

基于 Lenet-5 网络组合特征融合的水表数字识别

邱亮^a,涂俊翔^a,禹杰^b

(福州大学 a. 机械工程及自动化学院; b. 机电工程实践中心, 福建 福州 350108)

摘要:为有效识别水表数字区域的半字符和提高数字识别准确率,设计一种基于 Lenet-5 网络组合特征融合的水表数字识别算法。对水表原始图像进行预处理,获取完整字符和半字符数字组成的数据集;在 TensorFlow 深度学习框架下搭建卷积神经网络,选取卷积核为 3×3 、 1×3 、 3×1 的卷积层作为改进网络的卷积层,增加卷积层的层数,并对非对称卷积前后的组合特征进行融合,提高网络对半字符的识别能力;用图像集训练和检测改进网络的识别性能。实验结果表明,迭代次数为 10 000、学习率为 0.001 5 的改进网络模型在测试集上识别准确率达到 99.025%,识别准确率高于原 Lenet-5 网络和一些典型的水表数字识别算法。

关键词:半字符;数字识别;Lenet-5 网络;组合特征融合

中图分类号:TP391.4 **文献标识码:**A **文章编号:**1671-5276(2020)06-0189-04

Digital Recognition of Water Meter Based on Combined Features Fusion of Lenet-5 network

DI Liang^a, TU Junxiang^a, YU Jie^b

(a. Faculty of Mechanical Engineering and Automation; b. Mechanical and Electrical Engineering Practice Center, Fuzhou University, Fuzhou 350108, China)

Abstract: To effectively identify semi-characters of water meter in the digital region and improve the recognition accuracy, an identification algorithm based on combined features fusion of Lenet-5 network is proposed for recognition of water meter digits. The original image of the water meter is preprocessed to obtain the images of complete figures and half-digits, which are formed into the data set. Then, the convolutional neural network is built in TensorFlow deep learning framework. The number of convolutional layer is increased, the convolutional layers with convolution kernel of 3×3 , 1×3 , 3×1 are applied as the convolutional layers of improved network and the combined features before and after asymmetric convolution are fused for enhancing the recognition ability of network in half-characters. The performance of improved network is tested on the image set. The experimental results show that the improved network model with 10 000 iterations and learning rate of 0.001 5 accuracy achieves 99.025%, which is better than the original network and the current typical methods for digital identification of the water meter.

Keywords: semi-character; digital recognition; Lenet-5 network; combined feature fusion

0 引言

字符识别是自动抄表系统的核心技术之一。近年来,随着深度学习的发展,研究人员已经提出了多种基于卷积神经网络的字符识别算法,识别性能得到不断改善。卷积神经网络的优势在于卷积层具有良好的自学习能力,并且识别的泛化能力优于传统方法。

Lenet-5 是由 LECUN Y^[1] 提出的一个解决手写体数字识别问题的卷积神经网络,结构简单且识别性能良好。此后研究人员对原网络进行不断地改进,提高识别的准确率和速度,并且还将改进的 Lenet-5 网络应用于印刷体数字^[2-3]、汉字^[4]、单词^[5]以及植物^[6]等识别问题中。

水表数字的识别属于印刷体数字识别的一种,其特点是随着字轮的转动会在表盘数字区域出现半字符,而不完整数字的存在增加了识别的难度。目前,已有研究人员提

出基于 Lenet-5 网络的印刷体数字识别算法,例如街景门牌号^[2]、车牌号码^[3]的识别,但未在水表数字识别中获得有效的应用。相关学者对水表数字识别的研究主要是采用模板匹配^[7]、特征匹配^[8]以及神经网络^[9-10]等方法。高菊等^[7]提出基于模板匹配与特征匹配相结合的水表数字识别算法,有效提高了识别精度,并且通过上下半字连续的待识别图像与变字长模板进行匹配,实现半字符数字的识别。但文中算法过于繁琐,降低了识别效率。尹远等^[8]利用方向梯度直方图(histogram of oriented gradient, HOG)提取数字的特征信息,再用训练好的支持向量机(support vector machine, SVM)进行分类识别。但文中未对表盘出现的半字符数字进行识别研究,难以获得有效的应用。王长旭等^[10]提出基于神经网络的水表数字识别方法。通过两个卷积神经网络分别实现对数字完整字符和半字符的识别,识别准确率达到 98%。但是对数字完整字符和半字符的识别需要用不同的神经网络,过程较为冗杂。

基金项目:福州市科技局项目资助(2017-G-71, 2018-G-57)

第一作者简介:邱亮(1994—),男,辽宁沈阳人,硕士研究生,研究方向为深度学习和图像处理。

本文在现有研究基础上,为进一步改善水表半字符数字的识别问题,提高水表数字的识别精度,设计一种基于 Lenet-5 网络组合特征融合的水表数字识别算法。

1 卷积神经网络的基本原理

卷积神经网络(convolutional neural network, CNN)是具有一定深度的前馈神经网络,在训练过程中通过监督学习调整网络参数,实现对图像分类或图中目标的检测。CNN 的特点是通过卷积获取图像的局部特征,并在同一特征图内共享卷积参数,因此可以有效地减少网络参数,加速网络的学习。CNN 的基本结构主要包括输入层、卷积层、池化层、全连接层及输出层。

CNN 的输入通常为待识别的图像,尺寸为 $h \times w \times d$,其中 h 为图像高度, w 为图像宽度, d 为图像的通道数,灰度图像 d 为 1,彩色图像 d 为 3。

卷积层由多个卷积核组成,通过卷积核在输入图像上的滑动,提取图中的不同特征。其中卷积核的尺寸决定了输出特征图中每个像素点对应输入图像上映射区域的大小,该映射区域称为感受野。卷积计算公式为

$$x_{ij} = f(W_{ij} \otimes x_{i-1} + b_{ij}) \quad (1)$$

其中: x_{i-1} 表示上层特征图; x_{ij} 表示卷积核 j 对应的输出特征图; W_{ij} 表示第 j 个卷积核的权重参数; b_{ij} 表示偏置值; $\otimes$ 表示卷积运算; f 表示激活函数。

池化层对特征图进行信息过滤,实现缩减特征图的尺寸^[11],提高运算效率,增强特征的鲁棒性。极大池化是常用的池化方法,以局部接受域内像素的最大值表示该区域。池化计算公式如下:

$$x_i = P_{\text{ooling}}(x_{i-1}) \quad (2)$$

其中 P_{ooling} 表示池化函数。

全连接层对提取到的特征进行非线性组合,整合所有的局部特征信息。对于图像分类问题,通过归一化指数函数进行分类,由输出层输出不同类别的概率值,概率值最大的即为图像的分类标签。

2 水表数字识别的实现

本文提出的基于 Lenet-5 网络组合特征融合的水表数字识别算法主要包括原始图像预处理、神经网络的训练和测试,算法流程如图 1 所示。

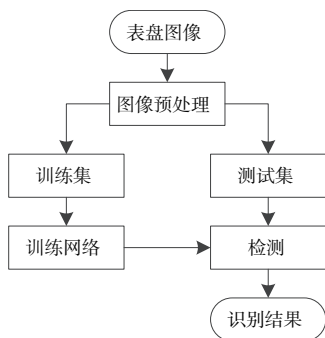


图 1 算法流程图

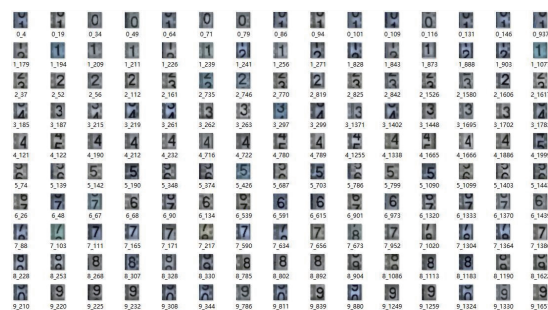
2.1 表盘图像预处理

本文以水表机芯代替水表,便于获取表中不同数字的图像,机芯表盘如图 2 所示。原始表盘图像可能存在噪声、污损等现象,而且数字区域周围的阴影以及数字间高低位置的不同,造成数字分割和二值化过程较为复杂,处理不当会在图中产生难以去除的噪声或数字结构的缺失,影响后续识别的准确性。因此本文为了简化表盘图像预处理,采用原始彩色图像作为数据集对卷积神经网络进行训练和测试。

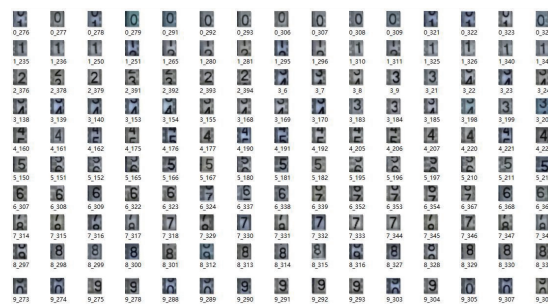


图 2 表盘图像

首先采用 Hough 圆检测提取原始图像中的表盘圆,将表盘与背景分割,然后通过 Hough 直线检测提取数字区域边框,根据边框顶点的坐标从表盘圆分割出数字区域。由于数字滚轮宽度一致,采用等分策略分离数字,并对字符图像进行仿射变换实现数据扩充,增加数据集的多样性,提高网络的鲁棒性和泛化能力。半字符图像表示流量示数介于二者之间,故以图中较小数字命名图像(9、0 以 9 进行命名)。最后将字符图像的尺寸统一调整为 32×28 ,部分彩色字符图像数据集如图 3 所示。



(a) 训练集



(b) 测试集

图 3 彩色图像集

2.2 改进网络模型的搭建

改进的网络模型如图4所示。输入层为字符图像,图像维度为 $32 \times 28 \times 3$;卷积层有5个,分别采用尺寸为 3×3 、

1×3 、 3×1 的卷积核进行特征提取;池化层采用尺寸为 2×2 的窗口进行特征降维;全连接层输入节点个数为660,输出节点个数为10,分别对应数字0~9。池化层输出经过线性整流激活函数(rectified linear unit, ReLU)去线性。

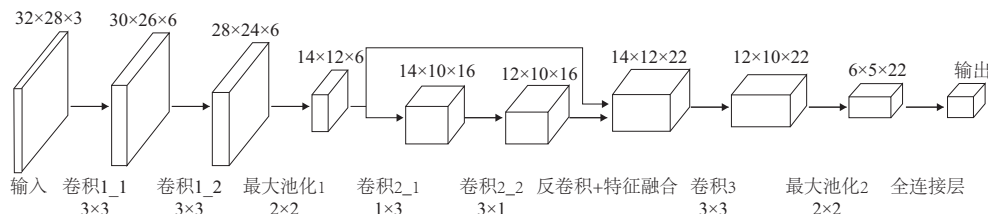


图4 改进的网络模型

根据文献[12-13]所述,采用小尺寸的卷积核可以有效减少训练参数,提高计算效率,因此本文以两个卷积核为 3×3 的卷积层替换原网络的第一卷积层。如图5所示,替换后,特征图中各点对应的感受野大小不变,而且前两层卷积核均为6个,卷积1_2层的卷积核深度为6,获取的特征图中的组合特征信息更加丰富。

此外,文献[13]还指出当特征图的尺寸在12~20之间时,用非对称卷积核代替 $n \times n$ 的卷积核可以进一步减少训练参数,同时保证网络的识别性能。本文在第一次池化后采用 1×3 和 3×1 的卷积层替换原网络第二卷积层,每层卷积核为16个。如图6所示,替换后卷积生成的特征图中每个像素点对应的感受野减小至 3×3 ,像素点包含的全局信息减少而局部细节增多,有利于获取更多的数字形状特征,保证网络准确地识别半字符数字。

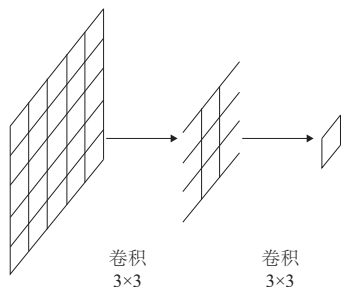


图5 3×3 卷积替换 5×5 卷积

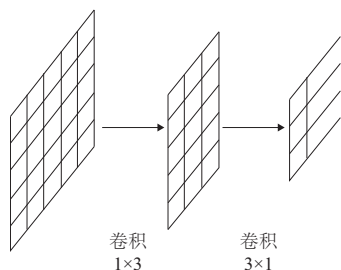


图6 非对称卷积替换 3×3 卷积

本文在网络结构的调整中发现,错误识别的图像多数为半字符图像,主要由于数字半字符的尺寸较小以及数字间部分结构的相似性所引起。在目标检测网络中常采用多尺度的特征图融合方法提高对小目标检测的准确性,本

文对非对称卷积前后的组合特征进行串接融合,并将融合特征图再次通过 3×3 的卷积层提取特征,提升网络对半字符的识别能力。

3 实验与结果分析

3.1 实验准备

本文算法实验环境为:Ubuntu操作系统、Python语言、Opencv等科学计算包,在TensorFlow深度学习框架下搭建神经网络。通过预处理得到24 000张彩色字符图像,其中完整字符和半字符图像各12 000张,每个数字约2 400张图像,并且按照5:1的比例分成训练集和测试集。训练过程中,学习率为0.001 5,采用Adam算法优化训练参数,迭代次数为10 000。

3.2 结果分析

本文通过训练过程交叉熵损失函数的收敛性和测试阶段识别的准确性评估网络的性能。在对原网络的特征提取层进行优化后,训练参数大大减少。从图7中可以看出,改进网络的损失函数收敛更快,迭代2 000次左右即可收敛至0值附近,并且随着迭代次数的增加,损失函数稳定趋于0值。

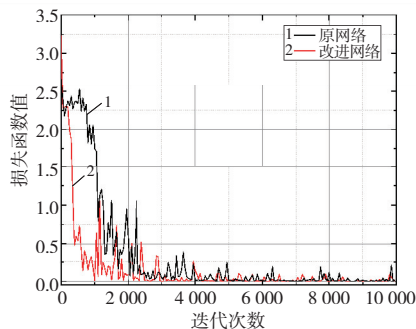


图7 损失函数

表1为原网络与改进网络的测试结果。原网络对完整数字和半字符数字均产生错误识别,而改进后的网络错误识别均为半字符图像。图8为改进网络两种典型的错误识别图像,图8(a)是将示数为9的图像识别为3,此类错误产生的

原因可能是网络的迭代次数不足,本文将该类错读图像用迭代次数为 11 000 的网络重新进行识别,识别结果全部正确;图 8(b)是将示数为 3 的图像识别为 4,其中数字 4 处于图像中间偏下位置,表明当前示数介于 3~4 之间,但图中上半字符结构较小,难以区分,从而引起错读。

表 1 Lenet-5 与改进网络的测试结果

模型	正确识别个数	错误识别个数
Lenet-5	3 885	115
本文算法	3 961	39

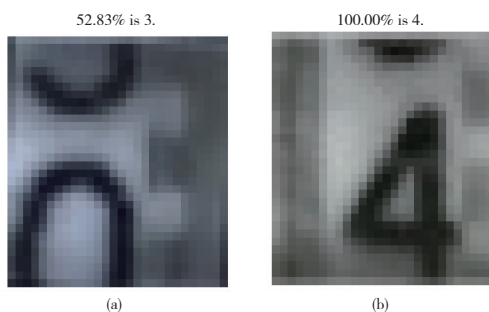


图 8 错误识别的图像

表 2 为本文算法与一些典型水表数字识别算法^[7-9]在识别性能上的比较。改进网络与原网络对水表数字的识别时间基本相同,但网络改进后识别准确率获得明显提高;模板匹配和特征匹配所设置的模板和特征个数少于 CNN 模型从数据集中提取的特征数目,BP 神经网络由于节点数和层数较少,无法获取图中较深层的特征,故识别准确率均低于本文算法。

表 2 不同识别方法的准确率和识别时间

识别方法	准确率/%	识别时间/ms
本文算法	99.025	2.107
Lenet-5	97.125	3.524
模板匹配	97.500	—
特征匹配	96.400	—
BP 神经网络	95.000	—

4 结语

本文设计的基于 Lenet-5 网络组合特征融合的水表

数字识别算法,采用字符原图进行训练,对原网络的卷积层进行优化,并且通过卷积特征图的融合提高识别精度,保证有效地识别半字符。实验表明,改进网络的检测效率高,迭代次数为 10 000、学习率为 0.001 5 的改进网络模型识别准确率达到 99.025%,识别精度明显高于其他水表数字识别算法,有效提高了自动抄表技术的可靠性。

参考文献:

- [1] 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.
- [2] 马苗,陈芳,郭敏,等. 基于改进 LeNet-5 的街景门牌号码识别方法 [J]. 云南大学学报(自然科学版), 2016, 38(2): 197-203.
- [3] 董峻妃,郑伯川,杨泽静. 基于卷积神经网络的车牌字符识别 [J]. 计算机应用, 2017, 37(7): 2014-2018.
- [4] LIN D Z, LIN F, LV Y, et al. Chinese character CAPTCHA recognition and performance estimation via deep neural network [J]. Neurocomputing, 2018, 288: 11-19.
- [5] SUEIRAS J, RUIZ V, SANCHEZ A, et al. Offline continuous handwriting recognition using sequence to sequence neural networks [J]. Neurocomputing, 2018, 289: 119-128.
- [6] LÓPEZ-JIMÉNEZ E, VÁSQUEZ-GÓMEZ J I, SANCHEZ-ACEVEDO M A, et al. Columnar cactus recognition in aerial images using a deep learning approach [J]. Ecological Informatics, 2019, 52: 131-138.
- [7] 高菊,叶桦. 一种有效的水表数字图像二次识别算法 [J]. 东南大学学报(自然科学版), 2013, 43(增刊 1): 153-157.
- [8] 尹远. 水表数字识别的研究与实现 [D]. 武汉: 华中科技大学, 2016.
- [9] 李云栋,孙长庆. 基于 BP 神经网络的实时水表自动识别系统的研究 [J]. 计算机测量与控制, 2013, 21(5): 1330-1332.
- [10] 王长旭. 水表读数识别系统的设计与实现 [D]. 北京: 北京邮电大学, 2018.
- [11] 张慧,肖蒙,崔宗勇. 基于卷积神经网络的 SAR 目标多维度特征提取 [J]. 机械制造与自动化, 2017, 46(1): 111-115.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Computer Science, 2014, 32(74): 2965-2978.
- [13] SZEGEDY C, VANHOUCHE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]// Computer Vision and Pattern Recognition. 2016: 2818-2826.

收稿日期: 2019-10-21