

基于 CNN 的车牌识别系统

徐胜舟 周煜

(中南民族大学 计算机科学学院 武汉 430074)

摘要 针对现有的车牌识别系统在遇到复杂条件,例如暗光、遮挡、多车牌、能见度低等情况时,难以有效地定位并识别车牌,提出了一种基于卷积神经网络的车牌自动识别系统.在车牌定位阶段综合应用 3 种定位方式对车牌进行初步定位检测,然后使用 CNN 模型对检测到的候选车牌进行判断;在车牌字符识别阶段,将分割出的字符输入到设计好的卷积神经网络模型中进行训练,得到的输出结果即为识别的车牌字符.在 5906 张车牌图像和非车牌图像以及 36261 张字符图片上的实验结果表明:提出的车牌识别系统对车牌和字符的识别率分别达到了 94% 和 96.4%,明显优于传统的车牌识别方法,具有极高的实用性,可以满足绝大多数场景的使用需求.

关键词 车牌定位; 车牌识别; 字符识别; 卷积神经网络

中图分类号 TP391.4 文献标识码 A 文章编号 1672-4321(2017)03-0125-06

License Plate Recognition System Based on CNN

Xu Shengzhou Zhou Yu

(College of Computer Science, South-Central University for Nationalities, Wuhan 430074, China)

Abstract The existing license plate recognition system is difficult to locate and identify the license plate effectively when it encounters complex conditions such as dim light, plate is blocked, multiple plates and low visibility. An automatic license plate recognition system based on convolution neural network (CNN) has been proposed in this paper. In the license plate location phase, three kinds of positioning methods are integrated for the initial locating of the license plate. Then, the CNN model is used to judge the selected license plate. In the license plate character recognition phase, segmented characters are input to a designed CNN model, and the output of the CNN model is the result of the recognized characters. The experiment is based on 5906 license plate images and non-license plate images, and 36261 characters images. The results of the experiment show that the recognition rates of the proposed system for license plate and character are 94% and 96.4% respectively, which is significantly better than that of traditional license plate recognition methods. It meets the needs of the vast majority of the use of the scene, with a high practicality.

Keywords license plate locating; license plate recognize; character recognize; CNN

进入 21 世纪以来,城市交通问题日益严峻.智能交通监管系统的出现为我们带来了高效可行的解决方案,可以在大范围内高效、针对性地进行识别、监督,大幅度降低人力成本,提高监管效率.

车牌识别技术是目前智能交通系统中的核心技术之一^[1].通过计算机软件来处理车牌信息,还可以将这些车辆信息存储到相关部门的数据库,为后续其他的管理系统提供参考.车牌识别技术在节省人力物力的前提下有针对性地对车辆进行监测并且不会对车辆带来任何不便.

现有的车牌识别系统没有良好的通用性,在遇到复杂条件,例如暗光、多车牌、角度倾斜等情况时,不能有效地定位并识别车牌.而本文的思路是,首先对程序进行“训练”,输入大量的车牌数据进行学习,这些数据不仅包括日常的图片,还包括上述的复杂条件下拍摄的图片.这样在实际应用时,已经有了训练的基础,在处理各种情况时的效果会明显优于直接识别的效果.文献[2]和文献[3]也采用了先定位后训练的方法,但在定位阶段的成功率不高,不能应对上述复杂条件;并且基于 SVM 和 BP 神经网络

收稿日期 2017-06-08

作者简介 徐胜舟(1982-)男,副教授,博士,研究方向:模式识别, E-mail: xushengzhou2008@163.com

基金项目 国家自然科学基金资助项目(61302192)

的车牌判断和字符识别率不高.文献[4]和文献[5]仅进行了基于CNN的字符识别研究,未实现预处理阶段的车牌定位和判断工作.

本文实现了从图片中定位车牌并识别车牌字符的功能.程序整体上分为两大模块:车牌定位和字符识别.在车牌定位模块,通过综合使用3种定位方法从图像中截取出候选车牌区域,然后使用CaffeNet卷积神经网络模型对候选车牌进行判断,即判断定位出的区域是否为车牌.在字符识别模块,先从车牌图片中分割出单个字符,然后使用训练好的CNN模型来识别字符,最后输出结果.

1 车牌定位

车牌定位是从图片中定位出候选车牌区域.本文使用3种定位方法:边缘检测法、颜色定位法和文字定位法,并综合使用3种方法来应对不同情况,最终输出最优定位结果.

(1) 边缘检测法.由于车牌均为矩形,因此可以检测图片中的矩形区域来定位疑似车牌的区域.具体步骤如下:程序首先读入图像,并对图像进行去噪^[6],随后将图像转为灰度图^[7].这两步都是为了能够更好地检测到图像中的边缘.在检测到垂直边缘^[8]后,对图像进行二值化^[9]和形态学操作^[9],这是为了将前面检测到的边缘的内部填满,即形成连通区域,这样才能够找到矩形车牌图块.接着画出连通区域的外接矩形^[11],这些矩形图块中就包含要定位的车牌.最后,对这些矩形图块进行筛选,剔除不符合条件的图块,输出并保存余下的矩形区域.

(2) 颜色定位法.上述边缘检测法虽然能够成功定位大多数车牌,但在处理有大量垂直边缘的图像时,难以准确的定位车牌,因此我们需要其他方法来弥补这方面的不足.由于我国的绝大多数车牌均为蓝色或黄色背景,这为通过颜色来定位车牌提供了条件.

首先将图片转换为HSV颜色模型^[12,13].在最常见的RGB颜色模型中,R、G、B三个分量各代表红、黄、蓝.HSV与之不同,HSV模型中的H、S、V分别代表是:色调、饱和度和亮度.H分量是唯一与色彩相关的分量.因此,只需要固定H的值,就会呈现基本稳定的颜色.当H值不变时,V和S的变化会分别导致饱和度和亮度的变化.当V和S都为最大值,1时,颜色是最纯正的.S越小,颜色越白.V越小,颜色越黑,当V=0时,颜色变为黑色.S和V的

值也会影响最终颜色的效果.因此,设定一个阈值,使得S和V都大于阈值时,颜色才属于H所表达的颜色.在通过颜色定位到车牌区域后,进行二值化、闭操作、取轮廓三个操作,将车牌区域截取出来.最后,对这块区域进行角度调整,对不同倾斜角度的图像进行旋转或偏斜扭转^[14]操作,将其调整为正视角.

(3) 文字定位法.边缘检测法和颜色定位法可以满足在正常光照下拍摄的图片中进行车牌定位,但在弱光条件下并不能很好的检测并定位到车牌区域,文字定位法可以弥补这一不足.文字定位法首先使用最大稳定极值区域(MSER)算法^[15]提取文字,然后使用种子生长法将这些图块连接起来,最后组合成车牌区域.

在本程序中的具体实现步骤如下.

1) 将图片转换为HSV颜色模型,并提取H通道.

2) 灰度化.将图片转换为灰度图.

3) 根据检测区域点分别生成MSER+和MSER-结果.MSER+检测黄色车牌上的黑色字符,MSER-检测蓝色车牌上的白色字符.

4) 判断MSER提取出的区域的尺寸,删除不符合车牌文字尺寸的区域.

5) 查找强种子(图1车牌区域中的左起第2~7个方框).将靠近的强种子进行集合,画出一条穿过它们的中心线(图1车牌区域中白色的线,连接第4~6个方块),这条线即是车牌的中轴线.

6) 在这条线的附近寻找弱种子(图1车牌区域中最左侧的方框).通常,这些强种子和弱种子的距离不会太远,并且尺寸基本一致.弱种子在强种子的左右进行查找.

7) 全部找完后,强种子和弱种子的总数即是车牌区域的文字个数.图1为字符定位结果图.

(4) 车牌定位结果.经过对比测试,以上3种定位方法分别适用于不同情况:边缘检测定位法适用于图片中的垂直边缘较少的图片;颜色定位适用于车身及背景中无蓝色和黄色的图片;文字定位法针对弱光条件下拍摄的图片的定位效果优于其他两种方法.因此,车牌定位整体步骤如下.

1) 首先使用边缘检测定位车牌,若定位到候选车牌,则直接输出所有候选车牌图块;若没有定位到车牌,则进入步骤2);

2) 使用颜色定位法定位车牌,若检测到候选车牌区域,则输出所有候选车牌图块;若检测到大量蓝

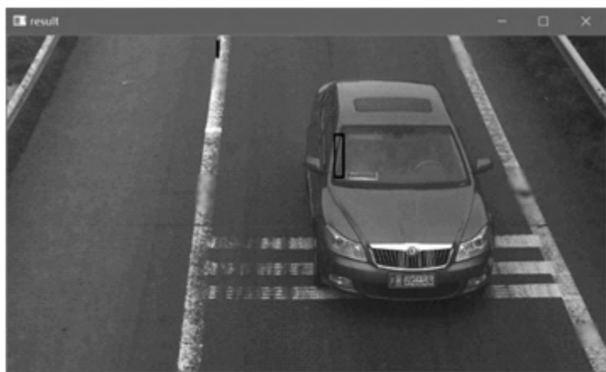


图 1 字符定位

Fig. 1 Character positioning

色或黄色区域,或无法检测到指定颜色,则不输出任何结果并进入步骤 3);

3) 使用字符定位法,最终输出定位结果.

2 车牌判断及字符分割

(1) 车牌判断. 通过定位得到一组候选车牌,如图 2 所示. 接下来通过卷积神经网络来判断车牌图块.



图 2 候选车牌

Fig. 2 Candidate plate

将候选车牌图块作为神经网络的输入,输出则为判断是车牌的图块.

(2) 字符分割. 在得到车牌图像后,接下来将车牌上的字符分割出来. 将车牌转化为灰度图像,然后进行字符分割. 具体步骤如下.

1) 灰度化. 将 RGB 图片转化为灰度化图片;

2) 颜色判断与二值化. 判断车牌颜色,对蓝色车牌使用正二值化,对于黄色车牌使用反二值化.

有时二值化会产生貌似字符“1”的图形,如图 3 所示. 由于中文车牌只有 7 个字符,所以我们把图块

从左到右排序,依次选择 7 个. 这样,被误判为“1”的缝隙就可以被排除.

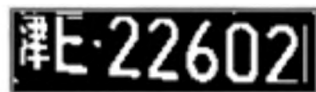


图 3 错误字符“1”

Fig. 3 Wrong character “1”

3) 取轮廓. 截取每个字符轮廓.

4) 截取图块. 把图中的矩形图块截取出来,为后续步骤做准备. 将截取的字符图片归一化为 20 × 32 的矩形图块,如图 4 所示.



图 4 截取到的字符图片

Fig. 4 Intercepted character picture

最后,将分割出的字符输入到卷积神经网络中进行字符识别,最终输出结果字符串. 识别结果图 5 所示.



图 5 字符识别结果

Fig. 5 Characters recognize result

(3) 识别训练. 卷积神经网络是一种高效的图像识别方法. 20 世纪 60 年代,Hubel 和 Wiesel 在研究猫脑皮层中用于局部感知和方向选择的神经元时发现其独特的网络结构可以有效地降低反馈神经网络的复杂性,继而提出了卷积神经网络(CNN).

卷积神经网络有 3 个核心:局部感受野、权值共享和池化. 局部感受野和权值共享的引入极大的减少了网络模型的参数,使得模型的配置和使用更加简便. 池化可以减小图像在网络间传输时的规模,从而降低了计算量,提高了运算效率. 这是卷积神经网络优于传统人工神经网络的原因.

本文使用 Caffe 深度学习框架^[16]并采用 CaffeNet 网络模型进行训练. Caffe 深度学习框架由 BVLC(Berkeley Vision and Learning Center)及其社区贡献者开发. 项目创建者是伯克利大学博士贾扬清. 由于 Caffe 清晰、高效、易于训练等特点,在深度学习中被广泛使用. 在车牌字符识别训练中,使用

Caffe 框架的优点在于训练速度快并且配置简便,且训练后的识别结果十分可观. CaffeNet 是基于 AlexNet 改进的网络模型,将 norm1、pool1 以及 norm2、pool2 互换了顺序,在相同迭代次数下准确率提升了 0.2%.

车牌判断与字符识别均采用相同的识别网络,

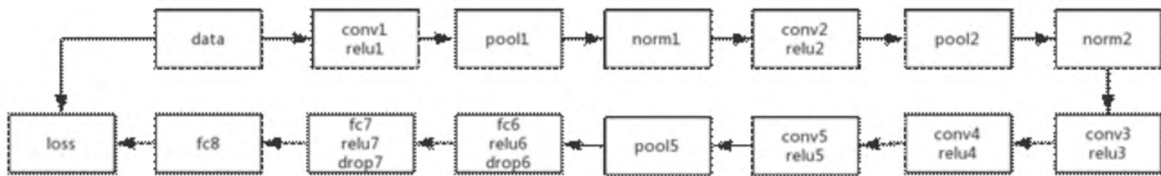


图6 CaffeNet 网络模型结构

Fig. 6 CaffeNet Network model structure

CaffeNet 模型共有 8 层,前 5 层为卷积层,后 3 层为全连接层. 识别训练具体步骤如下.

Step1. 将图像归一化为 $227 \times 227 \times 3$ 的 RGB 图像.

Step2. Conv1 层,使用 96 个大小为 11×11 的卷积核,使用 ReLU 激励函数,来确保特征图的值在合理范围内,以 4 个像素为一个单位的步长向右移动,最后进行池化(pool1)和归一化(norm1). 池化处理采用最大池化算法,得到 96 个 27×27 的特征图,然后再将这些特征图作为输入数据,进行第二次卷积.

Step3. Conv2 层,使用 256 个 5×5 的卷积核,扩充边缘为 2 个像素,保证进行卷积运算后的特征图尺寸不变. 卷积后会得到一个新的 256 个特征图,也就是会有 256 个 27×27 大小的特征图. 然后进行与 Step1 中相同的 ReLU 以及池化(pool2)和归一化(norm2)操作.

Step4. Conv3 层与 Conv4 层均使用 384 个 3×3 的卷积核,卷积后得到 384 个 13×13 的新特征图,这两层都没有池化层.

Step5. Conv5 使用 256 个卷积核,扩充边缘均为 1 个像素,使用池化层防止过拟合,得到 256 个 6×6 大小的特征图.

Step6. fc6 和 fc7 全连接层,使用 4096 个神经元,对 256 个大小为 6×6 特征图,进行一个全连接,即将 6×6 大小的特征图,进行卷积变为一个特征点. 再进行 dropout 随机从 4096 个节点中丢掉一些节点信息,同样是为了防止过拟合从而提高准确率,然后得到新的 4096 个神经元.

Step7. fc8 连接层,采用的是 65 个神经元,对 fc7 中 4096 个神经元进行全连接,然后通过高斯过

其识别训练过程基本一致,不同之处仅在于:车牌判断是将输入图片分为两类,即车牌与非车牌;而字符识别是将输入图片分为 65 类,其中包括 31 类中文字符、24 类英文字符和 10 类数字字符. CaffeNet 网络模型的结构如图 6 所示.

滤波器,得到 65 个 float 型的值,即为预测的可能性.

3 实验结果及分析

将候选车牌图块按照 7:3 的比例分为两类:训练集和测试集. 候选车牌图块来自于开源项目的数据集和搜索引擎搜索到的图片. 其中,车牌图块和非车牌图块共有 5906 张,选取 4160 张作为训练集,余下 1746 张作为测试集. 训练过程 loss 和 accuracy 曲线如图 7 所示,其中曲线 a、b、c 分别为测试集的正确率、训练集的 loss、测试集的 loss. 可以看出,在经过大约 700 此迭代后测试集的准确率稳定在了 94%.

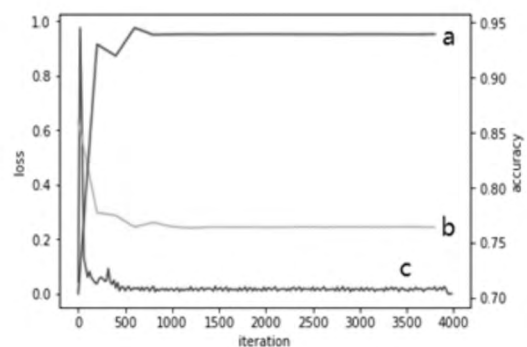


图7 车牌识别训练的 loss 和 accuracy 曲线

Fig. 7 License plate recognition training loss and accuracy curve

字符识别训练同样采用 CaffeNet 网络结构. 从车牌图片中分割出字符图片,共 65 组,其中 31 组中文字符,24 组英文字符(不含字母 I 和字母 O),10 组数字字符. 同样将 65 组字符分别按照 7:3 的比例分为两类训练集和测试集. 中文字符共有 13643 个,英文字符共 16510 个,数字字符共 6108 个,在迭代大约 2000 次后测试集的正确率趋于稳定,训练过

程的 loss 和 accuracy 曲线如图 8 所示,其中曲线 a 为测试集的正确率,曲线 b 为训练集的 loss,曲线 c 为测试集的 loss. 3000 次迭代的最终正确率为 96.4%.

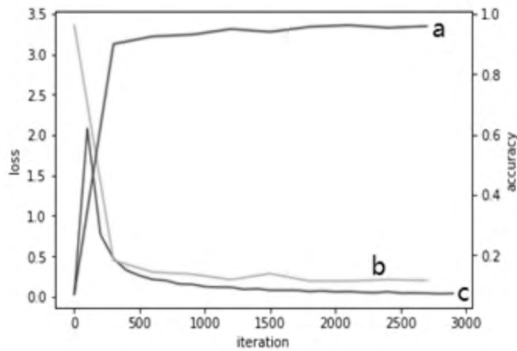


图 8 字符识别训练的 loss 和 accuracy 曲线

Fig. 8 Character recognition training loss and accuracy curve

本文的实验结果如表 1 所示,正确率优于模版匹配法和 BP 神经网络,略低于文献[4]中的 LeNet5 网络;识别用时优于 BP 神经网络和 LeNet5 网络,识别速度更快.其中,表 1 中的识别用时为 100 张字符图片的识别用时.虽然正确率略低于 LeNet5 网络,但本文旨在开发一种通用性较强的车牌自动定位和识别系统,且识别速度更快.由于本文所采用的数据集存在样本分布不均的情况,例如字符“藏”的数量远低于其他字符,导致测试集的正确率有所降低,但这并不会对实际使用造成影响.此外,文献[4-5]是仅仅对字符识别进行的测试,而本文则是对完整的车牌识别系统进行的测试,更具有实际可用性.

表 1 字符识别正确率比较

Tab. 1 Comparison of character recognition accuracy

网络	正确率/%	识别用时/s
模版匹配	81.2%	无
BP 神经网络	93.1%	0.785
LeNet5	97.8%	0.299
CaffeNet	96.4%	0.271

图 9~11 分别为暗光条件、倾斜角度和正常条件下拍摄图片的车牌识别结果图,本系统均可成功定位车牌并识别字符.从一张图片中定位并识别车牌的整体用时为 0.515 s.因此,本文设计并实现的车牌识别系统是一个高效可用的完整系统,可以应对绝大多数复杂条件下的车牌定位及识别工作,整体识别效果高于现有车牌识别系统.

参 考 文 献

[1] 耿庆田,赵宏伟.基于分形维数和隐马尔科夫特征的



图 9 暗光条件的识别结果

Fig. 9 The recognition result of dim light condition



图 10 倾斜角度的识别结果

Fig. 10 The recognition result of tilt angle condition



图 11 正常条件的识别结果

Fig. 11 The recognition result of normal condition

车牌识别[J]. 光学精密工程, 2013, 21(12): 3198-3204.
 [2] 曾泉,谭北海.基于 SVM 和 BP 神经网络的车牌识别系统[J]. 电子科技, 2016, 29(01): 98-101.
 [3] 吴聪,殷浩,黄中勇等.基于人工神经网络的车牌识别[J]. 计算机技术与发展, 2016, 26(12): 160-163

- + 168.
- [4] 赵志宏,杨绍普,马增强. 基于卷积神经网络 LeNet-5 的车牌字符识别研究[J]. 系统仿真学报, 2010, 22(3): 638-641.
- [5] 张立,朱玉全,陈耿. 基于卷积神经网络 SLeNet_5 的车牌识别方法[J]. 信息技术, 2015, (11): 7-11.
- [6] 丁怡心,廖勇毅. 高斯模糊算法优化及实现[J]. 现代计算机, 2010, (8): 76-77 + 100.
- [7] 张桂华,冯艳波,陆卫东. 图像处理的灰度化及特征区域的获取[J]. 齐齐哈尔大学学报, 2007, 23(4): 49-52.
- [8] 乔闹生,邹北骥,邓磊等. 一种基于图像融合的含噪图像边缘检测方法[J]. 光电子·激光, 2012, 23(11): 2215-2220.
- [9] 郭佳,刘晓玉,吴冰,等. 一种光照不均匀图像的二值化方法[J]. 计算机应用与软件, 2014, 31(3): 183-186 + 202.
- [10] 潘巍,杨娜菲,安荣. 基于梯度与形态学的低质量车牌定位[J]. 计算机工程, 2011, 37(13): 144-146.
- [11] Suzuki S, Be K. Topological structural analysis of digitized binary images by border following [J]. Computer Vision, Graphics and Image Processing, 1985, 30(1): 32-46.
- [12] 黄社阳,刘智勇,阮太元. 基于 HSV 颜色空间和 SVM 的车牌提取算法[J]. 计算机系统应用, 2014, 23(8): 150-153.
- [13] 杨涛,张森林. 一种基于 HSV 颜色空间和 SIFT 特征的车牌提取算法[J]. 计算机应用研究, 2011, 28(10): 3937-3939 + 3976.
- [14] 管焱然,管有庆. 基于 OpenCV 的仿射变换研究与应用[J]. 计算机技术与发展, 2016, 26(12): 58-63.
- [15] Nistér, D., Stewénius, H. Linear time maximally stable extremal regions. [C]//Springer. Computer Vision—ECCV 2008, Berlin: Springer, 2008: 183-196.
- [16] 杨楠. 基于 Caffe 深度学习框架的卷积神经网络研究[D]. 石家庄: 河北师范大学, 2016.