

基于投影法和卷积神经网络的手写汉字图像分割研究 *

张 莉¹, 孟范泽¹, 刘思霖¹, 冯 锐¹, 王 钢², 蔡 靖¹

(1. 吉林大学 仪器科学与电气工程学院, 吉林 长春 130026; 2. 北华大学, 吉林 吉林 132013)

摘要: 为提高手写汉字的识别率, 针对手写汉字的有效分割, 建立了卷积神经网络手写汉字体识别模型, 并对投影法和轮廓检测法的适用性进行了对比分析。实验结果显示, 相较于轮廓检测法, 投影法更适用于手写汉字识别中对文字图像的处理工作, 可以实现对所需文字的有效切分, 同时简化手写汉字识别网络的设置并提高识别准确率。

关键词: 手写汉字识别; 神经网络; 投影法; 轮廓检测法

中图分类号: TN919.82

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.201085

中文引用格式: 张莉, 孟范泽, 刘思霖, 等. 基于投影法和卷积神经网络的手写汉字图像分割研究[J]. 电子技术应用, 2021, 47(11): 73–75, 80.

英文引用格式: Zhang Li, Meng Fanze, Liu Silin, et al. Application of projection method in the recognition of handwritten Chinese characters[J]. Application of Electronic Technique, 2021, 47(11): 73–75, 80.

Application of projection method in the recognition of handwritten Chinese characters

Zhang Li¹, Meng Fanze¹, Liu Silin¹, Feng Rui¹, Wang Gang², Cai Jing¹

(1. College of Instrumentation & Electrical Engineering, Jilin University, Changchun 130026, China;

2. Beihua University, Jilin 132013, China)

Abstract: In order to improve the recognition rate of handwritten Chinese characters and for the effective segmentation of handwritten Chinese characters, this paper established a handwriting Chinese font recognition model based on convolutional neural network, and compared and analyzed the applicability of projection method and contour detection method. The experimental results show that, compared with the contour detection method, the projection method is more suitable for the processing of characters and images in handwritten Chinese character recognition, which can realize the effective segmentation of required characters, simplify the setting of handwritten Chinese character recognition network and improve the recognition accuracy.

Key words: handwritten Chinese character recognition; the neural network; projection method; contour detection method

0 引言

随着科技的发展以及人们日常生活工作中对手写汉字识别的需求与日俱增, 精确识别手写票据、手写试卷以及档案信息表等文件中的手写汉字, 将会为社会带来极大的便利。然而, 汉字类别繁多, 字形结构复杂, 一直是手写字体识别中的难点和热点^[1], 且个人手写汉字字体特点也不尽相同^[2]。从文献[3]可以看出, 随着所需识别汉字的数量以及神经网络复杂程度的提升, 相应的计算时间也会呈指数形式提升^[4-5]。由此可见, 实现对手写汉字图像的有效分割, 将会减少手写汉字的识别量, 相应地也降低了手写汉字识别的复杂度与计算时间。

为了达到精准分割、有效识别的目的, 本文建立了卷积神经网络(Convolutional Neural Networks, CNN)手写汉字识别模型。对投影法和轮廓检测法的适用性进行了

对比分析, 通过实验对投影法在手写汉字识别中的适用性进行了验证。

1 投影法

图像是由像素组成的, 一幅 $M \times N$ 个像素的数字图像, 各像素点的灰度值构成 M 行、 N 列的矩阵^[6]。图像被二值化后, 有色像素点的灰度值变为 1, 无色像素点灰度值为 0。二值图的灰度矩阵 A 如式(1)所示:

$$A = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ f_{21} & f_{22} & \cdots & f_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{MN} \end{bmatrix} \quad (1)$$

其中, $f(i, j) = \begin{cases} 1, & \text{有色像素点} \\ 0, & \text{无色像素点} \end{cases}, i=1, 2, \dots, M; j=1, 2, \dots, N$ 。

首先需要对文字图像进行二值化处理, 二值图像为

* 基金项目: 吉林大学“大学生创新创业训练计划”创新训练项目(201910183696); 吉林省教育厅“十三五”科学技术项目(JJKH20200964KJ); 吉林省科技发展计划项目(20190303038SF, 20190303043SF)

白底黑字^[7-8]。投影法基于这种二值图像的特点来对图片分割。用式(2)对每行的像素点进行横向行投影,求得每行的黑色像素点数记为 Sum(*i*)。

$$\text{Sum}(i) = \sum_{j=1}^N f(i, j), \quad i=1, 2, \dots, M \quad (2)$$

其中,Sum(*i*)中的首个非零元素即为文字图像的上边界,行号记为 Row(*k*)=*i*,*k*=1。从 Sum 中的 *i* 号元素起对 Sum 中的元素遍历,并对 *x* 号元素及 *x*-1 号元素进行异或逻辑运算,结果每出现一次 1,在 Row(*k*+1) 中记录 *x* 的值,并将 Sum(*i*) 中最后一个不为 0 的元素行号记录到 Row 中的最后一个元素,此为文字图像的下边界。遍历结束后,根据 Row 中记录的行号,对图像进行分割并输出,即可分割出文字图像中的每行文字。

基于同样的原理,对分割好的每行文字图像进行纵向投影,并重复上述步骤,即可得到每行文字图像中的每个文字图像。

2 对比分析

2.1 基于 CNN 卷积神经网络的手写汉字识别模型

本文使用基于 MATLAB 的 DeepLearning toolbox CNN 卷积神经网络神经网络模型,针对分割算法的适用进行了研究^[3], LeNet-5 经典 CNN 网络架构原理图如图 1 所示。

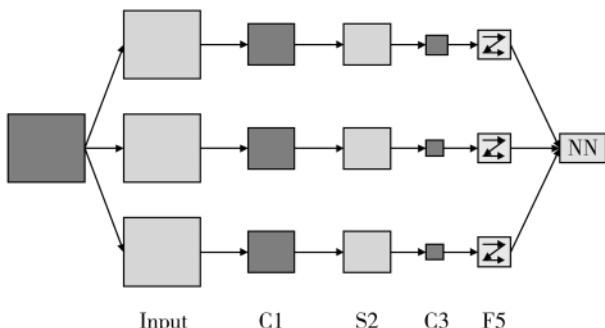


图 1 LeNet-5 经典 CNN 网络架构原理图

在图 1 中,Input 为输入图像;C1、C3 为卷积层,其功能是对输入数据进行特征提取,每一个神经元都进行局部的特征提取,在更高层将其整理合并;S2 为采样层(池化层),池化层对卷积层提取出来的特征图进行池化处理,可以在减少数据处理量的同时保留有用信息,即对图像由高分辨率向低分辨率的转换,通常卷积层池化层交替分布,使得特征图数目逐渐增多,分辨率逐渐降低;F5 为全连接层,全连接层在卷积神经网络中可以看作是一个“组合器”,把卷积层提取的局部特征重新通过权值矩阵组装起来,形成完整的图像^[9]。

针对手写体汉字的识别,本文采用的是 MATLAB toolbox 中 nnet/cnn 目录下的内置函数“Make predictions on data with network”,其运行结果界面如图 2 所示,原图像与其识别结果均显示在界面图中。

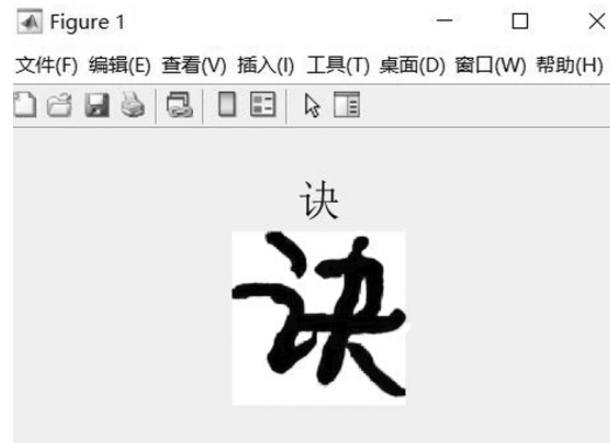


图 2 识别结果界面

2.2 对比分析

为了验证投影法的适用性,本文选用轮廓检测法^[10-11]与投影法进行对比分析。轮廓检测法的基本原理是对二值化后的图片进行顶部和底部轮廓检测,获得文字高度。之后分别进行轮廓线的凹检测和凸检测,记录凹、凸轮廓处的位置,确认是凹或凸轮廓后,对图片进行分割,并保存为 BMP 图片。本文选用如图 3 所示大小为 972×443 的手写汉字图片作为原始图像。

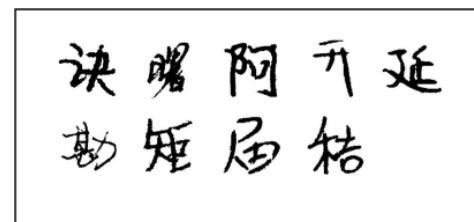


图 3 原始图像

以图 3 中的第一行文字为需要识别的目标汉字,分别使用投影法和轮廓检测法对行分割结果进行列分割,得到目标汉字图像,并对目标汉字图像进行归一化处理^[12-14],送入 CNN 神经网络模型进行识别,流程框图如图 4 所示。

原始图像图 1 使用投影法行分割后显现大小为

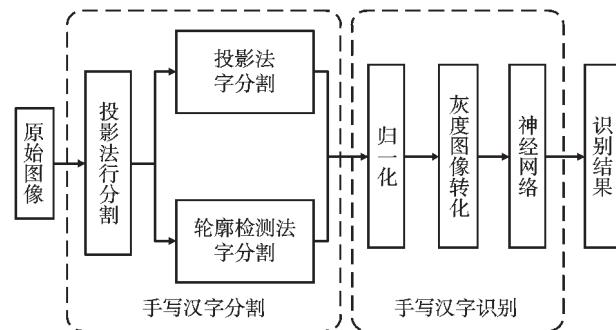
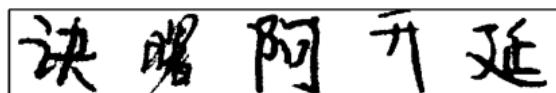


图 4 手写汉字识别流程框图

972×150的第一行图像(如图5(a)所示)、大小为972×150的第二行图像(如图5(b)所示)。



(a) 分割出的第一行图像



(b) 分割出的第二行图像

图5 分割结果

采用第一行文字图像(图5(a))进行列分割,得到每个汉字的图像如图6所示。



列 1-1 列 1-2 列 1-3 列 1-4 列 1-5

图6 分割出的第一行图像

从原始图像(图3)中可以看到,第一行文字从左至右依次为“诀、曙、阿、开、延”,图像大小依次为138×150、103×150、129×150、98×150、144×150。观察图6分割结果,投影法准确地分割出了原始图像第一行的“诀、曙、阿、开、延”5个汉字。将图6所示的5个汉字图像逐个进行归一化处理并转化为位深度为8的灰度图像,送入CNN卷积神经网络模型进行识别,识别结果如图7所示。



图7 神经网络识别结果图

从图7识别结果可以看出,分割出的“诀、曙、阿、开、延”的手写汉字图像识别结果为“诀、曙、阿、开、延”,分割后的汉字图像识别率为100%。采用相同原始图像,使用轮廓检测法对图5(a)进行列分割,分割结果如图8所示。

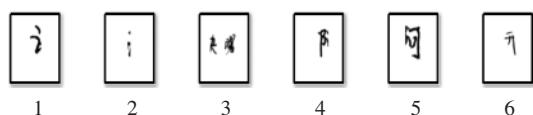


图8 轮廓检测法分割结果图

由图8可以看出,轮廓检测法将第一行的“诀、曙、阿、开、延”5个汉字拆分成了“讠、宀、夬、曙、阝、可、开”,仅有“开”字成功分割,且图像存在变形,“延”字在分割过程中丢失。将得到的单字图片归一化并转化为位深度为8的灰度图像后,送入神经网络模型识别结果如图9所示。



图9 神经网络识别结果图

由图9可以看出,分割出的“讠、宀、夬、曙、阝、可”的手写汉字图像识别结果为“井、况、块、诀、椰、刊”,分割后的汉字图像识别率为0%。

综上实验对比两种分割结果,投影法在手写汉字识别过程中的分割效果更好。

3 结论

手写汉字具有汉字类别繁多、字形结构复杂、书写随意性大等特点,提高手写汉字识别的准确度一直是研究的难点和热点。本文从提高识别准确度的目的出发,以CNN卷积神经网络模型的识别结果为依据,对比分析了投影法和轮廓检测法在手写汉字图像分割处理中的适用性。研究发现,无论是从图像分割的结果来看,还是从文字识别的结果来看,投影法都是更优解。此项研究在手写汉字识别研究工作者以及手写文件信息化处理工作者的工作实践中有着重要影响,同时也为文档管理类工作者在手写资料的整理方面提供了一个切实可行的思路。

参考文献

- [1] 李兰英,周志刚,陈德运.DBN 和 CNN 融合的脱机手写汉字识别[J].哈尔滨理工大学学报,2020,25(3):137-143.
- [2] 高灿.基于卷积神经网络的脱机手写汉字识别系统研究[D].淮南:安徽理工大学,2017.
- [3] 郑延斌,韩梦云,樊文鑫.基于二维主成分分析与卷积神经网络的手写体汉字识别[J].计算机应用,2020,40(8):2465-2471.
- [4] 孟冉洋.简述手写文字识别技术[J].中国高新区,2019(3):28,30.
- [5] 伍薇.基于 CNN 卷积神经网络的企业电子档案分类法研究[J].山西档案,2018(5):83-85.

(下转第 80 页)

且尽可能保存不需要增强部分的细节信息,避免最终图像增强过度,同时利用改进权值自适应融合细节部分与原始图像最终获得输出图像。仿真结果表明,相较于传统线性非锐化掩模方法,本文方法其图像评价指标 H 值、 P_{SNR} 值分别平均提高了0.2%、28.4%, A_{MBE} 值平均降低了88.4%, S_{SM} 的值更加接近1。最终在电润湿电子纸上通过实验证明,本文算法的增强性能更佳,在增强纹理细节的同时,尽可能减少噪声的放大,避免过度增强,取得了较好的显示效果。

参考文献

- [1] 卜倩倩,王丹,邱云,等.反射显示技术的研究进展[J].液晶与显示,2019,34(2):169-176.
- [2] JUNG G S,LEE J S,WON Y H.Effects of liquid property and substrate roughness on the response time of an electrowetting liquid lens[C]//MOEMS and Miniaturized Systems XVII,2018.
- [3] 林珊玲,李甜甜,曾素云,等.基于人眼视觉的电润湿电子纸显示器亮度非线性校正方法[J].光子学报,2019,48(8):33-41.
- [4] JOSEPH J,ANOOP B N,WILLIAMS J.A modified unsharp masking with adaptive threshold and objectively defined amount based on saturation constraints[J].Multimedia Tools and Applications,2019,78(8):11073-11089.
- [5] KANSAL S,PURWAR S,TRIPATHI R K.Image contrast enhancement using unsharp masking and histogram equalization[J].Multimedia Tools and Applications,2018,77(20):26919-26938.
- [6] 沈磊,苏建忠,郭肇敏,等.基于反锐化掩模技术的红外图像增强算法设计[J].南开大学学报(自然科学版),2019,52(1):29-35.
- [7] GUAN R,WAN Y.An improved unsharp masking sharpening algorithm for image enhancement[C]//Eighth International Conference on Intelligent Computing,Instrumentation and Control Technologies (ICICICT),2017:156-161.

(上接第75页)

- [6] 杨帆.数字图像处理与分析[M].北京:北京航空航天大学出版社,2019:1-38.
- [7] 高飞,汪敏倩,吴宗林.一种基于投影的车牌字符切割方法:中国,CN201611131250.4[P].2017-05-10.
- [8] 黄威生.图像文字识别方法:中国,CN201910952417.0[P].2020-01-17.
- [9] 俞洋,厉丹,马一丁,等.基于CNN卷积神经网络的手势识别系统[J].电脑知识与技术,2020(10):210-212.
- [10] 孙华魁.数字图像处理与识别技术研究[M].天津:天津科学技术出版社,2019.
- [11] 张晶.数字图像处理应用研究[M].长春:吉林大学出版社,2019.
- [12] 范春奇,任坤,孟丽莎,等.基于深度学习的数字图像修

Conference on Digital Image Processing(ICDIP 2016),2016.

- [8] 王仕女,孙文胜.基于非锐化掩模与Beta变换的图像增强研究[J].软件导刊,2019,18(9):207-210.
- [9] 朱莉.一种改进的自适应非锐化掩模图像增强算法[J].微电子学与计算机,2012,29(6):158-160.
- [10] 冯策,戴树岭,赵帅和.基于显著特性的保持边缘滤波算法[J].北京航空航天大学学报,2015,41(4):669-674.
- [11] HE K M,SUN J,TANG X O.Guided image filtering[C]//European Conference on Computer Vision,2010:1-14.
- [12] 许磊,崔光茫,郑晨浦,等.基于多尺度分解和显著性区域提取的可见光红外图像融合方法[J].激光与光电子学进展,2017,54(11):111-120.
- [13] 谢伟,周玉钦,游敏.融合梯度信息的改进引导滤波[J].中国图象图形学报,2016,21(9):1119-1126.
- [14] ZHUANG L Y,GUAN Y P.Adaptive image enhancement using entropy-based subhistogram equalization[J].Computational Intelligence and Neuroeence,2018(2):1-13.
- [15] SHEBA K U,RAJ S G.Adaptive fuzzy logic based Bi-Histogram equalization for contrast enhancement of mammograms[C]//2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT),2017:156-161.

(收稿日期:2020-11-12)

作者简介:

熊铃铃(1995-),女,硕士研究生,主要研究方向:电子纸显示和图像处理技术。

林珊玲(1991-),女,博士,讲师,主要研究方向:电子纸显示和平板显示器件及其驱动技术。

林志贤(1975-),通信作者,男,博士,教授,主要研究方向:电子纸显示和平板显示器件驱动以及图像处理技术,E-mail:1zx2005000@163.com。



扫码下载电子文档

复算法最新进展[J].信号处理,2020,36(1):102-109.

- [13] 朱世宇,孙令翠,张桥,等.基于深度学习的不完整的手写体数字图像识别研究[J].智能建筑与工程机械,2019,1(2):104-106.
- [14] 金连文,钟卓耀,杨钊,等.深度学习在手写汉字识别中的应用综述[J].自动化学报,2016,42(8):1125-1141.

(收稿日期:2020-11-09)

作者简介:

张莉(1987-),女,硕士,工程师,主要研究方向:本科实验实践教学、信号处理与模式识别。

孟范泽(1998-),男,本科,主要研究方向:电气工程及其自动化。

刘思霖(1999-),女,本科,主要研究方向:电气工程及其自动化。



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