

杭州电子科技大学

硕士学位论文

题目：基于神经网络的车牌识别技术研究

研究生 王 晶

专 业 电子与通信工程

指导教师 李训根 副教授

完成日期 2016 年 12 月

杭州电子科技大学硕士学位论文

基于神经网络的车牌识别技术研究

研 究 生：王晶

指导教师：李训根 副教授

2016 年 12 月

Dissertation Submitted to Hangzhou Dianzi University
for the Degree of Master

Research on License Plate Recognition Technology Based on Neural Network

Candidate: Wang Jing

Supervisor: Associate professor Li Xungen

December, 2016

杭州电子科技大学

学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明： 所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。申请学位论文与资料若有不实之处，本人承担一切相关责任。

论文作者签名： 日期： 年 月 日

学位论文使用授权说明

本人完全了解杭州电子科技大学关于保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属杭州电子科技大学。本人保证毕业离校后，发表论文或使用论文工作成果时署单位名称仍然为杭州电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。（保密论文在解密后遵守此规定）

论文作者签名： 日期： 年 月 日

指导教师签名： 日期： 年 月 日

摘 要

自动牌照识别系统（LPR）利用数字图像处理技术来定位和识别牌照号码上的字符，并将结果输出为文本字符串或其他类型的数据格式，其可以被操作者容易地理解。LPR 系统已经应用于需要通过其车牌号码自动控制机动车辆的存在和识别的各种监视应用，例如自动电子收费系统（ETC），自动泊车服务员。因为 LPR 应用给公路和城市交通的现代化管理带来了极大的便利，进一步促进了模式识别领域的发展，近几年来，神经网络可谓是模式识别领域最令人瞩目的方向，尤其是深度学习神经网络，其中卷积神经网络（CNN）是一种常见的深度学习架构，近几年在语音分析和图像识别领域非常火热，并且已经得到了应用。它的有两个优点：权值共享和稀疏连接，使网络结构与生物神经网络更加接近，权值共享降低了网络模型的复杂度，并大量地减少了权值的数量。

本文第一章论述了汽车牌照识别的主要几种应用技术和现阶段的国内外发展动向。第二章介绍了车牌图像处理方面的相关技术，包括图像灰度化、图像增强、以及二值化、边缘检测等前期处理方法；然后对于至关重要的车牌定位，本文综述了几种常用的车牌定位算法，包括纹理特征分析法、数学形态学方法、以及小波变换等；随后简要地介绍了基于投影特征、连通域特征的字符分割算法；最后是介绍了现有的车牌字符识别技术，最常用有模板匹配法以及 BP 神经网络识别法；第三章详细介绍了卷积神经网络的发展、结构、推导及实现、还有训练过程；第四章逐步实现了车牌识别的处理过程；第五章给出车牌识别系统实验结果及对结果的分析；第六章总结全文内容并展望车牌识别研究的未来发展方向。

本文是基于深度学习神经网络的车牌识别系统设计，与传统的浅层神经网络相比，深度学习模型具有更强的学习能力，可提高车牌字符的识别率。本文首先采用了一系列的图像预处理技术对车牌图像进行初步的车牌信息提取；由于单一的定位方法很难在复杂的情况下准确定位，故本文采用的是结合车牌的纹理和颜色特征，再通过粗定位和细定位相结合的方法对车牌进行定位；定位后的车牌经过前期处理后总免不了会包含一些噪声点，考虑到这一点本设计选用了结合先验知识的垂直投影分割方法对定位后的车牌进行字符分割；最后采用卷积神经网络的方法用于识别分割后的字符，得出识别结果。其中 CNN 的结构是在经典的 CNN 结构基础上，结合车牌字符的数目及汉字识别的难度，增加了卷积层的特征图数目和输出层的神经元个数。

从实验结果可以看出，多种预处理方法的结合可以更准确的定位、分割车牌，另外，本文采用的卷积神经网络的方法来识别字符，相比传统的神经网络字符的识别率提高了 5%。

关键词：车牌识别，字符识别，深度学习，卷积神经网络

ABSTRACT

The Automatic License Plate Recognition System (LPR) utilizes digital image processing techniques to locate and identify the characters on the license number and output the result as a text string or other type of data format that can be easily understood by the operator. The LPR system has been applied to a variety of surveillance applications, such as automated electronic toll collection (ETC), automatic parking attendants, which require the automatic control of the presence and identification of motor vehicles by their license plate numbers. Because LPR application brings great convenience to modern management of highway and urban traffic, and further promote the development of pattern recognition field, neural network is the most noticeable direction of pattern recognition in recent years, especially in depth learning Neural network, in which convolution neural network (CNN) is a common depth learning architecture, which has been applied in the field of speech analysis and image recognition in recent years, and has been applied. It has two advantages: weight sharing and sparse connection, so that the network structure and biological neural network closer to the weight sharing reduces the complexity of the network model, and a significant reduction in the number of weights.

In the first chapter of this paper, the author discusses the main application technologies of automobile license plate recognition and the development trend of domestic and foreign. The second chapter introduces the relevant technology of license plate image processing, including image grayscale, image enhancement, binarization, edge detection and other pre-processing methods; and then for the most important license plate location, this paper introduces several commonly used Including the texture feature analysis method, mathematical morphology method, wavelet transform and so on. Then we introduce the character segmentation algorithm based on the projection feature and the connected domain feature. At last, we introduce the characters recognition technology of the license plate, The third chapter introduces the development, structure, derivation and implementation of convolution neural network in detail, and the training process. Chapter 4 deals with the implementation of license plate recognition in a step-by-step way. The fifth chapter introduces the experimental results and the analysis of the results of the license plate recognition system. The sixth chapter summarizes the content of the paper and looks forward to the future development of the license plate recognition.

This paper is based on the deep learning neural network to design the license plate recognition system. Compared with the traditional shallow neural network, the depth learning model has stronger learning ability and can improve the recognition rate of the license plate characters. In this paper, a series of image preprocessing techniques are used to extract the license plate information from the license plate. Because the single positioning method is difficult to locate accurately under complex conditions, this paper adopts the combination of the texture and color characteristics of the license plate. The location of the license plate after positioning through the pre-processing will inevitably contain some noise points, taking into account this point in the design of a combination of prior knowledge of the vertical projection method of segmentation on the positioning of the positioning of the license plate, And then uses the method of convolutional neural network to recognize the segmented characters and obtain the recognition result. The structure of CNN is based on the classical

CNN structure, which increases the number of the feature map and the number of neurons in the output layer by combining the characters of the license plates and the difficulty of Chinese character recognition.

From the experimental results we can see that the combination of a variety of pretreatment methods can be more accurate positioning, segmentation license plate, in addition, the use of convolutional neural network method to identify characters, compared to traditional neural network character recognition rate 5%.

Key words: license plate recognition, character recognition, deep learning, convolutional neural network

目 录

摘 要.....	I
ABSTRACT.....	II
目 录.....	IV
第 1 章 绪 论.....	1
1.1 课题研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.3 本文研究内容.....	3
第 2 章 车牌图像处理技术.....	5
2.1 灰度化及图像增强.....	5
2.1.1 图像灰度化.....	5
2.1.2 图像增强.....	6
2.2 二值化.....	8
2.3 边缘检测.....	9
2.3.1 Robert 算子.....	9
2.3.2 Prewitt 算子.....	10
2.3.3 Sobel 算子.....	10
2.3.4 Canny 算子.....	10
2.4 车牌定位.....	12
2.4.1 基于纹理特征分析的车牌定位.....	12
2.4.2 基于数学形态学的车牌定位.....	14
2.4.3 基于小波变换的车牌定位.....	15
2.5 字符分割.....	16
2.5.1 基于垂直投影法的字符分割.....	16
2.5.2 基于连通域特征的字符分割.....	18
2.6 字符识别.....	19
2.6.1 模板匹配算法.....	20
2.6.2 基于 BP 神经网络的识别算法.....	20
2.7 本章小结.....	21
第 3 章 卷积神经网络.....	23
3.1 深度学习介绍.....	23
3.2 卷积神经网络概述.....	23
3.3 卷积神经网络推导及实现.....	25
3.3.1 卷积层.....	26
3.3.2 子采样层.....	26
3.3.3 完全连接层.....	27
3.4 卷积神经网络结构.....	27
3.4.1 稀疏连接.....	28
3.4.2 权值共享.....	29
3.4.3 池化层.....	30

3.5 卷积神经网络的训练.....	30
3.6 本章小结.....	31
第4章 基于卷积神经网络的车牌识别.....	32
4.1 车牌图像预处理.....	32
4.1.1 车牌图像的灰度化.....	32
4.1.2 车牌图像二值化.....	34
4.1.3 车牌边缘检测.....	35
4.2 车牌定位的实现.....	36
4.3 车牌字符分割.....	37
4.4 基于卷积神经网络的字符识别.....	38
4.5 本章小结.....	40
第5章 实验结果分析.....	41
5.1 实验结果测试.....	41
5.2 实验结果分析.....	43
5.3 本章小结.....	44
第6章 总结与展望.....	45
6.1 论文总结.....	45
6.2 研究方向展望.....	45
致 谢.....	47
参 考 文 献.....	48
附 录.....	51

第1章 绪论

1.1 课题研究背景及意义

现实的生活中,辨别重要的信息和对信息进行分析和处理,是现代智能交通的用途所在。随着经济的发展,车辆的生产和保持率全球迅速增长,对运输和车辆管理产生了很大的压力和挑战,因此自动管理车辆已迫切需要。

智能交通系统对现实生活中的车牌图像进行有效的解析,是一项利用我们的计算机来实现的图像识别技术。图像的识别技术的发展应用能给人们在日常生活中带来极大的方便,同时也能有利于社会的发展。由于这项技术的重要作用,人类投入了大量的精力对其进行了研究,在近阶段这项技术已经有了突破性的进展。在人类的日常生活中得到了成功而又广泛的应用在了很多的领域,其中特别在车牌识别和障碍检测中发挥了重要的作用。

在交通管理及应用中,复杂的视频车辆检测系统不断推出,车牌识别系统已经发现了各式各样的地方,来适应自身系统,而不仅仅是监测收费点或停车场等的寻访。它还能集成到视频车辆的检测系统,往往是安装在感兴趣的领域,如交叉路口控制、交通监控等。确定车辆违反交通法规或发现被盗车辆。

车牌是车辆的重要身份,也是区分车辆在国际范围内的具体身份。因此,车牌识别不仅可以减少由于预订的劳动造成的误差概率,而且还实现车辆的自动故障排除和跟踪。因此,作为智能交通系统的核心技术之一,车辆牌照识别技术在交通行业和公共安全管理部门得到了相当的重视和应用,如交通执法,收费,交通调查,车辆跟踪和监视。

在现实的生活中,车牌识别技术的主要的用途^[1,2]有如下几点:

第一、各个场所的停车点和小区门口进出

车牌识别技术应用在各个场所的停车点和小区门口进出处,它的主要功能如下几点:识别进出车辆的号码、时间等,例如,可以自动判断车辆是否属于该社区,对外来车辆进行计时收费等。在停车点与社区进出口的车辆管理若采用人为管理,手动记录号码,进出时间等,非常耗时又复杂,还容易出错,而且人力上也需要不断投入。而引入车牌识别系统,就可以很好地解决这些问题,能省去许多人力物力,以及避免人为记录错误等情况,对于车牌被遮或难以辨认这样情况。也可以通过自身技术的优势进行辨别,为社区的管理提供可靠的保障。

第二、高速公路上收费站点的实施应用

高速公路的建设,一直都是交通部门的重要的工作内容,因为它是我国交通建设与管理的重要枢纽,而各个进出口则设有许许多多的收费站。在各个收费站中,不仅需要进行收费管理,而且还需要进行道路交通以及车票的监测控制管理,因此将这项技术能够应用于高速公路收费站能大大地提高收费站的工作效率,包括车牌的识别,违法车辆的监控等情况。

有时候会因为天气的原因,或者车速过快,这对于车牌的识别难度会增加,这就需要有更高的车牌识别的技术来达到识别车牌的目的。传统的车牌识别技术,由于其自身的局限性导致了车牌识别率低的问题,然而现在的高清的车牌识别技术有效的解决了这个难题,为交通部门和其他相关的部分提供了清晰的画面和有效的数据,大大的提高了他们的工作效率。另外,车牌识别系统还可以将车牌识别的有效信息发送到指定的中心管理服务器上,并与盗窃嫌疑车辆车牌信息的比较,判断高速公路收费站的车辆入口和出口数据库中是否有盗窃嫌疑车辆,这样可以帮助工作人员很快的对相关车牌的信息做出有效的判断。

第三、应用在日常公路卡口中

随着物质生活水平的提高,现在的社会私家车越来越多,而个人驾车技术的差异也导致了在行驶的过程中事故频发,给交通管理带来了各种问题。对于在事故发生时要是能清晰的记录事故发生的整个过程,能有效迅速的处理这类问题,车牌识别技术在其中起了至关重要的作用。这项技术的引入使得交通管理变得更加的智能,减少了人力资源的投入,同时这使得整个系统更加的稳定可靠,收集数据也可以更加的直观方便。

车牌识别系统的使用已经遍及交通道路和人类的生活中。

1.2 国内外研究现状

车牌识别技术最初是在 20 世纪 80 年代的时候由国外研究者提出的,一开始只有很少的有关图像识别方面的技术用来处理车牌自动的识别,还没有形成完整的车牌自动识别体系,多数情况下是需要人工处理部分来帮助实现车牌识别。到了 20 世纪 90 年代后,由于应用需求,大量的研究者投入到车牌自动识别的研究中,同时也发表了很多关于车牌自动识别方便的文章,使得车牌识别技术的应用越来越成熟。

车牌自动识别有对静态图像处理的系统,后来由于需要又出现了对动态图像处理的系统,比如以静态图像为处理目标的有 PaoloComelli 提出的光学识别汽车车牌系统,该系统对汽车的速度有限制,因此它的应用是在收费站对汽车实时拍摄的静态图像,然后再进行识别,它的识别率接近 91%;以动态图像为处理目标的有英国 Alphatech 公司研究出的 ARGUS 车牌识别系统,该系统可以在车速可高达每小时 100 英里的情况下识别时间约为 100 毫秒;以色列的 Hi-TechSolutions 公司是自动车辆识别系统技术比较领先的研究者之一,它提出的一系列的解决方案深受推崇。并且已经投入开发如 SEE/CAR 系列,可以在汽车车速高达 80 公里/时的情况下每秒还能识别 2-3 辆汽车,SEE/CAR 系列的系统对于几个不同国家车牌识别都适用,See/Car Chinese 系统由于是国外研发的,所以还不能在我国广泛的推广主要由于其对汉字无法识别。美国的 ALPHATECH 公司研发飞 AVI 产品,因为其很好的容错性而在车牌识别领域有较高的地位。另外,新加坡 optasia 公司针对自己国家的车牌特征研发了 VLPRS 产品;澳大利亚 Auty 等建立了 Safe--Cam 系统用于交通监控的,这个系统能不间断的运行一整天,而且也能抓拍到高速的骑车静止图像。最主要的还能准确快速的处理图像。此外,许多

西方发达国家都自己研究出了适合于本国的车牌自动识别系统，比如：德国、英国、日本、意大利、加拿大等。

国内在 90 年代也开始了自己的车牌识别系统的研究，并取得了一定的成果，比如定位算法较好的有郭大、陈礼民等人提出的基于车牌底色识别的车牌定位方法^[3]，该文章中提出了结合汽车车牌纹理特征分析和颜色特征分析实现车牌定位的方法，这种方法明显地克服了单用纹理和形状分析时难以解决的车牌区域变大的问题；识别算法较好的有魏武、张起森等人提出的一种基于模板匹配的车牌识别方法^[4]，该方法可不需分割车牌中的字符而有效识别实际系统中低质量和模糊的车牌图像中的字符，识别率可达 95% ，而且识别时间不超过 1 s；还有吴聪、殷浩等人提出了基于人工神经网络的车牌识别，他们采用了遗传算法与神经网络相结合的方法，该方法具有时效性和鲁棒性。除此之外，很多国内知名的大学，像浙江大学、上海交通大学、西安交通大学、等知名学府的重点实验室^[5]，都在做有关这方面的探索。

目前比较成熟的有中科院自动化研究所北京汉王科技有限公司研究的智能交通产品“汉王眼”^[6]，还有浙江大学图形图像研究所潘云鹤等人研究的“车牌通”产品等。另外杭州也有很多视频监控公司，比如海康威视、浙江大华，都对车牌识别系统进行研究。这些公司所提供的对这项技术的识别率都高的惊人，基本都超过了 90%，不过这主要是因为他们的应用场景受到了限制，如果在现实复杂的环境下是达不到这么高的识别率的。

然而车牌识别系统业界一直都没有一个成熟的百分百适用的方案。一些 90% 以上的车牌识别系统都是跟高清摄像机做了集成，由摄像头传入的高分辨率图片进入识别系统，可以达到较高的识别率。但是如果图像分辨率一旦下来，或者图里的车牌脏了的话，那么很遗憾，识别率远远不如我们的肉眼。也就是说，距离真正的智能的车牌识别系统，目前已有的系统还有许多挑战。

1.3 本文研究内容

本文的主要研究内容，在于使用一种新的字符识别算法，与传统的识别算法相比，能提高在恶劣条件下采集到的车牌图像的识别率。本文的具体组成框架如下：

第一章，主要介绍了车牌识别技术的国内和国外的现状和发展，同时也阐述了其中多种应用技术。

第二章，重点解析了车牌图像处理的核心技术，同时表述了图像灰度化、二值化和边缘检测等前期处理方法；然后介绍了几种常用的车牌定位算法，包括数学形态学方法、纹理特征分析法、以及小波变换等；简要地介绍了基于投影特征、连通域特征的字符分割算法；最后是介绍了现有的车牌字符识别技术，最常用有模板匹配法以及 BP 神经网络识别法^[7]。

第三章，详细地介绍了卷积神经网络的发展、结构、推导及实现、还有训练过程。

第四章，对车牌识别的整个过程进行分析，从车牌的预处理、到车牌定位的方法、最后到车牌字符分割已经分割后的字符识别，并展示了每一步的处理结果。

第五章，介绍了车牌识别系统软件设计的实现方法和实验结果。

第六章，总结全文内容，归纳了一些值得进一步研究思考的问题，最后展望研究未来发展方向。

第 2 章 基于车牌识别的图像处理技术

图像的采集，图像的预处理过程，车牌定位模块，字符分割和字符识别模块组成了车牌图像识别中五个常用的部分^[8]。其中车牌定位技术与字符识别技术是车牌识别技术研究中至关重要的部分，同时也是难点。

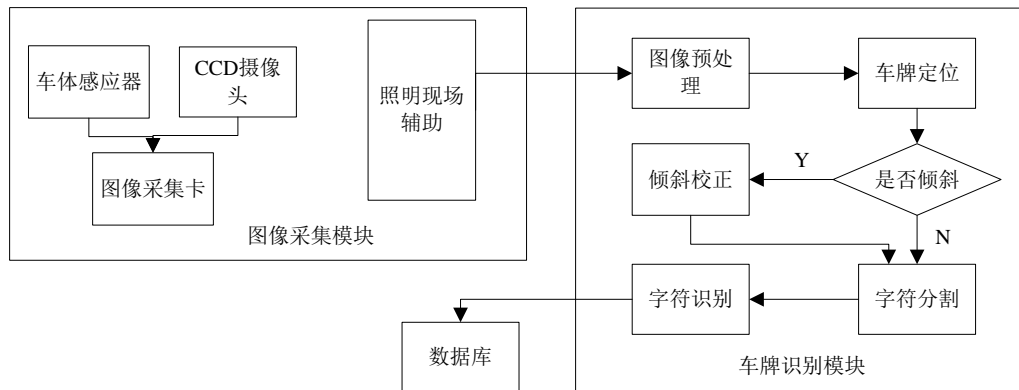


图2.1 车牌识别系统的构成

2.1 灰度化及图像增强

2.1.1 图像灰度化

灰度图是指不包含色彩信息只包含亮度信息的图像，例如平时看到的亮度连续变化的黑白照片就是很典型的一幅灰度图。灰度化处理的定义^[9]是把一幅彩色图像变为灰度图像。我们知道，彩色图像有三个分量为红（R）、绿（G）、蓝（B），每个分量取值的范围都为 0 至 255，利用排列组合计算也就相当于 1600 多万种颜色的变化范围。由于直接从图像采集设备得到的图像颜色分辨率高，包含的信息数据量大，极占存储空间会使图像处理时 CPU 的执行速度变慢，为了方便后续的图像分析处理，常将彩色图像转换为灰度图像，以加快系统处理速度。灰度化处理就是使彩色的红（R）、绿（G）、蓝（B）三个分量相等的过程。它对像素值进行了简练的处理，例如灰度值大的像素点用亮来表示，零为最小的像素值，灰度值小的像素点为黑色。以下是灰度处理的三种方法：

第一种是最大值法，令 R、G、B 的值等于这三个值中最大的一个，即

$$R = G = B = \max(R, G, B) \quad (2.1)$$

第二种是平均值法，取 R、G、B 三个的值和的平均值，即

$$R=G=B=\frac{R+G+B}{3} \quad (2.2)$$

第三种是加权平均值法，根据 R、G、B 不同的比重，选择合适的权值，并取 R、G、B 三个值的加权和平均，即

$$R=G=B=\frac{W_rR + W_cG + W_bB}{3} \quad (2.3)$$

其中 W_r 是 R 分量的权值、 W_c 是 G 分量的权值、 W_b 是 B 分量的权值。由于人眼对红、绿、蓝三种颜色的敏感度不同，敏感度最高的是绿色，所以 G 的比重最小，取 $W_c=0.11$ ；敏感度最低是蓝色，所以 B 的比重最大，取 $W_b=0.59$ ；对红色的敏感度居中，取 $W_r=0.30$ ，这样得到的灰色图像效果最佳。

2.1.2 图像增强

图像增强^[10]是对数字图像进行有效的调整的过程，为了更深入的对图像进行分析。图像增强的目的是改善图像中的信息使图像呈现出更好的解释性，使人类观看者便于清晰地感知，或者在其他自动图像处理技术中提供更清晰易于处理的输入图像，方便后续处理。例如，可以删除噪点，锐化或增亮图像，从而更容易识别关键特征。

图像增强是用于视觉检查或用于机器分析的数字图像质量的改进，而不知道劣化的来源。对于生物识别识别，图像增强的主要作用是减轻照明的影响。图像增强技术可以分为两大类：直接对像素进行操作的空间域方法和对图像的傅立叶变换进行操作的频域方法。广泛使用的空间域方法是直方图均衡^[11]，旨在通过使用归一化累积直方图作为灰度映射函数来增强图像的对比度。应当注意，许多图像增强方法是面向问题的：在一种情况下工作良好的方法对于另一种情况可能是完全不适当的。对于生物测定学应用，选择图像增强方法的优点是这些方法最终可以提高识别性能。

直方图均衡是用于增强图像的外观的常用技术。直方图均衡化处理的中心思想是把原灰度图像中灰度值比较集中的区域的值均匀地散开到整个灰度范围内，相当于把灰度图像的像素值重新进行分配，这样就使得一定范围内的像素值比较均匀的分布。

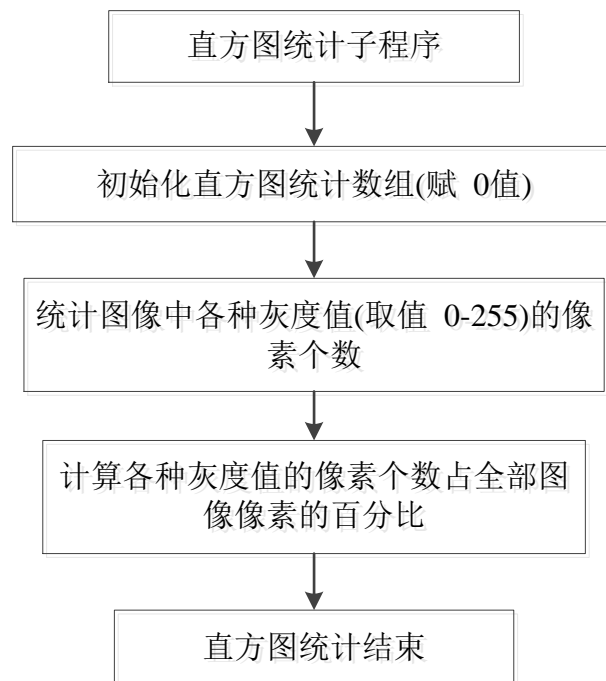


图 2.2 数字图像直方图统计过程

上图是对于一个灰度图像如何得到它的灰度值函数的统计过程。

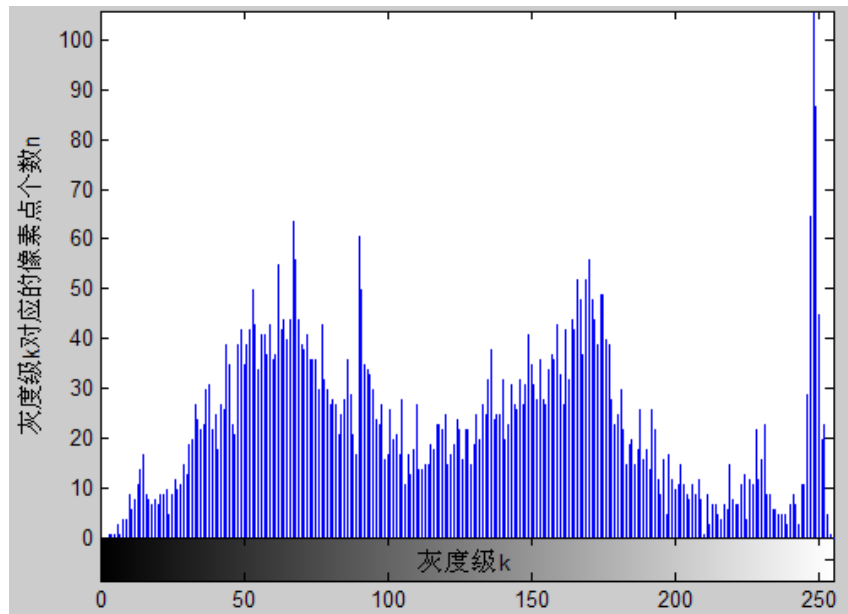


图 2.3 灰度级的像素的个数

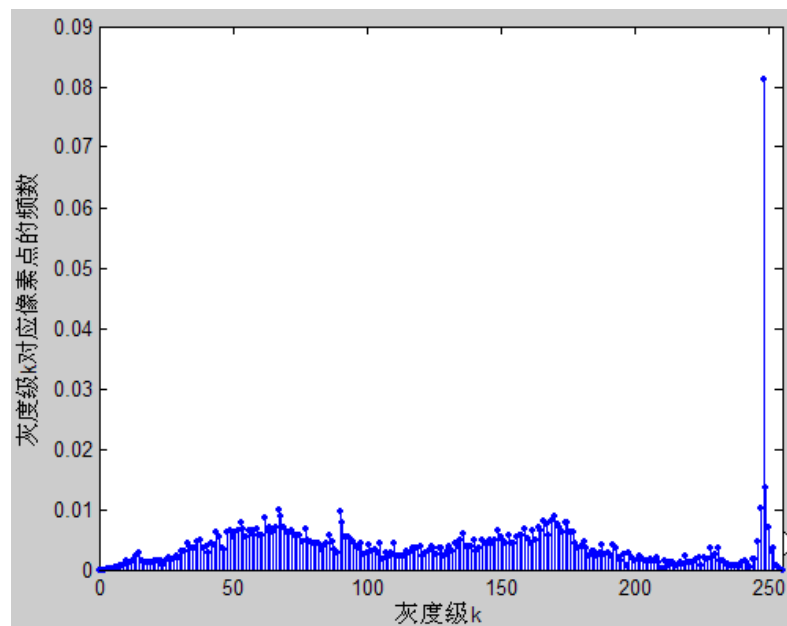


图 2.4 灰度级的像素的频率

依据定义在离散形式处理步骤如下：

运算						
原始图像灰度级 k	0	1	2	3	...	255
原始直方图 $p(k) = \frac{n_k}{n}$	0.19	0.25	0.21	0.16	...	0.02
计算累积直方图各项: $t_k = \sum_{i=0}^k \frac{n_i}{n} = \sum_{i=0}^k p(i)$	0.19	0.44	0.65	0.81	...	1.00
取整扩展: $t_k = \text{int}[(L-1)t_k + 0.5]$	48	112	166	207	...	255
均衡化后灰度级映射关系: $k \rightarrow t_k$	0 \rightarrow 48	1 \rightarrow 112	2 \rightarrow 166	3 \rightarrow 207	...	255 \rightarrow 255

图 2.5 直方图均衡化原理

但是直方图均衡化原理也有缺点：由于变换后图像的灰度级减少，可能会使某些细节丢失；还有一些图像，比如有高峰的直方图，再经过均衡化处理后，就会使对比度增强的非常不自然。

2.2 二值化

图像二值化^[12]是将图像上的像素点的灰度置成黑色 0 或白色 255，这样整个图像就变成了黑白效果的图。就是通过设定适当的阈值将 256 个灰色等级的图像转换成可以反映图像整体或局部特征的二值图像。二值图像的像素值只有 0 和 255，不再像彩色图像一样涉及到像素的多级值，二值化处理后会使得后续进一步的处理变得简单很多，而且数据的处理和压缩量都会变小。这样在实用系统中，处理的速度高、成本低、信息量大的图像处理的花销将大大减小。

在全局阈值法中，可以看到具有前景字符和背景的牌照图像两部分。当车牌上的光散射的均匀的时候，那么用二值化法可以设置合适阈值来分离背景和字符。当散射不均匀时，车牌二值化不能用全局阈值来设置。此时必须使用本地阈值方法。图像二值化处理的关键点是合理地选择阈值。当阈值设置太小时，噪声会变大；将阈值设置得过大时，它的分辨率将随之降低同时滤除非噪声信号。

设原灰度图像函数表示为 $f(x,y)$ ，二值化后的图像函数表示为 $g(x,y)$ ， t 为用来区分对象和背景的阈值，二值化的数学表达过程表示如下：

$$g(x,y) = \begin{cases} 0 & f(x,y) < t \\ 255 & f(x,y) > t \end{cases} \quad (2.4)$$

最简单的情况是阈值处理，其中强度分布被在选定阈值处有效的阶跃函数代替。在这种情况下，输入图像中具有低于阈值的灰度级的任何像素在输出图像中被映射到 0，其他像素映射到 255。

2.3 边缘检测

边缘检测^[13]是图像处理的关键手段，其是实现背景和前景的分离的有效方式，其从复杂背景中找到数字区域。传统的边缘检测器从图像的高频（噪声）分量提取边缘信息。形态是另一个重要的图像处理理论，它有效地消除了噪声并保留了图像的原始信息。

边缘检测包括各种数学方法，其旨在识别数字图像中图像亮度急剧变化或者更正式地具有不连续性的点。边缘是指图像的亮点发生巨变的点连接成一条曲线，这条曲线可以有效的区别图像的特征部分和背景部分。在特征提取和特征检测的领域内，边缘检测成为了计算机视觉和数字图像处理中的一个必不可少的工具。

在理想情况下，边缘检测后的图像中会形成一条连续的曲线，这个曲线是划分特征区域和背景区域的分割线。所以通过边缘检测算法分割图像后，大量的无用数据将会被过滤掉，这样就减少了我们后续的数据处理存储量，同时保留图像的重要结构性质。如果边缘检测步骤成功，则因此可以显著简化解释原始图像中的信息内容的后续任务。然而，并不总是可以从复杂的现实生活图像获得这样的理想边缘。

从图像提取的边缘通常受到碎片的阻碍，意味着边缘曲线不连接，缺失边缘段以及与图像中的感兴趣现象不对应的假边缘，从而使解释图像数据的后续任务复杂化。

到现在已经有许多的边缘检测方法被研究出来，下面介绍几种一阶导数表达式的算子：

2.3.1 Robert 算子

Roberts 算子^[14]是将任意一对互相垂直方向上的差分看成求梯度的近似方法，利用这种原理，Robert 算子采用的对角方向相邻两像素值之差代替该梯度值。

假设灰度图像 $f(x,y)$ 是具有整数像素坐标的输入。其梯度定义为：

$$\nabla f(x,y) = \frac{\partial f(x,y)}{\partial x} i_x + \frac{\partial f(x,y)}{\partial y} i_y \quad (2.5)$$

由数学中微分和差分知识，我们可以用下面差分来近似表示上述微分的过程，其中 $(x+1,y+1)$ 是 (x,y) 斜向上的像素点：

$$|\nabla f(x,y)| = \sqrt{(f(x,y) - f(x+1,y+1))^2 + (f(x+1,y) - f(x,y+1))^2} \quad (2.6)$$

一般情况下，可以将上式简化为如下两种形式：

$$|\nabla f(x,y)| = |f(x,y) - f(x+1,y+1)| + |f(x+1,y) - f(x,y+1)| \quad (2.7)$$

$$|\nabla f(x,y)| = \max(|f(x,y) - f(x+1,y+1)|, |f(x+1,y) - f(x,y+1)|) \quad (2.8)$$

公式(2.7)和(2.8)表示称模板的形式如下：

$$R_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad R_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (2.9)$$

Roberts 算子在斜向边缘的检测效果要低于在垂直边缘和水平边缘的效果，同时在边缘处的定位也比较准，但对于具有陡峭的低噪声的图像效果最好，倾斜的效果不理想，容易造成漏检，会产生一些伪边缘。因此，该算子适用于边缘明显且噪声较少的图像分割。

2.3.2 Prewitt 算子

Prewitt 算子^[15]主要分为两个模型样板，这两个模板分别是用于检验测量水平边缘的 R_h 样板，另一个是检验测量垂直边缘的 R_v 样板。如下所示：

$$R_H = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad R_V = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad (2.10)$$

Prewitt 算子的计算方法是利用它的像素点与其相邻的周围像素点之间的灰度差，此灰度差一般会在图像边缘处取得极限值，利用这个原理来检测图像边缘，它的计算方式决定了 Prewitt 算子对噪声具有平滑作用，缺点是定位精度不够高。

2.3.3 Sobel 算子

Sobel 算子^[16]对图像执行 2-D 空间梯度测量，因此强调对应于边缘的高空间频率的区域。通常，其用于在输入灰度图像中的每个点处找到近似绝对梯度大小。

在理论上，运算符由一对 3×3 卷积内核组成，一个内核简单地是另一个旋转 90° 。

$$R_H = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad R_V = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (2.11)$$

这些内核被设计为最大地响应相对于像素网格垂直和水平地延伸的边缘，两个垂直取向中的每一个的一个核。内核可以单独应用于输入图像，以产生每个方向的梯度分量的单独测量（称为这些 G_H 和 G_V ）。然后将它们组合在一起以找到每个点处的梯度的绝对幅度和该梯度的取向。梯度大小由下式给出：

$$G = \sqrt{G_H^2 + G_V^2} \quad (2.12)$$

产生空间梯度的边缘的取向角由下式给出：

$$\theta = \arctan\left(\frac{G_H}{G_V}\right) \quad (2.13)$$

在这种情况下，取向 0 意味着从黑到白的最大对比度的方向在图像上从左到右延伸，并且从其逆时针测量其他角度。

Sobel 算子的计算速度比较慢，但是它的较大卷积核在更大程度上能平滑输入图像，因此使得操作者对噪声不太敏感。

2.3.4 Canny 算子

Canny 边缘检测器^[17]是边缘检测算子，其使用多级算法来检测图像中的宽范围的边缘。Canny 算子的计算方式是这样的，它很巧妙的把数学上求单元函数极大值的问题和边缘检测问题相互联系，并且它具有高的检测精度和很好的信噪比。

Canny 算法对图像进行处理如下：

1) 高斯滤波器

由于所有边缘检测结果容易受到图像噪声的影响，因此必须滤除噪声以防止由噪声引起

的错误检测。为了平滑图像，应用高斯滤波器以与图像卷积。

$$G = \frac{1}{2\pi\sigma^2} e^{(-x^2/2\sigma^2)} \quad (2.14)$$

此步骤将稍微平滑图像，以减少边缘检测器上明显噪声的影响。

2) 计算梯度的幅值和方向

取 Canny 算子的一阶卷积差分模板为：

$$H_1 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \quad (2.15)$$

边缘检测算子返回水平方向（G_x）和垂直方向（G_y）上的一阶导数的值。由此可以确定边缘梯度方向 $\theta(x,y)$ ：

$$M(x,y) = \sqrt{E_x(x,y)^2 + E_y(x,y)^2} \quad (2.16)$$

$$\theta(x,y) = \arctan \frac{E_y(x,y)}{E_x(x,y)} \quad (2.17)$$

其中：

$$E_x(x,y) = f(x,y) * H_1(x,y) \quad (2.18)$$

$$E_y(x,y) = f(x,y) * H_2(x,y) \quad (2.19)$$

3) 对梯度幅值进行非极大值抑制

非极大值抑制是边缘稀疏技术，非极大值抑制应用于“细化”边缘。在应用梯度计算之后，从梯度值提取的边缘仍然相当模糊。非极大值抑制可以帮助将除了局部最大值之外的所有梯度值抑制为 0，其表示具有强度值的最明显变化的位置。梯度图像中的每个像素的算法是：将当前像素的像素和边缘强度在它们之间的正和负梯度方向上的边缘强度相互比较；假如当前像素的边缘强度与具有相同方向的掩模中的其它像素（即，指向 y 方向的像素）相比是最大的，则将其与垂直方向上的像素上方和下方的像素进行比较，并保留该值。否则，该值将被抑制。

在一些实施方式中，算法将连续梯度方向分类成一小组离散方向，然后在前一步骤的输出（即边缘强度和梯度方向）上移动 3×3 滤波器。在每个像素处，如果其幅度不大于梯度方向上的两个相邻像素的幅度，则其抑制中心像素的边缘强度（通过将其值设置为 0）。

4) 双阈值

在应用非极大值抑制之后，边缘像素相当准确地呈现真实边缘。然而，在这一点上仍然存在由噪声和颜色变化引起的一些边缘像素。为了摆脱来自这些干扰因素的杂散响应，必须用弱梯度值滤出边缘像素并且用高梯度值保持边缘。因此，设置两个阈值以阐明不同类型的边缘像素，一个称为高阈值，另一个称为低阈值。如果边缘像素的梯度值高于高阈值，则将它们标记为强边缘像素。如果边缘像素的梯度值小于高阈值并且大于低阈值，则将它们标记为弱边缘像素。如果像素值小于低阈值，则它们将被抑制。这两个阈值是经验确定的值，当应用于不同的图像时需要定义这些值。

2.4 车牌定位

通过图像处理技术识别车辆的问题可以分为两个子问题：定位车牌和识别车牌字符。牌照位置是一项具有挑战性的任务，因为它必须能够补偿各种户外条件，如照明，阴影和复杂的背景。在识别领域中长期以来一直在研究字符识别的问题，同时对牌照位置进行了积极的研究。粗略地说，每个研究通过提取牌照区域的候选者的方法和从这些候选者中选择正确的一个的方法来区分。接下来为了叙述方便我们把牌照区域的候选者被称为 LP 候选者。许多研究集中在颜色和边缘信息以提取 LP 候选，并已使用几何属性，例如牌照的宽度，高度和纵横比来选择正确的。牌照具有特定的颜色，因此它可以在特定颜色的区域找到。许多研究人员使用颜色特征提取 LP 候选。通常，在室外环境中颜色特征不稳定，因此基于颜色的方法^[18,19]只能应用于其中牌照的背景颜色的车辆图像。

在诸如照明，阴影和相机特性的可变条件下，边缘特征比颜色特征更鲁棒。车牌的区域具有高密度的边缘和纹理信息。多种研究试图提取含有这些特征的区域。基于在边缘上的水平和垂直投影的一些研究在车辆图像的背景简单的情况下提供了良好的结果。然而，当车辆图像位于具有复杂背景的区域中时，它们具有困难，因为车牌包括在投影中。为了克服这个缺陷，使用一个矩形窗口来查找局部区域中包含丰富边缘的区域。由于窗口尺寸和纵横比是固定的，因此该方法只能在所有牌照在尺寸和纵横比相似时使用。另一种方法是基于形态的方法通过在预定距离内合并相邻边缘来生成 LP 候选，它们不受牌照尺寸变化的影响，但它们严重依赖于牌照的纵横比。

2.4.1 基于纹理特征分析的车牌定位^[20]

车牌定位技术一直是车牌识别技术不可或缺的一个步骤，也是非常关键的一步，其位置性能直接关系到车牌识别系统速率，也影响识别系统的速度。同时它是一个经典的问题。由于中国车牌的颜色和形式多样化，在牌照位置的过程中存在许多困难：悬挂位置的不确定性、严重的许可证污损、光和环境因素等。由于这些原因，虽然许多团队已经开发了相关产品，但是结果在实际应用中不令人满意。

基于车牌统计直方图是一种较为典型也是市场上大量应用的定位方法。因此我们可以利用它的优势进行车牌的识别。可以分两步去实现这种定位方法^[21]：第一步，对车牌进行粗定位，第二步，对车牌进行精确定位。

1) 车牌粗定位

通过分析原始灰度图像 $f(i, j)$ ，我们可以知道原始图像中的牌照具有以下特征：牌照的灰色背景和字母字符之间的对比度大；牌照区域中的水平灰度变化频繁；车牌悬挂在车底，这个位置是整个图像的底部。

根据上述特征，我们使用沿水平方向的一阶差分运算，这可以使具有频繁灰度级变化的区域突出。一阶差分是：

$$g(i, j) = f(i, j) - f(i, j + 1) \quad (2.20)$$

其中, $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。 m 和 n 是图像的高度和宽度。操作之后, 大多数区域的图像的灰度值非常小, 在水平方向上出现大的灰度值, 然后二值化处理图像, 使得其可以车牌区域更突出并去除大部分背景干扰。

最后, 我们进行向下水平扫描, 因为牌照由 7 个字符组成, 通常, 每行的边缘点在水平区域上大于 14。根据中国牌照特征和试验结果, 我们确定边缘点为 15, 然后我们可以找到合格的牌平面, 最后在水平方向剪切原始灰度图像的子图像。

2) 车牌精确定位

在粗略定位之后, 我们可以发现牌照区域通常包括车辆边界。因此, 我们需要基于粗定位进行细定位。目前中国车辆牌照有四类, 蓝色背景为白色字符, 黄色背景为黑色字符, 白色背景为黑色或红色字符, 黑色背景为白色字符; 由于车牌区域布满了整个图片的大部分, 我们可以依据车牌颜色和颜色固有特性来精确定位牌照位置。

实际上, 车牌的最高频率是蓝色背景车牌中的白色字符, 后面是黄色背景车牌中的黑色字符, 白色背景中的黑色或红色字符, 黑色背景车牌中的白色字符。因此, 首先确定蓝色背景车牌中白色字符的精确定位, 如果有蓝色区域, 但是找不到满足该区域的要求, 根据频率从高到低在黑色字符在黄色背景, 黑色字符在白色背景和白色字符在黑色背景车牌区域定位判断。

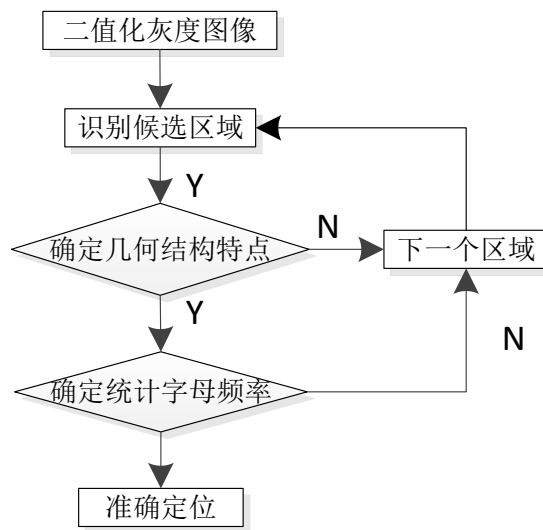


图2.6 精确定位算法的流程图

准确定位的详细过程如下:

(1) 二进制灰度图像: 需要处理的灰度级灰度值分配给 1, 其他灰度值分配给 0, 可以得到二进制图像 $m(x, y)$ 。

(2) 识别候选区域: 识别每个候选区域, 然后根据降序从最大区域进入 (3) 步骤对它们进行排序。

(3) 确定几何结构特征: 计算所选候选区域的长宽比 R , 如果 $R \in [5, 2]$, 进入步骤 (4)。

否则，处理下一个区域。

(4) 确定统计字母-频率特性：计算所选候选区域的平均跳跃频率 L 的水平方向的灰度值，如果 $L \in [15, 5]$ ，则牌照位置是选定的候选区域，许可证板的背景颜色是灰度代表的颜色；否则，返回步骤(3)以处理下一个区域。

2.4.2 基于数学形态学的车牌定位^[22]

数学形态学^[23]是用于基于集合理论、晶格理论、拓扑和随机函数的几何结构的分析和处理的理论和技术。数学形态学最常用于数字图像，但它也可以用于图形、表面网格、实体和许多其他空间结构。拓扑和几何连续空间概念，如尺寸，形状，凸度，连通性和测地距离，由数学形态学引入连续和离散的空间。形态学图像处理的基础是数学形态学，其由根据上述表征变换图像的一组算子组成。

数学形态学最初是为二进制图像开发的，后来扩展为灰度函数和图像。随后的完全晶格的泛化被广泛接受为数学形态学的理论基础。基本形态学算子是膨胀、侵蚀、开运算和闭运算。

1) 膨胀运算

数学形态学中的两个最基本的操作是侵蚀和膨胀。这两个操作符都接收两个数据作为输入：要被侵蚀或扩张的图像，以及结构化元素。两个输入数据各自被视为以对于二进制和灰度图像稍微不同的方式表示坐标集。

膨胀运算的过程是扩充物体的边界，并将其边界的一切背景点合并到物体中。膨胀的作用是滤除图像中存的细微小空洞，这种处理的结果是使得图像面积变大。结构元素 B 对 A 的膨胀定义为：

$$A \oplus B = \bigcup_{b \in B} A_b \quad (2.21)$$

膨胀是可交换的，也由下式给出：

$$A \oplus B = B \oplus A = \bigcup_{a \in A} B_a \quad (2.22)$$

如果 B 具有原点的中心，则 A 的 B 的膨胀可以被理解为当 B 的中心在 A 内部移动时由 B 覆盖的点的轨迹。

膨胀也可以通过以下方式获得：

$$A \oplus B = \{z \in E | (E^s)_z \cap A \neq \emptyset\} \quad (2.23)$$

其中 B^s 表示 B 的对称，即：

$$B^s = \{x \in E | -x \in E\} \quad (2.24)$$

2) 腐蚀运算

腐蚀是将物体的边界归纳为背景点，相对物体的边界就会被相应的缩小。去除图像中的离散点，就像滤除图像噪声。从而得到了图像的有效面积。

结构元素 B 对二值图像 A 的腐蚀由以下定义：

$$A \ominus B = \{z \in E | B_z \subseteq A\} \quad (2.25)$$

其中 B_z 是 B 通过向量 z 的平移，即：

$$B_z = \{b + z | b \in B\}, \forall z \in E \quad (2.26)$$

当结构元素 B 具有中心（例如， B 是圆盘或正方形），并且该中心位于 E 的原点上时，则 A 对 B 的侵蚀可以被理解为中心到达的点的轨迹。

A 对 B 的侵蚀也由下式给出：

$$A \ominus B = \bigcap_{b \in B} A_{-b} \quad (2.27)$$

3)开运算和闭运算

开运算和闭运算是数学形态学的两个重要运算符，它们都源于侵蚀和膨胀的基本操作。像那些操作符，他们通常应用于二进制图像。开运算的基本效果在某种程度上类似于侵蚀，因为其趋向于从前景像素的区域的边缘去除一些前景像素。然而，它比一般的侵蚀更具破坏性。对于二值化图像 A 和结构元素 B 来说， A 通过 B 的侵蚀获得 A 的开运算，随后通过 B 扩大所得图像：

$$A \circ B = (A \ominus B) \oplus B \quad (2.28)$$

开运算由下面被给出：

$$A \circ B = \bigcup_{B_x \subseteq A} B_x \quad (2.29)$$

A 通过 B 的闭合，通过 A 对 B 的膨胀，接着通过 B 对所得结构的侵蚀来获得：

$$A \cdot B = (A \oplus B) \ominus B \quad (2.30)$$

由上述公式可以发现，开运算和闭运算和之间是对偶的。除此之外闭运算对图像具有过滤功能，图像内部的细微小空洞和裂隙可以通过它填补。

2.4.3 基于小波变换的车牌定位^[24]

小波变换是一种新的变换分析方法，类似于具有完全不同的优值函数的傅里叶变换（或更多地是窗口傅立叶变换）。主要区别在于：傅立叶变换将信号分解成正弦和余弦，即局部化在傅立叶空间中的函数；同时它相比于傅里叶变换能够提供一个随频率改变的“时间-频率”窗口，而之前窗口大小是不随频率变化的。小波变换通过变换能够充分突出问题某些方面的特征，不论是时间频域还是空间频域都能对其进行局部化分析，并利用平移或者伸缩变换对信号一步一步的进行细化，最终可以让信号在高频处时间细分和低频处频率细分，并且能够对时频信号自动适应，这样对信号的任何细节都可以聚焦到。

对于函数 $\psi(x) \in L^2(R)$ ，且有：

$$C_\psi = \int_R \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (2.31)$$

上式中 $\hat{\psi}(\omega)$ 是 $\psi(\omega)$ 的傅里叶变换，再令

$$\psi_{a,b} = |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \quad (2.32)$$

则函数 $f(x) \in L^2(R)$ ，的小波变换定位为：

$$W_f(a, b) = \langle f, \psi_{a,b} \rangle = |a|^{-\frac{1}{2}} \int_{\mathbb{R}} f(x) \psi\left(\frac{x-b}{a}\right) dx \quad (2.33)$$

相应的小波变换的反变换公式为：

$$f(x) = C_\psi \iint_{\mathbb{R}^2} W_f(a, b) \psi_{a,b}(x) \frac{da db}{a^2} \quad (2.34)$$

基函数 ψ 可以被认为是已经过滤函数 $x(t)$ 的系统的脉冲响应。变换的信号提供关于时间和频率的信息。因此，小波变换包含类似于短时傅里叶变换的信息，但是具有小波的附加的特殊性质，其在基函数的较高分析频率下以时间分辨率显示。

在数学中，小波变换^[25]是由小波产生的某个正交序列的正方形可积分（实数或复数值）函数的表示。如今，小波变换是时间-频率变换中最受欢迎的之一。

上述车牌定位的这些方法各有优缺点，实际应用中由于车牌的各种特征不同，想用单一的特征定位方法往往很容易产生失败，所以要实现快速、准确地定位车牌可以结合车牌的不同特征，比如纹理颜色，同时结合车牌的各种特征对车牌进行定位可以很大程度地提高车牌定位的准确率。

2.5 字符分割

如果在前期的一连串的预处理得到了比较理想的情况，我们将得到的是一个清晰的仅含车牌的二值化图像，其中的字符紧贴上下边缘的水平线，车牌图像的背景像素为黑色，采用数字 0 表示；车牌的字符像素为白色，采用数组 1 表示。为了达到目的，我们必须将车牌图像中的字符一个一个的分割出来。

一般的分割方法是以一条垂直线从左到右对车牌图像进行扫描，当遇到白色像素点，则判断为有字符。但是在实际得到的车牌图像中由于前期处理不当或者图像本身拍摄的不清晰而产生很多随机干扰，这种复杂的情况下将字符间隔检测出来，同时不能误识别纵向间隔，在中文汉字中有很多左右偏旁结构的，会很容易造成这种情况的出现。

这里有两种最常用的主流方法的字符分割方法^[26]，垂直投影方法和连通域分析法。基于投影分析的字符分割算法 投影方法涉及在二值化之后垂直投影牌照图像，这意味着根据每个牌照字符沿着水平方向的像素数量计算每列，获得我们分割位置的字符之间的间隙的最小值。这种方法在环境噪声较小，边界干扰较小，但附着力和间隔适应性差，或者牌照点的字符有边界干扰的情况下工作良好。

2.5.1 基于垂直投影法的字符分割^[27]

最常用的字符分割方法是通过垂直投影法去实现。竖直方向上的字符块的投影在它们的间隙理想情况下是零的，但是由于噪声的干扰，字符间的投影不会为零，但是会在它们的间隙相当于一个区间最小值，以此类字符的具体的分割位置。

要具体分割出单个字符还需要进一步处理，下面将按行进行行字切分。由于车牌图像的

字符之间是存在一段距离的，经二值化后就是一段空白，使用行间的距离生成的垂直投影距离将字符区域按照每行切分。对于一个二值化后的字符区域图像，对车牌区域的二值码元从上至下逐行扫描形成数值平行投影，从其垂直投影的直方图可以很清楚地看出字符的边界。然后对单个字符切割：将行切分的得到的数值区域利用每个车牌的直线投影中支付之间的距离将每个独立的字车牌符切分出来。

对于一帧二值图像 $f(x,y)$ 在 x 轴上的投影为：

$$p_x(x,y) = p_x(x,0) = \int_{-\infty}^{\infty} f(x,y) dy \quad (2.35)$$

其分布函数 $H_x(x)$ 为：

$$H_x(x) = \sum_{i=0}^{M-1} f(i,x) \quad (2.36)$$

以下是垂直投影的算法流程图：

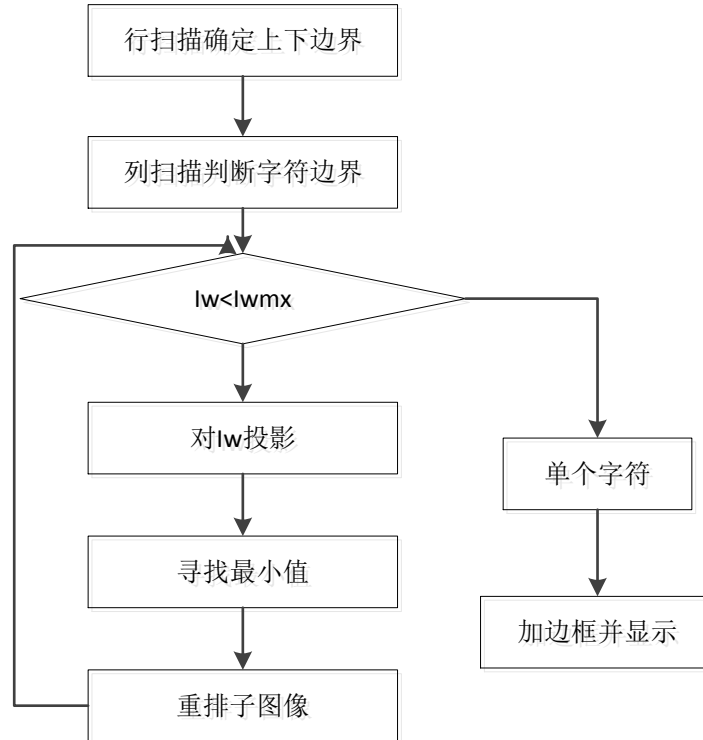


图2.7 垂直投影算法流程图

例如，对于二值化后的车牌图像：



图2.8 二值化后的车牌图像

其垂直投影图为：

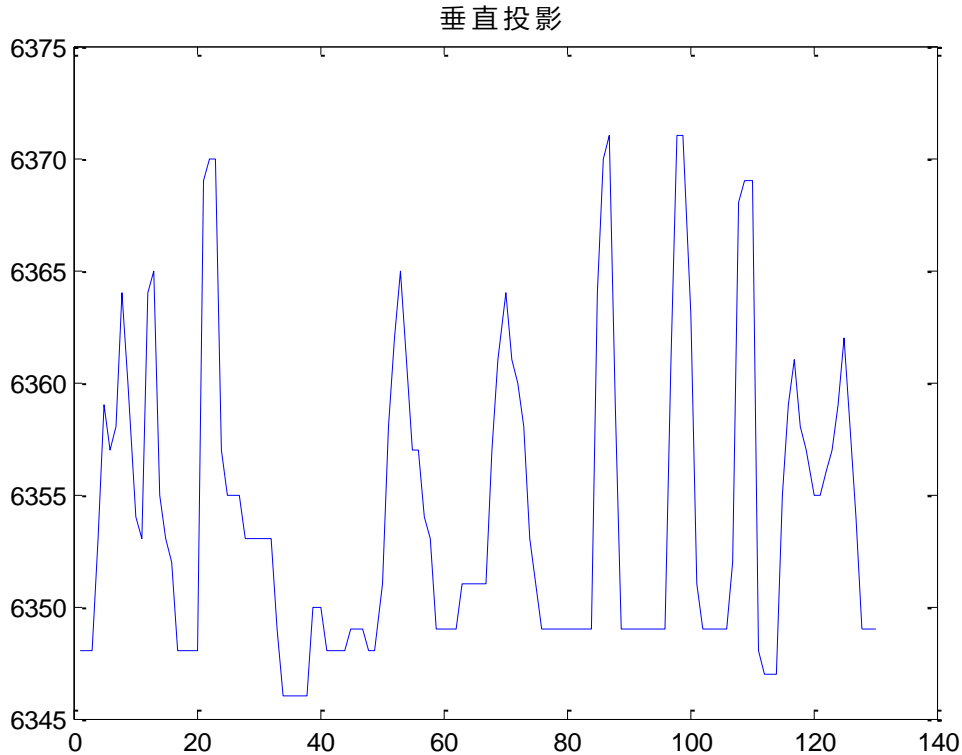


图2.9 车牌图像的垂直投影图

但是，光凭这一点我们不容易精确的分割出字符，还需要借助其他因素，例如通过利用车牌字符的尺寸和书写格式等特点，可以大大增加字符分割的正确性。虽然这种方法可以有效的分割断裂字符，但是在记录特殊环境下的汽车图片时，往往无法成功。综上所述，这种方法的优点是处理过程简单，但缺点是非常明显的，当字符之间存在粘连问题，那么利用垂直投影法来解决就不能奏效了；另外在解决一些汉字的不连通问题上也存在严重不足。

2.5.2 基于连通域特征的字符分割^[28]

这种方法的核心是利用连通域特征，所谓连通域，即数字和英文字母的字符笔画具有的共同区域。而我国的车牌号码有固定的排列的格式，包括其中的字母和数字都是根据规定排序的，因此可以利用连通域特征来切割车牌字符。

通常用车牌区域里的白色像素的数量来表征二值图像中连通域的面积，为了更好的实现字符的切割，需要保证字符区域的连通域面积占据较大的比例，这样才能降低噪声连通域的干扰，且当字符的连通域面积最大时，字符切割的效率最高。

不同于国外车牌，中国车牌中包含的第一个汉字需要特殊处理。部分左右偏旁的汉字中间会有一个微弱的间隙，所以需要依据字符宽度所占车牌宽度的比例，从右至左的切分。详细过程为：首先要采集车牌右侧五分之一的区域里最大的连通域面积，并将其作为车牌的第七个字符，接着通过得到的字符连通域的宽度来区别字符的左边界和右边界。接下来依照上述步骤，一直执行直到采集到第六个字符则停止采集，因为此时已经采集到车牌的汉字，它在第六个字符的左侧。具体流程如下图，其中 N 为已分割字符数。

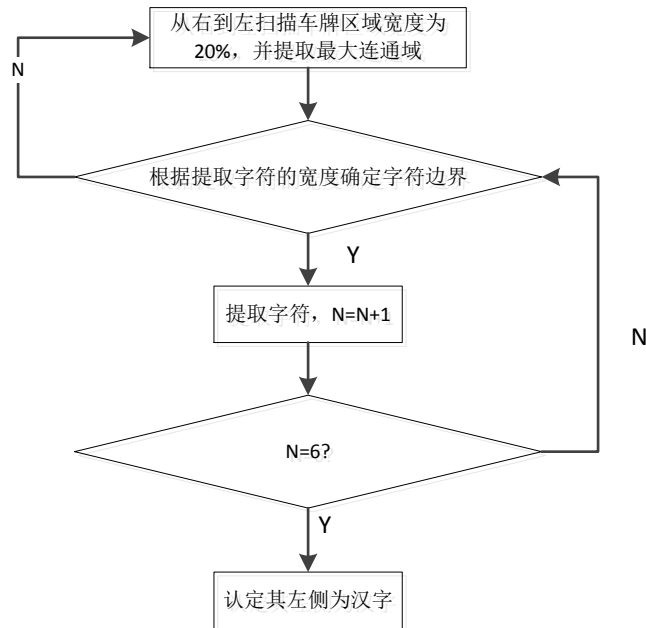


图2.10 连通域分割车牌的流程

基于连通域的字符分割法有一个显著的优点，它可以在车牌定位不准的条件下，依然保证有效的分割字符。相比于垂直投影法，对于车牌中出现的断裂字符无法精确处理，在现实生活中的实际应用意义不大。

通过上文对垂直投影法和连通域分割法的分析和优缺点的罗列，可以发现它们的应用领域比较单一，不适用于复杂环境中的车牌识别。为了扬长避短并充分利用两种方法的优势^[29]，近年来，清华大学的闫鹏等人首次将两种方法进行组合，使用聚类方法从候选车牌中提取连接到域的字符，以通过单个字符投影来筛选和分析并处理粘连和断裂的状况。

2.6 字符识别

字符识别^[30]其实就是把待识别字符图像转换成文本形式的文字，其原理是首先提取待识别字符的特征结构，把这个特征结构与已经存在的标准字符的特征结构相匹配，匹配成功则该标准字符就是待识别字符，最后输出识别结果。

字符识别的原理框图下图所示：

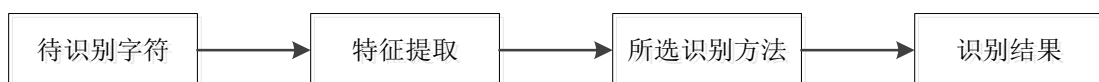


图2.11 字符识别原理

车牌字符识别与 OCR 系统类似。虽然它是一个小的，字符是模糊的，易于受到外来因素的影响。所以，这项技术的实现非常的困难。有两种识别牌照字符方法它们分别是基于模板匹配方法与基于神经网络的方法。第一种方法具有更快的识别速度，特别是对于二值图像，并且它可以满足实时要求，因此当车牌图像清晰并且预处理工作更好时，该方法可以获得更

高的识别率。因此，它现在被广泛使用。牌照字符集的特征包括匹配中文字符，英文字母和数字，与数字模板匹配的模板，然后语法分析结果，确认结果的合法性，最后，排除错误的识别结果。第二种方法具有大量的容错能力，但识别速度慢，难以满足实时性要求。

2.6.1 模板匹配算法

模板匹配^[31]是用于找到匹配模板图像的图像的小部分的数字图像处理中的技术。模板匹配的基本方法使用适合于我们想要检测的搜索图像的特定特征的卷积模板。该技术可以容易地对灰色图像或边缘图像执行。在图像结构匹配掩码结构的地方，卷积输出将最高，其中大图像值乘以大掩码值。

该方法通常通过首先挑选搜索图像的一部分以用作模板来实现：我们将称为搜索图像 $S(m, n)$ ，其中 (m, n) 表示搜索图像中的每个像素的坐标。我们将称为模板 $T(mt, nt)$ ，其中 (mt, nt) 表示模板中每个像素的坐标。然后，我们简单地在搜索图像中的每个 (m, n) 点上移动模板 $T(mt, nt)$ 的中心（或原点），并且计算 $S(m, n)$ 中的系数和 $T(mt, nt)$ 在模板跨越的整个区域上。由于考虑了模板相对于搜索图像的所有可能位置，具有最高分数的位置是最佳位置。公式描述如下：

搜索范围是： $1 \ll i \ll W - m, 1 \ll j \ll H - n$

通过比较 T 和 S_{ij} 的相似性，完成模板匹配过程。衡量模板 T 和子图 S_{ij} 的匹配程度，可用下列两种测度：

$$D(i, j) = \sum_{m=i}^N \sum_{n=1}^N [S^{i,j}(m, n) - T(m, n)]^2 \quad (2.37)$$

$$D(i, j) = \sum_{m=i}^N \sum_{n=1}^N |S^{i,j}(m, n) - T(m, n)| \quad (2.38)$$

将上面第一个式子展开得到下面的：

$$D(i, j) = \sum_{m=i}^N \sum_{n=1}^N S^{i,j}(m, n)^2 + \sum_{m=i}^N \sum_{n=1}^N T(m, n)^2 - 2 \sum_{m=i}^N \sum_{n=1}^N S^{i,j}(m, n) T(m, n) \quad (2.39)$$

从展开的公式中可以看出，中间第二项是一个常数，也就是只跟模板有关系，而第一个和第三个是和原图有关的，随着模板在原图上的移动，这两个值也在发生变化。当 $D(i, j)$ 的值最小时，说明找到了目标，这个算法的效率是非常的低下的，每移动一个像素就要进行一次计算，计算量是相当的大。

模板匹配算法的操作相当简单，这是它最主要的特点，但是这样的优点也造就了它的缺点，如果字符前期处理不清晰、或者字符图像有缺损的部分，模板算法的识别率就会受到影响，识别率可能会大大的降低。

2.6.2 基于 BP 神经网络的识别算法

神经网络是一种计算方法，其基于大量神经单元集合，松散地建模生物大脑解决由轴突连接的生物神经元的大群集的问题的方式。每个神经单元与许多其他神经单元连接，并且连接可以在它们对所连接的神经单元的激活状态的影响中实施或抑制。每个单独的神经单元可以具有将所有其输入的值组合在一起的求和功能。在每个连接和单元本身上可以存在阈值函数或限制函数，使得它必须在它可以传播到其他神经元之前超过它。这些系统是自学习和训

练而不是显式编程，并且在传统计算机程序中难以表达解决方案或特征检测的领域中是优异的。

BP 网络^[32]可以学习输入和输出数据之间的关系而不用事先知道两者之间的任何关系，这种能力在关系映射，数据拟合等方面有着巨大的作用，BP 网络使用反向传播算法，是一种将误差反方向调节网络的算法。BP 神经网络^[33,34]是一个三维建筑，其网络模型拓扑结构包括输入层、隐含层、输出层三部分如下图：

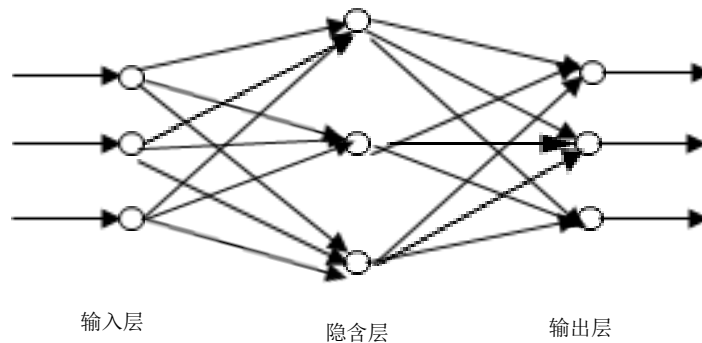


图 2.12 BP 神经网络结构图

BP 神经网络的建立主要是确定三层神经元数量：

第一是输入层节点数 BP 神经网络输入层是图像预处理后特征输出的维数。

第二是 BP 神经网络中的隐层节点数，隐层节点数的选取与网络中输入层节点数和输出层节点数都有关系，还跟需要解决的问题的复杂度有关，同时还跟网络中的传递函数的形式有关，而且还跟样本数据的特点等很多因素相关，隐含层节点数的选取受很多因素的影响，它的确定还需要是根据实验者的经验。一般的，隐含层神经元的数量越多能解决的问题复杂度越高，识别结果也就越准确，但是这样网络的训练时间就会变长。而且隐含层节点数并不选择越多越好，过多反而适得其反使识别率的急剧下降。

隐含层节点数确定的一些基本方法：首先我们的选择肯定要达到精度的要求，其次神经网络的结构要尽量紧凑些，这样在达到精度要求的情况下能使隐含层节点数尽可能减到最少。

第三是输出层的节点数。节点的输出层数由标准输出的数目决定的，也就是目标预期输出的数量。

不过 BP 算法依然存在一些问题：首先由于 BP 神经网络算法本身的问题造成网络梯度越往下越稀疏，这样从顶层慢慢到下层误差校正的信号就会越变越小，最后造成结果不准确；其次容易出现局部最小值，当值从远处慢慢逼近最优区域的时，就会出现在局部最小值处收敛；最后，我们训练数据的时候需要使用有标签的数据，而实际数据是没有标签的，需要经过人工处理，这个跟我们的大脑相比，显然学习能力不够。

2.7 本章小结

本章主要介绍了车牌图像处理方面的相关技术，第一小节包括图像灰度化、图像增强、

以及二值化这些前期处理方法的介绍，这些预处理的目的是为后面的字符识别做基础，有效的预处理可以分割出清晰真实的字符，尽可能地减小对后面的识别算法的影响；第三小节介绍了多种边缘检测的算子；第四介绍了几种常用的车牌定位算法，包括数学形态学方法、纹理特征分析法、以及小波变换等；第五小节简要地介绍了基于投影特征、连通域特征的字符分割算法；第六小节介绍了现有的车牌字符识别技术，最常用有模板匹配法以及 BP 神经网络识别法。

第 3 章 卷积神经网络

3.1 深度学习介绍

介绍卷积神经网络之前,我们必须先对深度学习做一些了解。深度学习^[35](也称为深度结构化学习,分层学习或深层机器学习)是基于一组尝试对数据中的高级抽象进行建模的算法的机器学习的分支。在一个简单的情况下,你可以有两组神经元:接收输入信号的那些和发送输出信号的。当输入层接收到输入时,它将输入的修改版本传递到下一层。在深层网络中,在输入和输出之间有许多层,允许算法使用多个处理层,由多个线性和非线性组成,线性变换。

深度学习是基于数据的学习表示的更广泛的机器学习方法的一部分。观察可以以许多方式表示,例如每个像素的强度值的向量,或者以更抽象的方式作为一组边缘,特定形状的区域等。在简化学习任务时,一些表示比其他表示更好。深度学习的承诺之一是用无监督或半监督特征学习和层次特征提取的有效算法替代手工制作的特征。

深度学习算法通过比浅层学习算法更多的层来变换它们的输入。在每层,信号由处理单元变换,像人工神经元,其参数通过训练“学习”。从输入到输出的变换链是信用分配路径(CAP)。CAP描述输入和输出之间的潜在因果连接,并且可以在长度上变化-对于前馈神经网络,CAP的深度(因此网络的深度)是隐藏层的数量加上一个(因为输出层也被参数化)但是对于递归神经网络,其中信号可以通过层传播多于一次,CAP在长度上潜在地是无限的。没有普遍认同的浅层深度学习与深度学习的阈值,但大多数研究人员认为深度学习有多个非线性层($CAP > 2$),Juergen Schmidhuber认为 $CAP > 10$ 是非常深的学习。

其中,卷积神经网络(CNN)已经成为处理视觉和其他二维数据的首选方法。CNN由在顶部具有完全连接的层(匹配典型的人工神经网络中的那些)的一个或多个卷积层组成。它也使用绑定的权重和池层。这种架构允许CNN利用输入数据的2D结构。与其他深层架构相比,卷积神经网络在图像和语音应用中都显示了出色的结果。他们也可以训练与标准反向传播。CNN比其他常规前馈神经网络更容易训练,并且具有许多较少的参数来估计,使得它们是使用的高度吸引人的架构。

3.2 卷积神经网络概述

过去几年,深度学习^[36]在解决诸如视觉识别(visual recognition)、语音识别(speech recognition)和自然语言处理(natural language processing)等很多问题方面都表现出非常好的性能。在不同类型的深度神经网络当中,卷积神经网络是得到最深入研究的。

卷积神经网络^[37]是一种常见的深度学习架构,受生物自然视觉认知机制启发而来。1959年,Hubel&Wiesel发现,动物视觉皮层细胞负责检测光学信号。受此启发,1980年Kunihiko

Fukushima 提出了 CNN 的前身—neocognitron。

20 世纪 90 年代, LeCunetal 等人发表论文, 确立了 CNN 的现代结构, 后来又对其进行完善。他们设计了一种多层的人工神经网络, 取名叫做 LeNet-5^[38], 可以对手写数字做分类。和其他神经网络一样, LeNet-5 也能使用反向传播算法训练。

CNN 能够得出原始图像的有效表征, 这使得 CNN 能够直接从原始像素中, 经过极少的预处理, 识别视觉上面的规律。然而, 由于当时缺乏大规模训练数据, 计算机的计算能力也跟不上, LeNet-5 对于复杂问题的处理结果并不理想。

2006 年起, 人们设计了很多方法, 想要克服难以训练深度 CNN 的困难。其中, 最著名的是 Krizhevsky et al.提出了一个经典的 CNN 结构, 并在图像识别任务上取得了重大突破。其方法的整体框架叫做 AlexNet, 与 LeNet-5 类似, 但要更加深一些。

AlexNet 取得成功后, 研究人员又提出了其他的完善方法, 其中最著名的要数 ZFNet, VGGNet, GoogleNet 和 ResNet 这四种。从结构看, CNN 发展的一个方向就是层数变得更多, ILSVRC 2015 冠军 ResNet 是 AlexNet 的 20 多倍, 是 VGGNet 的 8 倍多。通过增加深度, 网络便能够利用增加的非线性得出目标函数的近似结构, 同时得出更好的特性表征。但是, 这样做同时也增加了网络的整体复杂程度, 使网络变得难以优化, 很容易过拟合。

卷积神经网络已逐渐应用于各种领域, 手写数字识别^[39]、人体行为检测^[40]、人脸表情检测^[41]、人脸识别^[42]、身份证号码识别^[43]等模式识别领域。

在机器学习中, 卷积神经网络 (CNN 或 ConvNet) 是一种前馈人工神经网络, 其中它的神经元之间的连接模式受动物视觉皮层的组织的启发。单个皮质神经元响应受限制的空间区域中的刺激, 称为接受场。不同神经元的接受场部分重叠, 使得它们平铺视野。单个神经元对其接受场内的刺激的响应可以通过数学上的卷积近似运算。卷积网络受生物过程的启发, 并且是被设计为使用最少量的预处理的多层感知器的变体。它们在图像和视频识别、推荐系统和自然语言处理中具有广泛的应用。

卷积神经网络也称为移位不变或空间不变人工神经网络, 其基于其共享权重结构和转换不变性特征来命名。CNN 由多层接收场组成, 这些是处理输入图像的部分的小神经元集合, 然后将这些集合的输出进行平铺, 使得它们的输入区域重叠, 以获得原始图像的更好的表示, 这对于每个这样的层重复, 平铺允许 CNN 容许输入图像的平移。

卷积网络包括卷积层和完全连接的层的各种组合, 在每层的末尾或后面应用逐点非线性。引入对输入的小区域的卷积运算以减少自由参数的数量并改进概括。卷积网络的一个主要优点是在卷积层中使用共享权值, 这意味着对于层中的每个像素使用相同的滤波器 (权重组), 这既减少了内存占用又提高了性能, 另外 CNN 中的自由参数的数量不随着输入尺寸与标准多层神经网络的情况成比例地增加。

3.3 卷积神经网络推导及实现

(1) Feedforward Pass 前向传播^[44]

因为卷积神经网络是多类问题，我们先设有 c 个类、 N 个训练样本，用平方差函数推导公式为：

$$E^N = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^c (t_k^n - y_k^n)^2 \quad (3.1)$$

上式中 t_k^n 表示的是第 n 个样本对应的第 k 维标签。 y_k^n 表示的是第 n 个样本对应的网络输出的第 k 个输出。

训练样本的误差值的总和是由所有训练集上的误差组成的，所以我们先考虑一个比较熟悉的神经网络，比如 BP 网络。我们可以用下列方法来表示第 n 个样本所具有的误差：

$$E^n = \frac{1}{2} \sum_{k=1}^c (t_k^n - y_k^n)^2 = \frac{1}{2} \|t^n - y^n\|_2^2 \quad (3.2)$$

第 1 层的输出表示方式为：

$$x^l = f(u^l), \text{ 和 } u^l = W^l x^{l-1} + b^l \quad (3.3)$$

上式中 $f(\cdot)$ 是输出激活函数，常用的激活函数有 sigmoid 函数或者 tanh 函数。

(2) Backpropagation Pass 反向传播^[45]

反向传播算法是用链式求导法则推导出来的，其来回的误差也就是误差对基的变化率，即误差对基的导数，称它为每个神经元的基的灵敏度，定义如下：

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial u} \frac{\partial u}{\partial b} = \delta \quad (3.4)$$

$$\text{因为 } \frac{\partial u}{\partial b} = 1 \quad (3.5)$$

$$\text{所以 } \frac{\partial E}{\partial b} = \frac{\partial E}{\partial u} = \delta \quad (3.6)$$

$$\text{基的灵敏度为：} \frac{\partial E}{\partial b} = \delta \quad (3.7)$$

上式是和 $\frac{\partial E}{\partial u}$ 是相等的，这也是为什么高层的误差可以反向传播到底层的原因。

反向传播第 1 层的灵敏度为：

$$\delta^l = (W^{l+1})^T \delta^{l+1} \circ f'(u^l) \quad (3.8)$$

式中“ \circ ”表示每个元素相乘。神经元的灵敏度在输出层上是不一样的：

$$\delta^L = f'(u^L) \circ (y^n - t^n) \quad (3.9)$$

最后，运用 δ 规则对每个神经元进行权值更新：

$$\frac{\partial E}{\partial w^l} = x^{l-1} (\delta^l)^T \quad (3.10)$$

$$\Delta W^l = -\eta \frac{\partial E}{\partial w^l} \quad (3.11)$$

在上式中 η_{ij} 是学习率，并对应于每一个权值 w_{ij} 。

CNN 包括神经元层，并且其被优化用于二维模式识别。CNN 具有三种类型的层，即卷积层，子采样层和完全连接层。

3.3.1 卷积层

卷积层由称为特征图的几个二维平面的神经元组成。来自特征图的每个神经元连接到来自前一层的神元邻域，导致所谓的接受场。计算输入特征图和相应内核之间的卷积。然后将这些卷积输出与可训练的偏差项一起求和，然后将其传递给诸如双曲正切的非线性激活函数，以获得新的特征值。权重在卷积矩阵中共享，使得可以利用减少的权重集合来处理大图像。卷积层用作特征提取器，其使用本地接受场和共享权重的概念来提取其他信号中的输入的突出特征，例如角、边缘、端点或非视觉特征。共享权重的特点有以下几个，它减少了训练的自由参数的数量，简化了机器的复杂性，并减少了测试误差和训练误差之间的差距。卷积层的一个有趣的特性是，如果输入图像被移动，特征映射输出将被移位相同的量，但是否则将保持不变。该属性是 CNN 对输入的移位和失真的鲁棒性。层中的一个特征映射将具有相同的权重向量。完整的卷积层由若干特征图（具有不同的权重向量）组成，使得可以在每个位置处提取多个特征。

该层的功能是使用滤波器组对输入图像执行滤波操作，从而产生另一组图像。连接表用作输入-输出对应关系的轨迹记录。来自连接到同一输出图像的输入的滤波器响应被线性组合。该层执行由以下公式给出的映射：

$$x_j^l = f(\sum_{i \in M_l} x_i^{l-1} * k_{ij}^l + b_j^l) \quad (3.12)$$

其中*表示 2D 有效卷积。对于特定层，滤波器 k 的大小相同。它与输入的大小一起定义输出图像 x_j 的大小，之后，将非线性激活函数应用于类似于标准多层网络的 x 。

3.3.2 子采样层

降低精度的简单方法是降低特征地图的空间分辨率。这可以通过执行局部平均的所谓的子采样层来实现。由于特征映射对输入中的平移敏感，因此子采样层降低了图像的分辨率，从而降低了平移（移位和失真）效应的精度。该层通过将相邻神经元的平均值转换为单个值来减少前一层（通常为 2×2 ）的相邻神经元的输出。接下来，将该值乘以可训练的权重（可训练系数），加上偏差，然后将结果传递到非线性激活函数，例如双曲正切。如果系数小，则单元以准线性模式操作，并且子采样仅模糊输入。如果系数较大，则取决于偏差的值，子采样单元可以被视为执行“噪声 OR”或“噪声 AND”函数。

子采样层 l 中的特征映射 j 的输出 x_j^l 由下式给出：

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l) \quad (3.13)$$

上式中 $\text{down}(\cdot)$ 是一个下采样函数。上式显示出了输出图像和输出图像之间的关系，即输出图像是对输入图像的不同模块的所有神经元进行求和加偏置所得到的。

3.3.3 完全连接层

最后，在几个卷积层和子采样层之后，神经网络中的高级推理通过完全连接的层完成。完全连接的层中的神经元具有与前一层中的所有激活的完全连接，如在常规神经网络中所见。因此，它们的激活可以用矩阵乘法后跟偏置偏移来计算。

值得注意的是，完全连接层和卷积层之间的唯一区别是卷积层中的神经元仅连接到输入中的局部区域，并且卷积体积中的许多神经元共享参数。然而，两个层中的神经元仍然计算点积，因此它们的函数形式是相同的。因此，事实证明，可以在完全连接和卷积层之间进行转换：

对于任何卷积层，存在实现相同正向功能的完全连接层。除了在某些块（由于局部连接性），其中许多块中的权重相等（由于参数共享），权重矩阵将是大部分为零的大矩阵。

相反，任何完全连接层都可以转换为卷积层。例如，正在观看大小为 $7 \times 7 \times 512$ 的一些输入音量的具有 $K = 4096$ 的完全连接层可以等价地表示为具有 $F = 7$, $P = 0$, $S = 1$ 的卷积层， $K = 4096$ 。换句话说，我们将滤波器大小设置为正好为输入音量的大小，因此输出将简单为 $1 \times 1 \times 4096$ ，因为只有一个深度列“匹配”于输入音量，给出与初始完全连接层相同的结果。

层的数量取决于应用，并且层中的每个神经元是以下层的输入。此外，当前层仅接收来自前一层输入加上通常为 1 的偏差。每个神经元对它的每个输入应用权重并且加总所有加权的输入。总加权值受非线性函数的影响，例如 **sigmoid** 函数来将神经元的输出限制到一个范围的值。在每个层中使用多个平面，使得可以检测多个特征。

3.4 卷积神经网络结构

卷积神经网络（CNN）源于我们的生物系统，它基本上是多层感知器（MLP）的变体。CNN 的主要优点是它只需要很少的学习参数，因为它是由相同基函数的翻译版本组成的。因此，CNN 具有产生相当多的识别性能的能力，仅有少数样品。CNN 具有构建对平移不变性不敏感的网络的能力。CNN 是采用监督方式学会的，CNN 的网络结构^[46]主要有稀疏连接和权值共享的两个特点。

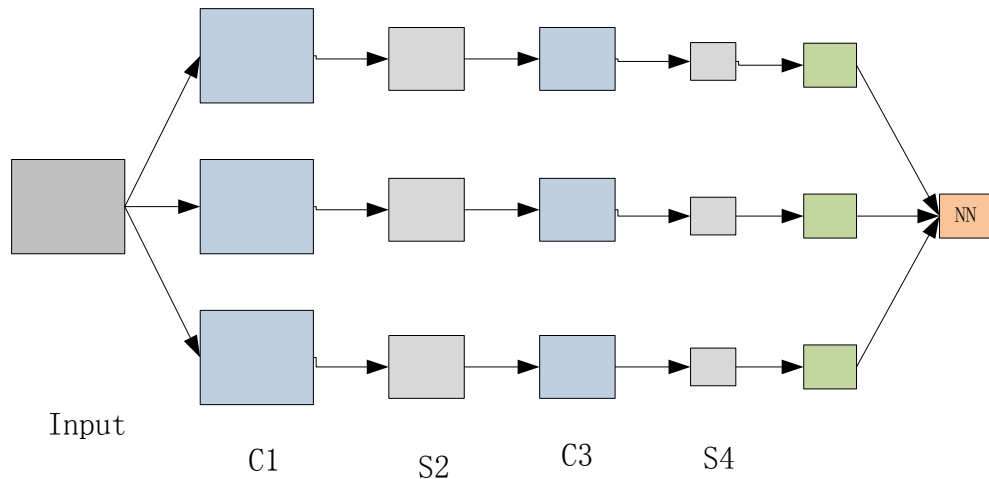


图 3.1 卷积神经网络的概念示范

在上图中，输入图像分别和三个滤波器（也即卷积核）进行卷积，这三个滤波器是可训练的，在 C1 层分别产生一个特征图像，在 C1 中每四个像素点经过公式 3.12 的运算，然后在 S2 层将会得到一个特征映射图，这些特征映射图再通过滤波器进行卷积得到 C3 层的特征图像，之后 S2 到 C3、C3 到 S4 与之前的输入图像到 C1、C1 到 S2 的处理过程是一样的，最后是经过一个全连接网络得到输出。

C 层是特征提取层，在这层的神经元和所在（S 层）的局部感受相互连接着，当输入样本时，每个神经元负责提取样本的一个局部特征，提取后就形成了它与 S 层的位置关系；S 层所属特征映射层，它是 C 层的后面一层，网络中的计算层是由大量的特征映射构成的，在平面层中能看到被映射在上面的图像特征。在相同平面上的每个神经元的值是一样的。

CNN 中的特点提取层（C 层）和特征映射层（S 层）是相间存在的，网络中对于样本的容错能力是通过独特的特征提取架构来达到的。

3.4.1 稀疏连接

卷积神经网络在相邻的两层之间采用的是稀疏连接^[47]，它是通过强制使用局部的连接模式，这样只利用局部空间图像特征，第 m 层的隐层单元不是和第 $m-1$ 层的输入单元的全部区域连接，而是只和它的局部区域有连接，这样就只接受第 $m-1$ 层的局部区域的特征，这个区域就被称做空间连续的接受域。下面进一步对其结构进行描述：

设第 $m-1$ 层为输入单元，第 m 层的空间接受域的宽度为 3，换句话说就是 m 层的每个单元只和 $m-1$ 输入层的 3 个相邻的神经元相连，同样的第 m 层与第 $m+1$ 层也是这种连接规则，如下图所示。

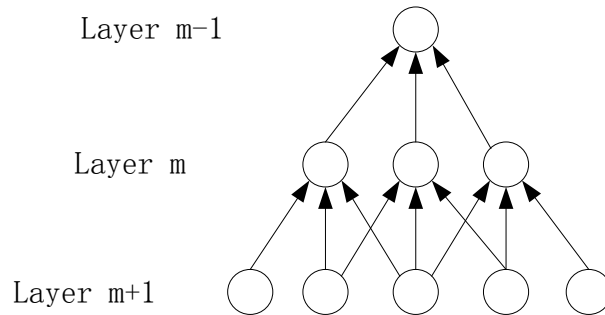


图 3.2 CNN 中相邻神经元之间的稀疏连接

上图中可以看到相对于第 m 层，第 $m+1$ 层的神经元的接受域的宽度也是 3，但相对于输入层 $m-1$ 层的接受域为 5，过滤器将会学习这种结构的局部空间模式。同时从上图也可以看出，如果存在多个这样的层堆叠起来后，会使得过滤器逐渐覆盖到了更大的视觉区域。

当权重是稀疏的时候，我们可以从矩阵中精确地恢复第一层的权重矩阵。有人认为，稀疏连接是一种自然约束，可以导致实践中的性能提高。我们显示权重可以使用有效的优化方法正确恢复。这样的方法早已用于诸如字典学习和主题建模的线性模型。

3.4.2 权值共享

参数共享方案用于卷积层以控制自由参数的数量。它依赖于一个合理的假设：如果一个补丁特征对于在一些空间位置处的计算是有用的，则它也应当在不同位置处进行计算是有用的。换句话说，将单个 2 维深度切片表示为深度切片，我们约束每个深度切片中的神经元以使用相同的权重和偏置。

由于单个深度切片中的所有神经元共享相同的参数化，因此 CONV 层的每个深度切片中的正向通过可以被计算为神经元权重与输入体积（因此名称：卷积层）的卷积。因此，通常将权重集称为与输入卷积的滤波器（或核）。

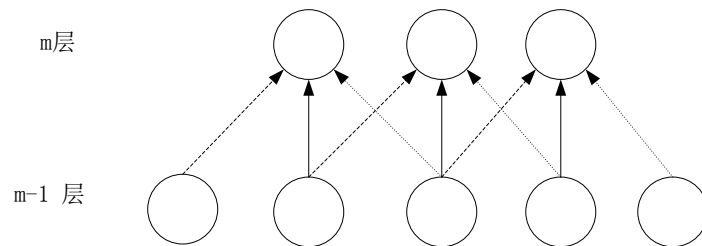


图 3.3 CNN 中神经元之间的连接权值共享

在 CNN 中，每个滤波器在整个视野中复制。这些复制单元共享相同的参数化（权重向量和偏置）并且形成特征图。这意味着给定卷积层中的所有神经元检测完全相同的特征。以这种方式的复制单元允许检测特征，而不管它们在视野中的位置，从而构成平移不变性的性质。

总之，这些特性使得卷积神经网络来实现对视觉问题的更好推广。由于权重共享减少网络学习中大量的自由参数，这样可以极大的减少识别算法对网络运行的存储器需求。减少内存占用允许培训更大，更强大的网络。

3.4.3 池化

CNN 的另一个重要概念是池化^[48]，池是非线性下采样的一种形式。有几个非线性函数来实现池，其中最大池是最常见的。它将输入图像分割成一组非重叠矩形，并且对于每个这样的子区域，输出最大值。一旦找到特征，其确切位置不如其相对于其它特征的粗略位置重要。池化层的功能是逐渐减小表示的空间大小以减少网络中的参数和计算量，并且因此也控制过度拟合。通常在 CNN 架构中的连续的转换层之间周期性地插入池层。池化操作提供了一种形式的平移不变性。

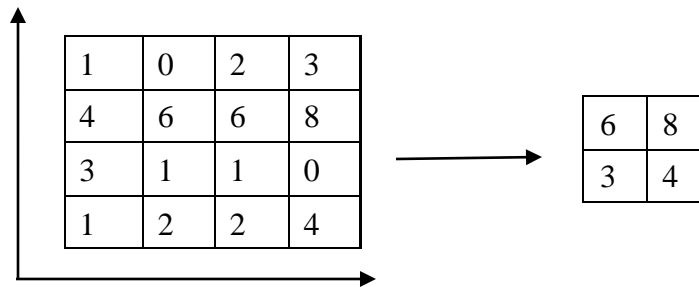


图3.4 最大池与 2×2 过滤器

池化层在输入的每个深度切片上独立操作并在空间上调整其大小。最常见的形式是具有大小为 2×2 的滤波器的合并层，在输入中的每个深度切片上施加 2 个下采样的跨度，宽度和高度都为 2，丢弃 75% 的激活。在这种情况下，每个 MAX 操作将获取最多 4 个数字。深度尺寸保持不变。

总之，卷积网络的核心思想有三大特点：一是局部感受野（也即稀疏连接）、第二是权值共享、第三是对空间或者时间进行亚采样，三者的完美的融合成功的达到了特征映射中一定程度的尺度、位移、和形变的不变性。

3.5 卷积神经网络的训练

从卷积神经网络的结构，我们很容易看出，卷积网络实质上就是数学上输出与输入一对一的映射关系，只是输入和输出之间的映射表达式是有卷积网络自己学习的，这个也是所谓的有指导学习模式，卷积网络的训练就是给网络输入大量的输入样本和输入所对应标准输出样本，让他们之间产生固定的映射关系，当我们使用这个已经训练好的网络模型时，网络会根据输入的特征，并按照之前已经训练好的映射关系产生对应的输出。在开始训练之前，所有权重应该选用多个不同的小随机数来初始化。小随机数的特点所在是减小网络体积，避免陷入饱和状态，从而使得训练没有成功。选择小随机数的目的是用来确保网络能够正常的学习。事实上，如果我们初始化权重矩阵时把所有的权值用相同的数字，网络将失去学习的能力。

卷积神经网络的训练算法如下。其中包括 4 个步骤，每 2 步骤被分为一个阶段：

首先是前向传播训练的一个阶段：它的过程是取一个样本 (X, Y_p) ，把 X 设置为输入；接

着算出对应输入样本的输出量 O_p 为多少。这个过程是信息经过提取变换，将最终的信息传到输出层。最后结果的得出是计算输入的量与与每一层的权值矩阵相乘，得到最后的输出结果：

$$O_p = F_n(\dots(F_2(F_1(X_p W^{(1)})W^{(2)})\dots)W^{(n)}) \quad (3.14)$$

然后是传播训练阶段：这个过程要求我们对输出样本值 O_p 和标准样本输出值 Y_p 他们之间的差值进行计算；最后，通过此差值按照极小化误差的方法反向传播到前面的层次中用以调整矩阵的权值。

3.6 本章小结

本章的第一小节简单介绍了深度学习的由来、定义以及特点，第二小节详细介绍了卷积神经网络的发展及应用，第三、四小节详细介绍了卷积神经网络推导及实现，并对其结构进行深入的分析，第五小节对神经网络的训练步骤进行了介绍，这一章对卷积神经网络进行详细地介绍，目的在于将卷积神经网络应用于车牌图像识别中，下一章将介绍基于卷积神经网络的字符识别。

第4章 基于卷积神经网络的车牌识别

车牌识别的步骤除了图像采集主要还有四步：车牌图像的预处理，车牌定位以及车牌字符的分割识别。

车牌识别的过程框图如图 4.1 所示：

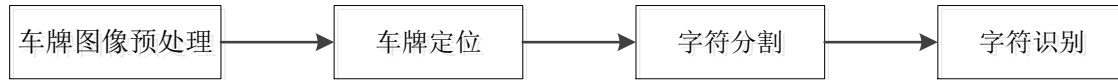


图4.1 车牌识别过程

4.1 车牌图像预处理

车牌图像预处理是图像处理分析的前期准备工作。一般来说是对图像的一个变换过程，在图像分析之前对图像进行一个简单的处理，使得图像要分析的内容特征更加突出，内容更加简练，信息量更少且唯一。因此在图像分析之前必须进行预处理，包括灰度化、二值化、去除噪音、边界增强等等。

4.1.1 车牌图像的灰度化

前面第二章已经说明过，因为人眼对红、绿、蓝三种颜色的敏感度不同，敏感度最高的是绿色，所以 G 的比重最小，取 $W_c=0.11$ ；敏感度最低是蓝色，所以 B 的比重最大，取 $W_a=0.59$ ；对红色的敏感度居中，取 $W_r=0.30$ ，这样得到的灰色图像效果最佳，故本设计采用加权平均值的灰度处理图像。



图4.2 原图像



图4.3 灰度化后的图像

本文是对汽车车牌的识别，由于其应用场合对准确度要求高，对物体的边缘清晰度要求较高，所以先用中值滤波法对图像进行平滑处理。为提高图像的车牌与背景对比度，使车牌图像与背景的反差色度增强，需要对图像进行增强处理，本文采用了直方图均衡的方法。



图 4.4 灰度增强化后的图像



图4.5 灰度图像均衡化后的直方图

上图是没有经过图像增强处理的直方图，下面是经过图像增强后的直方图：

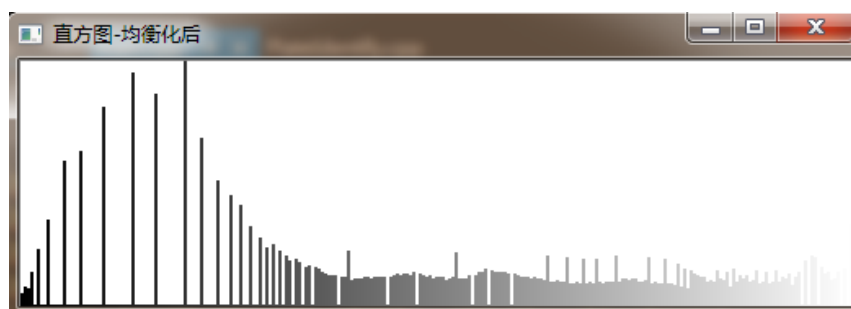


图4.6 灰度图像增强后均衡化的直方图

4.1.2 车牌图像二值化

本设计采用了全局阈值法对图像进一步处理，选取不同的阈值，二值化处理的图像差别很大，选取合适的阈值才能得到有完整有用信息的图像，本系统取阈值 $t=100$ ，即使灰度化后像素点的值大于 100 的等于 255（白色），小于 100 的像素点等于 0（黑色）。

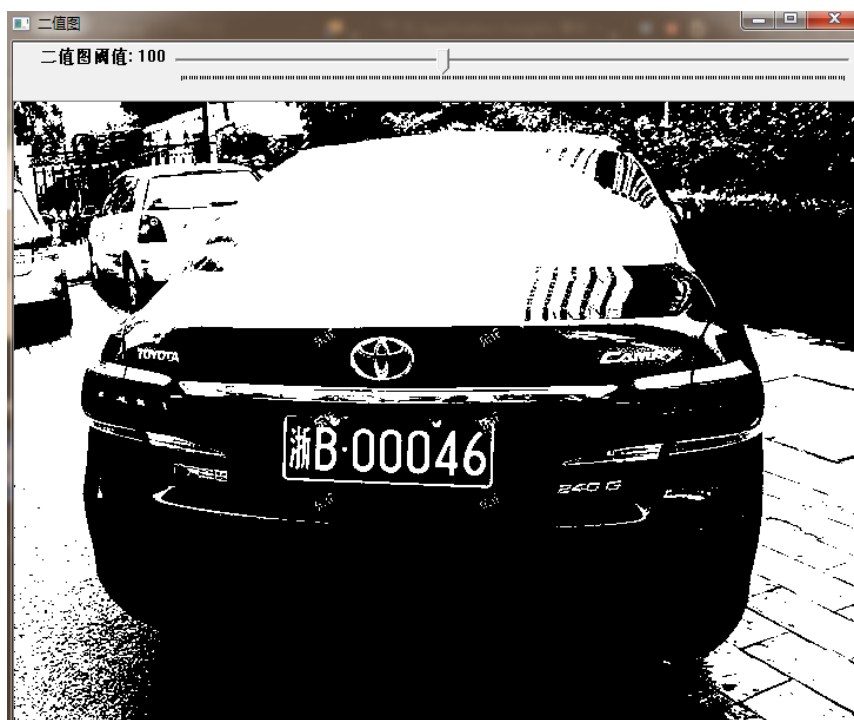


图4.7 二值化后的图像

4.1.3 车牌边缘检测

本设计对比了 Sobel 算子和 Canny 算子的处理效果，最终决定采用 Canny 算子对车牌图像进行增强处理。

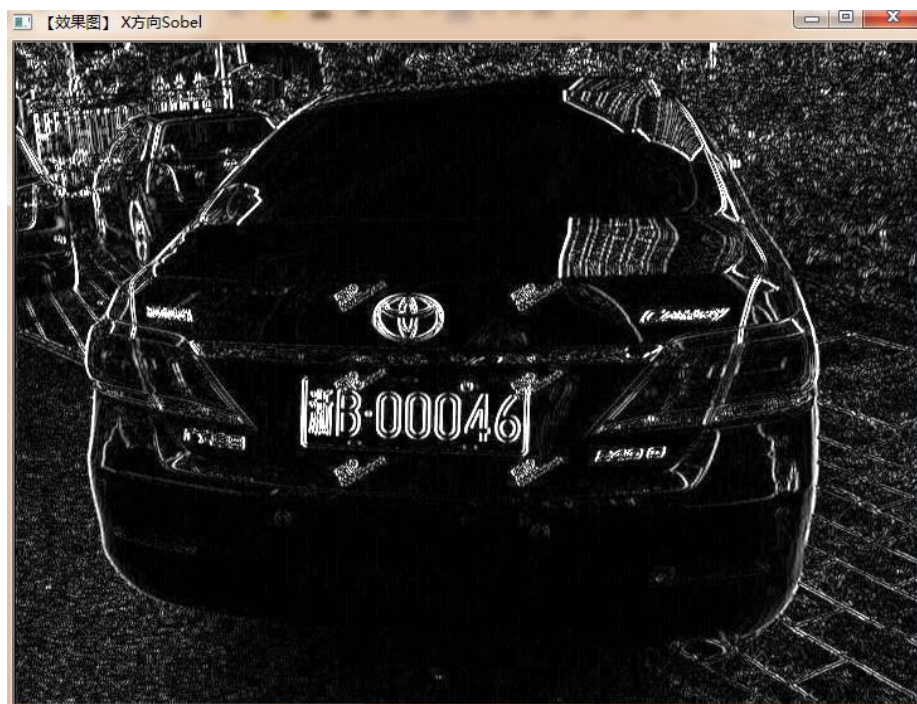


图4.8 X方向Sobel的图像

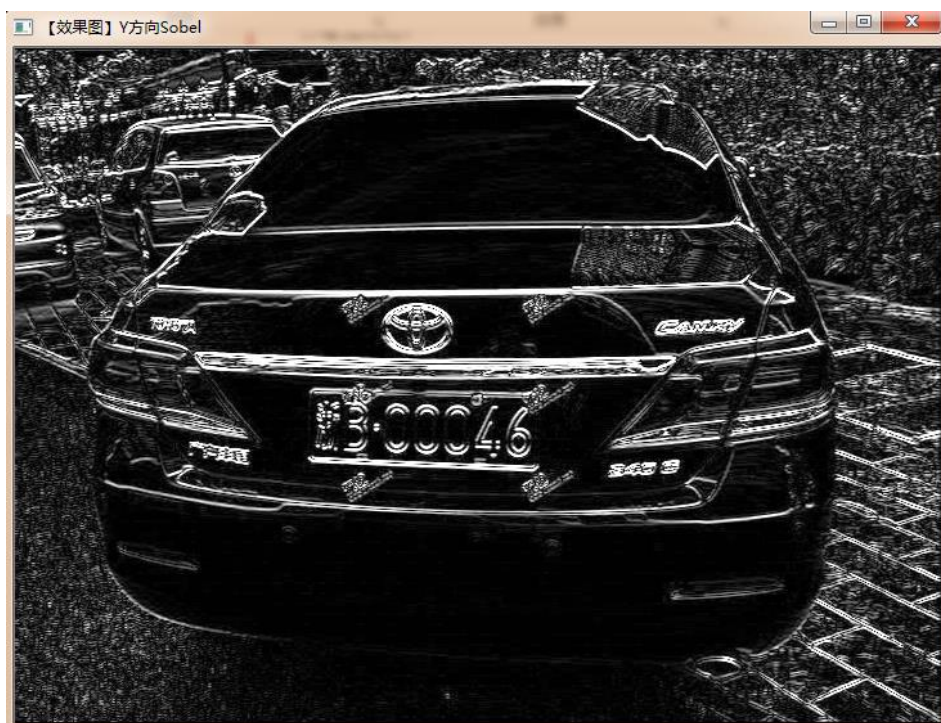


图4.9 Y方向Sobel的图像

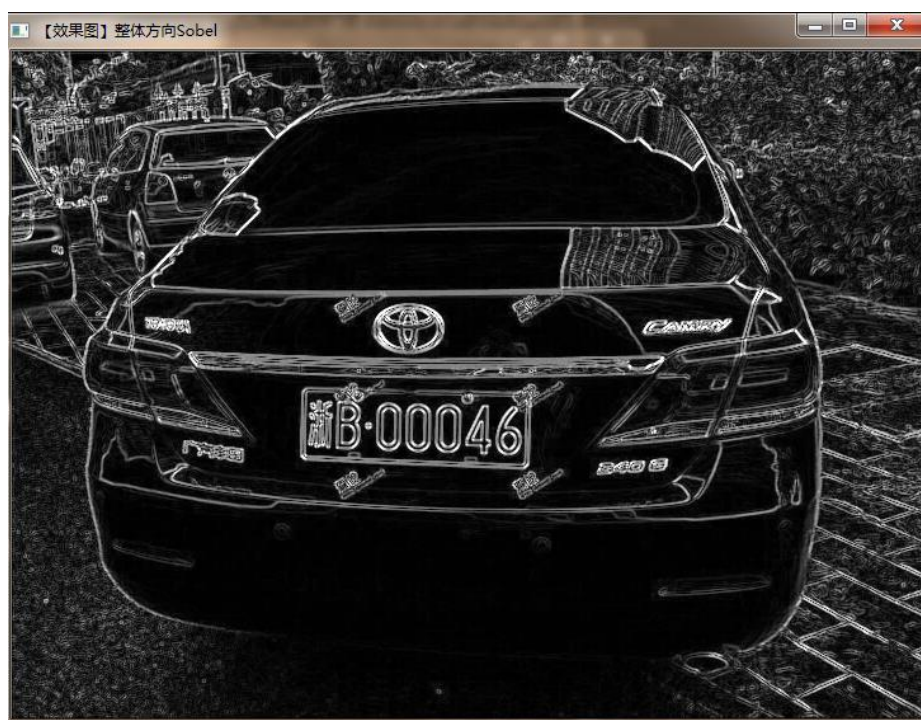


图4.10 Sobel后的图像

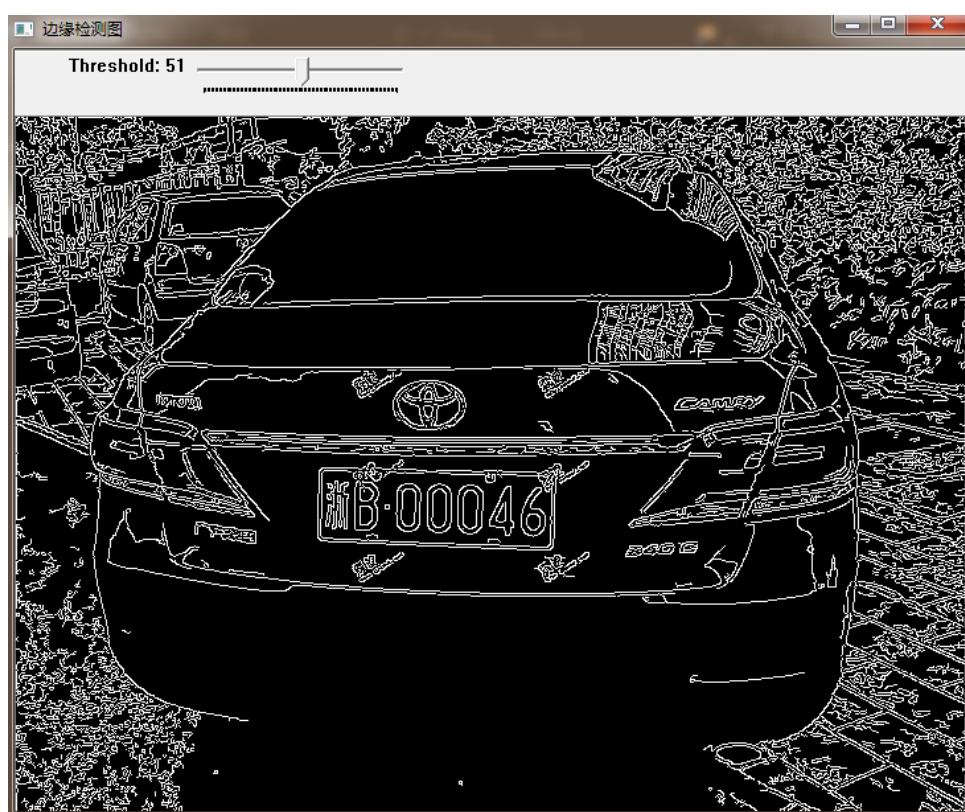


图4.11 Canny后的图像

4.2 车牌定位的实现

本文结合了车牌的颜色特征和形态特征采用了两种方法定位车牌，首先在第一阶段中采

用了基于数学形态学的定位方法定位图像后再进行细化定位，在进一步的定位中又采用车牌颜色特征的方法定位车牌图像，这样可以获得更好的定位效果。



图4.12 定位的车牌图像

下面为分割出的车牌图像：



图4.13 切割出的车牌图像

4.3 车牌字符分割

本实验采用结合先验知识和垂直投影的方式对车牌字符进行分割的，由于只采用单一的分割方法在不确定的复杂情况下，分割容易不奏效，因此，本文结合车牌字符本身的特征，采用垂直投影的方法分割。下图是几个车牌分割的结果显示。



图4.14 切割出的车牌字符

4.4 基于卷积神经网络的字符识别

在实验中使用的卷积神经网络结构不包括计算的输入层,所有层都包含可训练的参数(权值)输入是一个 32×32 像素图像这显著大于最大字符数据库(以 28×28 字段为中心的至多 20×20 像素),这样做的原因是期望潜在的独特特征被提取。在 LeNet 中,该组中心的最后卷积层的接收场在 32×32 输入的中心形成 20×20 区域。

在以下卷积层中标记为 C_x , 子采样层被标记为 S_x , 其中 x 是层索引。

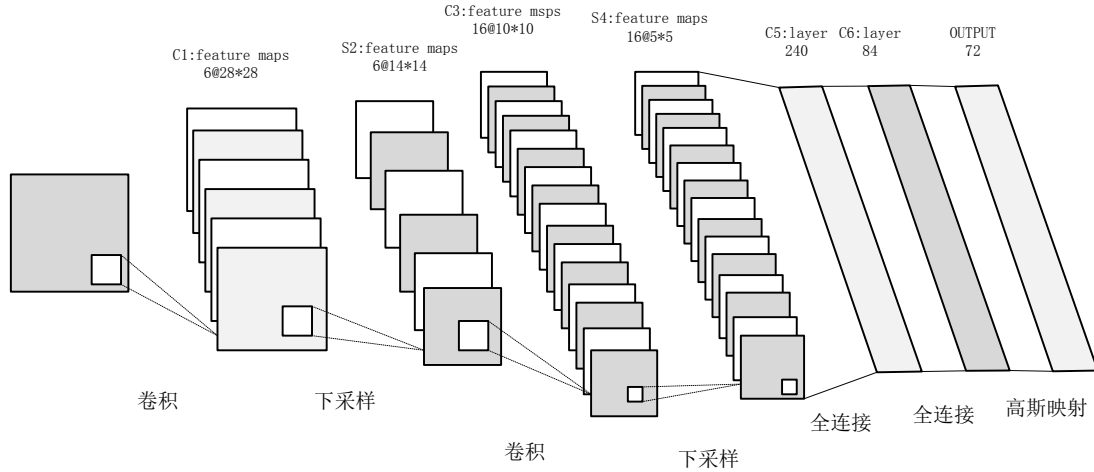


图 4.15 卷积神经网络结构

每个层(即卷积层和下采样层)有多个特征 Map^[49,50], 一个特征 Map 负责提取输入的一种特征, 另外一个特征 Map 是由多个神经元组成的。

C1 层是具有 6 个大小为 28×28 的特征图的卷积层。每个特征映射中的每个神经单元连接到输入层中的 5×5 邻域, 映射到 C1 层的大小为 28×28 。由上图 C1 层有 6 个滤波器, 根据上一章理论计算公式, C1 层的可训练参数的个数为:

$$(5 \times 5 + 1) \times 6 = 156$$

与输入层的连接个数为:

$$156 \times (28 \times 28) = 122304$$

S2 层是具有 6 个大小为 14×14 的特征图的子采样层。每个特征图中的每个神经单元与 C1 中的对应特征图中的 2×2 邻域相连接。S2 层具有 12 个可训练参数和 5580 个连接。

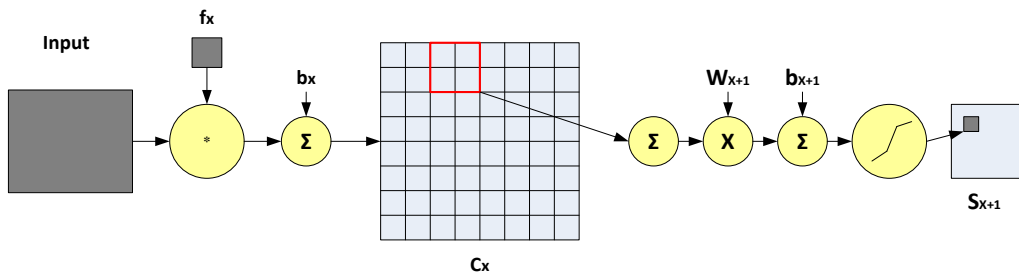


图 4.16 卷积和子采样过程

上图是卷积层与子采样层卷积的过程, 先对 C1 层的四个输入求和, 然后乘以可训练的系

数 W 并且添加偏置 b 得到的结果再通过激活函数传递到 $S2$ 层，其中 $C1$ 层中作为同一次输入的 2×2 感受野是不重叠的，所以得到的 $S2$ 中的特征图数目是 $C1$ 中特征图的数目 $1/4$ ，其中行和列的各占一半。

$C3$ 层也是一个具有 16 个大小为 10×10 的特征图的卷积层，每个特征图中的每个单元在 $S2$ 的特征图的子集中的相同位置处连接到几个 5×5 邻域。 $C3$ 层具有 1516 个可训练的参数和 151600 连接。这一层也是通过 5×5 的卷积核去与上一层进行卷积，然后得到含有 10×10 个神经元的特征 map，因为 $C3$ 层有 16 种不同的卷积核，这样对应的就有 16 个特征 map。

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X				X	X	X			X		X	X
4			X	X	X				X	X	X		X	X		X
5				X	X	X				X	X	X		X	X	X

图 4.17 $S2$ 层特征图与 $C3$ 层特征图的连接方法

上图表给出了由每个 $C3$ 特征图与 $S2$ 特征图组合的连接方法，从中可以看出， 5×5 的卷积核与 $S2$ 进行卷积得到 16 个特征 map 的方法。

$S4$ 层是具有 16 个大小为 5×5 的特征图的子采样层。每个特征图中的每个单元以类似于 $C1$ 和 $S2$ 的方式连接到 $C3$ 中的对应特征图中的 2×2 邻域。 $S4$ 层有 32 个可训练参数， $S4$ 与上一层之间的可训练连接数有 2000 个。

$C5$ 层是具有 240 个特征映射的卷积层。与 $S2$ 到 $C3$ 的连接方式不同， $C5$ 层的每个单元与 $S4$ 的所有特征图上的 5×5 邻域都相连接。因为 $S4$ 和 $C5$ 这两层的特征图大小都是 5×5 ，所以它们之间的连接方式是完全连接。

$F6$ 层有 84 个单元，与上一层（ $C5$ ）之间是完全连接。 $F6$ 层中计算它们的输入向量和它们的加权向量之间的点积，偏差被添加到该向量。然后将用于单元 i 的表示为 a_i 的该加权和通过 sigmoid 激活函数，以产生由 x_i 表示的单元 i 的状态：

$$x_i = f(a_i) \quad (4.1)$$

压缩函数是一个按比例缩小的双曲正切函数：

$$f(a) = A \tanh(Sa) \quad (4.2)$$

其中 A 是函数的幅度， S 确定其在原点处的斜率函数 f 与 $+A$ 和 $-A$ 处的水平渐近线是奇数。

最后，输出层由欧几里得径向基函数单元 RBF 组成，有 72 个神经元，每个神经元对应一个字符类别，每类一个单元，每个有 84 个输入。每个 RBF 单元 y_i 的输出计算如下：

$$y_i = \sum_j (x_j - w_{ij})^2 \quad (4.3)$$

4.5 本章小结

本章主要介绍了车牌识别的过程，并使用了一种新的字符识别算法——将卷积神经网络算法作为车牌字符识别的实现方案。该算法在第三章进行了详细地介绍，在第二小节详细介绍了它应用于车牌字符识别中的结构组成，各个层的设计原理以及实现方法细节，卷积神经网络算法用于字符识别，相比于传统的 BP 神经网络有非常明显的优点，接下来第五章将给出实验的最后结果。

第 5 章 实验结果测试与分析

5.1 实验结果测试

该系统的界面是用 MFC 做的，使用的语言是 C 和 C++^[51,52,53]，实验结果呈现：打开图像。



图 5.1 打开图像

点击“车牌检测”按钮，车牌被定位，并用红色框标出，显示如下：



图 5.2 车牌检测

点击“车牌分割”按钮，车牌字符就被分割，显示如下：

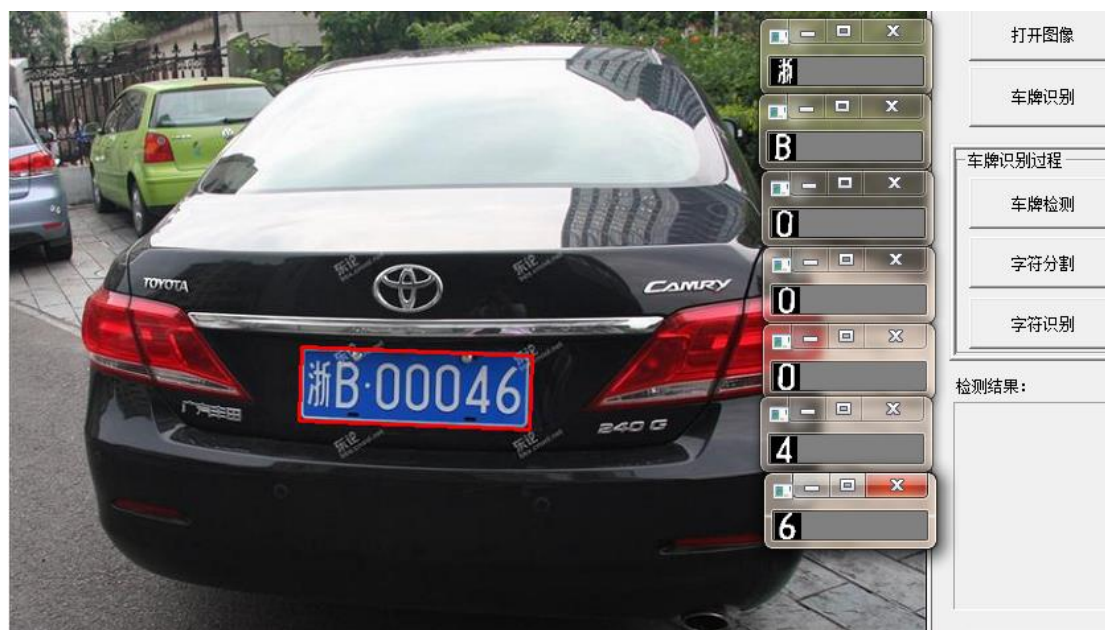


图 5.3 字符分割

点击“字符识别”按钮，弹出车牌识别的结果，显示如下：



图 5.4 字符识别

点击“车牌识别”按钮，会在检测结果窗口显示出车牌号码，显示如下：



图 5.5 车牌识别

该系统是对静态车牌图片进行识别的，对于正常清晰的车牌图片都是能识别的，但是对于拍摄不清晰的车牌图片依然会有错误，甚至无法识别。

5.2 实验结果分析

首先，需要收集实验所需字符图片，中国的车牌号码是由 10 个阿拉伯数字、25 个英文字母(不包括字母 I)，还有全国各省市、自治区的简称 31 个汉字组成，另外我们忽略车牌中的一些其它字符，比如“.”。本实验收集了 600 幅实际车牌照照片，先对其中的车牌字符进行分割（也可直接网上找别人已经分割好的字符样），用这些字符分别建立两个数据集：训练数据集与测试数据集，可用于训练的数据被分为两个不同的集合：训练集合和验证集合。为了提高神经网络的泛化能力，这两个数据集之间不应有任何重叠，这种技术称为交叉验证。

由于数字和英文字母的笔划数量少，而且车牌字符中的数字与字母形状规则，一般如果拍摄清晰、前期处理良好，识别起来比较容易。但是汉字识别相比于数字与字母的识别就稍微困难点，因为汉字字符较为复杂，在前期处理的时候也更容易受到噪声的干扰，导致分割出来的汉字字符模糊不清，造成后期识别困难。所以我选择了不同清晰度、不同大小的字符，以满足实际场景中采集环境的复杂性。

网络的真实性能仅在网络使用测试数据测试以测量网络对训练期间未见到的数据的性能时才显示。所涉及的字母数字字符是除了“ I “之外的所有字母，数字字符从 0 到 9，汉字全国各省市、自治区的简称有 31 个，因此总字符为 66 个。图 5.6 给出了图像的训练样本集，以” 6 “和” E “为例。



图 5.6 训练样本集

本研究中使用的训练和测试图像的数量为 600 和 256，分别具有 1.5% 的错误分类率或 98.5% 的准确性。该方法与 LeNet-5 类似，只是 LeNet-5 是用来识别手写数字的，而本设计中 CNN 结构与 LeNet-5 不同的在后面特征层 C5 的数目，由于识别字符的范围及难度的增加，本设计增加了 C5 的数目和输出层神经元的数目。另外在训练数据上，本实验的数据样本不多，有的研究训练数据和测试数据集的数量超过 2000 个样本，远远高于本文中提出的样本。但是就本实验而言，实验效果还好。

另外，实验还对比了 BP 神经网络与卷积神经网络的测试结果。试验结果显示采用卷积神经网络对机动车辆牌照进行识别具有较高的识别准确率。

表 5.1 识别结果对比

	传统 BP 神经网络		卷积神经网络	
	训练集(%)	测试集(%)	训练集(%)	测试集(%)
汉字	98.20	92.1	100	96.6
字母/数字	100	94.0	100	98.8
综合	99.5	93.6	100	98.5

本实验仍然有个别图像无法识别，经分析，这些车牌图像基本上都是在拍摄光线比较暗，导致图片模糊，而经过前期处理并没有改善多少，以至于无法定位出车牌或者是由于前期处理不完善使车牌不完整，在分割的时候无法分割出字符，也有的是分割出的字符不清晰或者存在噪声导致最后无法识别。

5.3 本章小结

本章主要是实验结果的展示、以及对实验结果进行了分析，并对比了传统的 BP 神经网络的识别效果，CNN 结构的识别率提高了 5%，分析总结了造成识别失败的原因，最后对识别失败的数据进行了分析，识别失败的车牌多数是由于拍摄不清晰致使车牌定位困难或者分割不完全导致字符不清晰无法识别。

第 6 章 总结与展望

6.1 总结

车牌自动识别系统是现代化交通管理深度学习是近年来模式识别领域的最火热的方向，大受研究者的欢迎，而且深度学习也具有良好的应用价值。本文首先对 CNN 原理进行了深入的理解，然后将其应用到车牌识别系统中，得到了很好的识别率。本文对车牌识别过程进行了处理：首先是对采集的车牌图像进行预处理，包括灰度化，图像增强，二值化，这些都是图像处理最基础的处理方法，然后对图像进行边缘检测，本文考虑到各种环境的复杂性选择了多种方法综合处理的车牌定位方法，在字符分割处理上是结合先验知识采用垂直投影分割方法对定位后的车牌进行分割的，最后采用卷积神经网络的方法用于识别分割后的字符，得出识别结果。

论本文车牌处理的过程，是本人一步一步学习并编写代码来实现的，每一步的处理都并不是百分之百理想的，总会有因为图像本身的不清晰或者被破坏而导致失败的，但是由于自己慢慢学习的过程的已经熟悉了车牌识别的过程，后续可以尝试其它方法进行处理，或者结合多种方法对车牌进行处理，毕竟我们采集车牌图像的环境是复杂的，存在各种因素会造成图像的不清晰而不利于识别。

车牌字符识别是车牌识别中至关重要的一步，本文对典型的卷积神经网络结构进行改进，实验表明，改进后的卷积神经网络对字符的识别率有一定的提高，对收集的车牌字符识别率达到 98.5%，之后的工作可以进一步增加样本的数量，以及训练次数，同时也可改变卷积神经网络的网络层的数量来观察识别率。不过个人觉得这些还是有点难度的，还需要在深度学习方面进一步学习与研究，毕竟足够的理论知识才能支撑起实践的执行。

6.2 展望

车牌识别的处理本身由于图片采集条件的复杂性，受光线、角度等因素影响大，现在很多车牌识别系统也都是有适用范围的，需要用硬件能力保证图片采集的良好环境，因此业内没有任何一个解决方案是对任何条件下都适用的，故对于多变环境下的车牌图像，我们仍然有很大的进步空间，未来我们是否可以研究出一个算法尽可能地处理各种环境下车牌图像。

另外，本文是基于深度学习神经网络的车牌识别，对于深度学习神经网络，要从数百万个图像中了解数千个对象，我们需要一个具有强大学习能力的模型，这就造成了识别任务的巨大复杂性，在 CNN 结构中，当网络层次不断增加，输入信号在网络中通过特征提取层再到特征映射层，这样一层一层的映射下去，对输入图像不断进行分解和转换，会形成大量的中间层特征图像与最后的目标图像，这样导致训练时间大大的增加，单独的 CPU 已经无法满足我们对时间速率的要求了，不过由于硬件平台的快速发展，我们可以采用 GPU 来训练神经网络。

络，这样可以大大缩短卷积神经网络的训练时间，提高效率。

当然，卷积神经网络作为一种深度映射的网络结构，还有很多东西值得我们研究，比如我们能否找到 CNN 的理论极限，或者我们能不能构建一种卷积网络，它的层数是固定的，让它在图像特征处理方面能表现出跟人脑一样的图像识别效果，这才是真正的人工智能。

致 谢

两年半的硕士研究生学习生涯转瞬即逝，很快我也将和身边众多的同学一道走出校园，走上工作岗位，开始新的人生阶段。此刻回首，我发现自己在这过去的两年多时间里确实成长了许多，从一个青涩稚嫩的本科生变成了一个具有一定专业能力的工科硕士研究生，知识技能日益丰富，视野心胸愈发开阔，世界观、人生观和价值观都变得更加成熟，个人的品味得到了显著提升。这些都是我的导师、我的同学以及亲爱的母校赐予我的！

首先，我要感谢我的导师——李训根教授。李老师知识渊博、技能精湛、视野宽广、学术思维极其活跃，他对自己方向上的学术研究充满了热忱，对各种顶尖前沿的学术成果如数家珍，与专业同行交流密切，对专业发展的方向把握准确。在读期间，我正是在李老师的悉心点拨与孜孜不见的教导下，才能够顺利的完成了专业知识与技能的研修，做出了令自己满意的成果，进而提升了自己的个人价值。即使毕业了，我也一定会牢记李老师的谆谆教诲，时刻鞭策着自己不辱师门，奋发向前。

其次，我要感谢我身边朝夕相处的同学，大家团结友爱，互帮互助，一同走过了这两年宝贵的学习时光。我要感谢我的师兄焦鹏飞和蔡允彬，正是在两位的指点与关怀下，我才能够迅速上手新的知识与技能，找准自己的特长以及将来真正属于自己的发展方向；我要感谢同一届的同门汪泓才和朱诗榜，大家能够同时拜入李老师门下，并能够齐心协力，在协助老师完成科研教学任务的同时攻克自己的学业难关，实属缘分；我还要感谢我的师弟们，正是你们的到来为实验室注入了新鲜的活力，让大家在学习工作忙碌之余多出了许多欢声笑语。

最后，我必须要感谢亲爱的母校——杭州电子科技大学。两年前，是您给了我一个继续深造、继续自我提升的机会，让我能够在宽敞明亮的教室里用心学习，在设备先进的实验室中专注工作，在鸟语花香、生机勃勃的校园间惬意徜徉。在杭电的两年半时光已经成为了我人生发展的重要拐点，您让我走上了一条诸多杭电优秀前辈曾经走过的道路，也必将在我今后的人生岁月中不断地鼓舞激励着我。我在此衷心希望母校今后能够愈发强盛，愈发兴旺，为祖国的发展培养越来越多杰出的人才，让我等毕业生能够以一名杭电人的身份骄傲的自居于各自的工作岗位上！

参 考 文 献

- [1] 文洁玲,上官明灿,孙文侠, 赵勇. 浅谈车牌识别技术应用现状[J]. 公路交通科技应用(技术版).2006(7):159-160.
- [2] Sheng ZHANG .License Plate Recognition Technology Development Research and Improvement[J]. Management Science and Engineering, 2013, 7 (2):56-59.
- [3] 郭大波,陈礼民,卢朝阳.韩丽萍. 基于车牌底色识别的车牌定位方法[J].计算机工程与设计,2003(05)
- [4] 魏武, 黄心汉, 张起森,等. 基于模板匹配和神经网络的车牌字符识别方法[J]. 模式识别与人工智能, 2001, 14(1):96-98.
- [5] 李志强, 李永斌. 车牌识别技术的发展及研究现状[J]. 科技信息, 2012(5):110-110.
- [6] Xin N, Lansun S. Research on License Plate Recognition Technology[J]. Measurement & Control Technology, 1999, 44-47:3667-3671.
- [7] 赵先军.基于神经网络的车牌识别技术研究[D]. 西安: 西安电子科技大学, 2005.11
- [8] Sheng ZHANG. License Plate Recognition Technology Development Research and Improvement[J]. Management Science and Engineering, 2013, 7 (2):56-59
- [9] 周科伟. Matlab 环境下基于神经网络的车牌识别[D]. 西安电子科技大学, 2009.
- [10] Wei Qi, Jing Han, Yi Zhang, Lian-fa Bai. Hierarchical image enhancement.Infrared Physics and Technology[J]. 2016, 76:704-709.
- [11] 吴成茂.直方图均衡化的数学模型研究[J].电子学报, 2013, 41(3):598-602.
- [12] 李凌.一种改进的车牌图像二值化方法[J].人工智能及识别技术,2006(35):163-164.
- [13] 付希金.自然背景下车牌识别关键技术研究[D].长春: 东北师范大学, 2008.
- [14] 胡宝清.模糊理论基础[M]. 武汉: 武汉大学出版社, 2004.
- [15] 陶文兵,金海.一种新的基于图谱理论的图像阈值分割方法 [J].计算机学报. 2007(1):110-119.
- [16] 袁春兰,熊宗龙,周雪花,彭小辉. 基于 Sobel算子的图像边缘检测研究[J]. 激光与红外, 2009, 39(1):85-87.
- [17] 梅跃松,杨树兴,莫波.基于 Canny 算子的改进的图像边缘检测方法[J]. 激光与红外, 2006, 36(6):501-503.
- [18] Gang Li, Zheng Liu, Chang-tao Chen, Feng-qing Han. A Method Based on Character Edge Color for Quick Locating Vehicle License Plate[J]. Journal of Multimedia, 2009, 4(6):356-362.
- [19] Yuanmei Tian, Juan Song, Xiangdong Zhang. An algorithm combined with color differential models for license-plate location[J]. Neurocomputing, 2016, 212:22-35.

- [20] 杨俊,戚飞虎.一种基于形状和纹理特征的车牌定位方法[J].计算机工程,2006(2):170-171.
- [21] M. -K. Kim. License plate location method unaffected by variation in size and aspect ratio[J]. International Journal of Automotive Technology. 2010, 11 (5):751-758.
- [22] 李刚,曾锐利,林凌等.基于数学形态学的车牌定位方法[J].仪器仪表学报.2007,28(7):1323-1327.
- [23] S. H.M. Kasaei,S. M.M. Kasaei,S. A. Monadjemi. A Novel Morphological Method for Detection and Recognition of Vehicle License Plates[J]. American Journal of Applied Sciences, 2009, 6(12):2066-2070.
- [24] 索迪.基于小波分析的图像处理在车牌识别中的应用[D].长春: 长春理工大学, 2008.
- [25] Wei Huang,Xiaobo Lu,Xiaoqing Ling. Wavelet packet based feature extraction and recognition of license plate characters[J]. Chinese Science Bulletin, 2005, 50(2):97-100.
- [26] 刘明军,谢宏霖,孙雪松,由枫秋.车牌字符分割算法的比较研究[J]. 济南大学学报(自然科学版). 2006(03)
- [27] 迟晓君,孟庆.基于投影特征值的车牌字符分割算[J]. 计算机应用研究.2006(07):256-257.
- [28] 甘 玲, 林小晶.基于连通域提取的车牌字符分割算法[J].计算机仿真.2011, 128 (04).
- [29] 王永杰,裴明涛,贾云得.多信息融合的快速车牌定位[J].中国图象图形学报, 2014.19(3):471-475
- [30] Guangmin Sun, Gang Li, Lei Xu, Jing Wang.The Location and Recognition of Chinese Vehicle License Plates under Complex Backgrounds[J].Journal of Multimedia, 2009, 4 (6):442-449
- [31] 魏武,张起森,王明俊,黄中祥.一种基于模板匹配的车牌识别方法[J]. 中国公路学报, 2001, 14(1):104-106.
- [32] 汤茂斌,谢渝平,李就好.基于神经网络算法的字符识别方法研究[J]. 微电子学与计算机, 2009, 26(8):91-93.
- [33] Wei Qi Yan, Jia Wang. BP-Neural Network for Plate Number Recognition[J]. International Journal of Digital Crime and Forensics. 2016, 8 (3):34-45
- [34] 欧阳俊,刘平.基于 BP 神经网络的车牌字符识别算法研究[J]. 光学与光电技术,北京, 2012,10(5):67-71.
- [35] 刘建伟,刘媛,罗雄麟.深度学习研究进展[J].计算机应用研究. 2014,31 (7):1921-1930.
- [36] Daniel Weimer, Bernd Scholz-Reiter, Moshe Shpitalni. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection[J]. CIRP Annals-Manufacturing Technology, 2016.
- [37] 胡正平,陈俊岭,王蒙,赵淑欢.卷积神经网络分类模型在模式识别中的新进展[J].燕山大

- 学学报,2015 ,39 (04): 283-291.
- [38] 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(03):638-641.
- [39] Rashad Al-Jawfi. Handwriting Arabic character recognition LeNet using neural network[J]. International Arab Journal of Information Technology, 2009, 6(3):304-309.
- [40] Earnest Paul Ijjina,C. Krishna Mohan. Human action recognition using genetic algorithms and convolutional neural networks[J]. Pattern Recognition, 2016, 59(11):199-212.
- [41] Andre Teixeira Lopes, Edilson de Aguiar, Alberto F. De Souza, Thiago Oliveira-Santos. Facial Expression Recognition with Convolutional Neural Networks: Coping with Few Data and the Training Sample Order[J]. Pattern Recognition,2016, 61:610-628.
- [42] Răzvan-Daniel Albu. Human Face Recognition Using Convolutional Neural Networks[J]. Journal of Electrical and Electronics Engineering, 2009, 2(2):110.
- [43] 郑永森.基于卷积神经网络的身份证号码识别研究与实现[J]. 计算机光盘软件与应用, 2015(3):13-14.
- [44] F Bastien,Y Bengio,A Bergeron,N Boulangerlewandowski. Deep Self-Taught Learning for Handwritten Character Recognition[J]. Arxiv Preprint Arxiv, 2016. 01.007.
- [45] Lécun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [46] Céile Barat, Christophe Ducottet. String representations and distances in deep Convolutional Neural Networks for image classification[J]. Pattern Recognition, 2016, 54(C):104-115.
- [47] Markus Thom. Sparse Activity and Sparse Connectivity in Supervised Learning[J].Journal of Machine Learning Research,2013, 14(1):1091-1143
- [48] 汪济民.基于卷积神经网络的人脸检测和性别识别研究[D].南京: 南京理工大学, 2015.
- [49] Szarvas Mate, Yoshizawa Akira, Yamamoto Munetaka, Ogata Jun.Pedestrian detection with convolutional neural networks [C].IEEEIntelligent Vehicles Symposium Proceedings. USA: IEEE, 2005:224-229.
- [50] A Krizhevsky,I Sutskever,GE Hinton. ImageNet Classification with Deep Convolutional Neural Networks[J].Advances in Neural Information Processing,2012, 25(2):2012
- [51] 谭浩强. C 程序设计 (第三版) [M]. 北京: 清华大学出版社, 2005: 1-390.
- [52] Prata S, 云巅工作室译. C Primer Plus (第五版) 中文版[M]. 北京: 人民邮电出版社, 2005: 1-626.
- [53] 严蔚敏, 吴伟民. 数据结构: C 语言版[M]. 北京: 清华大学出版社, 2007: 1-321.

附 录

作者在读期间发表的学术论文及参加的科研项目

发表的学术论文：

1. 王晶. 基于深度学习神经网络的车牌字符识别技术的研究[J]. 工业控制计算机, 2016. (已录用)