

复杂背景图像的字符识别算法研究

张红霞¹, 王 灿¹, 刘 鑫¹, 白志城¹,
付秀娟¹, 王 刚², 梅天灿², 王学华¹

(1. 武汉工程大学 材料科学与工程学院, 武汉 430205;

2. 武汉大学 电子信息学院, 武汉 430070)

摘要: 为了解决字符识别过程中的局部曝光、印刷字符的断裂以及变形和自然环境下的背景污染等问题, 提出了一种分块处理与卷积神经网络 (CNN) 相结合的字符图像识别算法; 首先利用 OpenCV 机器视觉库, 结合分块处理、伽马运算、参数调整等方法对产品零件表面印刷字符进行预处理, 初步解决图像局部曝光和字符断裂问题; 其次为了获得单个字符图像, 利用数学形态学算法对局部曝光处理后的二值化图像进行分步分割, 进而去掉字符间的无用信息; 最后利用 Keras 模块为字符识别提供的 API 搭建 CNN 模型, 经过对 100 多张字符的识别训练, 准确率高达 96.9%, 为某汽车零部件自动化生产中的字符识别提供了可靠的依据。

关键词: 局部曝光; 字符分割; 字符识别; 数学形态学; CNN 卷积神经网络

Automatic Recognition Algorithms Research for Printed Characters with Local Exposure

Zhang Hongxia¹, Wang Can¹, Liu Xin¹, Bai Zhicheng¹, Fu Xiujian¹,
Wang Gang², Mei Tiancan², Wang Xuehua^{1*}

(1. School of Materials Science and Engineering, Wuhan Institute of Technology, Wuhan 430205, China;

2. School of Electronic Information, Wuhan University, Wuhan 430070, China)

Abstract: To solve the problems of local exposure, fracture and distortion of printed characters, and background pollution in natural environment during character recognition, a character image recognition algorithm combining block processing and convolution neural network (CNN) is proposed. Firstly, using the OpenCV library, the printed characters on the surface of product parts were preprocessed by block processing, gamma operation and parameter adjustment, and the local exposure and character breaking of images were preliminarily solved. Secondly, in order to obtain a single character image, the image processed in the early stage is multiple segmented by the mathematical morphology algorithm, so as to remove the useless information between characters. Finally, the API provided by Keras for character recognition was used to build CNN model. After over 100 characters recognition training, the accuracy rate was as high as 96.9%, the model provides a reliable basis for character recognition in automatic production of an auto parts.

Keywords: local exposure; character segmentation; character recognition; mathematical morphology; convolution neural network (CNN)

0 引言

随着信息通信技术的迅猛发展, 人工智能涉及生活圈的方方面面, 其中智能化生产就是最常见的。字符刻印是产品的唯一标识, 在大规模的产品生产、存储、检测和装配线上, 产品的字符标记显得尤为重要。通过智能化的设备对现实背景下的产品进行身份识别是智能化生产的基本功能需求。这种字符识别的系统一般包括 3 个方面: 产品图像的获

取与预处理, 字符分割和字符识别。基于字符图像的预处理过程最大程度的优化图像, 获取清晰连贯的字符边缘是字符分割与识别的首要条件, 但是在自然光下获取的图像信息存在部分不满足智能识别要求的缺陷, 如噪声污染, 局部曝光, 字符部分缺失、模糊和变形等, 如何提高这些缺陷字符的识别质量是研究人员一直关注的热点问题^[1-2]。

字符识别常用模板匹配和深度学习^[3-5], 对于复杂环境下的不同产品而言, 字符识别系统对算法的鲁棒性有着更高要求, 所以单一地运用某种方法不能解决整个问题^[6], 字符识别系统各个环节的改进与创新都是当前人工智能的热点。本文通过伽马增强矫正字符图像的对比度, 设计了一种获取曝光区域并进行像素均衡化的局部曝光图像处理算法, 对其进行二值化、膨胀腐蚀、旋转矫正、开闭运算等预处理操作获得较为完整的字符边缘。最后由分步分割获取的字符样本图像构建训练集, 利用 Keras 模型提供的 API 接口, 构

收稿日期: 2019-01-02; **修回日期:** 2019-01-29。

基金项目: 国家自然科学基金 (51575406); 武汉工程大学研究生创新基金 (CX2017019)。

作者简介: 张红霞 (1993-), 女, 湖北武汉人, 硕士研究生, 主要从事机器视觉方向的研究。

王学华 (1976-), 男, 湖北襄阳人, 博士, 教授, 主要从事材料加工及其过程控制方向的研究。

建了一个三层卷积神经网络, 通过字符识别训练实现了对局部曝光字符的可靠识别。

1 图像预处理

图像预处理能有效控制图像的噪声干扰, 矫正字符断裂, 扭曲变形, 甚至局部曝光等问题, 对于后期字符分割和识别起到了至关重要的作用。本文字符识别系统的整体结构如图 1 所示。

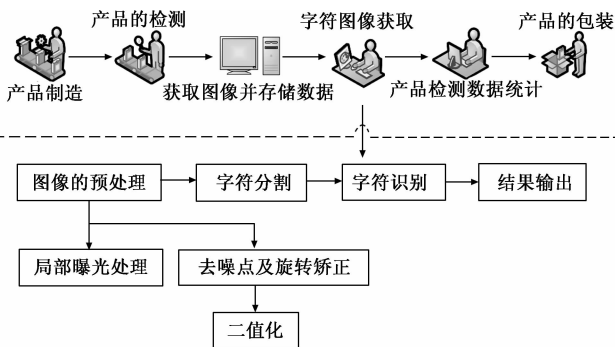


图 1 产品零件表面的字符识别系统简易架构

图 2 为拍摄某产品零件获取的印刷字符图像, 图像中字符信息包括产品标记符, 字母, 阿拉伯数字以及一些间隔符号等, 且字符中并没有“1”, “2”, “0”的字母, 降低了与“1”, “2”, “0”在字符切割和识别过程中的复杂度。由于拍摄角度和现场环境原因导致图像曝光不均匀, 存在局部曝光过度的情况, 其次背景材料为金属材料, 在空气中容易氧化锈蚀, 导致局部位置存在显著的“暗斑”和“亮斑”, 增加了图像预处理过程中的噪声; 另外, 字符在印刷过程中由于振动或者速度原因导致部分字符断裂, 部分字符粘连等现象。为了实现这些字符信息的准确识别, 需要对图像进行合理的前处理过程。

1.1 图像伽马增强

不同光照强度下获得的图像背景灰度不一致和细节丢失问题, 采用伽马增强运算 $s = cr^\gamma$, 取修正参数 $\gamma = 0.8, c = 1$, 将字符图像的背景灰度级即 r 都增强到一定的像素值范围内即 s , 既增强了图像的对比度和亮度, 也保证了后期局部曝光过程的灰度变换等操作。利用图 2 (a) 所示的图像进行伽马增强之后得到的灰度图像如图 2 (b)。

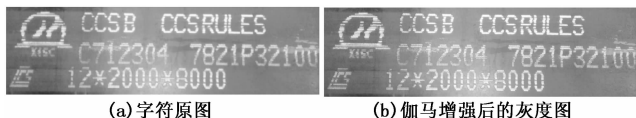


图 2 产品零部件印刷字符图像及灰度图

调整 $\gamma = p/10, p = 1, 2, \dots, 10$, 随着 γ 的增大, 图像的背景颜色不断加深, 二值化字符图像边缘断裂情况越明显。当 $\gamma < 0.6$ 时, 字符局部曝光区域与背景像素值相近, 字符边缘不易识别; 当 $\gamma = 0.8$ 时, 如图 3 背景像素值集中在 $130 \sim 200$ 之间, 使得像素值 $i \in (0, 130)$ 的点得到减弱, $i \in (130, 200)$ 的点得到增强, $i \in (200, 255)$ 的点保持不变。伽

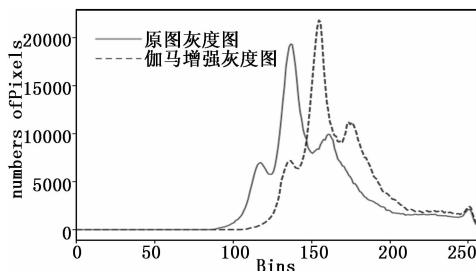


图 3 伽马增强前后图像的灰度直方图

马增强拉伸了图像中灰度级较低的区域, 同时压缩了灰度级较高的部分, 从而保留了字符边缘点的像素值, 确定了所有字符图像局部曝光区域的及处理。

1.2 局部曝光处理

图像的灰度直方图表明背景灰度值明显存在分布不均匀现象^[7-9], 伽马增强后图像像素值大部分集中在在 $130 \sim 200$ 范围内, 若像素分布相差较大直接二值化会导致噪声影响字符的边缘检测, 轮廓被噪声覆盖或者断裂。如图 4 (a) 为阈值 $T = 170$ 时字符图像的二值化效果图, 可以看出背景噪声与字符区域相互作用形成字符联通或断裂, 难以进行字符的自动分割。需要对图像进行分块操作, 解决局部曝光造成的分割问题^[10]。



图 4 字符图像二值化

1.2.1 图像分块操作

图 2 (a) 可以看出单个字符所在的局部矩形方块内的像素点个数是在一定的范围内分布。过少的像素点数则表示较暗的区域, 需要保留其原像素值或者增大像素值以达到背景均衡化; 过多的像素点数则表示噪声聚集区, 需要分离背景与感兴趣的字符区域, 降低背景像素值。为了找出感兴趣区域 (ROI) 即局部曝光区域, 可以对整体图像 (1895×417 像素) 进行分块处理, 图 5 为分块原理图。

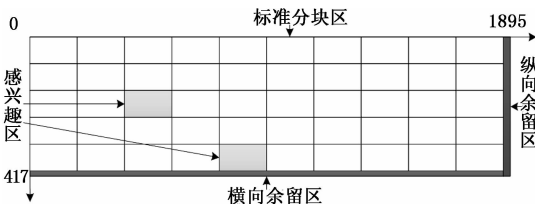


图 5 字符图像分块原理图 (单位: 像素)

从字符大小和噪声区域观察分析判断块高 H 在一个字符高度, 块宽 $W \approx 2H$ 时, 能最有效保证完整的字符边缘和局部曝光操作, 对噪声的处理也比较细致。对不同块的个数

裁边 2 个像素值截取保留。

确定粘连字符的最小外接矩形的宽度 w :

$$m = \lceil w/w_s \rceil, \quad m = 1, 2, 3 \dots \quad (2)$$

已知粘连外接矩形的点 (x, y) , 宽 w , 和高 h , 分割第 i 个最小矩形的参数:

$$\begin{cases} x_i = x + \frac{w}{m} \times i \\ y_i = y \\ w_i = \frac{w}{m} \\ h_i = h, i \in [1, m] \end{cases} \quad (3)$$

由于最小外接矩形中字符“1”的宽度最小, 其他字符的印刷宽度是一致的。以除“1”外最小宽度的字符为标准既排除了其他字符因印刷不规范或拍摄角度导致的宽度变形, 又保证了分割过程中 m 取整的标准性。整体分割后的部分单个字符图像如图 7 (a) 所示。

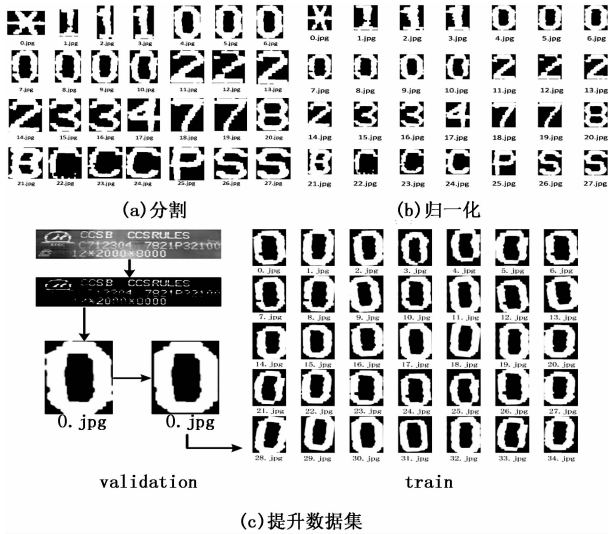


图 7 字符图像分割效果 (部分字符)

3 字符识别

字符识别常用模板匹配和神经网络两种算法。模板匹配算法计算量大, 匹配过程要求严格; 卷积神经网络 (CNN) 对高维数据处理高效, 特征分类效果好^[13-15], 可用于识别各种物体。由于实验待识别的图片数量少, 分割之后待识别字符的数据集小, 实验表明 CNN 网络也适用于小数据集的训练, 有较高的识别率。

3.1 准备训练集、验证集与测试集

统计 7 张待识别图片的分割字符, 部分字符“1”, “Z”, “O”等因字符间相似不存在印刷和识别过程, 整理把数据集分为 27 类, 形成每类含有 7~20 张不等的测试集。为了尽量利用有限的训练数据, 通过 Keras 提供的 ImageGenerator 函数对原始样本图像实现一系列旋转、移动、剪切变换、放大缩小, 水平变换等随机变换, 将数据集进行提升, 如图 7 (c)。训练模型中不存在两张完全相同的图片, 有利于抑

制过拟合, 使得模型的泛化能力更好。提升后形成每类有 110~200 张不等的训练集, 将训练集和验证集的比例控制在 10: 1 左右。

整理所有字符文件, 将图片格式保存成单通道的二值化图片, 白字黑底, 并将所有图片归一化到 60×40 像素值大小的格式如图 7 (b), 降低计算复杂度, 保证后期模型训练。

3.2 搭建与训练模型

卷积神经网络可对图像进行正确的分类, Keras 模块是基于 Tensorflow、Theano 等后端的高层神经网络框架, 可在 CPU 和 GPU 之间无缝切换, 执行效率高。由于 27 类小数据集的应用, 所以模型使用简单的三层卷积网络结构, 数据模型采用 Sequential 模型。

Sequential 模型是单输入和单输出简单模型, 层与层之间只有相邻关系。在本实验中搭建的 3 层卷积网络上加上激活函数 ReLU, 最后接池化层 max-pooling 和 2 个全连接网络 Dense, 以单个神经元和激活函数 Softmax 结束模型, 这种配置会产生多分类的结果, 为了与这种配置相适应, 使用 categorical_crossentropy 作为损失函数。利用 flow_from_directory 函数从 jpgs 图片中直接产生数据和标签, 并根据这个生成器来训练网络, 网络结构如图 8。

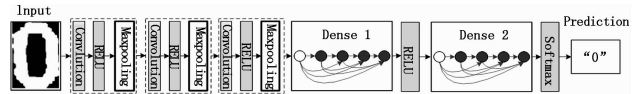


图 8 Keras 卷积网络结构

4 数据分析

训练集 4490 个字符文件, 验证集 230 个字符文件。模型训练反复调整参数, 定义 batch_size = 256, epoch = 70。在 CPU 上训练这个模型的 loss、acc、value_loss 和 value_acc 与 epoch 的关系如图 9 所示。loss 和 value_loss 随着 epoch 的增大依次减小, 不断趋近于 0; acc 和 value_acc 随着 epoch 的增大依次增大, 不断趋近于 1。

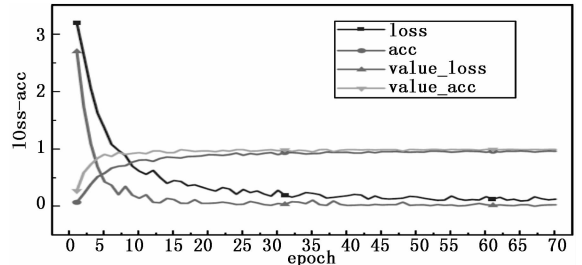


图 9 字符识别准确率

随机连续取 5 次 epoch 的数据如表 2 所示, 当 epoch = 65 时, 准确率高达 96.91%, 损失值低为 0.0969, 实验表明随着 epoch 的增大, 损失值越来越小, 准确率越来越大, 且训练集和测试集近似相等。记录得到每个 epoch 在 CPU 上耗时 80~150 s。训练的模型稳定良好, 可以保证字符全部识别, 速度快且准确率高。

表 2 epoch 为 5 次的记录

epoch	T/s	Loss	Acc	Value_loss	Value_acc
65	95	0.0969	0.9691	0.0304	0.9913
66	105	0.1042	0.9637	0.0306	0.9913
67	89	0.1361	0.9609	0.0229	0.9913
68	100	0.1167	0.9677	0.0287	0.9957
69	110	0.1282	0.9643	0.0295	0.9957

5 结束语

系统由 VS2015 开发平台采用 C++ 编码, 在 CPU 主频 3.2 GHz, 8 G 内存的 Windows 7 (64 位) 电脑平台上开发实现并测试识别效果。结合 OpenCV 机器视觉库和 Keras 的高层神经网络 API, 本文对局部曝光的工业零部件印刷字符进行取图, 曝光预处理, 单个字符的分割, 图像归一化和 CNN 神经网络识别处理, 从像素值的参数调整上解决了局部曝光和自然环境下背景污染问题, 并以小数据集的提升构建神经网络模型, 实现了小数据集下的 96.9% 的正确识别率。

参考文献:

[1] Maofu L, Ya L, Huijun H, et al. Genetic algorithm and mathematical morphology based binarization method for strip steel defect image with non-uniform illumination [J]. Journal of Visual Communication and Image Representation, 2016, 37.

[2] 孙 强. 低质量印刷体字符分割与识别研究 [D]. 南京: 南京理工大学, 2014.

[3] Huaigu C. Machine-printed character recognition [M]. Springer London: 2014-06-15.

[4] 高 聪, 王福龙. 基于模板匹配和局部 HOG 特征的车牌识别

算法 [J]. 计算机系统应用, 2017, 26 (1): 122-128.

[5] Li X P, Lv, X X, Wang S Z, et al. Research on the recognition algorithm of the license plate character based on the multi-resolution template matching [A]. 2010 4th International Conference on New Trends in Information Science and Service Science (NISS) [C]. 2010.

[6] 曾 超, 陈 雨. 车牌超分辨率重建与识别 [J]. 计算机测量与控制, 2018, 26 (3): 244-249.

[7] Kota M, Takashi K, Akira T. A novel method for design of gray-level transformation functions for image contrast improvement based on the histogram equalization technique [J]. Electronics and Communications in Japan, 2013, 96 (7).

[8] 余旺盛, 田孝华, 侯志强. 基于区域边缘统计的图像特征描述新方法 [J]. 计算机学报, 2014, 37 (6): 1398-1410.

[9] 李立源, 龚 坚, 陈维南. 基于二维灰度直方图最佳一维投影的图像分割方法 [J]. 自动化学报, 1996 (3): 314-322.

[10] 孟 梦. 高清摄像机自动白平衡和自动曝光算法研究及实现 [D]. 西安: 西安电子科技大学, 2014.

[11] 洪 涛, 梁伟建, 卢玉凤. 标牌粘连字符自适应定位分割重建与识别 [J]. 中国图象图形学报, 2014, 19 (6): 886-895.

[12] Bahram R. Implementation of a high speed technique for character segmentation of license plate based on thresholding algorithm [J]. International Journal of Image, Graphics and Signal Processing, 2012, 4 (12).

[13] 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.

[14] Hui L, et al. Reading car license plates using deep neural networks [J]. Image and Vision Computing, 2018, 72.

[15] 胡 洋. IC 芯片印刷字符识别算法研究与应用 [D]. 武汉: 华中科技大学, 2015.

(上接第 161 页)

参考文献:

[1] 付 平, 常德功, 密封设计手册 [M]. 北京: 化学工业出版社, 2011.

[2] 郝木明. 机械密封技术及应用 [M]. 北京: 中国石化出版社, 2010.

[3] 魏 龙, 顾伯勤, 孙见君, 等. 机械密封端面摩擦特性参数及其测试技术 [J]. 润滑与密封, 2006 (11): 198-202.

[4] 耿荣生. 声发射技术发展现状: 学会成立 20 周年回顾 [J]. 无损检测, 1998, 20 (6): 151-154.

[5] 耿荣生, 沈功田, 刘时风. 声发射信号处理和分析技术 [J]. 无损检测, 2002, 24 (1): 23-28.

[6] 沈功田, 戴 光, 刘时风. 中国声发射检测技术进展: 学会成立 25 周年纪念 [J]. 无损检测, 2003, 25 (6): 302-307.

[7] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [A]. Proceedings of the Royal

Society of London A: Mathematical, Physical and Engineering Sciences [C]. 1998: 903-995.

[8] Wang J, Li Z C, Wang D Y. A method for wavelet threshold denoising of seismic data based on CEEMD [J]. Geophysical Prospecting for Petroleum, 2014, 53 (2): 164-172.

[9] Yeh J R, Shieh J S, Huang N E. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method [J]. Advances in Adaptive Data Analysis, 2010, 2 (2): 135-156.

[10] 杨 孟, 王 瑾, 周西峰, 等. 基于 CEEMD 和小波包的降噪方法研究 [J]. 南京邮电大学学报 (自然科学版), 2018, 38 (2): 41-47.

[11] 刘 莹, 韩 焱, 郭亚丽, 等. 基于 CEEMD 的爆破振动信号自适应去噪 [J]. 科学技术与工程, 2015, 15 (32): 54-58.

[12] 张尔卿, 傅 攀, 等. 基于声发射主轴机械密封端面开启状态识别技术田. 润滑与密封, 2014, 39 (4): 67-73.