



浙江工业大学

硕士学位论文

论文题目：基于深度学习的车牌识别算法研究

作者姓名 王恒

指导教师 潘翔教授

学科专业 计算机科学与技术

所在学院 计算机科学与技术学院

提交日期 2017年5月16日

浙江工业大学硕士学位论文

基于深度学习的车牌识别算法研究

作者姓名：王恒

指导教师：潘翔 教授

浙江工业大学计算机科学与技术学院

2017 年 5 月

**Dissertation Submitted to Zhejiang University of Technology
for the Degree of Master**

**Research on the license plate recognition based on deep
learning**

Candidate: Wang Heng

Advisor: Prof. Pan Xiang

**College of Computer Science and Technology
Zhejiang University of Technology
May 2017**

浙江工业大学

学位论文原创性声明

本人郑重声明：所提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科研成果。除文中已经加以标注引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的研究成果，也不含为获得浙江工业大学或其它教育机构的学位证书而使用过的材料。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人承担本声明的法律责任。

作者签名: _____ 日期: _____ 年 _____ 月 _____ 日

学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权浙江工业大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于

- 1、保密☐，在_____年解密后适用本授权书。
- 2、不保密☐。

(请在以上相应方框内打“√”)

作者签名: 日期: 年 月 日

导师签名: 日期: 年 月 日

基于深度学习的车牌识别算法研究

摘 要

对于智能交通系统，车牌识别技术是最为基础也是非常关键的一个功能。近年来特定环境下的车牌识别技术已经达到了很高的精度。但是车牌识别是一种对外部光照环境比较敏感的技术。在一些光照环境不好、拍摄距离远且角度不正、夜间大灯等复杂环境影响下，会严重影响车牌图像质量。从而导致车牌字符退化，影响识别率。论文结合近年来的深度学习理论展开车牌识别研究，主要工作如下：

1. 对于车牌相似字符识别问题，提出了基于深度学习的特征提取和识别方法。该方法首先对字符图像进行归一化处理。然后以归一化后的图像为输入，构建五层深度网络对相似字符由低层到高层的特征表达。在特征提取层定义上采用对字符边缘特征敏感的卷积函数，从而能够对相似字符的局部差异进行分析。实验分析表明，和支持向量机等传统方法相比较，论文算法能够显著提高相似字符识别准确率。

2. 针对车牌多字符识别效率问题，提出了集成深度学习的多层次车牌识别系统。算法首先采用支持向量机对字符进行统计分类。然后对相似字符时采用卷积神经网络进行识别优化，得到车牌识别结果。实验结果表明该方法不仅能够保证车牌字符整体识别率，而且避免了全部采用深度网络所导致的识别速度慢问题。

3. 设计和实现车牌识别系统。在上述算法研究基础上，设计实现了车牌识别系统。系统通过动态链接库技术完成车牌识别核心算法的定义，方便二次开发和应用接口。然后构建车牌识别应用，实现图像的导入和识别。

关键词：车牌识别，深度学习，相似字符识别，卷积神经网络

RESEARCH ON THE LICENSE PLATE RECOGNITION BASED ON DEEP LEARNING

ABSTRACT

License plate recognition is the most basic and important part for intelligent transportation system. In recent years, the license plate recognition technology has reached a very high precision. But the license plate recognition is a kind of technology which is sensitive to the external light environment. Some cases like brights, long shooting distance and night light and other complex environmental impact will seriously affect the quality of the license plate image. As a result, the recognition rate will be decreased due to the degradation of license plate characters. In this paper, we study the license plate recognition based on the theory of deep learning in recent years. The main work is as follows:

1. For license plate similar character recognition, a feature extraction and recognition method based on deep learning is proposed. Firstly, the character image is normalized. Then the normalized image is used as input, and the five layer depth network is used to express the similar characters from low level to high level. Based on the definition of the feature extraction layer, we use the convolution function which is sensitive to the edge features of the characters. The experimental results show that compared with the traditional methods such as support vector machine, the algorithm can significantly improve the accuracy of similar character recognition.

2. For the problem of multi character recognition of license plate, this paper proposes a multi-level license plate recognition system. Firstly, SVM is used to classify the characters. Then, the similar character is further recognized by using the convolution neural network. The experimental results show that the proposed method can not only guarantee the overall recognition rate of the license plate character, but also avoid the problem of slow recognition speed caused by the depth network.

3. Design and implementation of license plate recognition system. On the basis of the above algorithms, the license plate recognition system is designed and implemented. The system uses dynamic link library technology to complete the definition of the core algorithm of license plate recognition, which is convenient for application interface. Then the license plate

recognition is applied to complement the image import and recognition.

Key Words: license plate recognition deep learning similar character recognition
convolutional neural network

目 录

摘要.....	i
第 1 章 绪论.....	3
1.1 研究背景及意义	3
1.2 国内外研究现状	4
1.2.1 深度学习理论研究.....	4
1.2.2 车牌识别技术研究.....	5
1.2.3 车牌识别相关产品.....	7
1.3 存在的问题及分析	8
1.4 本文的研究目标和主要内容	8
1.5 本文组织与结构	9
第 2 章 基于深度网络的车牌相似字符识别.....	10
2.1 引言	10
2.2 相似字符图像研究	10
2.3 图像预处理	11
2.4 算法概述	11
2.5 深度卷积神经网络结构	12
2.5.1 卷积层.....	13
2.5.2 非线性变换.....	14
2.5.3 下采样层.....	15
2.6 训练过程及特征的可视化	17
2.6 实验结果与分析	19
2.6.1 深度学习参数选择.....	19
2.6.2 算法性能比较.....	20
2.7 本章小结	22
第 3 章 集成深度学习网络的多层次车牌识别算法.....	23

3.1	引言	23
3.2	车牌定位	23
3.2.1	国内车牌几何特点	23
3.2.2	车牌图像预处理	24
3.2.3	车牌定位算法	26
3.3	车牌字符分割	29
3.4	多层次车牌识别	32
3.4.1	字符归一化和二值化	33
3.4.2	支持向量机粗识别	35
3.4.3	卷积神经网络相似字符识别	36
3.5	实验结果与分析	37
3.5.1	车牌区域训练数据预处理	37
3.5.2	实验结果比较与分析	38
3.6	本章小结	40
第 4 章	车牌识别系统设计与实现	41
4.1	系统开发环境	41
4.2	系统设计与实现	41
4.2.1	系统设计	41
4.2.2	关键数据结构	42
4.2.3	核心算法动态链接库设计	45
4.2.4	车牌识别展示原型系统	46
4.3	本章小结	47
第 5 章	总结与展望	48
5.1	总结	48
5.2	展望	49
	参考文献	50
	致谢	53
	攻读学位期间参加的科研项目和成果	54

第 1 章 绪 论

1.1 研究背景及意义

随着我国改革步伐加快和工业化进程深入以及国人的生活水平不断提高,汽车已经成为人们日常出行不可或缺的交通工具,这使得汽车在百姓生活中逐渐得到普及。截止 2015 年底,全国机动车保有量达 2.79 亿辆,新增汽车 1781 多万,汽车保有量增长创历史新高^[1]。由于城市交通的发展进度满足不了快速增长的机动车数量,这给城市交通带来了巨大的压力:如早晚高峰和雨天的道路拥挤、违章及交通事故频发等。面对这些问题,智能交通系统(Intelligent Transportation System, ITS)是目前最有效的解决手段。智能交通系统综合应用信息技术、计算机视觉技术、图像处理技术、数据通讯传输技术,使得系统能够自动对当前车辆进行监控并进行数据分析,形成一个实时、准确、高效的综合交通运输管理系统。目前,海康卫视、大华等企业都提供了智能交通的一体化解决方案,明显提高了车辆管理效率。

对于智能交通系统,车牌识别技术(License Plate Recognition, LPR)最为基础也是非常关键的一个功能。车牌作为汽车唯一的身份标识,为了便于对车辆的管理和监控,车牌识别系统便应用而生了。作为识别车辆信息的重要工具,它的出现主要目的是解决日益复杂多变的交通环境并保障安全驾驶。车牌识别系统的关键在两个方面,首先如何在全天候下利用摄像头获取车辆的清晰图像,再用什么算法才能获得更好的识别效果。随着对车牌识别系统研究的深入,车牌识别的技术变得越来越成熟并且识别速度也不断提高。如果车牌识别错误,会导致后续车辆大数据分析、违法数据检索等方面存在一系列方面的问题。车牌识别通过摄像头采集车辆的图像或者视频,利用数字图像处理、模式识别等技术提取和识别汽车牌照的号码信息。在光学字符识别领域基础上发展起来的车牌识别技术,经过近 10 多年的研究和发展,在特定场合已经达到了非常高的精度。目前,车牌识别系统做的越来越成熟,已经广泛应用在现实生活中,并向着智能化、自动化和无人化管理方向发展。随着技术的成熟、硬件成本的降低和人力成本的升高,现在路过停车场、车库、交通路口、小区的进出口等场合,都会见到许多地方已经装上了车牌识别系统,如图 1-1 (a) 所示停车场管理系统,采用车辆识别技术替代了传统的刷卡手段。不仅能提高入场效率,而且降低了人力成本。但是车牌识别是一种对外部光照环境比较敏感的技术。对于停车场

系统，由于摄像头和车辆距离短、角度小，从而使得车牌清晰度非常高，因此字符的特征区分度非常高，保证了车牌识别的高精度。在一些光照环境不好、拍摄距离远而且角度不正、夜间大灯等复杂环境影响下，会严重影响车牌图像质量（如图 1-1(b)）。从而导致车牌字符退化，影响识别率。

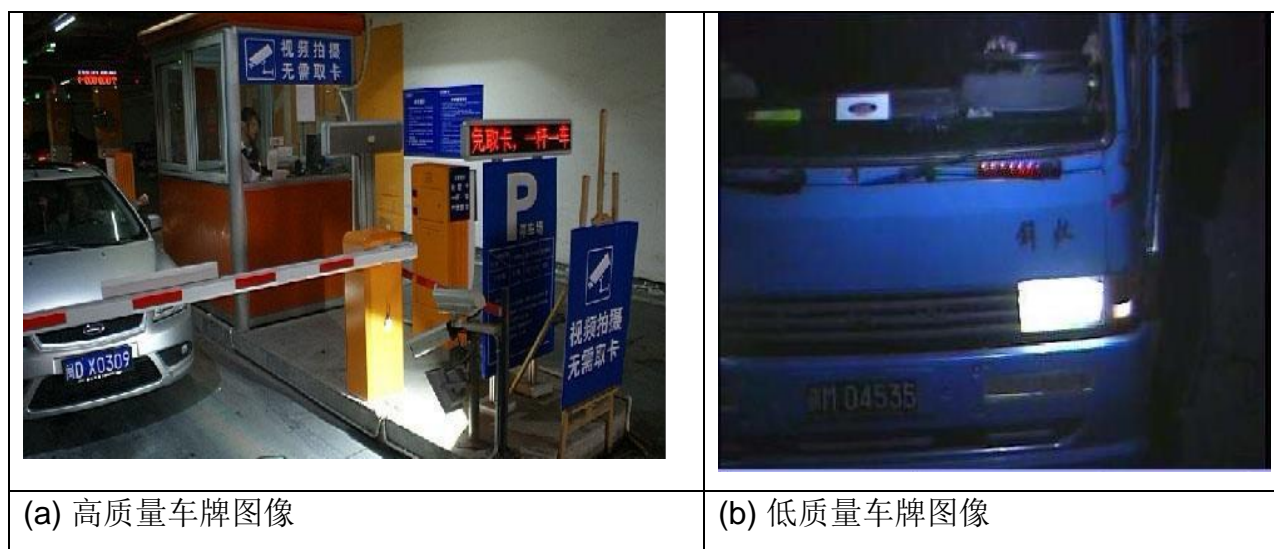


图 1-1 不同环境下的车牌识别

论文针对车牌识别中的相似字符展开研究。主要是在光照不好的前提下，退化的相似字符特征非常不明显，而且如果直接对图像采用二值化处理会导致特征丢失。因此，论文结合近年来的深度学习算法展开相似字符识别研究。通过卷积网络直接在灰度图像上提取特征，对相似字符构造分类器，实现相似字符的高精度识别。为了避免深度学习网络由于复杂结构导致的识别效率降低，论文提出支持向量机和深度学习网络相结合，构造多层次网络。最后，论文采用动态链接库技术完成车牌识别核心算法的封装，并给出了演示系统。

1.2 国内外研究现状

本节对论文相关研究进行分析和总结。首先阐述了深度学习理论的相关研究，然后对车牌识别相关技术进行了回顾，最后给出了国内外一些比较典型的车牌识别系统产品。

1.2.1 深度学习理论研究

深度学习的概念在 2006年由多伦多大学的教授 G. E. Hinton等对人工神经网络的研究时提出，从此拉开了深度学习在学术界和工业界的序幕。它是机器学习领域里新出现的一种方法，是一种含有多隐含层的深层的机器学习模型而且具有特定的网络结构。通过层层相连的方式模拟人类大脑的神经结构并应用于不同领域^[2]。目前深度学习的理论还处于起

步初始的研究阶段，但是在语音识别、图像分类和识别、自然语言等领域的广泛应用，尤其在大数据的背景下，利用海量的数据取得了巨大的成功。

深度学习算法最先在图像领域被尝试和应用并掀起了一股热潮。1998 年，多伦多大学教授 Yann LeCun 等人提出卷积神经网络 (Convolutional Neural Networks^[3])，将该方法应用到手写数字图像的识别中，并取得非常良好的效果。在卷积神经网络中含有关键的卷积层来提取图像有用信息，一般含有隐含层的数量在 3 个以上。卷积层模型的设计受到 D. H. Hubel 等^[4]基于对猫视觉皮层细胞研究的启发，它能有效降低网络复杂性的特殊网络结构，因此 LeCun 等人将他们提出的局部感受野概念应用到了卷积层，现在我们又称之为卷积核。卷积神经网络在尺度较小的图像上的应用取得了非常突出的成绩，但是在尺度图像上的应用效果一直不佳。

直到 2012 年 10 月，Krizhevsky 等人提出 AlexNet 网络模型^[5]，在著名的 ImageNet 大规模视觉识别挑战赛 (ImageNet Large Scale Visual Recognition Challenge, ILSVRC)^[6] 取得世界最好的成绩，他们在自己参与的组里获得第一名的成绩。他们构建的神经网络模型是在卷积神经网络结构的基础上做了改进，完全没有使用任何人工设计的算法对特征进行提取，并且首次将权重衰减的概念引入到在网络模型的训练中，有效地控制权值大小，防止网络出现拟合的现象。在随后两年的比赛中，深度学习算法被广泛使用，大多数是基于卷积神经网络的改进及其变形方法。尤其到了 2014 年的比赛，几乎所有参加大规模视觉挑战赛的队伍都采用了深度学习算法^[7]。

从深度学习首次在 ILSVRC 挑战赛上取得成功之后，到后来全部的参赛队伍都使用了深度卷积神经网络或者及其延生的算法，可以发现深度学习算法避免了人为花费大量时间设计特征提取方法的工作，因此它比传统的人为设计提取特征方法在图像识别领域具有更明显的优势。更关键的是随着计算机处理能力不断提升，GPU 加速技术的快速发展，并且 GPU 技术在训练数据中的应用，加快了网络的训练的速度。因此在大规模的训练数据集下，花费更短的时间从而获得更好的拟合训练模型^[8]。

1.2.2 车牌识别技术研究

车牌识别技术最早可以追溯到 20 世纪 80 年代，它的概念是由国外学者提出的并且那时候已经展开了广泛研究，这个阶段一般采用简单的数字图像处理技术来识别车牌。利用分辨率较低的摄像机获取车辆地图像，然后计算机对获取的图像进行简单地运算处理。进入 90 年代后，车牌识别研究取得了比较明显的进展。车牌识别一般的流程：首先

利用成像设备获取高质量的车辆图像，定位车牌位置并获取车牌图像，然后利用分割字符技术，进而对字符进行识别，最后得到车牌的号码。在车牌识别的问题上，国内外学者都提出了适用不同环境下的方法。但是车牌识别技术中最关键步骤在于车牌定位、车牌字符分割和车牌字符识别。



图 1-2 车牌识别系统算法流程

随着计算机性能的不不断提升，在图像处理上的速度也相应加快，以及有关成像设备的更新，A. S. Johnson^[9]通过使用数字图像处理技术完成了对车牌的识别。首先利用了图像分割寻找车牌位置，再利用先前的构造字符模板进行匹配进行字符识别，最后对车牌进行自动识别。AEun Ryung^[10]使用了彩色图像处理的办法，他使用的 HLS 彩色模型的方法识别率最高为 91.25% Paolo Comell^[11]使用了光学字符识别技术(Optical Recognition)对车牌的识别，通过梯度分析的方法检测出车牌的位置，之后用模式匹配和模板以及启发式的算法识别车牌号码。在实际应用中，这个时候车牌识别准确率和识别速度有了相当大提高，但是还没做实时性的要求。上海交通大学赵雪春等^[12]提出了一种采用彩色分割及多级混合集成分类器的车牌自动识别方法，车牌定位正确率达 98.6% 字符识别正确率达到 95%以上。华中科技大学黄心汉等^[13]提出了一种基于模板匹配和神经网络的车牌识别方法，该方法集成了改进的神经网络和模板匹配进行字符识，并且具有较高的识别率和较快的识别速度。另外，还有许多高校也做过类似的研究，浙江大学图形图像研究所、清华大学人工

智能国家重点实验室等。

从车牌定位上，比较常用的有基于灰度图像的车牌定位^{[14][15]}和基于彩色图像的车牌定位^{[16][17][18]}这两类方法。基于灰度图像信息的方法，通常根据车牌图像的投影特征、纹理特征、边缘灰度直方图统计特征、角点特征和形态学特征等灰度特征。这些算法都是基于车牌的一些具体特征来定位车牌的，虽然算法有针对性，效率高，运算速度较快，但存在局限性，遇到复杂的环境时不能很好的适用。基于彩色图像的定位算法一般利用颜色模型将图像中的颜色分割或转换为其他颜色空间，然后利用彩色图像边缘检测与区域生长相结合或者形态学计算，最后利用纹理信息及结构特征等方法进行分析和进一步判断，并确定最终车牌的区域。虽然基于彩色图像的定位方法能适应复杂的环境，但是相应的计算量较大，速度慢，因此对于实时性较高的场景就很难适用。

在字符分割上，效率最高的方法是基于垂直投影的车牌字符分割方法^[19]。通过字符在竖直方向上的投影，那么就会出现有规律的波峰波谷，而波谷处可能就是字符间或字符内的间隙处，因此字符的正确分割位置应该在波谷的附近，再将直方图的数值与设定的阈值比较，得到字符的一些潜在的始末位置，然后根据车牌字符之间的距离和字符的宽高比确定车牌字符的位置，但是该方法不能够很好地处理字符粘连问题。基于模板匹配的车牌字符分割方法^[20]，主要根据车牌上字符的距离特征和字符宽高特征设计分割模板，然后对车牌的二值化图像进行垂直投影，通过车牌字符宽度、排列规则等特征信息进行匹配分割。有效地解决了车牌受磨损、污染以及剥落等影响带来的字符噪声造成切分不准确问题，对于相邻字符连通问题也能得到解决，但是在边框去除不完全的情况下，很容易误将边框当成字符，最终导致分割出错。

在字符识别环节，常见的使用基于模板匹配的方法^{[21][22]}、基于神经网络的方法^{[23][24]}，基于支持向量机的方法^{[25][26][27]}等。基于模板匹配的算法原理简单，计算复杂低，识别速度快，但是对图像质量具有较高的要求。基于神经网络的匹配算法，有很强的自学能力，但是存在训练时间长、数据表达能力差、网络结构复杂等问题，而且会出现局部最优和过拟合的现象。基于支持向量机的算法 also 具有很强的学习能力，但是对字符平移和旋转敏感。

1.2.3 车牌识别相关产品

车牌识别技术目前已经达到商业应用的性能，国内外很多公司都已经成功地把车牌识别算法应用到产品中。例如新加坡 Optasia 公司针对本国车牌推出了 VLPRS 系统；澳大利

亚 Auty 公司的 Safe-Cam 系统。此外日本、英国、加拿大、俄罗斯等各个西方发达国家和地区在人力和物力上投入了大量的资金和人才，致力研究适用于符合本国车牌规则的车牌识别系统。现在国内也存在许多研究车牌识别产品企业，其中比较成熟的产品有北京汉王科技有限公司研发的汉王眼，香港亚洲视觉科技有限公司(Asia Vision Technology Ltd)的慧光车牌号码识别系统，四川川大智胜公司的车牌自动识别系统等系统。这些公司的这些产品已经在停车场、小区管理、高速公路等多个领域投入使用。

1.3 存在的问题及分析

虽然国内的研究越来越多，车牌识别的算法也在不断改进和提高，而且获取车辆图像的设备越来越先进，图像也越来越清晰，但是车牌识别过程中仍然会遇到比较难处理的问题，尤其我国的情况更为复杂，主要表现在：

- (1) 中国车牌上的字符由汉字、英文字符和数字组成。字母和数字的笔画少且结构简单，因此比较容易识别，识别率高。但是汉字的笔画多且结构复杂，通常在获取车牌的时候遇到汉字中笔画粘连的情况，因此在识别时难度较大。
- (2) 我国车牌颜色种类较多，民用机动车主要可以分为蓝底白字的蓝牌和黄底黑字的黄牌，但是军车和警车用的白底黑字组合的白牌。由于车牌颜色多，也易于周围环境混合。通常蓝色和黄色车牌的首个字符均为汉字，但是警车的首位字符为汉字，军车字符排序又不一样。还有车牌的悬挂方式也不一致。
- (3) 道路的状况相对复杂，我国的车牌污损也比较严重，字迹模糊和退色，车牌的倾斜角度较大。以及其它字符区域的干扰，如车辆上粘贴各种字符，车牌部分被遮挡和车牌变形。这些情况都会导致车牌难以定位。退化的相似字符特征非常不明显，且相似字符在局部存在特征差异。
- (4) 高速运动会出现图像的模糊失真。在现实中我们会经常遇到车辆的速度过快，这样导致获取的图像会出现模糊的现象。进而使车牌字符的也变模糊，不易于字符的识别。因此需要应用去模糊的算法才能做到图像复原。

1.4 本文的研究目标和主要内容

车牌识别系统按照识别的流程主要可分为车辆图像获取、车牌定位、车牌字符分割和车牌字符识别四个模块。除了车辆图像获取外，车牌定位、分割和识别是车牌识别技术的主要研究对象。车牌定位首先对车辆图像获取模块输入的车牌图像进行必要的预处理

理，包括灰度化、灰度拉伸、弱化背景等操作使汽车牌照更加容易被发现，再根据图像的边缘特征、颜色特征、几何特征等进一步对车牌进行定位和识别工作，最终获得车牌的图像。分割模块对上一步输出的车牌图像进行倾斜矫正并二值化，然后根据垂直投影信息对车牌图像进行字符分割，输出字符图像。字符识别是对分割出的每个字符图像进行识别，经过分类器判断后决定字符的类别，最终字符按顺序组合在一起得到车辆牌照号码。

字符识别是本论文的核心章节。第二章介绍是基于深度学习的相似字符识别算法，为准确识别车牌打下基础。本文选择卷积神经网络作为应用算法，用 **Softmax** 分类器作为字符的分类器，从而分类识别出车牌字符，目前在字母和数字的识别分类上效果还不错。比较了卷积神经网络与 **SVM** 的性能上的比较，证明了卷积神经网络在相似字识别上具有较好的性能。为了进一步提高识别效率，因此设计了 **SVM** 和卷积神经网络相结合的字符识别分类器。最后完成了系统的实现。

1.5 本文组织与结构

本文在对上述主要步骤进行了详细地研究和挖掘，本文共分为五章，内容结构如下：

第一章绪论，介绍论文的研究背景和意义，在国内外研究现状中分别对深度学习的理论研究、车牌识别技术研究和车牌识别相关产品进行介绍，还讨论了车牌识别的问题和难点，最后介绍本文的研究工作和主要内容。

第二章深度卷积神经网络的车牌相似字符识别，介绍对字符识别的相关研究和深度学习算法，重点分析了卷积神经网络在图像识别上优势。然后构建了用于相似字符识别的网络结构，最后将训练好的分类器用于相似字符识别。

第三章集成深度学习网络的多层次车牌识别算法，首先介绍车牌定位，然后利用 **SVM** 对车牌真假的判断。车牌字符分割的算法。在字符识别方法，提出了支持向量机与深度学习相结合的车牌字符算法，即首先利用支持向量机对字符的粗识别，然后利用深度学习对相似字符进行更进一步的识别和判定。

第四章车牌识别系统设计与实现，介绍系统设计，系统结构以及各模块的主要功能，然后进行库的动态封装。

第五章总结和展望，对本文的研究方法和成果进行简要总结，最后针对本文存在问题和不足，对接下来的工作做出后续的展望。

第 2 章 基于深度网络的车牌相似字符识别

2.1 引言

设计车牌字符识别算法的目的是能够快速准确地识别车牌号码，它的识别结果的好坏将决定整个车牌识别的性能。一般字符识别包括字符预处理、字符特征提取以及分类判别。关键在于字符特征提取的算法和分类器的设计。基于此本章将采用深度学习算法对车牌字符识别问题的研究，首先介绍了深度学习的优势及提取特征，重点介绍卷积神经网络的网络结构和训练过程。通过实验对比分析，发现卷积神经网络适用相似字符的识别。

2.2 相似字符图像研究

对于相似图像识别研究基本上多以研究相似字符识别上展开的，常用二次识别法为主。一般第一次采用简单快速的方法对字符进行粗识别，判定字符是否属于相似字符集。如果判断为是则进行第二次识别，通过设计特征算法区分相似字符间的局部差异，这样能更好地对字符进行精确识别。

瞿中^[28]等人采用改进版的模板匹配法，首先建立相似字符集合，然后采用模板匹配法对字符进行第一次识别，如果识别结果在某个相似字符集内，采用局部特征匹配算法对相似字符进行二次识别，并且选取相似字符特征差距最大的区域作为特征区域（比如字符“0”和“D”差异最大的区域在字符的左上角和左下角），再与模板库中模板字符对应区域进行相似度计算，这样结合分级分类的思想有效地提高识别率。童剑军等^[29]人先对车牌字符进行二次立方插值法后，能使字符的边缘方向特征更加明显，然后对字符进行归一化后把字符分为 6 个子区域，根据不同相似字符类，对字符模板的子区域赋予不同的权值，进行进一步的精判别识别，比传统方法的识别率提高了 9%。黄文琪和吴炜等^[30]人利用贝叶斯网络分类器对二分类问题效果好的原因，并根据相似字符图像间差异存在于局部区域的先验条件，所以针对不同字符设计了不同的特征提取方法，最后通过对不同分类器作对比试验，证明贝叶斯分类器对车牌的相似字符识别的准确率更高且稳定性更好。

以上这些方法都利用了先验知识和个人经验，且需要花费大量时间设计提取特征的

算法。传统的特征提取方法只针对某一类字符，因此识别的算法不具备通用性，同时对分类器性能和特征提取的算法要求都较高。

2.3 图像预处理

图像预处理是字符识别之前一个重要的步骤。由于受外界因素的影响，采集到的字符图像可能存在噪声、光照等因素干扰，会直接影响字符识别过程，因此需要在此之前做相应的预处理工作。

将分割下来的单个字符进行归一化，即按照一定尺寸大小进行缩放操作，归一化后有利于对字符进行识别。在归一化之前还需要做一些预处理。在字符分割时，字符分割的边框是根据整个车牌的高和宽决定的，因此分割后的字符与图像边界之间可能存在一定的间隙。所以要去除这些边框，必须先找到字符的上下左右边界，然后进行裁剪。完成边框去除之后，接着按比例对字符图像进行缩放操作。如果已经知道了当前字符的大小和目标字符的大小，因此就可以计算出水平方向和垂直方向的缩放因子。最终采用插值法对字符进行缩放操作。

2.4 算法概述

深度学习最大的特点就是在于其模型的深度，就是它的网络结构含有多个隐含层，并且隐含层的数量要大于等于 3。它与传统方法依赖人工设计特征相比，避免了靠个人经验和知识设计特征算法。它通过分层学习的思想，将每上一层的输出特征作为下一层的输入，这样信息被分层表示，最终输出的特征就可以被分类了。例如以图像识别为例，将图像作为输入信息，第一层学习到的可能是一些点线或者边缘特征，第一层作为第二层的输入特征，因此可能学习更抽象的特征，可能是角点、线段等更高级的信息。分层思想还能将一个复杂的复杂函数进行简单化表示，通过层与层之间非线性变换，利用多层的优势使用较少的参数表示较为复杂的函数，比如函数 $\log(\sin(\exp^2(\cos(\theta))))$ ，用多层结构很容易实现该函数，如图 2-2 所示。深度学习的多层模型设计模拟人脑的神经结构，在识别物体的过程，逐层学习有用的数据特征，最终用机器语言对数据进行描述。总之，多层的学习模型对于数据具有更强的表达能力。

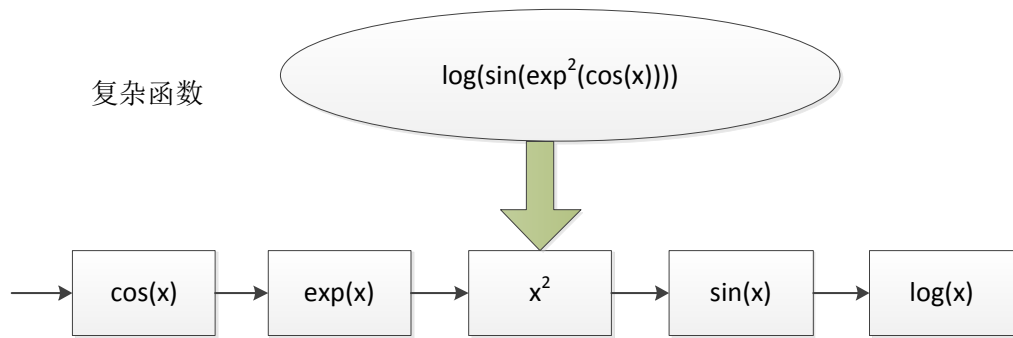


图 2-2 深度网络的分层表示

目前深度学习中最常用的有深度卷积神经网络(CNN, convolutional neural network)^[31]和深度信念网络(DBN, deep belief network)^{[32][33]}。其中深度信念网络未考虑到图像的二维结构,而卷积神经网络可以直接将图像作为输入,因此它更适合于二维结构的数据,自动学习特征,提取具有平移和形变不变性的特征表示^[34]。

卷积神经网络是一种包含卷积层和下采样层的神经网络。卷积层的设计源于研究猫视觉皮层细胞时发现的一种特殊神经结构——局部感受野^[4]。卷积层可由若干二维特征提取平面组成,每个平面上含有很多神经元。输入的数据通过卷积核和偏置进行卷积后,再经过一个转换函数输出该层的特征图。下采样层是对上层输出的特征映射图进行池化操作。经过上述的几次操作之后,最终这些像素值被光栅化,并将这些像素值连成一个一维向量得到输入。通常情况下,先对数据进行特征提取,每个神经元与上一层的局部区域相连,并对该区域进行提取特征,接着移动局部区域,使该局部区域与新的神经元相连,这样对所有区域进行特征提取。一旦完成特征提取,它们之间的位置关系就被确定下来。特征提取之后一般要进行特征映射操作,特征映射结构采用非线性函数作为卷积神经网络的激活函数,再经过池化后特征映射具有平移不变性。每个特征映射为一个二维平面,平面上的所有神经元的权值是共享的,因此网络的参数数量大幅减少,降低了网络的复杂度,使得网络的训练速度更快。下采样层将多个像素变为一个像素,这样对特征进行降维,有利于特征的分类。

卷积神经网络将局部感受野、权值共享、激活函数、下采样层等这些结构结合起来,使网络具有位移、尺度、形变不变性。

2.5 深度卷积神经网络结构

由论文采用深度卷积神经网络进行相似车牌字符识别。该网络采用反向传播结构,直

接对相似字符图像进行特征提取。

本文使用的深度卷积神经网络利用局部感知和权值共享降低了网络的复杂度，也加快了网络的学习和训练的速度，并且使用反向传播算法使得网络更加收敛，得到更好的训练模型。其实，具有卷积结构的深度神经网络的最基本组成单元为单层卷积神经网络（如图 2-3 所示）。首先在输入层之后连接多个单层卷积神经网络，这样完成了卷积神经网络最核心的组成部分，最后在它之后连接全连接层和分类器，就组成了一个深度卷积神经网络。

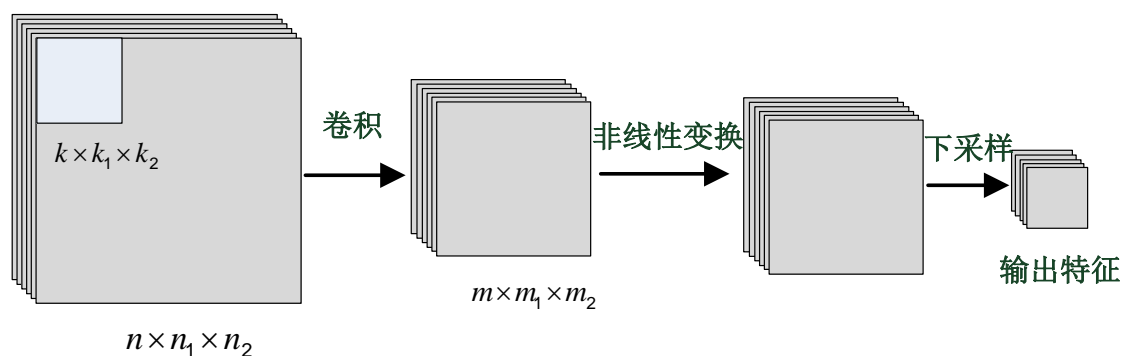


图 2-3 单层卷积神经网络结构

2.5.1 卷积层

卷积层的主要作用是特征的提取，它通过模拟生物的视觉皮层细胞，可对输入信号进行局部感知，这样和前一层的输入数据进行局部感受野相连，可以进行提取特征，这种局部感受野相当于特征的提取器。每个卷积核与前一层的特征图经过卷积操作后，能实现同一个映射面上的输入数据的权值共享。同时这种权值共享的方式减少了参数的数量，也简化了模型的复杂度。局部感受野的作用类似于一个 $n_1 \times n_2$ 大小的滑动窗口，并按一定的步长在字符图像上进行滑动，并且遍历了图像的所有位置。车牌的字符一般由粗的直线或曲线组成，形状特征明显，因此经过卷积之后，可以检测到字符图像的边缘、线段、角点等初级的视觉特征。因为卷积过程是对字符图像的所有位置上的特征提取并映射到下一层，避免了字符局部特征的丢失，也使得不同的字符特征存在差异性，有利于字符的识别。

卷积过程是由 n 个大小为 $n_1 \times n_2$ 的二维数据作为输入 x_i ，得到的输出结果 y_i 。 y_i 由 m 个 $m_1 \times m_2$ 大小的二维数据构成。该过程使用的卷积核 k 大小为 $k_1 \times k_2$ （通常卷积核的大小

采用 5×5)。计算的形式如下:

$$y_j = f(\sum x_i * k_{ij} + b_j) \quad (2-1)$$

其中 $*$ 为卷积运算符, k_{ij} 表示卷积核, b_j 为偏置。

2.5.2 非线性变换

非线性变换是将卷积层输出的特征映射图作为输入 x , 然后得到输入与输出之间的映射函数 $y = R(x)$ 。由于线性模型的表达能力有限并且有些数据是线性不可分, 所以它的引入使网络解决了线性不可分的问题。LeCun Yan 提出的 LeNet-5 网络中采用非线性激活函数 **sigmoid**, 它将变量映射到 0 到 1 之间。还有 **sigmoid** 变体版 **tanh**、**softsign** 等激活函数^[35], 但是目前深度学习多采用改进版 **ReLU**(rectified linear units)^{[36][37]}。传统神经网络一般采用 S 型曲线函数 **sigmoid** 作为激活函数, 但是使用传统激活函数在增加或者减少的时候梯度容易出现弥散效应, 但是对于 **ReLU** 而言并不会^[38]。从图 **sigmoid** 函数曲线可以发现, 函数曲线在两侧特别平缓, 而在中间却非常陡峭, 因此 **sigmoid** 函数对中间区域的信号增益较大, 对两侧的信号则抑制作用较大。在训练过程中, 它的两侧梯度值会非常小, 甚至趋于 0。而在深度神经网络中, 大多数神经元分布在 **sigmoid** 函数的两侧, 当梯度值趋于 0 的时候, 在反向传播过程中, 梯度值会越来越小, 训练速度也会大大减慢, 这种情况很容易导致信息丢失, 从而导致无法完成深度网络的训练。选择改进版的非线性激活函数 **ReLU** 之后, 从图中的曲线图可以发现, y 轴左侧的部分的值被压缩成 0, 而右侧的值保持不变, 这样对一部分神经元进行单侧抑制, 造成了网络的稀疏性, 能防止过拟合现象发生。又因为使用 **ReLU** 激活函数的网络梯度值经常为一个常数, 所以使网络的收敛速度更快。

以下是 4 种常见的非线性操作函数公式以及激活函数曲线。

$$\text{sigmoid: } f(x) = \frac{1}{1 + e^{-x}} \quad (2-2)$$

$$\text{tanh: } f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2-3)$$

$$\text{softsign: } f(x) = \frac{x}{1 + |y|} \quad (2-4)$$

$$\text{ReLU: } f(x) = \max(0, x) \quad (2-5)$$

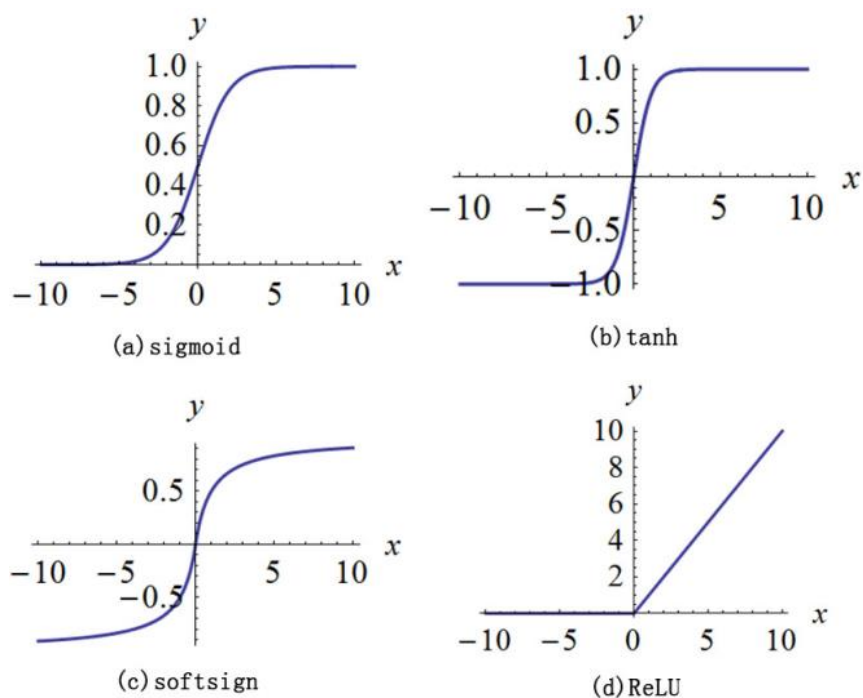


图 2-4 4 种常见的非线性函数曲线图

2.5.3 下采样层

下采样通过对上一层输出的数据做池化操作。池化操作之后，图像的分辨率降低，网络能抵抗图像的缩放、平移等形变问题，同时还能减少特征的数量。由于拍摄角度的问题和图像分割的原因，获得车牌字符的位置和角度不一，通过本文采用的最大值池化后，解决了上述的问题，因此算法对字符图像的旋转，平移等具有较好的鲁棒性。

经过下采样之后，输出与输入相同数量的特征图。下采样的计算形式如下：

$$x_i = f(\beta_i \text{down}(x_i) + b_i) \quad (2-6)$$

其中 $\text{down}()$ 表示下采样函数， β_i 表示下采样系数， b_i 为偏置。

本文采用的是 5 层结构的网络如图 2-5 所示。

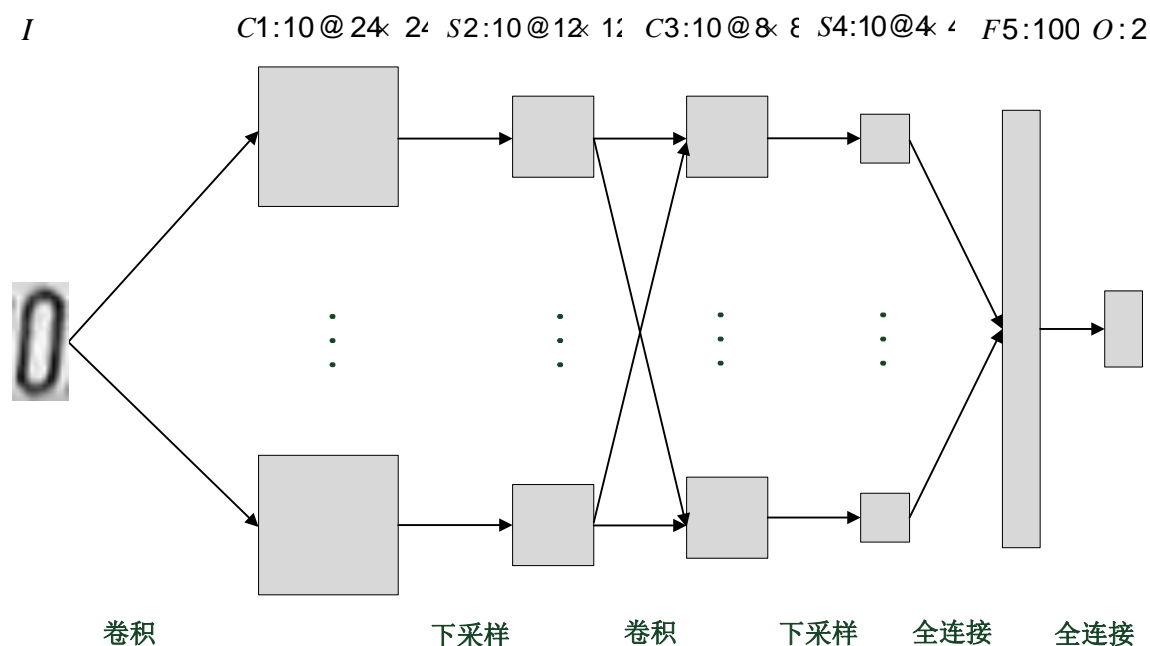


图 2-5 深度卷积神经网络

网络的输入是基于 28×28 的灰度图,即对于不同尺度的输入样本需先通过归一化处理得到指定大小的灰度图。

C1 层是个卷积层,对输入层经过大小为 5×5 卷积核和可加偏置 b 的卷积操作之后获得 10 个 24×24 大小的特征图,其中特征图中的每个像素点与前一层的输入层相对应一个 5×5 区域局部相连,且卷积步长为 1。每个卷积核含有 $5 \times 5 = 25$ 个权值参数和一个偏置参数,并且由于每个特征平面内的权值参数是共享的,因此该层共有 $(5 \times 5 + 1) \times 10 = 260$ 个可训练参数。

下采样层 S2 是对上一层 C1 做下采样处理, S2 层上的特征图与 C1 层上的特征图一一对应,同样含有 10 个特征平面,特征平面中的每个神经元与 C1 层的一个大小为 2×2 的“感受野”相连接,每个特征图包含 12×12 组成的下采样层。这样 S2 层的图像特征大小比 C1 层的特征图长宽各减一半, S2 层含有与特征图数量相同的权值和偏置,因此该层共有 20 个可训练参数。

网络层 C3 是由 10 个 8×8 大小的特征平面组成的卷积层,这里网络通过特征平面的每个神经元与 S2 网络层的若干个特征平面的 5×5 大小“感受野”连接来实现,感受野窗口移动步长为 1。C3 中的每一个特征图都是 S2 中几个特征图映射后叠加得到的,每个特征图包含 $10 \times 5 \times 5 = 250$ 个权值和一个偏置,所以 C3 层包含 $10 \times (250 + 1) = 2510$ 个可训练参数。

网络层 **S4** 是一个下采样层，下采样的方式同 **S2** 层一样，由 10 个 4×4 大小的特征平面组成。特征平面的每个神经元与 **C3** 层的一个 2×2 大小的“感受野”相连接，有 20 个可训练参数。

网络层 **F5** 为全连接层，由 100 个神经元组成，每个神经元与 **C5** 层进行全连接，有 $4 \times 4 \times 10 \times 100 = 16000$ 个可训练参数。最后，输出层 **O** 由 2 个神经元组成，每个神经元表示一个数字类别的置信度，由 **Softmax** 距离分类器连接 **F5** 层。

如果一个网络模型的层数增加后，那么输出层之前会连接多个全连接层，以保留更多的有用信息。这是因为在网络的前几层通过卷积层和池化层之后已经去掉了绝大部分的无用信息。

2.6 训练过程及特征的可视化

在训练模型的时候，通常由前向传播和反向传播^{[39][40]}两个过程构成。前向传播是将二维结构的图像作为输入数据，数据在网络的各层传播之后，最终获得数据的实际输出。算法的流程图如下：

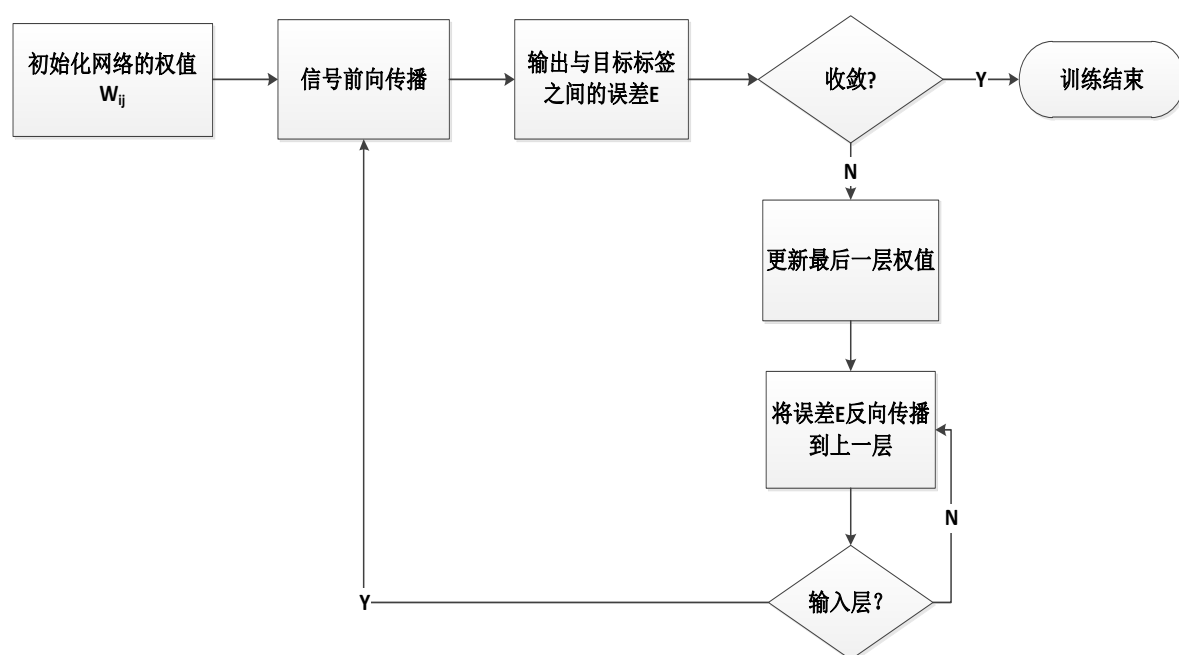


图 2-6 卷积神经网络训练过程

反向传播主要目的使网络更收敛。一组输入数据 **X** 在网络中传播之后，获得最后一层的特征图 **Y**。将输出特征图 **Y** 与预设的标签 **T** 进行对比之后，计算得到两者的误差 **E**。对于共 **c** 个类别，**N** 个训练样本，误差为：

$$E = \frac{1}{2} \sum_{k=1}^c (t_k^n - y_k^n)^2 = \frac{1}{2} \|t^n - y^n\|_2^2 \quad (2-7)$$

t_k^n 表示第 n 个样本第 k 维的标签值, y_k^n 表示第 n 个样本对应网络输出的第 k 维。
我们用 l 来表示当前层, 那么网络结构表示为:

$$x^l = f(u^l) \quad \text{其中 } u^l = W^l x^{l-1} + b^l \quad (2-8)$$

其中 $u^l = W^l x^{l-1} + b^l$, b 代表偏置数, W 代表权重, $f()$ 代表激励函数。

反向传播的误差代表了神经元的灵敏度, 它是求误差对偏置和权重的偏导, 用公式表示如下:

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial u} \frac{\partial u}{\partial b} = \delta \quad (2-9)$$

因为 $\partial u / \partial b = 1$, 得到 $\partial E / \partial b = \partial E / \partial u = \delta$, 那么就可以推出隐含层的灵敏度:

$$\delta^l = (W^{l+1})^T \delta^{l+1} \circ f'(x^l) \quad (2-10)$$

这里的 “ \circ ” 表示每个元素相乘。而输出层的灵敏度为:

$$\delta^l = f'(u^l) \circ (y^n - t^n) \quad (2-11)$$

那么对于第 l 层的神经元权值可以运用该层的灵敏度进行更新。通过求该层的误差对权值的导数, 即该层的输入与该层的灵敏度的叉乘, 然后乘以一个负学习率 η 就是该层的权值更新了。

$$\frac{\partial E}{\partial W^l} = x^{l-1} (\delta^l)^T \quad (2-12)$$

$$\Delta W^l = -\eta \frac{\partial E}{\partial W^l} \quad (2-13)$$

经过这样, 将误差 E 通过网络的反向路径传播到网络的每个节点上, 并更新对应的卷积核的权值 W 。随着网络的训练不断进行, 权值的误差值也随着迭代次数的上升而逐步减小, 最终使该值收敛在一个稳定范围内。

在输入层中已经定义了输入图片的大小为 28×28 , 但是我们获得实验数据都是彩色图片, 为了提高训练和测试的速度, 因此我们首先要对图片进行灰度化处理。

如图展示的是输入字符分别在深度卷积神经网络的卷积层和下采样层中输出的特征图, 并且特征由低层到高层变得越来越抽象。从图的左到右依次表示为输入层, 卷积层 C1, 下采样层 S2, 卷积层 C3 和下采样层 S4。从各层提取的特征可以发现, 深度卷积神经网络能有效的提取字符的特征, 在底层能够获得字符的点、线段等轮廓特征, 经过逐

层提取之后，在高层获得抽象的语义特征。

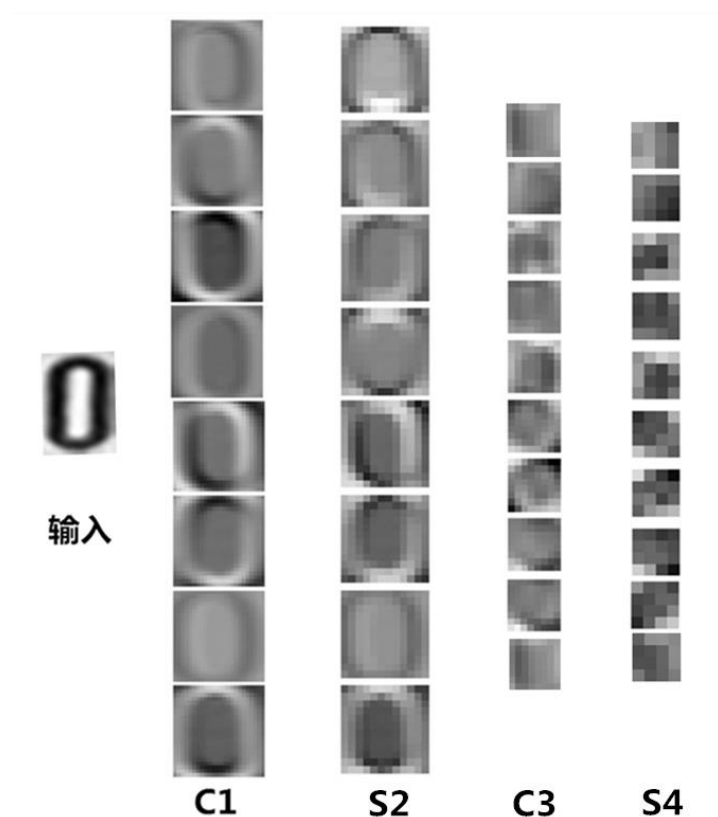


图 2-7 字符“0”在各层中输出的特征

2.6 实验结果与分析

2.6.1 深度学习参数选择

权值初始化是指在网络训练过程之前采用一定的策略对网络的权值和偏置进行初始化设置并使它们获得一个初始值。在网络训练过程中，一个初始权值的大小将直接影响网络的收敛速度和最后的分类性能，而一个理想的初始值可以使网络以较快的速度进行收敛并得到最优解。

根据目前研究权值初始化方法，可以将这些方法划分成两大类：非随机初始化和随机初始化。非随机初始化权值主要根据训练样本和网络结构本身的特点来选取相应的初始值。随机初始化权值一般在一个区间内随机选取一个较小的值。通常情况下，人们从均值为 0，方差为 0.1 的高斯分布中随机选取一个值赋给权值。较小的权值能保证敏感区有一定的宽度，但是网络的收敛速度变缓。较大的初始值，可以减少训练时间，但是如果与网络模型的最优解差距很大，那么网络训练速度变得很慢，甚至网络只能收敛到局部极值上，这样训练后的网络模型分类性能较差。Glorot 和 Bengio 在^[41]研究中使用了

Xavier 初始化的方式，信息可以再网络中均匀分布。何凯明等^[42]人在 Glorot 和 Bengio 研究的基础上提出了更适用于 ReLU 激活函数的初始化方法 MSRA，它是基于均值为零，方差为 2 除以输入的个数的高斯分布。本文选择偏置 b 的初始值为 0，权值 w 以 MSRA 方式随机初始化。

2.6.2 算法性能比较

我们收集了大量的车牌之后，从中分割出的字符作为本文的实验数据。主要采集了在结构和形状上相似并且识别率不高的字符，例如像“0”与“D”，“8”与“B”。如图所示，展示了部分的实验的字符图像，有的字符有噪声，有的字符过度分割，有的字符位置发生旋转。

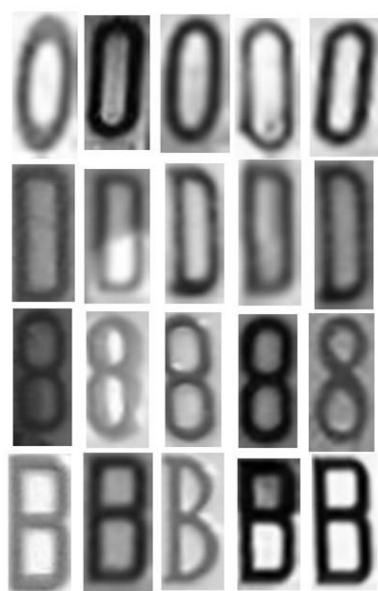


图 2-8 相似字符“0”、“D”、“8”和“B”

为了测试深度卷积神经网络在字符识别上的表现，本文选择支持向量机作为字符识别的分类器。由于从车牌中分割出来的字符中，获取的字符图像大小不一，因此需要对所有的字符进行预处理。相对于 SVM，首先将所有的样本转化为灰度图像并归一化。而对于卷积神经网络，统一设置所有字符的图像大小为 28×28 ，并且使字符的位置位于图像的中心。

我们使用了林智仁教授的 LIBSVM^[43]作为 SVM 的分类器，而卷积神经网络则使用图 2-5 所示网络结构。为了验证不同大小的训练集对分类结果的影响，本文对每组相似字符

各采用了 5 组实验，得到了如表 2-1 和表 2-2 所示的结果。

表 2-1 0 和 D 正确率对比

训练集	测试集	SVM	卷积神经网络
1000	300	95.3%	96.7%
3000	600	96.0%	97.8%
5000	1000	95.2%	98.9%
7000	1500	94.7%	97.5%
10000	2000	94.2%	98.5%

表 2-2 8 和 B 正确率对比

训练集	测试集	SVM	卷积神经网络
1000	300	97.0%	97.1%
3000	600	97.6%	98.5%
5000	1000	96.4%	97.8%
7000	1500	93.4%	98.7%
10000	2000	94.9%	98.9%

通过表 2-1 和表 2-2 的数据显示，当训练集的数量不断增加时，卷积神经网络的错误率下降明显，并且比 SVM 的识别率更高。这表明“深度模型+大数据”的组合，对卷积神经网络识别性能提升起了非常大作用，因为海量的数据能使卷积神经网络学习到更丰富的内在信息。而 SVM 作为两层的浅层模型，在提取特征过程中丢失了相似字符之间局部差异的特征。随着训练和测试样本数量逐渐地上升，SVM 的准确率明显降低，由于 SVM 在大数据下获取图像的特征出现了过拟合的情况，在第四组“8”和“B”的字符识别中，发现将许多“8”的字符被误识别成了“B”，但是对“B”字符的识别率较高，是由于两层结构的 SVM 模型在特征表示上的有限性。

为了验证深度卷积神经网络是否出现过拟合现象，我们选择了“0”和“D”中训练

集为 5000，测试集为 1000 的一组数据，训练误差随着实验迭代次数的上升而逐渐变小，最终该值稳定在可接受的范围内。

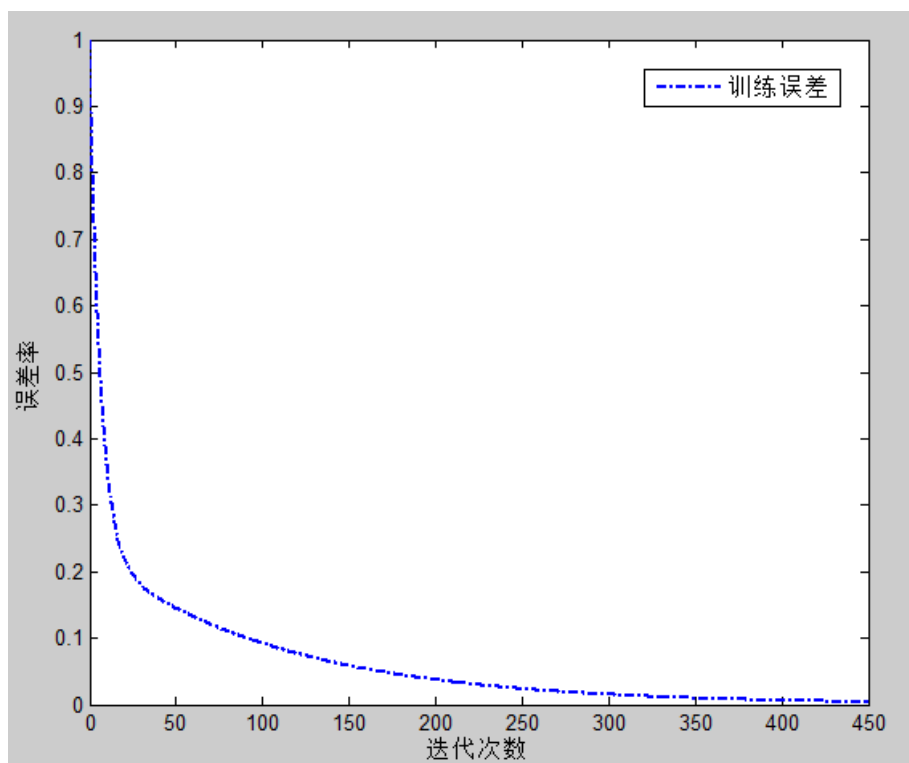


图 2-9 训练误差曲线

2.7 本章小结

本章将深度学习应用到相似字符识别中，构建了自动学习特征提取和识别的深度卷积神经网络。通过实验对比，利用卷积神经网络算法模拟人脑神经连接结构，使其在自动学习过程中能更好地获得有用特征，避免了传统方法人为设计特征的缺点，另外权重共享和下采样层一定程度上减少了训练时间。在面对平移、旋转、有噪声的字符图像时，深度卷积神经网络依然能取得较好的识别效果，体现了高准确率和良好的稳定性。随着训练样本数量不断地增加，深度卷积神经网的准确率逐渐提高，表明通过大数据训练，能学习到更丰富的内在特征信息。但是深度学习由于网络层次复杂，在识别和训练中需要花更多的时间，同时需要大量样本作为训练的支持。

第3章 集成深度学习网络的多层次车牌识别算法

3.1 引言

在上一章基础上，本章讨论如何通过支持向量机和深度学习网络完成一套完整的车牌识别系统，主要包括车牌定位、字符分割以及字符识别三个模块组成。其中定位模块包括对车牌目标候选区域的提取，然后对车牌真伪作进一步的判别和分析工作，最后输出分割后的车牌图像；车牌字符分割是对车牌定位中输出的车牌图像进行倾斜校正、二值化等预处理为字符分割提高了良好的条件，然后根据车牌字符的排列规则和几何特征对车牌图像进行分割，并输出字符图像；字符识别模块是系统中的最后一个环节，它的作用尤其重要，影响着整个系统的性能。因此本章对字符识别问题提出了多层次的车牌字符识别算法。它的主要思想是针对字符中存在结构相似的字符，通过 SVM 对字符初次识别，然后用深度学习算法对容易被 SVM 误识别的相似字符进行二次识别，最后组合输出车牌号码。这种方法不仅考虑了时间和效率，而且考虑了识别率和系统性能。通过实验结果表明该方法在综合考虑了时间和识别率上，对系统有一定的应用价值和研究意义。

3.2 车牌定位

3.2.1 国内车牌几何特点

我国车牌种类繁多且规则多样，主要根据最新颁布的《中华人民共和国机动车牌》安全标准（GA 36-2014）^[44]。但是生活中最常见的有蓝底白字和黄底黑字这两类车牌，并针对这两类车牌的单行车牌进行研究。

如图所示，这两类车牌的外廓尺寸均为 $440mm \times 140mm$ 。其中每个字符的大小为 $90mm \times 45mm$ ，分割符是一个直径为 $10mm$ 小圆点。除了第二个和第三个字符之间的距离较大以外，其它相邻字符之间的距离都是相等的，它们之间的距离均为 $12mm$ 。但是也存在有些字符比如字符“1”，没有占满整个字符区域的宽度，因此它与其他字符之间距离大于 $12mm$ 。

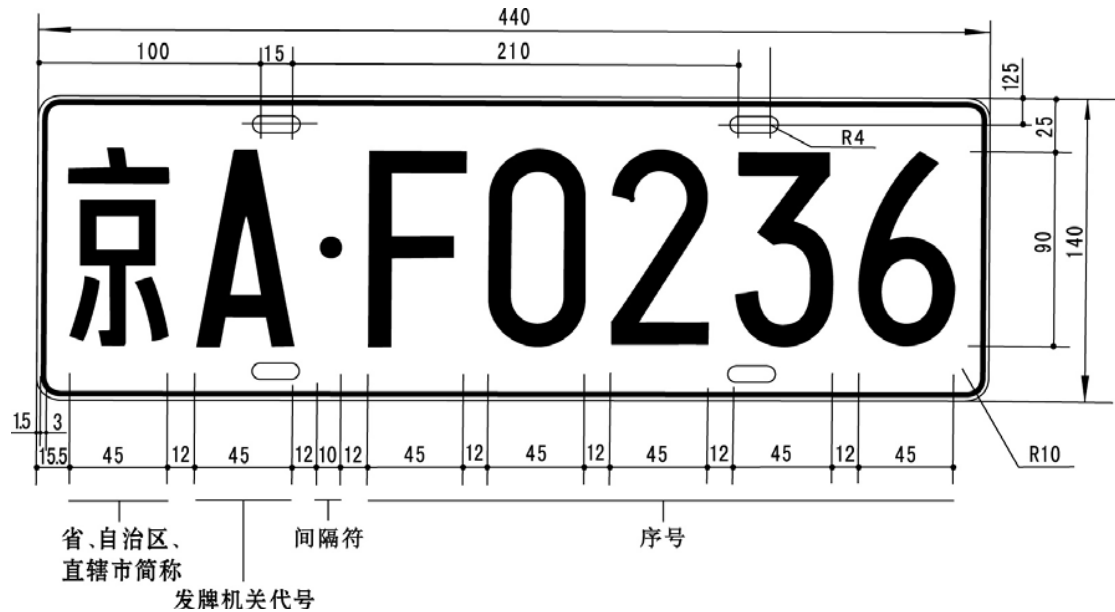


图 3-1 标准单行车牌

我国标准的单行车牌共有 7 个字符，包括汉字、大写字母和数字。其中第一个字是汉字代表各个省、自治区等代表地点，还有可能是“使”字。第二个字符为大写字母代表发牌机关代号。第三个到第六个字符是数字和大写字母的组合，第七个字符除了可能是大写字母和数字外，还可能警车“警”、教练车“学”等汉字。另外，因为大写字母“O”与数字“0”、字母“I”与数字“1”非常相似，为了避免结构太相似的字符干扰，所以车牌上将不使用字母“O”和“I”。

3.2.2 车牌图像预处理

在实际环境中，车牌图像会由于光照不佳、拍摄环境等原因导致质量下降，会对车牌识别精度造成影响，因此需要对图像进行预处理操作。预处理操作主要包括两块，首先把彩色图像转化为灰度图像，然后通过灰度拉伸算法对图像进行对比度增强，提高车牌质量。

1) 图像灰度化

图像灰度化处理是将彩色图像转化为灰度图像的映射过程。一张彩色图像包含亮度信息和色彩信息，并占用较大的计算机内存。通过图像灰度化处理能够大大减少后续图像处理的计算量，这样可以提高计算机的处理速度。彩色图像中的每个像素颜色由 R、G、B 三个分量决定，灰度化将 RGB 三个分量按一定方式计算产生了最大值法、平均值法和加权平均值法。加权平均值灰度化按照人的视觉特点来合成的，生成的灰度图像更符合

人眼的视觉感受，因此本文采用此灰度化方式。此时灰度值计算公式如下：

$$gray = r * w_r + g * w_g + b * w_b \quad (3-1)$$

其中 w_r w_g w_b 分别表示对 R、G、B 分量的权值，且满足 $w_r + w_g + w_b = 1$ 。常用的一组系数是 $w_r = 0.299, w_g = 0.588, w_b = 0.11$ 。

2) 灰度拉伸

车牌图像会由于各种原因存在太亮或太暗的情况，只采用灰度化处理并不能满足实际需求。为了增加图像的对比度，论文采用灰度拉伸的方式来突出图像中车牌区域。灰度拉伸可以通过分段线性变换函数来拉伸灰度图中一些重要的区域，也可以压缩一些不重要的区域。如果图像灰度值整体偏小导致图像较暗，我们可以拉伸这些图像的灰度区间；如果图像灰度值整体偏大导致图像较亮，我们可以压缩这些图像的灰度空间。通过对这些区域的拉伸或压缩处理，可以有效地提高图像的对比度。

$$g(x, y) = \begin{cases} c & 0 \leq f(x, y) < a \\ \frac{d-c}{b-a} f(x, y) + c & a \leq f(x, y) < b \\ d & b \leq f(x, y) < 255 \end{cases} \quad (3-2)$$



(a)



(b)



(c)

图 3-2 灰度拉伸效果 (a) 原图 (b) 灰度图像 (c) 灰度拉伸

灰度拉伸前后的图像如图 3-2 所示，可以发现经过灰度变化之后车牌区域更加清晰了。

3.2.3 车牌定位算法

车牌定位是车牌识别技术的首个环节，它的好坏直接影响系统的准确率。目前主流的定位算法有基于边缘特征的方法^[45]、基于模板匹配方法、基于颜色特征的方法^[46]等。不同的方法都有各自的优缺点，由于车辆环境的不确定性因素很多，如果单纯使用一种特征实现车牌定位具有一定局限性。其中基于边缘特征的方法受车牌污损、背景复杂等因素干扰，车牌很难被定位，基于颜色特征的方法受光照、车牌和车身同色等因素影响而准确率不高。本文在对以上方法进行分析后，考虑采用颜色特征和边缘特征作为检测车牌的依据，比单一特征定位更具鲁棒性。为了进一步提高车牌定位准确率，最后通过分类性能较优的支持向量机对候选区域进行车牌和非车牌区域的区分，以确保定位结果的准确性。下图是算法流程图：

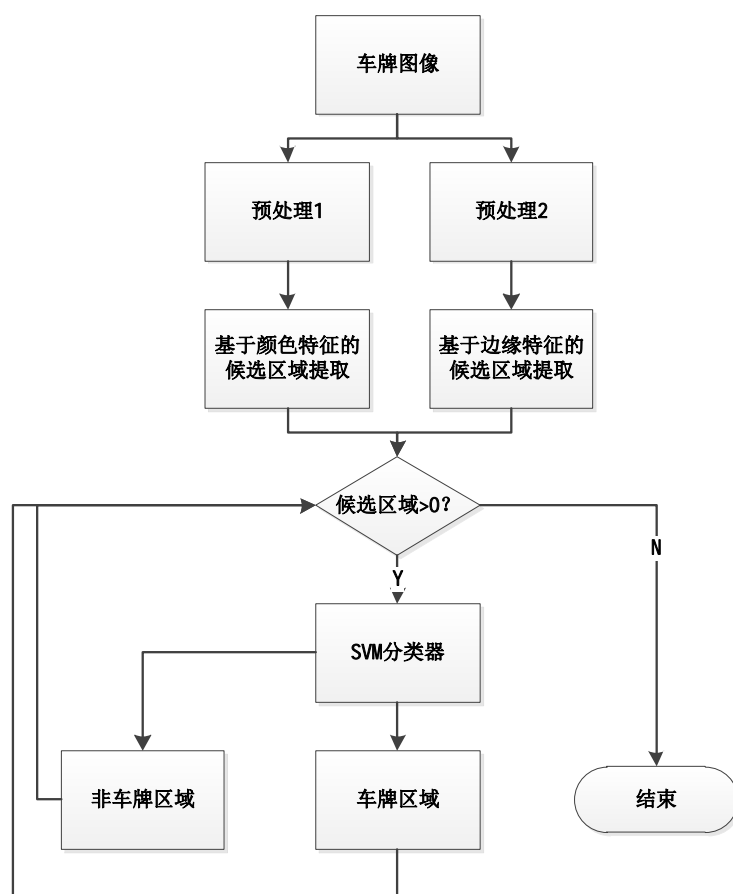


图 3-3 车牌定位流程图

为了能够检查出图像中的蓝色和黄色区域, 首先将彩色图像转化为 HSV 颜色空间的图像。根据 H (色度) 分量、S (饱和度) 分量对蓝色和黄色标记成前景并保持颜色不变, 其他区域设置成黑色背景。通过前景和背景的设定, 车牌蓝黄区域就被标记出来了。如下图所示:

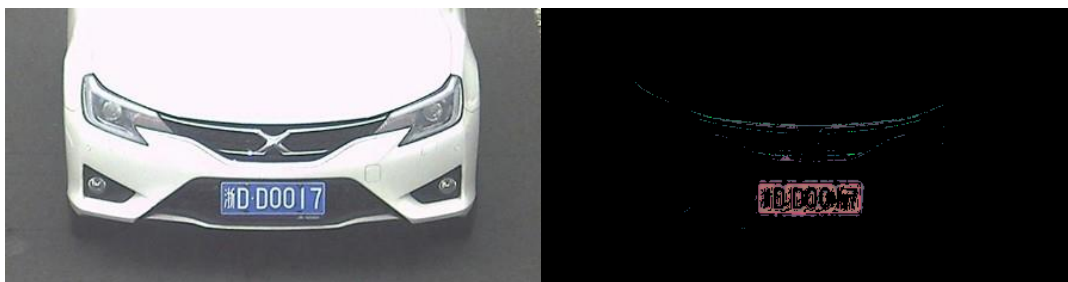


图 3-4 根据 H 和 S 分量处理后的图像

再用基于边缘特征的候选区域提取, 将彩色图像进行灰度化和灰度拉伸。为了消除图像的背景噪声干扰并突出车牌信息, 又对图像进行了中值滤波。经过这些处理后, 采用 Sobel 算子检测垂直方向的边缘特征。然后用闭运算来填充图像边缘之间的横向, 最后通过轮廓求最小外接矩形进行筛选, 清除较小矩形区域和倾斜角较大矩形区域。



图 3-5 基于边缘特征处理后的图像

上述基于颜色特征和边缘特征提取的候选区域中含有非车牌的区域, 因此本文利用了 SVM 分类器对于线性可分的二分类问题具有良好的泛化能力, 对这些区域进行真伪的判别。它的原理如下:

设样本集为 $(x_i, y_i), i = 1, 2, \dots, n$ $x_i \in R^N$ $y_i \in \{+1, -1\}$, 共有两个类别组成。如果 x_i 属于第一类, 则为 $y_i = +1$, 如果 x_i 属于第二类, 则为 $y_i = -1$ 。N 维空间中线性判别式和分类面方程分别为:

$$g(x) = w \bullet x + b \quad (3-3)$$

$$w \bullet x + b = 0 \quad (3-4)$$

并且对于任何样本 (x_i, y_i) 都满足

$$\begin{cases} w \bullet x + b \geq +1 & y = 1 \\ w \bullet x + b \leq -1 & y = -1 \end{cases} \quad i = 1, 2, \dots, n \quad (3-5)$$

将判别函数归一化,使两类样本都满足 $|g(x)| \geq 1$ 且到分类超平面的距离为 $d = |w \bullet x + b| / \|w\|$,

其中 $\|w\| = \sqrt{w^T \cdot w}$, 如果所有的样本都满足 $|g(x)| \geq 1$, 并且存在最近的样本 $g(x) = \pm 1$, 则超平面的分类间隔为 $\frac{1}{2}\|w\|$ 。如果存在使分类间隔最大, 则该平面为最优分类超平面。寻找最优分类超平面问题可以看成是二次规划的问题, 即求函数的最小值:

$$\min(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w \bullet w) \quad (3-6)$$

$$y(w \bullet x_i + b) \geq 1, i = 1, 2, \dots, n \quad (3-7)$$

利用 **Lagrange** 优化方法将最优分类面问题转换为对偶最优化问题, 形式如下:

$$\begin{cases} \max W(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=0}^n a_i a_j y_i y_j (x_i \bullet x_j) \\ s.t. \quad \sum y_i a_i = 0 \\ a_i \geq 0 \quad i = 1, \dots, n \end{cases} \quad (3-8)$$

其中 a_i 为每个样本对应的 **Lagrange** 乘子。根据 **Karush-Kuhn-Tucher** 互补条件^{[47][48]}可知,

最优解满足:

$$a_i (y_i (w \bullet x_i + b) - 1) = 0, \quad i = 1, 2, \dots, n \quad (3-9)$$

显然, 对于多数样本支持向量机的系数 a_i 为 0, 少数不为 0 的样本就是支持向量。最后得到最优分类判别函数:

$$f(x) = \text{sgn}\{ (w \bullet x) + b \} = \text{sgn}\left\{ \sum_{i=1}^n a_i * y_i (x_i \bullet x) + b^* \right\} \quad (3-10)$$

其中 x_i 代表支持向量, x 代表待测试样本。而 b^* 是个分类阈值, 可以通过任意一对支持向量用式 (3-7) 求得。

如果遇到线性不可分的问题，可以增加一个松弛项 ξ_i ，得到如下

$$y(w \bullet x + b) - 1 + \xi_i \geq 0 \quad (3-11)$$

或者通过一个非线性函数 $\phi(x)$ ，将训练集映射到一个高维特征空间。

$$w \bullet \phi(x) + b = 0 \quad (3-12)$$

本文提出了一种基于颜色和边缘的车牌定位以及 SVM 的车牌真伪判别方法。利用数字图像处理快的特点对图像进行候选区域的提取，减少了时间花费，多种特征结合则避免了其他因素影响漏选了一些区域，再根据一些先验知识如车牌的大小、位置、长宽比等，将重叠的候选区域和非车牌区域进行一系列的分析 and 筛选。然后利用 SVM 方法在小样本学习的情况下具有较好的泛化能力，并通过 SVM 分类器对候选区域的识别，大大提高了识别的准确性，并最终确定车牌的位置。

3.3 车牌字符分割

字符分割是车牌识别系统的中间环节，它主要的任务是将分割后的 7 个字符图像送入字符识别模块进行识别，因此字符分割的好坏将影响字符识别的结果。标准单行车牌格式为： $X_1X_2 \bullet X_3X_4X_5X_6X_7$ ，其中在 X_2 与 X_3 之间有一小圆点， X_1 为汉字， X_2 是英文字母， $X_3X_4X_5X_6X_7$ 是英文字母或阿拉伯数字。由于拍摄角度、外部环境等因素干扰，使得提取到的车牌可能出现倾斜、字符污损等现象，因此需要对车牌目标图像进行预处理操作。比如在字符分割前判断车牌是否倾斜，如果发生了倾斜那么进行必要的校正处理。而直接对图像进行分割的话容易出现字符粘连甚至会导致字符分割失败，不利于后续模块进行识别工作。如图 3-6 所示是本文的分割算法：

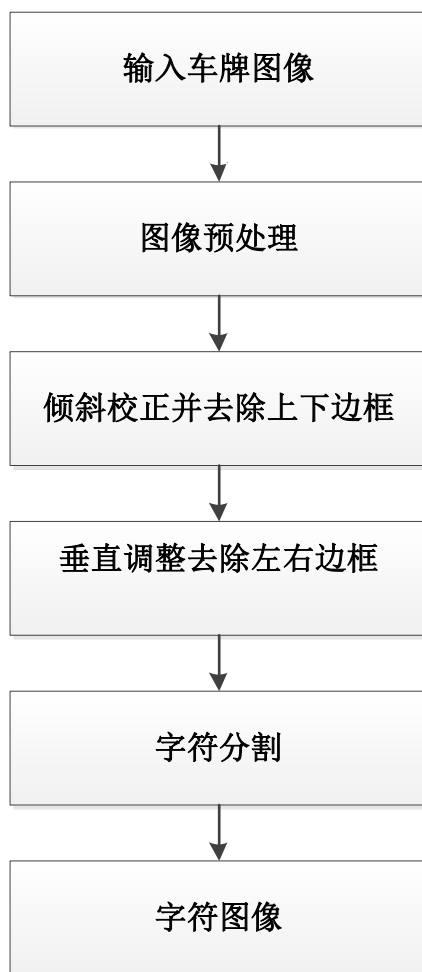


图 3-6 字符分割流程图

字符分割输入的图像是车牌定位输出的彩色图像，因此除了倾斜校正外，还进行了其他的预处理操作，因为一个好的预处理算法有利于提高分割的准确率^[49]。首先对车牌种类判别，本文主要针对蓝色和黄色车牌进行识别，由于这两种车牌经过二值化后会得到不同结果：黄色车牌得到白底黑字的图像，而蓝色车牌得到黑底白字的图像。所以本文利用车牌定位中颜色特征信息对黄色车牌进行反色处理。然后对车牌进行二值化、去噪、目标增强等处理以突出字符的像素点。

本文利用 Hough 变换^[50]实现图像的倾斜校正，该方法将图像直线上的任意一点 (x_0, y_0) 根据变换公式 $r = x \cdot \cos\theta + y \cdot \sin\theta$ 转换到极坐标空间上的一条曲线。由转换公式可知所有曲线必经过极坐标的同一点 (θ_0, r_0) ，于是 Hough 变换可以在二值图像中检测直线。当车牌完成定位后，利用 Hough 找出车牌的倾斜角度。考虑到其他因素的干扰，本文寻到车牌中最长的一条直线作为车牌的倾斜角度并予以校正。



图 3-7 车牌校正前后图像

二值化在车牌识别中有着重要的意义，无论在定位中还是在字符分割之前。因此在得到倾斜校正的车牌后，本文将使用 Otsu^[51]提出的最大类间方法对车牌进行二值化，以达到凸显字符并去除车牌噪声的目的。



图 3-8 车牌二值化图像

从矫正后的二值化图中可以明显看出，车牌的像素主要集中在车牌字符区域，其他区域则比较分散且变化小。所以对二值化车牌做水平投影，然后由上而下对水平投影扫描，查找离上边框最近的谷底，并记入该方向多行的跳变次数和像素信息。如果有多行的信息满足设定的阈值，并根据上宽与图像总宽比值认为已经找到了字符边缘区域，将该位置作为上边框的水平分割线。同理对下边做由上而下的扫描，去除下边框。如图3-9所示为去除上下边框后的二值图像。



图 3-9 去上下边框后图像

由去除上下边框的车牌可以确定字符的高度记做 `char_height` 根据车牌的背景知识得知车牌具有一定比例，字符也有固定的长宽比。然后对七个字符按照车牌中排序规则，得到车牌的宽度，这样由上步求得高度和车牌长宽比例可以估算目标车牌的宽度，即为 `plate_width` 对二值化车牌进行垂直投影寻找图像垂直投影的谷底，作为候选垂直分割线

的位置。根据投影信息、估算的宽度以及先验的知识，去除垂直的边框和确定第二和第三个字符之间最大分割间隔的位置。



图 3-10 去左右表框后图像

最后字符分割的过程主要是根据车牌的字符之间的距离和字符的宽高比特征进行分割。根据最大间隔位置和利用先验知识进一步验证垂直分割线位置，将字符图像分割成左右两个区域。左边两个字符区域分割的方法相对简单些，取左边区域投影长度的中间值作为分割点，再根据字符的宽高比验证字符的宽度是否符合正常字符的宽度，满足要求就作为分割点。以下是右边区域五个字符分割的流程：

（1）从左向右逐个计算出垂直投影块的数目，直到图像到达最右端为止，并且记入每个投影块的起始位置和终点位置为 $S[n]$ 和 $E[n]$ ， $n=1,2,\dots,n$ ， n 为分割得到的区域数。

（2）如果投影块数 $n>5$ ，那么计算每个相邻投影块中线之间的距离 L 。如果 L 小于设定的最小字符宽度，那么将这两个字符块投影进行合并，并且投影块个数减少一个。继续重复步骤（1），直到投影块个数等于 5。

（3）如果字符投影块个数小于 5，那么在投影块中查找投影宽度最大投影进行第二次分割。首先找到该区域的起始位置和结束位置，然后在该投影的中间区域查找最小的投影值。将最小值点视为分割点对该区域进行分割，并且投影块个数加上一个。继续重复步骤（1），同时每个投影块的宽度均满足字符的宽度阈值。

至此完成七个字符的分割，图 3-11 为字符分割的结果。



图 3-11 字符分割结果

3.4 多层次车牌识别

字符识别是车牌识别系统的最后一个环节，相比于车牌定位和字符分割，它对可靠性和准确性的要求是最高的。基于 SVM 的识别算法速度较快，但是在环境复杂和光照变化的情况下，识别率不高，主要是它对相似字符如 ‘8’ 和 ‘B’，‘0’ 和 ‘D’ 等存在误

识别问题。深度学习的识别率高，但是识别时间长。如果能将两种算法相结合，便可以找出一种非常优秀的算法。本文第二章基于深度学习的相似字符识别方法研究，就是为了提高对字符识别的整体准确率。因此本章提出了 **SVM** 多分类器对字符的第一次分类。又考虑 **SVM** 分类器对相似字符识别存在误识别的问题，因此在能够提高准确率和不影响系统性能的前提下，本文结合了大数据集和深度学习，建立了一个相似字符训练集，通过卷积神经网络对相似字符的二次识别。

根据车牌字符的种类和从左到右的出现次序，对基于 **SVM** 的字符识别方法设计了三组分类器，并以一定的准则来训练样本从而实现分类识别车牌字符的目的。该算法的输入是切分后的单个字符对 **SVM** 分类器进行识别，如果遇到相似字符那么就通过卷积神经网络进行识别，最后将输出车牌号码。字符识别模块的流程图如下所示：

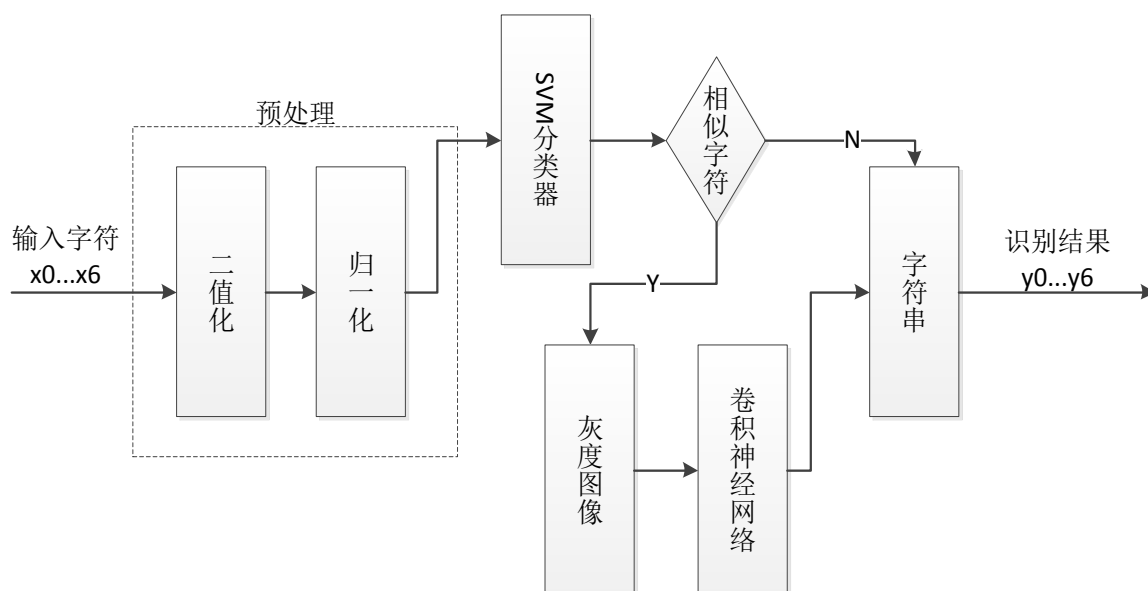


图 3-12 多层次字符识别算法

3.4.1 字符归一化和二值化

为了保证分类器的准确率并减少识别难度，因此需要对输入字符图像做预处理，使字符在识别之前具有相同的格式，减少非人为因素对字符识别造成的影响。首先对字符进行归一化操作，即对图像进行缩放，使所有字符图像的尺寸大小都为 28×28 。本文采用支持向量机对字符图像进行第一层识别，需要对字符图像进行二值化处理。目前比较流行图像缩放方法有最邻近插值法、双线性插值法、三次内插法等。邻近插值法计算量最小，处理速度最快，但是插值的效果较差。因为完成插值后的图像在灰度上会出现不

连续或者在灰度变化区域会出现明显的锯齿。三次内插法效果最好，但是计算量最大，对于实时性要求高的系统实用性低。双线性插值法的加权平均算法能改善锯齿效应并且算法复杂度适中，所以对于放大倍数较小且质量不高的图像使用双线性插值法。

$$g(x, y) = f(x', y') = f[a(x, y), b(x, y)] \quad (3-13)$$

其中 $f(x, y)$ 表示输入图像， $g(x, y)$ 表示输出图像。函数 $a(x, y)$ 和 $b(x, y)$ 分别描述了 x 到 x' 和 y 到 y' 的空间映射。

图像的二值化处理的目的是寻找一个合适的阈值 t 将目标和背景分离，这样有利于字符的识别。若输入的灰度图像为 $f(x, y)$ ，输出的二值化图像为 $g(x, y)$ ，将图像中像素分割成大于阈值 t 和小于阈值 t 两部分，二值化的过程表示如下：

$$g(x, y) = \begin{cases} 1 & f(x, y) \geq t \\ 0 & f(x, y) < t \end{cases} \quad (3-14)$$

其中 $g(x, y) = 1$ 表示目标，反之则表示背景。因此二值化的关键是寻找一个合适的阈值。本文采用最大化阈值法 (Otsu) ^[51] 对图像进行二值化处理，它是典型的全局阈值法，对阈值是自动获取的。其基本思想是选取一个初始化阈值 t 对图像的灰度值进行分割，于是图像像素值被分成了大于等于 t 的部分 **C1** 和小于 t 的部分 **C2**，然后计算 **C1** 和 **C2** 的类间方差 σ_B^2 ，其公式如下：

$$\sigma_B^2 = W_1 W_2 (M_1 - M_2)^2 \quad (3-15)$$

其中类 **C1** 的像素点总数为 W_1 ，类 **C2** 的像素点总数为 W_2 ，然后计算得到它们的平均灰度值分别表示 M_1 和 M_2 。它们的类内方差 σ_A^2 ，公式如下：

$$\sigma_A^2 = W_1 \sigma_1^2 + W_2 \sigma_2^2 \quad (3-16)$$

最后计算两个方差比 σ_B^2 / σ_A^2 ，找出图像最大的阈值 t 对图像进行二值化。

部分归一化和二值化的图像如图 3-13 所示。可以发现，对于大部分图像，二值化后的图像特征轮廓比较明显。但是对于部分相似字符，例如第一行的 0 和 D，具有非常相似的外部轮廓。



图 3-13 部分训练样本

3.4.2 支持向量机粗识别

通常支持向量机是用于二类别的分类，而在实际使用中经常会遇到多分类的问题，因此需要将二分类方法扩展成多分类方法。目前存在以下的四种方法：一对一、一对多、支持向量机决策树法以及多类支持向量机^[52]。其中“一对多”方法相对比较简单并且容易实现，所以本文采用一对多的方法来构造多分类器对车牌字符进行识别。

一对多的方法即把若干个二分类器合在一起变成一个多分类器，如图 3 所示。构造一个多分类器的基本过程如下：假如一个多分类问题共有 n 个类别， $m=\{1,2,\dots,n\}$ ，训练样本 $(x_i, y_i), i=1,2,\dots,n$ ，其中 $y_i \in m$ 。一对多方法就是用 SVM 构造出 n 个子分类器。对于第 i 个子分类器时，把属于第 i 类的数据记作正类，不属于第 i 类的数据记作负类，然后进行训练。

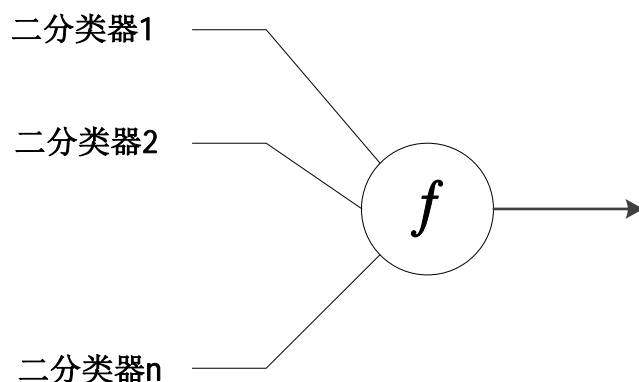


图 3-14 一对多分类器结构

首先建立样本的训练集，将样本归一化后转换成 N 维向量，如果训练集的大小为 M 个，那么生成 $M \times N$ 维的训练样本矩阵，并对应一个 $M \times 1$ 维的标签矩阵。在训练之前选择合适的核函数 K 。为了区分待测样本和其他样本，将待测样本的标签设置成+1，剩

下所有样本的标准设置成-1。把训练集样本矩阵以及相应的标签矩阵输入，在训练好一个子分类器后，设置新的标签矩阵再训练下一个子分类器。重复上述步骤，直到完成所有子分类器的训练。进行测试样本识别时，将样本送入分类器，循环检测所有分类器的输出，若检测到的输出是正的，则判定该样本为该类相对应的样本类型。如果遇到多个分类器输出正值，那么比较该样本到各分类器所确定超平面的距离，选择距离最大的分类器作为该样本的分类器。

根据最新制定机动车号码序号编码规则，本文将分类器定义为三组：汉字分类器、字母分类器、英文字母数字分类器。汉字字符共有 34 个，因此设计了 34 个二分类器。类似的，对于数字，构造了 10 个二分类器。但是对于字母数字字符构造了 34 个二分类器，因为 26 个字母字符中字符“l”和“O”不能使用。

3.4.3 卷积神经网络相似字符识别

为了提高相似字符的识别率，本文采用卷积神经网络针对相似字符的精识别。本文通过上述分类方法先将类似字符归为一类，再采用卷积神经网络相似字符进行二次识别。本类相似字符归类如表所示：

表 3-1 相似字符集

编号	相似字符集
1	8B
2	0D
3	2Z

由本文第一章可知，深度学习算法是机器学习的一个新分支，它是集特征提取与分类判决于一体的数据驱动算法。结合实际情况加以改进实现本文的字符识别算法，提出适用于本文算法输入、输出条件的改进的网络模型。考虑到车牌字符在提取过程中受环境因素较大，因此车牌字符识别比手写字符的识别更大。与第二章研究相似字符相同，本章对相似字符进行了训练。首先收集了大量的车牌图片，共 6 个相似字符。在深度学习模型基础上完成分类工作，模型的第一层的输入大小为 28×28 单个字符图片，为了训练后模型的稳定性，大约共收集 18000 张字符图片，平均每个类别有 3000 个左右，分别给每个相似字符类别绑定标签，每每两个一组送入 Caffe 框架中进行有监督训练，设置

训练的迭代次数为 10000 次。学习率为 0.001，调优的学习速率为 0.002 以及选用的权重衰减因子为 0.0002

在对传统卷积神经网络的研究基础上，本文的池化方式采用最大池化(max-pooling)，它能达到降采样和局部位移不变性的目的外同时还能保留更多的纹理信息。将卷积层与下采样层之间的激活函数由 sigmoid 改为 ReLU，由文献[38]可知 ReLU 具有备单侧抑制的能力，使网络具有稀疏性，因此分类的效果比使用其他激活函数更好。以及将 F5 层与输出层间连接由径向基函数分类器 (RBF) 改为 Softmax 回归分类器。

3.5 实验结果与分析

3.5.1 车牌区域训练数据预处理

注意到论文采用支持向量机提高车牌定位的准确率，因此需要构造训练数据用于真假车牌的判断。因此搜集了大量车牌与非车牌两类图像用于训练 SVM 分类器。以各式各样的国内车牌作为训练样本。对训练样本，选取了各 1000 幅车辆车牌图像和非车牌图像，其中车牌中包含汉字、数字和字母。为了提高 SVM 分类器的分类性能，对所有车牌样本进行归一化处理：将图像进行缩放，并统一设置成大小为 180×60 ，经过缩放后的样本作为车牌的训练样本。同样将非车牌样本也设置成大小为 180×60 子块。经过预处理后，如果直接输入车牌图像作为特征不利于 SVM 训练，而且分类性能也一般。因此本文选择 LBP (Local Binary Pattern) 算子描述车牌和非车牌的纹理信息，它具有旋转不变性和灰度不变性等显著的优点。通过对所有正负样本车牌区域提取 LBP 特征，并将它们组成一个特征矩阵作为输入的训练样本。对应的要有一个指定每个特征向量的标签的向量。将车牌区域的标签设为 +1，非车牌区域标签设为 -1。如图 3 所示为部分训练样本。然后根据其中的支持向量和参数生成，最后将训练好的 SVM 分类器保存为 XML 格式的特征模型。



(a)



(b)

图 3-15 部分训练样本 (a)正样本; (b)负样本

3.5.2 实验结果比较与分析

算法实现工具是 Matlab R2014a和 Visual Studio 2012 实验运行的电脑环境是 Intel (R) Pentium(R) CPU G2030 4.00GB 内存。

为了测试卷积神经网络的性能, 本文以 0 为例, 对训练好的网络模型进行测试能输出第二层和第三层的特征图, 如下图所示:

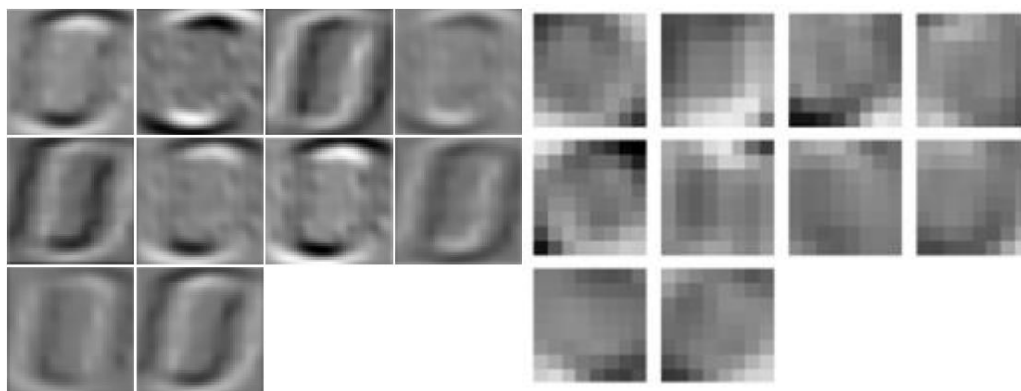


图 3-16 第一层和第三层特征图

特征图可以看出，网络的第一层主要提取的是字符的低层次特征，如角点、边缘和轮廓等。下采样层对于前一层的卷积层几乎没有有效信息的损失，但是特征维度却缩小了 4 倍，大大提高了整个网络的效率。随着网络层数的增加，其特征图随着网络层数的增加，其特征图描绘了更多的图像细节信息，网络最后送入全连接的层能够提取丰富的图像特征，使得最后能够得到较高的准确率。

本文选取 500 个车牌作为本实验的测试样本分别对 SVM、SVM+卷积神经网络进行了测试，测试结果如下表所示：

表 3-2 后 5 位车牌字符

不同方法	字符准确率	平均时间 (ms)
卷积神经网络	97.2%	183.2
SVM	95%	14.1
SVM+卷积神经网络	97%	50.87

由表可以发现，本文基于多层次的识别算法在识别准确率方面比仅使用 SVM 分类器提高了 2%左右。在时间性能方面，由于卷积神经网络是个深层的网络结构，数据输入之后需要按层顺序进行特征提取，每一层特征映射需要经过卷积以及激活函数两个运算，因此比浅层网络花费更多的时间。

观察分析发现，SVM 识别效果比 SVM+卷积神经网络差主要是相似字符的误识别，如字符“0”和“D”以及“8”和“B”等。本章对 SVM 容易误识别的问题，通过设置了卷积神经网络对字符再识别，字符的准确率有了明显提高，但是在时间运行上比 SVM

增加了，但是比仅使用卷积神经网络的时间少很多。经统计相似字符占车牌字符的比例不高，因此采用卷积神经网络进行二次识别，不仅保证了识别的准确率，还保证系统的整体性能。

分析原因，主要有以下几点：

第一，卷积神经网络在字符识别上取得较好的效果依赖于它的深层网络结构，层层相连的网络结构能够表示更复杂的数据。在训练样本充足的情况下，能够得到较好的训练模型，因此比浅层网络结构具有更优异的分类性能。**SVM** 是一种分类性能较优的方法，只需使用有限少量的训练样本就能得到较好的分类器，由于 **SVM** 是一种两层的浅层网络结构，在区分相似字符时表达能力弱。

第二，**SVM** 在运算效率方面较优，因为 **SVM** 属于两层模型，第一层是输入数据与支持向量的内积，输出是上一层节点的线性组合，所以识别快慢很大程度上取决于支持向量的数量。卷积神经网络使用的卷积层具有很强的特征提取能力，采用有监督的训练方法，通过反向传播使网络模型到达最优的收敛值，因此它能更好的拟合样本的分布，获得较好的识别效果。

3.6 本章小结

本章首先介绍国内车牌的几何特点，选择了标准单行蓝色和黄色车牌作为研究对象。采用基于颜色特征和边缘特征的方法最大限度地寻找车牌的候选区域，再用 **SVM** 对车牌和非车牌区别的车牌定位方法。字符分割使用水平投影和垂直投影的方法，并利用先验知识完成分割工作。本章重点在研究字符识别性能上的提高。根据深度学习的特点和优势，设计了多层次的识别方法，利用 **SVM** 对字符进行初次识别，再利用卷积神经网络对相似进行二次识别。算法经过测试表明，**SVM** 和卷积神经网络结合的识别具有较强的识别性能，并且花费的时间也不多。

第 4 章 车牌识别系统设计与实现

4.1 系统开发环境

在系统开发环境方面，考虑到车牌识别系统是一个较为复杂的图像处理程序，它包括了车辆图像的预处理、车牌的定位、车牌倾斜校正、字符分割、字符图像预处理和字符识别等一系列图像处理和图像识别的操作，同时众多的图像处理过程需要消耗大量的内存且图像处理函数的使用会占用 CPU 处理时间。因此对系统的稳定性和可靠性要求高。因为图像数据量较大，图像处理与分析对速度有较高要求，而 C++ 语言作为开发图像处理的最佳语言，同时还有很多开源的图像处理库，如 CxImage OpenCV 等，为系统开发提供许多图像处理函数，节省了开发时间，并且提高了开发效率。因此选择面向对象的编程语言 C++。

在开发平台上，采用了 Microsoft 公司开发的 Visual Studio 2012 集成开发环境，Qt 作为跨平台的 C++ 图形用户应用程序开发框架，已经支持 VS 环境开发的插件，并且 Qt 是一个极容易上手的 GUI 面向对象的框架。它提供应用程序开发者建立图形界面所需的所有功能，并且初学者能迅速上手和操作。因此代码开发难度相对较小、开发工作量也相对较少，减少了不必要的开发时间，提高了开发效率。

4.2 系统设计与实现

本系统主要有车牌定位、字符分割和字符识别三个功能模块组成。为了能使各模块相互独立，后期更容易维护，因此我们对这三个功能模块进行了单独的设计和实现。

4.2.1 系统设计

该系统的整体设计过程为：首先，由高清摄像机采集到车辆图像，对图像进行一系列预处理；其次，采用彩色模型和边缘特征检测对车牌候选区域的提取，再利用 SVM 对车牌候选区域进行判别；校正环节采取一种基于 Hough 倾斜车牌校正方法并去除车牌四周的边框；字符分割环节根据车牌字符先验知识把水平分割和竖直投影相结合；最后，结合 SVM 和卷积神经网络字符识别方法对字符识别，先利用 SVM 快速准确识别，再利用卷积神经网络对相似字符进行二次识别。

在整个系统的识别过程中，为了满足系统的实时性和准确性要求，因此本文在细节

上进行了修改和提升。比如：利用数字图像处理速度快的优势，根据车牌的颜色特征和边缘特征以及几何特征对车牌区域进行提取，避免了对整幅图像的扫描和检测。为了避免对所有车牌进行倾斜校正，在获取倾斜角度之后根据校正规则判断车牌是否需要作校正处理。在二值化过程中，采用最大化阈值法（Otsu）寻找最合适的阈值 t 分离目标和背景，在二值化后可以明显看到字符与背景分离；在字符分割过程，增加了对粘连字符和分裂字符的判断，分割后的结果有一定程度的提升。本文车牌识别系统的整体算法设计流程如图 4-1 所示：

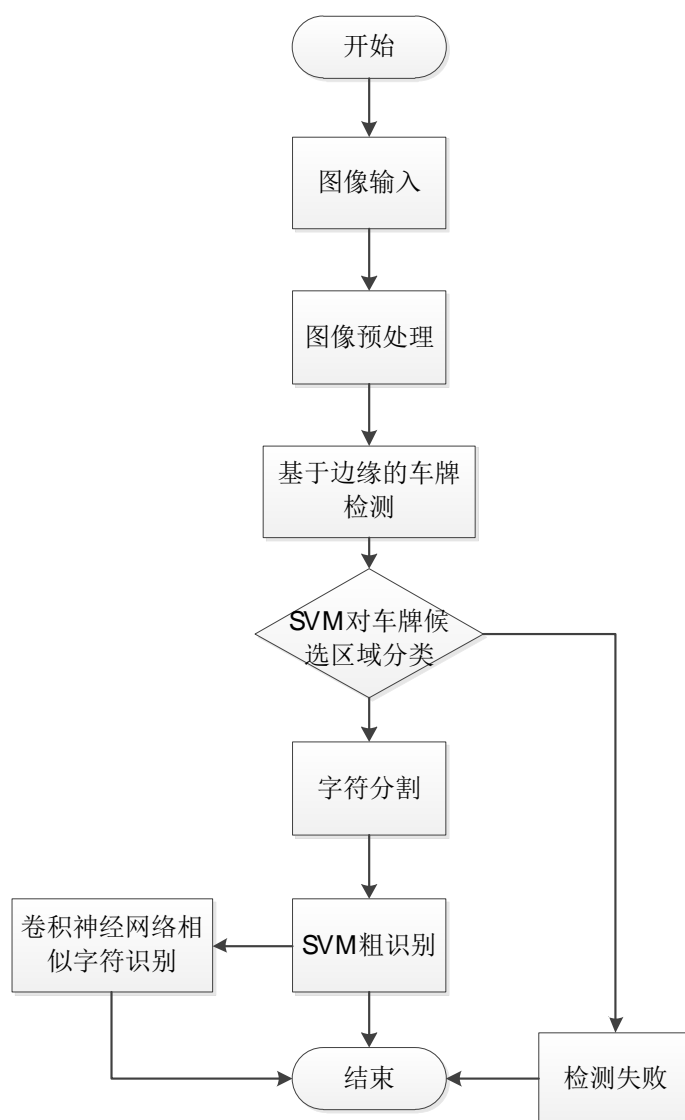


图 4-1 车牌识别系统算法流程图

4.2.2 关键数据结构

根据车牌的设计规则和几何特征，本文对车牌识别中需要用到的车牌信息和字符信

息作了定义，主要定义了车牌大小特征和最大倾斜角度、字符的大小和位置。为了方便用户对参数的修改和程序读取，将这些信息定义在配置文件中。

首先定义了车牌的数据结构，能够在车牌定位中去除倾斜角度较大和矩形区域较小的候选区域。

表 4-1 车牌信息

plate_width	车牌的宽度
plate_height	车牌的高度
min_plate_size_width	车牌的最小宽度
min_plate_size_height	车牌的最小高度
max_plate_angle	车牌最大倾斜角度

然后定义了字符的数据结构，主要定义了字符的大小、位置等信息，能够为字符分割提供必要的先验信息，有利于字符的分割。并对分割后的字符进行归一化处理，有利于提高字符识别的准确率。

表 4-2 字符信息

char_height	字符的高度
char_width	字符的宽度
char_whitespace_top	字符离车牌上边框距离
char_whitespace_bot	字符离车牌下边框距离
min_characters_num	分割后字符最少个数
max_characters_num	分割后字符最多个数
seg_char_height	归一化后字符高度
seg_char_width	归一化后字符宽度

车牌识别系统的最关键在于对 **Plate**类和 **Character**类的这两个类实现，具体的类图如图 4-2 和图 4-3 所示：

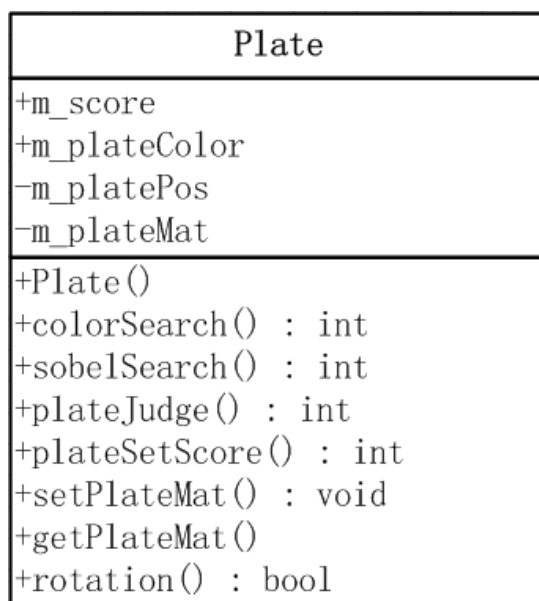


图 4-2 车牌定位类图

Plate类用于车牌定位，包含使用颜色特征的 **colorSearch()** 函数和使用 **sobel** 边缘特征的 **sobelSearch()** 函数对车牌候选区域检测。**plateJudge()** 函数则是先提取图像的 **LBP** 特征，再通过 **SVM** 分类器对车牌候选区域进行分类。另外还有车牌倾斜矫正、图像读取和设置等操作的成员函数。

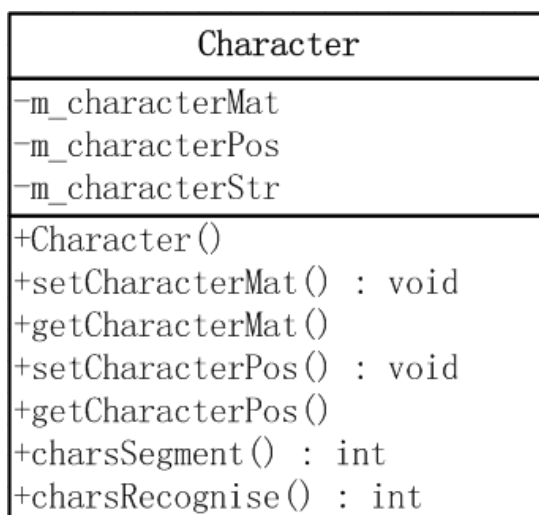


图 4-3 字符分割和字符识别类图

Character 类用于字符分割和字符识别。车牌图像首先通过 **charsSgment()** 函数获得分割后的字符图像 **m_characterMat** 和位置 **m_characterPos**，然后利用 **SVM** 和卷积神经网络对字符进行识别，并将识别结果保存在 **m_characterStr** 中。

4.2.3 核心算法动态链接库设计

动态链接库 (DLL, Dynamic Link Library) 是一种 Windows 操作系统中实现不同应用程序共享函数库的技术, 其中封装了一些可执行的函数和资源。当其他程序需要调用动态链接库时, 那么相应的 DLL 文件被载入和链接。因此一个动态链接库文件里的函数和资源可以被多个应用程序所共享。使用动态链接库有很多优点: 当多个进程同时使用同一个动态链接库, 内存中共享该动态链接库, 节省了内存开销。若有多个应用程序共享一个动态链接库, 可以节省磁盘空间。当动态链接库中的函数发生改变时, 不需要重新编译或者建立与该 DLL 文件进行链接。使用动态链接库可以实现不同编程语言编写的程序调用同个的 DLL 文件, 只要调用方式遵循函数的规则, 这样有利于模块化程序的开发。使用动态链接库还可以用于一些特殊的目的, 隐藏函数实现的细节。

为了方便系统的开发, 本文将一些关键代码封装成动态链接库文件 (LicensePlate.dll, 并开放了车牌识别核心接口, 包括车牌的定位、分割和字符识别。

以下是分别是车牌定位、字符分割和字符识别函数定义和参数说明:

`plateLocate(cv::Mat image, vector<Mat> &resultVec, int numb)`

主要参数说明:

`Image` 输入车辆图像

`resultVec` 输出车牌图像

`numb` 车牌个数

`charsSegment(cv::Mat image, vector<Mat> & resultVec, Color color)`

主要参数说明:

`Image` 输入车牌图像

`resultVec` 输出分割字符

`color:` 车牌颜色

`CharsRecognize(std::vector<Mat> regionsOfCharacter, string result)`

主要参数说明:

`regionsOfCharacter` 输入字符图像

`result` 输出识别结果

指在应用程序中用 `LoadLibrary` 函数来显示的方式加载, 得到动态链接库的模块句柄, 然后从句柄中获得所需调用函数的地址, 如下图代码所示:

```
HINSTANCE hinstDLL=NULL;
hinstDLL=LoadLibrary("LPR.dll");
CharsRecognize(connect)GetPmcAddressminstDLL" CharsRecognize");
最后将参数传入到函数, 得到字符识别的结果。
CharsRecognize(image,resultVecnumb);
FreeLibrary(hinstDLL) //卸载 dll
```

4.2.4 车牌识别展示原型系统

如下图所示为车牌识别系统的软件界面, 该界面是基于 Qt 框架编写, 左侧区域显示车辆图像, 右侧显示车牌识别的结果包括车牌定位和字符识别。

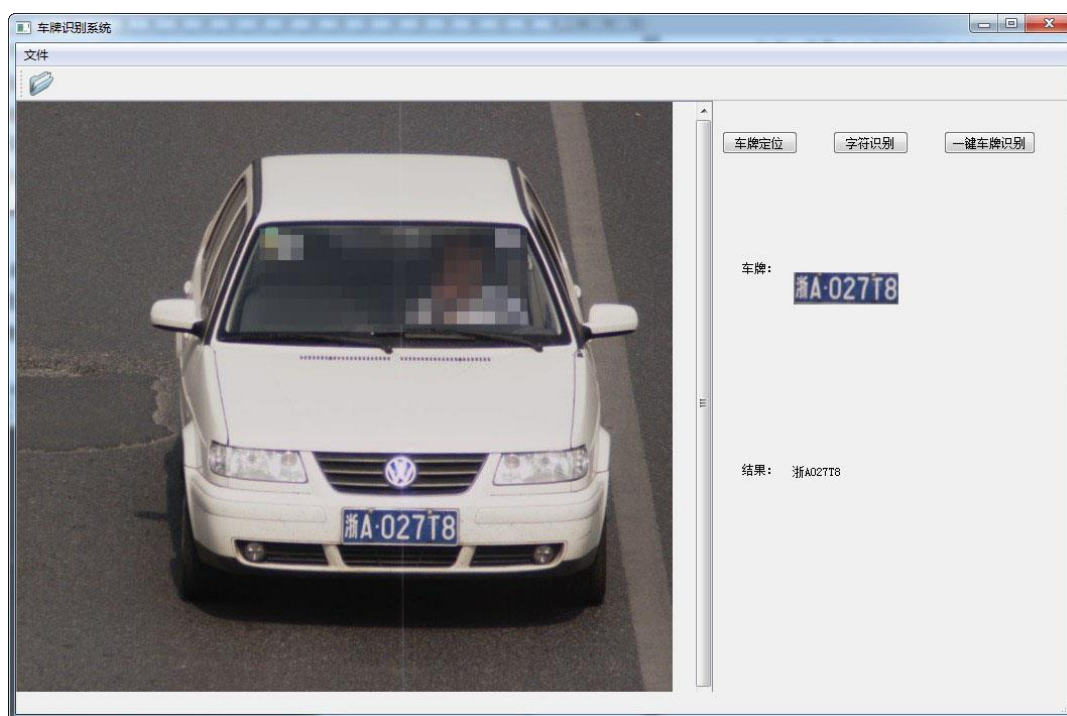


图 4-4 系统软件界面

按本文的车牌识别系统流程, 对该系统的识别效果进行了测试, 测试样本图片主要来源于实际车辆图片。系统测试的软件运行在以下环境:

(1) CPU主频 2.5Ghz 以上, 硬盘剩余空间 20G 以上, 内存 4G。

(2)操作系统 Windows 7。

本文的车牌定位方法在测试中所用的时间一般在 0.4 秒左右,在图像尺寸较小和图像质量较高的时候最快速度达到了 0.2 秒。当车牌出现倾斜的情况时,通过寻找最大的倾斜角度,用最大外接矩阵获取车牌所在的区域。

本文统计的是后 5 位车牌字符,通过对没有识别出来的车牌图像进行分析,发现其主要原因在于车牌部分污损、图像模糊不清以及光照因素等,尤其是在晚上,光照因素导致获取的车辆图像出现曝光、反光等现象,使整个车牌的亮度提高,因此在预处理之后字符会变粗甚至出现粘连的情况,不利于后期的字符分割和识别过程。通过对识别错误的车牌字符统计可知,除去汉字首字识别时,对于其他字符识别时,基本上可以达到准确识别车牌号码的效果。

4.3 本章小结

本章介绍了车牌识别系统所涉及的开发环境并对系统的性能做了优化,接着对参数配置的设置和各个模块的封装作了说明。最后,利用 VS 开发环境和 Qt 图形图像框架实现程序的可视化应用程序。

第5章 总结与展望

5.1 总结

本文给出的是一套完整的基于图像处理方面的车牌识别系统研究方案。首先介绍了车牌识别系统的整体框架，接着详述了国内外相关学者和企业在这篇领域的研究现状和发展状况。然后详细介绍了对相似字符识别的研究，利用深度学习在图像识别上的优势，提出采用卷积神经网络的识别方法。研究了车牌定位算法和字符分割算法，在此基础上提出了多层次的车牌识别系统，并提出创新点给出验证结果。主要工作总结如下：

(1) 针对车牌相似字符识别问题，提出了基于深度学习的字符识别，构建了自动学习特征提取和识别的深度卷积神经网络。随着训练样本数量不断地增加，使用卷积神经网络的准确率也相应的提高，因为深层的网络模型能够学习到更有用的内在特征信息。但是卷积神经网络比传统的浅层网络更为复杂，因此在识别和训练中需要花费更多的时间，同时在训练时使用了大量样本有利于对相似字符进行识别。

(2) 关于车牌字符识别中存在相似字符的问题，本文的第二章详细介绍了利用深度学习作为相似字符识别算法，为字符识别作了铺垫。接着对分割后的字符作预处理，采用插值的方式对所有字符进行归一化操作。又考虑到车牌多字符的识别效率问题，介绍了 SVM 多分类器，提出了采用基于 SVM 方法对字符进行第一识别，再结合卷积神经网络对相似字符进行二次识别。最后根据车牌字符的排列特点，构造了三个分类器。最后实验结果显示该方法在时间和性能上有一定优势。本文采用了 ReLU 激活函数和 Softmax 分类器对车牌相似字符图像进行识别的结果表明，深度学习算法在识别车牌字符上具有一定的应用价值。

(3) 设计了整套车牌识别系统，主要包括设计车牌定位、字符分割和字符识别三个模块，并将封装成动态链接库便于系统的开发。最后本文采用 C++语言在 VS 开发平台和 Qt 框架下，以及使用 OpenCV库实现了整个车牌识别的应用系统。

5.2 展望

在实际使用过程操作，由于受到光照环境、拍摄距离和角度、夜间大灯等复杂环境因素的影响，本文的一些方法仍存在不足之处，因此在未来工作中可以从以下几个方面进行改进和研究：

（1）本文实验的对象主要为蓝色和黄色这两类车牌，而且算法只针对这些车牌中的单行车牌，所以以后需要在其他颜色和双排车牌上进行研究和探索。

（2）在车辆图像受到光照影响导致图像质量下降或者拍摄问题导致车牌区域较小时，会存在车牌漏检的现象直接影响接下来的几个处理过程，还需要考虑其他有效特征来提取车牌区域。本文的实验所使用都是单一的车牌，对于同一图片的多车牌识别情况，本文还未做深入探究。

（3）对于字符识别过程中，样本集的汉字类别少，还需要收集其他各种类别的车辆图像，丰富实验的训练集和测试集。在汉字提取过程中容易出现字符迷糊和笔画粘连的现象，可以考虑其他的有关深度学习的识别方法。

（4）车牌识别系统在实际应用过程中对时间要求较高，还需要对车牌定位和字符识别作进一步优化。

参 考 文 献

- [1] 沈后功, 周凯. 2015年全国机动车和驾驶人迅猛增长 新增汽车 1781多万汽车保有量增长创历史新高[J]. 汽车与安全, 2016(2):4141.
- [2] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks.[J]. Science, 2006, 313(5786):504507.
- [3] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):227224.
- [4] Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. Journal of Physiology, 1962, 160(1):1061.
- [5] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:10971105.
- [6] Deng J, Dong W, Socher R, et al. ImageNet: A large hierarchical image database[C]// Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. 2009:248255..
- [7] Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3):211252.
- [8] Gupta V, Gavrilovska A, Schwan K, et al. GVIM: GPU accelerated virtual machines[C]// Proceedings of the 3rd ACM Workshop on System-level Virtualization for High Performance Computing. ACM, 2009:1724.
- [9] Johnson A S, Bird B M. Number plate matching for automatic vehicle identification[C]// IET, 1990:4/14/8.
- [10] Lee E R, Kim P K, Hang J K. Automatic recognition of a car license plate using color image processing[C]// Image Processing, 1994. Proceedings-94 IEEE International Conference. IEEE, 1994:301305 vol.2.
- [11] Comelli P, Ferragina P, Granieri M N, et al. Optical recognition of motor vehicle license plates[J]. IEEE Transactions on Vehicular Technology, 1995, 44(4):799.
- [12] 赵雪春, 戚飞虎. 基于彩色分割的车牌自动识别技术[J]. 上海交通大学学报, 1998(10):49.
- [13] 王敏, 黄心汉, 魏武, 等. 一种模板匹配和神经网络的车牌字符识别方法[J]. 华中科技大学学报(自然科学版), 2001, 29(3):450.
- [14] 郭琳琳, 黄仲瑾. 基于边缘检测的车牌定位方法[J]. 信息与电脑:理论版, 2010(4):2021.
- [15] 刘广起, 郑晓势, 张晓波. 基于图像纹理特征提取的车牌定位算法[J]. 中国图象图形学报, 2005, 10(11):14191422.
- [16] 张引, 潘云鹤. 彩色汽车图象牌照定位新方法[J]. 中国图象图形学报, 2001, 6(4):374377.
- [17] 郭捷, 施鹏飞. 基于颜色和纹理分析的车牌定位方法[J]. 中国图象图形学报, 2002, 7(5):472476.
- [18] Zhu W G, Hou G J, Jia X. A Study of Locating Vehicle License Plate Based on Color Feature and Mathematical Morphology[C]// International Conference on Signal Processing. IEEE, 2002:748 vol.1.
- [19] 冉令峰. 基于垂直投影的车牌字符分割方法[J]. 通信技术, 2012, 45(4):8991.
- [20] 陈寅鹏, 丁晓青. 复杂车辆图像中的车牌定位与字符分割方法[J]. 红外与激光工程, 2004, 33(1):2933.

- [21] 魏武, 张起森, 王明俊, 等. 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):104106.
- [22] 马俊莉, 莫玉龙, 王明祥. 一种基于改进模板匹配的车牌字符识别方法[J]. 小型微型计算机系统, 2003, 24(9):1670672.
- [23] 杨凡, 赵建民, 朱信忠. 一种基于 BP 神经网络的车牌字符分类识别方法[J]. 计算机科学, 2005, 32(8):192195.
- [24] Tindall D W. Application of neural network techniques to automatic licence plate recognition[C]// European Convention on Security and Detection. IEEE Xplore, 1995:81.
- [25] Chi X, Dong J, Liu A, et al. A Simple Method for Chinese License Plate Recognition Based on Support Vector Machine[C]// International Conference on Communications, Circuits and Systems Proceedings. IEEE Xplore, 2006:2142145.
- [26] Kim K K, Kim K I, Kim J B, et al. Learning-based approach for license plate recognition[C]// Neural Networks for Signal Processing X, 2000. Proceedings of the 2000 Signal Processing Society Workshop. IEEE, 2000:614623 vol.2.
- [27] 张宏烈. 支持向量机在字符识别中的应用研究[J]. 微计算机信息, 2006, 22(11):24247.
- [28] 瞿中, 常庆丽, 王永昆, 等. 改进的车牌相似字符分级分类识别算法研究[J]. 计算机工程与设计, 2013, 34(4):1330340.
- [29] 童剑军, 邹明福. 车牌字符的一种精判别识别方法[J]. 计算机工程与应用, 2005, 41(2):223226.
- [30] 黄文琪, 吴炜, 苏力思, 等. 基于贝叶斯网络分类器的车牌相似字符识别[J]. 四川大学学报自然科学版, 2013(4):775780.
- [31] Cun Y L, Boser B, Denker J S, et al. Handwritten digit recognition with a backpropagation network[C]// Advances in Neural Information Processing Systems. Morgan Kaufmann Publishers Inc. 1990:465.
- [32] Salakhutdinov R, Hinton G. Deep Boltzmann Machines[J]. Journal of Machine Learning Research, 2009, 5(2):19672006.
- [33] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7):1527.
- [34] Dahl J V, Koch K C, Kleinhan E, et al. Convolutional networks and applications in vision[C]// International Symposium on Circuits and Systems. DBLP, 2010:2563.
- [35] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[J]. Journal of Machine Learning Research, 2010, 9:249.
- [36] Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Networks[J]. Learning/statistics & Optimisation, 2010.
- [37] Nair V, Hinton G E. Rectified Linear Units Improve Restricted Boltzmann Machines[C]// International Conference on Machine Learning. DBLP, 2010:807.
- [38] Dahl G E, Sainath T N, Hinton G E. Improving deep neural networks for LVCSR using rectified linear units and dropout[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013:86098613.
- [39] Hinton G E. How neural networks learn from experience[J]. Scientific American, 1992, 267(3):144.
- [40] Bouvrie J. Notes on Convolutional Neural Networks[J]. Neural Nets, 2006.
- [41] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[J]. Journal of Machine Learning Research, 2010, 9:249.
- [42] He K, Zhang X, Ren S, et al. Delving Deep into Rectifiers: Surpassing Human Performance on ImageNet Classification[J]. 2015:102634.
- [43] Chang, Chih-Chung, Lin, et al. LIBSVM: A library for support vector machines[J]. Acm Transactions on Intelligent Systems & Technology, 2011, 2(3):3896.

- [44] 中华人民共和国公安部. GA 36-2014中华人民共和国机动车牌安全标准[P]. 2014.
- [45] 罗帆, 陈晟, 王敏,等. 一种基于边缘特征的汽车牌照定位算法[J]. 华中科技大学学报(自然科学版), 2004(S1):108-110.
- [46] 沈勇武, 章专. 基于特征颜色边缘检测的车牌定位方法[J]. 仪器仪表学报, 2008, 29(12):2673-2677.
- [47] Vapnik V N. The Nature of Statistical Learning Theory[J]. IEEE Transactions on Neural Networks, 1995, 8(6):1564-1564.
- [48] 瓦普尼克. 统计学习理论的本质[M]. 清华大学出版社, 2000.
- [49] 张引, 潘云鹤. 面向车辆牌照字符识别的预处理算法[J]. 计算机应用研究, 1999(7):85-87.
- [50] Hough P V C. Method and means for recognizing complex patterns[J]. 1962.
- [51] Otsu N. A Threshold Selection Method from Gray-Scale Histograms[J]. IEEE Trans on SMC, 1979, 9:62-66.
- [52] 余辉, 赵晖. 支持向量机多类分类算法新研究[J]. 计算机工程与应用, 2008, 44(7):185-189.

致 谢

时间就像奔流的大河一刻不停地流淌着，三年的研究生生涯转瞬即逝，我依然能够记得刚踏进校园时满怀热情地求知欲望。在这期间，我得到了导师无微不至的教导，也得到了家人和亲戚的支持，此外我还遇见了许多同学并且结交了一些热爱搞科研的朋友。在此，我将由衷感谢帮助过和关心过我的人。

首先，由衷的感谢我的导师潘翔教授。你的严谨的治学态度，开阔的思想，循循善诱的指点给了我很大的帮助，尤其是在我写学术论文思绪感到迷茫时，你为我理清思路，指点迷津。在读研究生期间，潘老师不仅在学术和科研上给了我很多的帮助，而且你的超强的编程能力使我的养成了良好的编程习惯。你为人师表的作风，给我们留下了一笔宝贵的财富。在此谨向潘老师致以我最诚挚的敬意和感谢。

其次，感谢我实验室的师弟师妹们，他们与我共同度过美好的三年。在这三年的时间里，我们有过甜美的回忆和喜悦，也遇到过困难和挫折。感谢我的伙伴们林俊勉硕士、焦吾振硕士、吴佩洁硕士、刘阳圣龔硕士、程志豪硕士、金怡硕士、过伟龙硕士、朱兆元硕士、洪之渊硕士、夏洋硕士等。从我刚接触深度学习到今天完成论文撰写，他们为我提供了无私的帮助，没有他们的帮助就没有这篇论文的顺利完成。

最后，特别感谢我的父母和亲人，正是他们在我学习生涯给予的支持和包容，给我提供了有利的学习条件，才能全身心地投入学习和科学研究并顺利完成硕士研究生阶段的学习和研究任务。

谨以此文献给所有关心帮助过我的人，祝你们一切顺利，幸福美满！

王恒

2017-3-7 浙江 杭州

攻读学位期间参加的科研项目和成果

录用和发表的论文

- [1] 基于深度学习的车牌相似字符识别[J]. 计算机科学. (2017年 6 月专辑). 署名排序 2