

基于卷积神经网络LeNet-5的车牌字符识别研究

赵志宏^{1,2}, 杨绍普², 马增强^{1,2}

(1. 北京交通大学机械与电子控制工程学院, 北京 100044; 2. 石家庄铁道学院, 河北 050043)



摘要: 将卷积神经网络 LeNet-5 引入到车牌字符识别中。为了适应目前中国车牌字符识别的需要, 对传统的卷积神经网络 LeNet-5 的结构进行了改进, 主要是改变输出单元的个数与增加卷积层 C5 特征图的个数。研究表明, 改进后的 LeNet-5 比传统的 LeNet-5 的识别率有所提高, 识别率达到 98.68%。另外, 与 BP 神经网络进行了比较研究, 从实验中可以看出在字符识别的正确率和识别速度上都优于 BP 神经网络。卷积神经网络在车牌识别中具有很好地应用前景。

关键词: 字符识别; 车牌识别; 卷积神经网络; LeNet-5

中图分类号: TP391.4

文献标识码: A

文章编号: 1004-731X (2010) 03-0638-04

License Plate Character Recognition Based on Convolutional Neural Network LeNet-5

ZHAO Zhi-hong^{1,2}, YANG Shao-pu², MA Zeng-qiang^{1,2}(1. School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China;
2. Shijiazhuang Railway Institute, Shijiazhuang 050043, China)

Abstract: The application of convolutional neural network LeNet-5 was proposed in license plate character recognition. To fit with the Chinese license plate character recognition problem, the traditional LeNet-5 was modified. The unit number of output layer was changed and the feature map number of C5 layer was added. Experimental results show that the recognition rate of modified LeNet-5 reaches 98.68% and is better than that of LeNet-5. The results are also compared with the BP neural network, which indicates that the modified LeNet-5 is superior in both recognition rate and speed. This method has a great potential for license plate recognition.

Key words: character recognition; license plate recognition; convolutional neural network; LeNet-5

引言

车牌识别在交通信息控制与智能交通中发挥着重要作用, 如: 高速公路收费管理、超速违章处罚、停车场管理、交通数据采集等。通常, 车牌识别系统主要包括三个部分: 车牌定位、字符分割与字符识别。本篇论文主要研究字符识别。

已经提出很多种字符识别方法^[1-4]。字符识别方法大致可以分为两类: 基于结构的方法^[1,2]和基于统计的方法^[3,4]。基于结构的方法的优点是符合人的直觉, 可以容易地处理局部变换, 其主要缺点是抗噪声能力差。基于统计的方法比较成熟, 具有抗干扰、噪声的能力, 识别性能好, 识别速度快的优点, 它的难点是特征提取, 因此, 在统计字符识别中, 特征提取的质量是影响系统性能的关键因素。

卷积神经网络可以直接从原始图像中识别视觉模式, 它需要的预处理工作非常少。卷积神经网络已成功应用于手写字符识别^[5,7], 人脸识别^[6], 人眼检测^[8], 行人检测^[9], 机器人导航^[10]中。卷积神经网络可以识别有变化的模式, 具有对简单几何变形的鲁棒性。卷积神经网络较一般神经网络在图像

处理方面具有如下优点:(1)输入图像和网络的拓扑结构能很好的吻合;(2)特征提取和模式分类同时进行;(3)权值共享可以减少网络的训练参数, 使神经网络结构变得更简单, 适应性更强。

卷积神经网络在国外的研究较多, 但目前在国内的研究与应用还刚刚起步。LeNet-5^[5]是一种卷积神经网络, 最初用于手写数字识别中, 本文将卷积神经网络 LeNet-5 用于中国车牌字符识别中。由于车牌字符识别的类别比手写数字识别的类别数目多, 不能直接应用传统的 LeNet-5 进行车牌字符识别。因此, 对传统的 LeNet-5 的结构进行了改进。

本文首先简单介绍卷积神经网络和 LeNet-5 的结构, 进而对 LeNet-5 的结构进行了改进, 较好地解决了车牌字符识别类别比手写数字识别的类别多的难题。设计了实验方案, 实验对比了改进后的 LeNet-5 与 BP 神经网络性能, 得出了一些有益的结论, 并对进一步的研究工作进行了展望。

1 卷积神经网络

卷积神经网络是近年发展起来, 并引起广泛重视的一种高效识别方法。现在, 卷积神经网络已经成为众多科学领域的研究热点之一, 特别是在模式分类领域, 由于该网络避免了对图像的复杂前期预处理, 可以直接输入原始图像, 因而得到了广泛的应用^[5-9]。

卷积神经网络通过结合三个方法来实现识别位移、缩放

收稿日期: 2008-05-28

修回日期: 2008-06-23

基金项目: 国家杰出青年科学基金 (50625518); 教育部科学技术研究重点项目 (205019)。

作者简介: 赵志宏(1972-), 男, 河北人, 博士生, 讲师, 研究方向为智能交通、神经网络; 杨绍普(1962-), 男, 河北人, 博导, 教授, 研究方向为车辆工程; 马增强(1975-), 男, 河北人, 博士生, 讲师, 研究方向为智能交通。

和扭曲不变性^[5]:局域感受野、权值共享和次抽样。局域感受野指的是每一网络层的神经元只与上一层的一个小邻域内的神经单元连接,通过局域感受野,每个神经元可以提取初级的视觉特征,如方向线段,端点,角点等。权值共享使得卷积神经网络具有更少的参数,需要相对少的训练数据。次抽样可以减小特征的分辨率,实现对位移、缩放和其它形式扭曲的不变性。

一般地,卷积层后有一个次抽样层来减少计算时间和建立空间和结构上的不变性。通常采用小的次抽样因子,这个方法非常简单和有效。

1.1 卷积层

在卷积层,前一层的特征图与一个可学习的核进行卷积,卷积的结果经过激活函数后的输出形成这一层的特征图。每一个输出的特征图可能与前一层的几个特征图的卷积建立关系。一般地,卷积层的形式如式(1)所示:

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

其中, l 代表层数, k 是卷积核, M_j 代表输入特征图的一个选择。每个输出图有一个偏置 b 。

1.2 次抽样层

一个次抽样层对输入进行抽样操作。如果输入的特征图为 n 个,则经过次抽样层后特征图的个数仍然为 n ,但是输出的特征图要变小(例如,变为原来的一半)。次抽样层的一般形式如式(2)所示:

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l) \quad (2)$$

其中, $\text{down}(\bullet)$ 表示次抽样函数。次抽样函数一般是对该层输入图像的一个 $n \times n$ 大小的区域求和,因此,输出图像的大小是输入图像大小的 $1/n$ 。每一个输出的特征图有自己的 β 和 b 。

2 LeNet-5 及其改进

2.1 LeNet-5 的结构

卷积神经网络 LeNet-5 的结构如图 1 所示。输入图像要经过大小归一化,每一个神经元的输入来自于前一层的一个局部邻域,并被加上由一组权值决定的权重。提取的这些特征在下一层结合形成更高一级的特征。同一特征图的神经元共享相同的一组权值,次抽样层对上一层进行平均。

卷积神经网络 LeNet-5 不包括输入,由 7 层组成,每一层都包括可训练的参数(权值)。该网络的输入是 32×32 的图像,其中 C 层是由卷积层神经元组成的网络层, S 层是由次抽样层神经元组成的网络层。

网络层 C1 是由 6 个特征图组成的卷积层。每个神经元与输入图像的一个 5×5 的邻域相连接,因此每个特征图的大小是 28×28 。

网络层 S2 是由 6 个大小为 14×14 的特征图组成的次抽

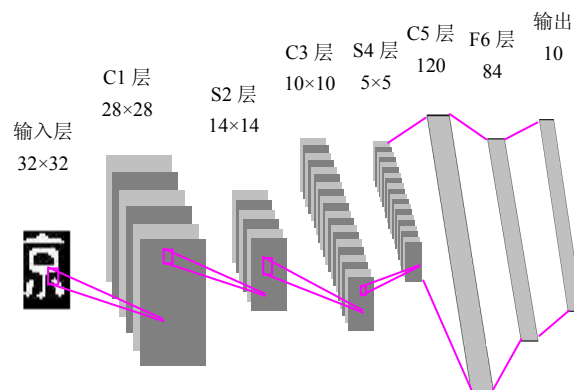


图 1 LeNet-5 神经网络的结构,每个面是一个特征图

样层,它是由 C1 层抽样得到。特征图的每个神经元与 C1 层的一个大小为 2×2 的邻域连接。

网络层 C3 是由 16 个大小为 10×10 的特征图组成的卷积层。特征图的每个神经元与 S2 网络层的若干个特征图的 5×5 的邻域连接。表 1 显示了 S2 层的特征图如何结合形成 C3 的每个特征图,其中的每一列表示 S2 层的哪些特征图结合形成 C3 的一个特征图,例如,从第一列可知,S2 层的第 0 个,第 1 个,第 2 个特征图结合得到 C3 的第 0 个特征图。

表 1 S2 层特征图与 C3 层特征图的连接方法

	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	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

网络层 S4 是由 16 个大小为 5×5 的特征图组成的次抽样层。特征图的每个神经元与 C3 层的一个 2×2 大小的邻域相连接。

网络层 C5 是由 120 个特征图组成的卷积层。每个神经元与 S4 网络层的所有特征图的 5×5 大小的邻域相连接。

网络层 F6,包括 84 个神经元,与网络层 C5 进行全连接。

最后,输出层有 10 个神经元,是由径向基函数单元(RBF)组成,输出层的每个神经元对应一个字符类别。RBF 单元的输出 y_i 的计算方法如式(3)所示:

$$y_i = \sum_j (x_j - w_{ij})^2 \quad (3)$$

2.2 LeNet-5 的改进

卷积神经网络 LeNet-5 最初是用于手写数字识别中,输出的类别数目为 10,中国的车牌识别与手写数字识别相比,需分类的类别数目要多很多。除了 10 个阿拉伯数字外,还有 25 个英文字母(不包括字母 I),全国各省市、自治区的简称 31 个,还有其它一些字符。

针对中国车牌字符识别与手写数字识别的不同,对传统

的 LeNet-5 进行了改进。主要有以下两点:

(1)将传统的 LeNet-5 的输出层由 10 个神经元改为 76 个神经元。收集到的车牌字符数据库有 76 类车牌字符,因此,将输出神经元的个数改为 76。

(2)增加传统的 LeNet-5 网络层 C5 层的特征图的数目,将 C5 层特征图的数目由 120 个增加到 360 个。由于车牌字符识别分类的类别数目增多,因此,考虑增加 C5 层特征图的个数,以增加整个网络的识别性能。

3 实验研究

为了验证改进的卷积神经网络 LeNet-5 的性能,收集了大量真实车牌数据,并建立了训练与测试用的数据集。车牌图像经过定位,字符分割后,存入车牌字符数据库,没有对车牌图像进行去噪处理。为了应用卷积神经网络 LeNet-5,所有的车牌字符都归一化为 32×32 大小。

图 2 显示了输入字符图像在卷积神经网络 LeNet-5 前面几层的处理结果,最左边是输入图像,从左到右依次是输入层, C1 层, S1 层, C2 层, S2 层。C1 层的输出代表相应地 6 个特征图。从图 2 中可以看出,神经网络 LeNet-5 可以很好地提取字符的特征。

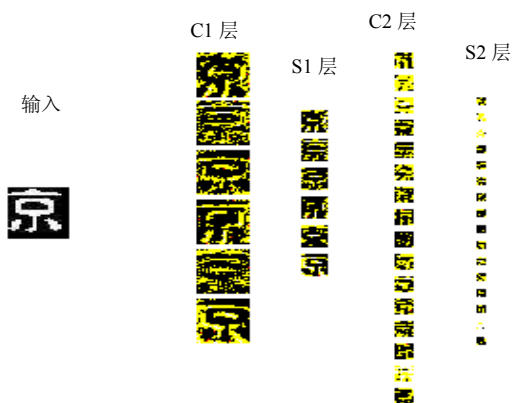


图 2 神经网络 LeNet-5 的输入图像及中间层神经元输出的例子。

3.1 实验数据

本文实验所用的车牌图像均由数码相机在实际环境中获取的分辨率为 640×480 的真实车牌图像,收集到实际车牌图像数据共 650 个,利用其中的字符分别建立训练数据集与测试数据集,训练数据集与测试数据集之间没有重叠。图 3 是实验中用到的一些车牌数据。



图 3 实验中用到的一些车牌数据

3.2 训练次数与识别率的关系

表 2 显示了当 C5 网络层特征图的个数为 240 时训练次数与识别率的关系,从表中可以看到,随着训练次数的增加,训练样本与测试样本的正确率都随着增加。并且当训练样本集的正确率达到 100%后,继续训练,测试样本集的正确率仍然得到提高,基本不出现过训练的情况。这体现了卷积神经网络良好的性能。

表 2 训练次数与识别率的关系

训练次数	400	500	600	700	800	900	1000	1100
训练样本集	92.93	98.62	99.83	100	100	100	100	100
测试样本集	89.65	96.30	97.49	98.02	98.15	98.46	98.59	98.59

3.3 C5 层改变对性能的影响

下面是改变 C5 层特征图的个数对卷积神经网络性能的影响。

3.3.1 对识别率的影响

C5 层特征图的个数与 LeNet-5 卷积神经网络的识别率的关系如表 3 所示。

表 3 C5 层特征图个数与识别率的关系

C5 层特征图数目	60	120	240	360	480
测试样本识别率	97.84	98.43	98.59	98.68	98.68

从表 3 中可以看到, C5 层特征图个数对车牌字符识别率有直接影响,随着 C5 层特征图个数的增加,车牌字符的识别率也有所提高。这说明,在类别数增加的情况下,增加 C5 层特征图的个数,可以提高 LeNet-5 卷积神经网络的分类能力。但当提高到一定程度后,再增加 C5 层特征图的数目,识别率将不再提高。

3.3.2 对识别速度的影响

网络层 C5 层特征图数目的改变对 LeNet-5 卷积神经网络的速度也有影响。不同 C5 层特征图个数的识别速度如表 4 所示。

从表 4 中可以看到,随着 C5 层特征图个数的增加,车牌字符识别的速度将变慢。这是由于增加 C5 层特征图的数目后,提高了 LeNet-5 卷积神经网络的复杂程度,因此,识别速度有所降低。这说明,可以通过降低网络层 C5 层的特征图的个数,来提高字符识别的速度。因此,在某些要求实时性高的应用场合,可以适当降低 C5 网络层特征图的个数。

表 4 不同 C5 层特征图数目的识别速度(单位:秒)

C5 层特征图数目	60	120	240	360	480
识别 100 个字符的时间	0.241	0.274	0.299	0.336	0.379

从上面的实验可知,改进后的 LeNet-5 在识别率方面优于传统的 LeNet-5,但识别速度有所降低。

3.4 与 BP 神经网络的比较

为了测试改进后卷积神经网络 LeNet-5 的性能,与常用

的三层 BP 神经网络进行了比较。BP 神经网络隐含层取 450 个节点,学习率采用 0.01,实验结果的比较如表 5 所示。从表 5 中可以看到,改进后的 LeNet-5 卷积神经网络的识别率比 BP 神经网络的识别率高大约 5%。

表 5 LeNet-5 与 BP 网络识别性能的比较

网络类型	LeNet-5	BP 神经网络
识别正确率	98.68%	93.37%

另外, BP 神经网络识别 100 个字符的时间为 0.785 秒,改进后的卷积神经网络 LeNet-5 的识别速度也优于 BP 神经网络。

3.5 错误数据分析

识别错误的 38 个车牌字符如图 4 所示。识别错误的 38 个车牌字符或者是由于预处理效果不好(字符分割的效果不理想),或者是由于噪声的影响(在字符的关键地方有很强的噪声),还有部分是由于倾斜的影响(倾斜角度很大)。识别错误的字符,可以通过有效的预处理过程来进一步减少。另外,可以通过扩大训练字符集来进一步提高车牌字符识别的正确率。



图 4 38 个识别错误的车牌字符模式

4 结论

本文对卷积神经网络 LeNet-5 用于车牌字符识别进行了各种研究。为了适应中国车牌字符识别的需要,对传统的卷积神经网络 LeNet-5 的结构进行了改进。研究结果表明,改进后的 LeNet-5 比传统的 LeNet-5 识别率有所提高,对收集到的车牌字符样本识别的正确率达到 98.68%。并且从实验中得到 LeNet-5 的分类性能与卷积层 C5 层特征图的数目有关。

(上接第 637 页)

- [10] Guo S, Huang L. Periodic oscillation for a class of neural networks with variable coefficients [J]. Nonlinear Anal. (S0362-546x), 2005, 6(3): 545-561.
- [11] Guo S. Bifurcation analysis in a discrete-time single-directional network with delays [J]. Neurocomputing (S0925-2312), 2008, 71(7/9): 1422-1435.
- [12] Cao J. New results concerning exponential stability and periodic solutions of delayed cellular neural networks [J]. Physics Letters A (S0375-9601), 2003, 307(2): 136-147.
- [13] Sun C, Feng C. Exponential periodicity and stability of delayed neural networks [J]. Mathematics and Computers in Simulation (S0378-4754),

另外,通过与三层 BP 神经网络的对比,可以看到改进后的卷积神经网络 LeNet-5 无论在字符识别的正确率还是识别速度上都优于 BP 神经网络。

进一步研究工作地展望:

- (1) 试验不同的卷积神经网络。除了 LeNet-5 之外,试验是否存在更好的适应车牌字符识别的卷积神经网络结构。
- (2) 设计和实现基于视频的车牌识别系统。实验用卷积神经网络建立一个比较实用的车牌识别系统,包括车牌定位、字符识别的功能。

参考文献:

- [1] 贾婧,葛万成,陈康力. 基于轮廓结构和统计特征的字符识别研究[J]. 沈阳师范大学学报: 自然科学版, 2006, 24(1): 43-46.
- [2] 任柯昱,唐丹,尹显东. 基于字符结构知识的车牌汉字快速识别技术[J]. 计算机测量与控制, 2005, 13(6): 592-594.
- [3] 廉飞宇,付麦霞,张元. 基于支持向量机的车辆牌照识别的研究[J]. 计算机工程与设计, 2006, 27(21): 4033-4035.
- [4] Al-Hmouz R, S. Challa. Intelligent Stolen Vehicle Detection using Video Sensing [C]// Proceeding of Information, Decision and Control. Adelaide, Qld., Australia. USA: IEEE, 2007: 302-307.
- [5] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition [C]// Proc. IEEE, 1998. USA: IEEE, 1998: 2278-2324.
- [6] Steve Lawrence, C Lee Giles, Ah Chung Tsoi, Andrew D Back. Face Recognition: A Convolutional Neural Network Approach [J]. IEEE Trans. on Neural Networks (S1045-9227), 1997, 8(1): 98-113.
- [7] Lauer F, C Y Suen, Bloch G. A trainable feature extractor for handwritten digit recognition [J]. Pattern Recognition (S0031-3203), 2007, 40(6): 1816-1824.
- [8] Tivive, Fok Hing Chi, Bouzerdoum, Abdesselam. An eye feature detector based on convolutional neural network [C]// Proc. 8th Int. Symp. Signal Process. Applic. Sydney, New South Wales, Australia. USA: IEEE, 2005: 90-93.
- [9] Szarvas Mate, Yoshizawa Akira, Yamamoto Munetaka, Ogata Jun. Pedestrian detection with convolutional neural networks [C]// IEEE Intelligent Vehicles Symposium Proceedings. USA: IEEE, 2005: 224-229.
- [10] Y Le Cun, U Muller, J Ben, E Cosatto, B Flepp. Off-road obstacle avoidance through end-to-end learning [M]. Advances in Neural Information Processing Systems. USA: MIT Press, 2005.
- [11] Zhou J, Liu Z, Chen G. Dynamics of periodic delayed neural networks [J]. Neural Networks (S0893-6080), 2004, 17(1): 87-101.
- [12] Cao J, Li Q. On the Exponential stability and periodic solutions of delayed cellular neural networks [J]. Journal of Mathematical Analysis and Applications (S0022-247x), 2000, 252(1): 50-64.
- [13] 周立群,胡广大. 具分布延时细胞神经网络的指数周期与稳定性[J]. 系统仿真学报, 2007, 20(10): 2511-2514. (ZHOU Li-qun, HU Guang-da. Exponential Periodicity and Stability of Cellular Neural Networks with Distributed Delays [J]. Journal of System Simulation (S1004-731X), 2007, 20(10): 2511-2514.)