component analysis) 算法分层多次计算得到目标卷积层所需的卷积核组,借助该卷积核组完成隐层间神经元的映射,生成该层的特征向量。计算得到卷积核组的步骤如下:

(1) 假使输入图像大小为 $n \times n$, 且有 N 张图像 $\{I_i\}_{i=1}^N$, 卷积核的大小假设为 $k \times k$, 对于每一张图像进行全覆盖分块, 得到 $s = (n-k+1) \times (n-k+1)$ 个子图像, 排列为列向量 $x_{i,1}, x_{i,2}, \dots, x_{i,s} \in R^{k \times k \times 1}$, 即得到输入训练图像的切片样本集 $x_i = [x_{i,1}, x_{i,2}, \dots, x_{i,s}]$, 以此方法将输入的张图像进行切片变换, 得到输入样本集的分块数据为 $Y = [x_1, x_2, \dots, x_N] \in R^{k \times k \times N \times s}$ 。

(2) 将切片变换得到的 Y 送入训练网络中, 先对其进行 PCA 主成分分析, 得到输入层的特征向量组 $V_{kk} = [v_1, v_2, \dots, v_{kk}] \in R^{k \times k \times N \times s}$, 在特征向量组中选取 L 个特征向量作为输入样本集的主要特征向量, 选取的 L 个特征向量即为 C1 层的卷积核组 V_1 。选取主成分特征向量应满足以下条件:

$$\begin{cases} \min \|Y - V_1 V_1^T Y\|_F^2 \\ V_1^T V_1 = I \end{cases}$$

其中: I 是单位矩阵, V_1 是 PCA 分析得到的协方差矩阵 XX^T 中选取的前 L 个特征向量, 表征着输入图像的主成分。

(3) 按照卷积核大小 $k \times k$, 将得到的特征向量 V_1 中的每个列向量进行重新排列, 即得到对应卷积层的卷积核初始值。

(4) 对每一层卷积层得到的特征图像进行同样处理, 即可得到每一层的初始化卷积核值。

改进的最终模型结构如图 5 所示, 结构设计为三层改进 CNN 层和两层全连接层。其中, CNN 层包含一次卷积核初始化以及一次卷积、激活函数和一次最大池化层。第一层 CNN 输入为切割得到的灰度图; 第二层 CNN 输入为第一层池化完后的张量; 依此类推。卷积完成后得到四维张量的特征向量, 大小为 $\text{batch} \times 8 \times 12 \times 64$, 接下来进行全连接层搭建。全连接输入层大小为 $\text{batch} \times 8 \times 12 \times 64$, 第一层全连接层的神经元个数为 $\text{batch} \times 550$, 第二层全连接层的神经元个数为 $\text{batch} \times 220$, 导出层神经元个数为 11, 对应分类数。由于在切割时可能会将银行卡号中的空格部分切割出来, 为提升模型预测的准确性, 将空格作为一类进行训练, 故分类总数为 11。利用交叉熵衡量目标和预测值之间的

差距,Adam 优化器控制学习速度。选用 relu 激活函数相比较于传统 tanh 等激活函数在梯度消失也有更好的表现。模型结构如图 5 所示。

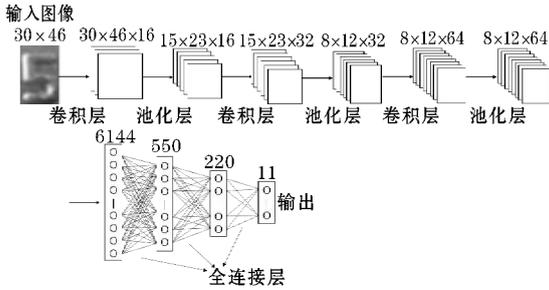


图 5 数字识别网络结构

3 实验及结果

3.1 文本检测结果

图 6 展示了采用调整参数前后的 CTPN 模型进行文本检测的效果,可以看出,调整参数后对背景中出现的垂直高度较大的文本产生了较强的抗干扰能力,能准确检测出银行卡号所在区域。且在未展示的复杂背景银行卡号中,背景中花纹类干扰较大的图像也有较好的检测效果。文本检测后,通过长宽比和像素点个数以及提取指定区域的长方形框,取出包含银行卡号的文本框。采用针对银行卡号进行调整参数后的 CTPN 模型,能极大地抑制银行卡号的复杂背景对银行卡号区域检测的干扰,约 150 张真实银行卡号测试图像中,平均 2.16 s 可检测出银行卡号区域。



(a) 调整参数前识别效果



(b) 调整参数后识别效果

图 6 CTPN 参数调整识别结果对比

3.2 字符切割结果

图 7 展示了银行卡号的分割结果。通过观察真实银行卡号图像,发现不同银行卡类型有不同格式,如 4-4-4-4、6-12、6-13 等。图 7 展示的是 4-4-4-4 类型分割结果,可以看到该分割方法对于复杂背景干扰下的银行卡号有较好的分割效果,能完整保留字符的同时,对背景影响不能彻底剔除的部分,将其切割为空格区域。

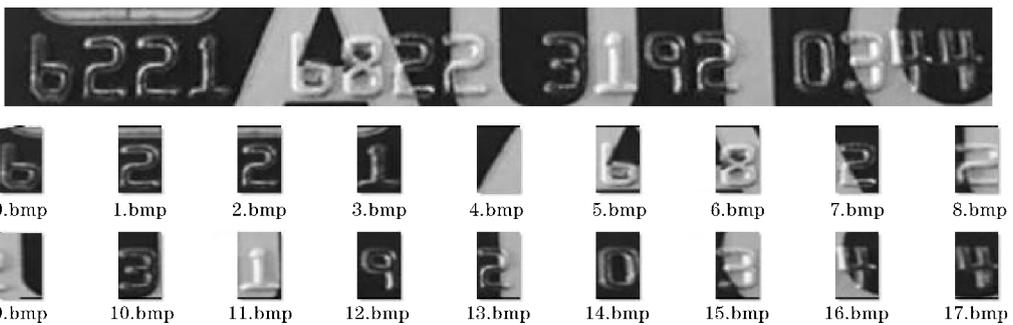


图 7 不同银行卡号的分割结果

3.3 数字识别结果

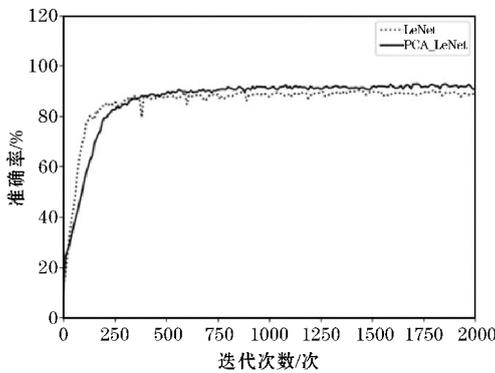
使用采集的真实银行卡数据集中的 80% 作为训练集,20% 作为测试集。其中训练集包含约 38400 张单个字符,空格部分约 6400 张,测试集中单个字符约 9600 张,空格部分约 1600 张。由于训练样本中单个字符数量不均衡,经过部分数据增强的方法使训练数据集均衡分布。将其分为三类背景干扰数据集,浅色背景、深色背景和强干扰背景的占比约为 1.5 : 1 : 1。

不同类型复杂背景干扰下的数据集采用同一个网络结构,但每个数据集对应的模型中的参数是独立的。由表 1 可知,利用 PCA 算法进行初始化卷积核值的 PCA_LeNet 模型的准确率在所有训练数据集上均有提高。浅色背景数据集上的准确率达 98.87%,相较于 LeNet 网络结构有较大提升,说明采用该方法对原网络进行改进可以较好应用在银行卡号识别任务上。在深色背景和强干扰背景数据集上的模型准确率也有提升,说明该模型能在复杂背景干扰下提取更精确的目标特

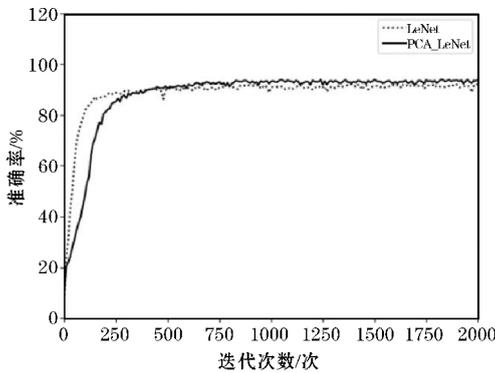
征。图7展示了不同数据集下 LeNet 模型和 PCA_LeNet 模型训练次数和准确率之间的关系。由图7可知,PCA_LeNet 模型很好地避免了训练过程中准确率的大幅波动,表现较为平稳,相对于原模型训练过程中出现的波动更少。由于改进模型中多了一层卷积,收敛速度较原模型慢,但是仍然可以平稳收敛且准确率有所提升。

表1 不同类型复杂背景干扰数据集下模型准确率对比 单位:%

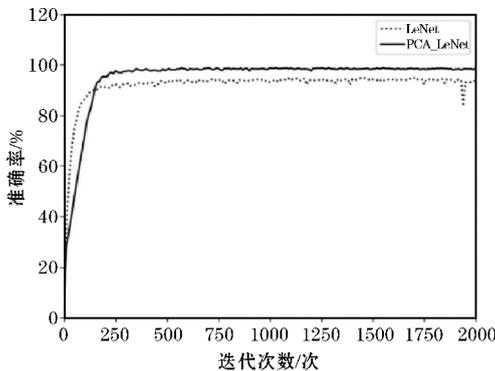
	强干扰背景	深色背景	浅色背景
LeNet 准确率	90.11	91.59	93.85
PCA_LeNet 准确率	92.66	94.15	98.87



(a) 强干扰背景



(b) 深色背景



(c) 浅色背景

图7 不同复杂背景数据集下训练迭代次数和准确率之间的关系

3.4 系统时效统计

对整个系统进行银行卡号识别时效测试,从真实银行卡号图像测试数据(约200张)统计结果来看,该系统从输入银行卡号图像到识别出银行卡号平均每张用时6.3 s左右。对于银行卡号的整体识别是有效的,但由于不能使字符切割百分百准确,使整个银行卡号的识别结果不理想,会出现个别空格识别出错或者部分训练数据集较少且复杂背景干扰极大的单个字符识别不准确。从表2的统计结果来看,采用深度学习和神经网络训练的模型在识别时用时较少,而系统在字符切割模块的时效性还有提高的空间。

表2 卡号识别系统的各模块时效 单位:s

子模块	时效
文本检测	2.16
字符切割	3.74
数字识别	0.4

4 结论

该方法综合运用了深度学习、图像处理、计算机视觉等多个领域的技术。针对复杂背景下银行卡号的识别,提出一种结合深度学习的识别方法,有效提升了识别准确性。主要工作如下:

(1)改进文本检测模型参数使其更加适用于检测银行卡号并极大抑制图像中复杂背景造成的干扰信息。

(2)针对文本检测得到的银行卡号目标图像进行预处理,实现自动旋转校正、边缘检测、背景花纹消除、提取掩膜等特定功能。

(3)采用PCA算法对经典LeNet神经网络的卷积核进行改进,提升了其在复杂背景下银行卡号数字的识别率。

通过多次实验及统计结果,该方法能较好地解决复杂背景下银行卡号识别问题。但在许多方面仍有不足,需要进一步改进和完善,主要是由于银行卡图像数据的特殊性、相关实验数据集数量有限。虽然通过图像增强算法能扩展数据集提升部分模型性能,但系统的识别效果还需要大量的真实银行卡图像进行验证。字符识别中,除了考虑图像采集导致的形变及倾斜,对于含有磨损等不同干扰的真实银行卡仍存在少许识别错误,需进一步研究。

参考文献:

- [1] 易尧华,申春辉,刘菊华,等. 结合 MSCRs 与 MSERs 的自然场景文本检测[J]. 中国图象图形学报,2017,22(2):154-160.
- [2] 蒋人杰,戚飞虎,徐立,等. 基于连通分量特征的文本检测与分割[J]. 中国图象图形学报,2006(11):1653-1656.
- [3] 徐婷. 图像文本检测与识别[D]. 北京:北京邮电大学,2017.
- [4] Lee J J, Lee P H, Lee S W, et al. Adaboost for text detection in natural scene[C]. 2011 International Conference on Document Analysis and Recognition. IEEE, 2011:429-434.
- [5] Wang K, Babenko B, Belongie S. End-to-end scene text recognition[C]. 2011 International Conference on Computer Vision. IEEE, 2011:1457-1464.
- [6] ChengYang Fu, Wei Liu, AnanthRanga, et al. Dssd:Deconvolutional single shot detector[C]. International Conference on Computer Vision Systems, 2017.
- [7] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]. European conference on computer vision. Springer, Cham, 2016:21-37.
- [8] Zheng Zhang, Chengquan Zhang, Wei Shen, et al. Multi-oriented text detection with fully convolutional networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016:4159-4167.
- [9] He P, Huang W, Qiao Y, et al. Reading Scene Text in Deep Convolutional Sequences[J/OL]. <http://arxiv.org/abs/1506.04395> arXiv: 1506.04395, 2015.
- [10] 张彤,肖南峰. 基于 BP 网络的数字识别方法[J]. 重庆理工大学学报(自然科学版), 2010(3):12.
- [11] 王娜,胡超芳. 基于客观聚类的手写数字识别方法[J]. 复杂系统与复杂性科学, 2019, 16(2):77-84.
- [12] 涂亚飞. 银行卡号字符的分割与识别算法研究[D]. 北京:北京交通大学, 2017.
- [13] 刘永雪,李海明. 卷积神经网络的优化在车牌号识别上的运用[J]. 上海电力大学学报, 2020, 36(4):351-356.
- [14] Tian Z, Huang W, He T, et al. Detecting Text in Natural Image with Connectionist Text Proposal Network[C]. European Conference on Computer Vision. Springer, Cham, 2016:56-72.
- [15] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.

Bank Card Number Recognition System under the Complex Background based on Deep Learning

TAN Shiyu¹, YANG Ling¹, SHI Chunxiang², XU Zixin¹

(1. College of Electronic Engineering, Chengdu University of Information Technology, Chengdu, 610225, China; 2. Meteorological Data Research Office, National Meteorological Information Center, China Meteorological Administration, Beijing, 100081, China)

Abstract: Aiming at the problem of bank card number extraction and recognition in a complex background, this paper proposes an algorithm for automatic extraction and recognition of card numbers based on the combination of deep learning and traditional optical character recognition(OCR) methods. We use the algorithm of an improved deep learning text detection model to detect the text content, then we use the OCR method to segment the digital part of the image, and finally use an improved neural network to recognize the number to obtain a continuous bank card number. The experimental results show that the neural network with improved convolution kernel can significantly improve the character recognition of bank card numbers in complex backgrounds and can better extract the character feature information in the image, and the accuracy can reach 98.87% under the light background dataset. The system can effectively identify bank card numbers in complex backgrounds, and its average time limit is about 6.3 s.

Keywords: bank card number recognition; deep learning; OCR; neural network