

卷积神经网络在车牌识别中的应用研究

刘华春

(成都理工大学 工程技术学院 电子信息与计算机工程系, 四川 乐山 614007)

摘要: 为了改善传统车牌识别方法中过分依赖车牌特征和鲁棒性不强等问题, 将具有良好图像识别性能的卷积神经网络 LeNet-5 模型引入到车牌字符识别中, 并对其结构进行改进以满足需要。设计了 2 个网络分别进行汉字和数字/字母识别, 将输出层类别由 10 增加到 31 和 34; C_5 卷积层的特征面数目增加到 480, 输入图像像素增加到 64×64 。对改进后的网络进行了实验, 并分别与 3 层 BP 神经网络和支持向量机 (SVM) 进行对比测试。实验结果表明, 该卷积神经网络避免了传统车牌字符识别方法中复杂的特征提取, 增强了鲁棒性, 提高了准确率。改进后的 LeNet-5 相比 BP 神经网络在识别准确率上可提高约 6%, 识别速度也更快; 与 SVM 相比较, 汉字分类准确率可以提高约 7%, 字符/数字准确率可以提高约 4%。

关键词: 车牌识别; 卷积神经网络; 支持向量机; 改进 LeNet-5 卷积网络; 深度学习

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2019)04-0128-05

doi: 10. 3969/j. issn. 1673-629X. 2019. 04. 026

Research on Application of Convolutional Neural Network in License Plate Recognition

LIU Hua-chun

(Department of Electronic Information and Computer Engineering, Engineering & Technical
College of Chengdu University of Technology, Leshan 614007, China)

Abstract: In order to improve the excessive dependence on license plate features and lack of robustness in traditional vehicle license plate recognition methods, the convolutional neural network LeNet-5 model with great image recognition performance is introduced into the license plate character recognition and its structure is improved to meet the needs. The main improvements are the design of two networks for Chinese character and digit/letter recognition, and increasing the output layer class from 10 to 31 and 34, and the number of feature faces in the C_5 convolutional layer to 480, and the input image pixels to 64×64 . The improved network is tested and compared with a 3-layer BP neural network and a support vector machine (SVM). The experiment shows that the convolutional neural network avoids the complex feature extraction in the traditional license plate character recognition method, and enhances the robustness and accuracy. Compared with BP neural network, the improved LeNet-5 accuracy can be improved by about 6% and the recognition speed is faster. Compared with SVM, Chinese character classification accuracy can be improved by about 7%, and character/digit accuracy can be increased by about 4%.

Key words: license plate character recognition; convolutional neural network; support vector machines; improved LeNet-5 convolutional network; deep learning

0 引言

随着中国汽车保有量的持续增加, 发展智能交通已经成为一个社会共识, 而车牌识别是智能交通的一个重要环节。车牌识别在道路收费管理、超速违规、电子警察、停车场管理、交通数据采集等交通信息控制中发挥着重要的作用。通常的车牌识别系统包括车牌定位模块、车牌校正和预处理模块、字符分割模块、字符

识别模块。文中着重研究车牌的字符识别模块部分。

深度学习是一种特定类型的机器学习技术^[1], 是实现人工智能的一种有效方式, 也是一种使计算机系统从经验和数据上得到改进的技术, 具有很强的识别能力和灵活性^[2]。深度学习可以发现和表征问题的复杂结构特征, 因此可以大大提高识别性能, 还避免了统计机器学习中关于特征提取的一系列问题^[3]。目前有

收稿日期: 2018-04-28

修回日期: 2018-08-30

网络出版时间: 2018-12-20

基金项目: 四川省教育科研重点项目 (A22017003); 四川省乐山市科技局重点项目 (16GZD050)

作者简介: 刘华春 (1966-), 男, 硕士, 副教授, 研究方向为机器学习、人工智能。

网络出版地址: <http://jns.cnki.net/kcms/detail/61.1450.TP.20181220.1035.026.html>

大量的证据表明,在计算机视觉、图像识别领域,深度学习的识别精度高于传统的图像处理和统计机器学习^[4]。卷积神经网络(convolutional neural network, CNN)是神经网络(neural network, NN)的发展,是源于人工神经网络的一种深度机器学习方法,近年来在图像识别领域取得了巨大的成功。CNN 由于采用局部连接和权值共享,保持了网络的深层结构,同时又大大减少了网络参数,使模型具有良好的泛化能力,又较容易训练,很好地解决了 NN 中网络训练时的梯度消失(vanishing gradient problem)和梯度爆炸(exploding gradient problem)等问题^[5]。目前,卷积神经网络已经成为众多科学领域的研究热点之一。由于卷积网络避免了对图像复杂的特征提取,可以直接输入原始图像,因而得到了更为广泛的应用。卷积神经网络在图像识别领域具有良好的适应性,目前在计算机视觉任务中应用广泛,并在手写数字识别、人脸识别等图像领域的应用中取得了很好的效果^[6]。

文中将改进的卷积神经网络 LeNet-5 引入到车牌字符识别中,以提高传统车牌识别方法的性能。

1 车牌特点和传统的识别算法

1.1 车牌的特点

1.1.1 车牌形状规格

国内车牌种类较多,文中主要研究常用的民用车牌,车牌类型为单排车牌,如图 1 所示。



图1 标准车牌规格

车牌尺寸为长 440 mm,宽 140 mm 的矩形,车牌上的字符由 1 个汉字字符和 6 个大写英文字母/数字组成。中间一个小圆点分隔符将车牌分成区域和编码两部分,分隔符前面为 2 个字符,后面为 5 个字符,小圆点分隔符的直径为 10 mm。车牌中的每个字符大小为 45 mm×90 mm 的矩形区域^[7-8]。字符之间有 12 mm 的标准间隔(小圆点间隔符与前后字符间隔也为 12 mm),但实际车牌中,可能存在字符间隔大于 12 mm 的情况,因为有些字符不能填满整个矩形空间的宽度,例如字符“1”和其他字符。车牌左右边框与第一个字符和最后一个字符距离为 15.5 mm,上下边框与字符距离为 25 mm。

1.1.2 编号规则

车牌字符是由汉字、英文大写字母和数字组成。编号规则为,第一个字符是中国内地各省、自治区、直

辖市的汉字简称,共有 31 个,第二个字符是发证机构代码字符,使用英文大写字母 A-Z,第三到第七位由英文大写字母和数字组成,不包含英文大写字母 I 和 O,因为容易与数字 1 和 0 混淆。所以,大写字母为 24 个,数字为 10 个,共计 65 个字符。

1.2 传统的车牌字符识别方法

目前,传统的车牌字符识别方法主要有字符模板匹配法、字符特征统计法和统计机器学习方法等^[9-12]。

1.2.1 模板匹配法

该方法是在车牌识别前,预先获取各种车牌的标准字符模板,构建一个库。然后将待识别的车牌字符逐一与标准字符模板库比较,计算待识别字符与每个字符模板的匹配程度,匹配度计算采用字符图像对应的像素进行逻辑运行,匹配度最高的模板即为该字符的类别。由于该方法采用像素逻辑运算进行相似度的匹配,对于相似度较小的字符,如 Q 与 0,8 与 B 不能很好地分别。另外,在光照条件不好,较为模糊或磨损、断裂等情况下,识别率变低,所以,鲁棒性低。

1.2.2 特征统计法

字符特征统计方法比模板匹配方法检查更多的特征,该方法是提取车牌字符的一系列统计特征作为识别字符的判别特征,然后设计某种判别规则和决策函数,对待识别的字符进行分类。该方法的缺点与模板匹配法的缺点相似,即抗干扰能力低,性能不稳定。

1.2.3 机器学习方法

机器学习方法是将大量的车牌字符训练学习模型,学习输入输出映射模型,让分类器进行有监督学习。机器学习算法有 SVM、贝叶斯、决策树等。模型训练完成后,将待识别的字符输入模型,输出字符所属的类别。其缺点是这类机器学习方法通常需要进行大量车牌特征的抽取,抽取的特征不同,选用的分类器算法不同,往往结果也不完全相同。

综上,由于这些方法都各有其自身的不足,有的鲁棒性不强,有的需要抽取复杂的车牌特征,对经验依赖性高。为了改善这些问题,进一步提高识别准确率,文中将深度学习中的 LeNet-5 卷积神经网络模型进行改进后应用于车牌字符识别。

2 卷积神经网络

卷积神经网络是采用卷积运算进行特征提取的一类深度神经网络,其结构通常由卷积层和子采样层交替组成前端,后端输出采用全连接层结构。

2.1 卷积层

卷积层的主要任务是进行特征提取,运算方式是卷积运算,连接方法是局部连接。卷积层通常由多个特征平面组成,共同完成前一层的特征提取,每个特征

平面内神经元的连接权值相同(权值共享)^[13]。在特征平面中,每个神经元连接到前一个特征平面的特定区域中的神经元(局部感受野)。卷积计算过程是首先对输入的特征平面像素进行边界填充,以保证卷积核与输入像素尺寸匹配,然后将卷积核在整个特征平面上滑动一个固定的步长,从而构建新的特征平面^[14]。卷积层的连接关系表示为:

$$x_j^l = f(\sum_{i \in p_j} x_i^{l-1} k_j^l + b^l) \quad (1)$$

其中, p_j 是与 L 层中的神经元 j 相连的输入特征平面的局部感受野,为一矩形区域。每个卷积核仅是对输入局部感受野区域进行卷积运算,不改变像素之间的空间相关性; x_j^l 是第 L 层特征平面上的第 j 个神经元的值; x_i^{l-1} 是第 $L-1$ 层特征平面上的第 i 个神经元的值; k_j^l 是 L 层卷积核的权重; b^l 是 L 层的卷积偏移量; $f(\cdot)$ 是激活函数。

卷积运算完成后,通常采用 sigmoid 或 tanh 激活函数对结果进行变换。

2.2 子采样层

子采样层也称为池化层(pooling layer)或下采样层(subsampling)。该层的作用是降维,即降低输入图像的灵敏度。由于输入数据经过卷积层的卷积运算后,根据卷积核形成多个特征平面,是一个升维过程,需要随后进行适度的降维处理,否则,可能造成“维数灾”。因此,子采样层的功能是适当地对卷积层的输出进行降维处理^[15]。子采样层的每个特征平面上的神经元连接到卷积层的共享值区域,特征平面的数量

是不变的,通过子采样层的映射缩小比例,子采样过程可以由式 2 表示。采用子采样层处理,使得网络降低了图像位移翻转的灵敏度。

$$x_j^l = f(\text{pool}_{i \in p_j}(x_i^{l-1}) + b^l) \quad (2)$$

这里类似于卷积层, x_i^{l-1} 是 $L-1$ 层特征平面上神经元 i 的值, x_j^l 是 L 层特征平面上神经元 j 的值; b^l 是偏移量; $\text{pool}(\cdot)$ 是采样函数。

2.3 全连接层

卷积神经网络的前端是由若干卷积层和子采样层构成,后端输出部分由全连接层组成。全连接层完整地连接当前层和前一层的所有特征平面神经元^[16]。由于卷积层和下采样层是二维特征平面,而全连通层输出一维类别向量,所以网络中首个全连接层需要将二维特征平面转换为一维向量形式。全连接层通常采用多层感知器网络来实现,最后输出识别类别。

3 LeNet-5 网络和改进

3.1 LeNet-5 网络结构

LeNet-5 网络最早是由 YannLeCun 于 1998 年提出的一个卷积神经网络模型,主要用于手写数字识别,具有卷积神经网络的典型结构,是深度神经网络中最有代表性的实验系统之一,曾广泛用于美国银行手写数字的识别。网络结构由 7 层组成,每层包含训练参数,LeNet-5 中主要有卷积层、下抽样层、全连接层 3 种连接方式。如图 2 所示,其中 C_x 层代表卷积网络, S_x 层代表子采样层网络, F_x 层代表全连接层网络。

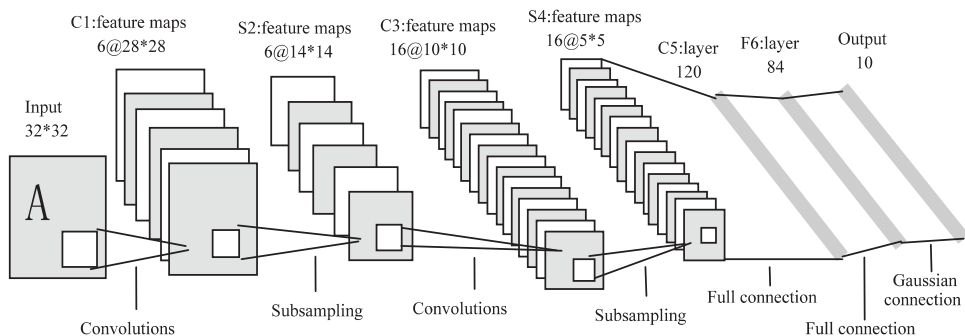


图 2 LeNet-5 卷积神经网络结构

C_1 层是卷积层,包含 6 个特征平面,每个特征平面由一个卷积核映射而形成。每个特征平面中的单元连接到输入层中的一个 5×5 的区域,即卷积的输入区域大小是 5×5 。该层的卷积运算使用 6 个 5×5 的卷积核与输入区域进行卷积运行。每个特征图平面内只使用一个共同卷积核,即共享权值。

S_2 层是下采样层,采用的是 2×2 的输入域, C_1 层的 6 个特征平面分别以 2×2 的区域进行采样,得到 6 个 14×14 的下采样平面。 C_1 层的 4 个单元作连接到 S_2 层的 1 个单元,作为 S_2 层的输入,输入区域滑动运

行,每次滑动 2 个像素。在这里,使用最大池化(maximum pool),大小为 2×2 的采样核。

C_3 层是一个由 16 个 10×10 大小的特征平面组成的卷积层。卷积运算与 C_1 层相同,不同之处在于, C_3 层的每个神经元与 S_2 层中的多个特征平面相连,即该层的每个神经元连接到 S_2 层多个特征平面的 5×5 大小的输入区域。 C_3 层输出是 16 个 10×10 的特征平面。

S_4 层是下采样层,采样核大小为 2×2 。 C_3 层的 16 个 10×10 的特征平面分别作为输入,以采样核为 2×2 的下抽样,仍然使用最大池化, S_4 层得到 16 个 5×5 的

特征平面。

C₅层是由 120 个卷积结果构成的卷积层。由于 S₄层的特征平面大小为 5×5,与本层的卷积核大小相同,这样经过卷积后,形成特征平面的大小为 1×1,所以就形成一维的卷积结果。这里采用 120 个神经元,每个神经元都连接到 S₄层的 16 个特征平面,这样就得到 120 个卷积结果。

F₆层是由 84 个神经元组成的全连接层,它与 C₅层的 120 个神经单元完全连接。

输出层 OUTPUT 是由 10 个神经元组成的全连接层,与 F₆层完全连接。每个神经元代表输出的一个类别,在这里分别代表数字 0 到 9。采用径向基 RBF 分类器连接 F₆层,RBF 的映射关系如式 3 所示。

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

其中, x_j 为 F₆层上的第 j 个神经元; y_i 为输出层的第 i 个神经元的输出值; w_{ij} 为神经元 i 到神经元 j 之间的偏置。输出层的每个节点 y_i 表示一个数字类别的置信度。

3.2 LeNet-5 模型的改进

在车牌识别中,由于车牌分割的输出由汉字和 6 位数字/字母组成,而传统 LeNet-5 网络主要用于数字 0-9 的手写数字识别。这明显是不同的,所以 LeNet-5 模型不能直接应用,需要根据中国车牌中字符类别等特点进行改进。

3.2.1 输出层的改进

由于在传统的 LeNet-5 模型中输出的数字只有 10 个,而车牌对象共有 65 个类别(31+24+10=65),考虑到车牌中的字母/数字与中文字体相差很大,所以设计两个分类器,一个用于汉字分类,一个用于字母数字的分类。这样,模型输出层的单元个数分别设计为 31 个和 34 个。

3.2.2 调整特征平面数

由于模型的识别数量修改为 31 和 34,比原来模型的输出类别大,而且有较为复杂的汉字识别,为了更全面地提取图像特征,需要调整一些层的特征平面的数量。因此,将 C₁层和 S₂层的特征平面数量由 6 个增加到 24 个,将 C₃和 S₄层的特征平面数量由 16 个增加到 52 个,将 C₅层的特征平面数量由 120 个增加到 480 个。调整各层参数,将对输出的结果产生影响。

3.2.3 改进模型输入大小

由于传统模型中为 32×32 的输入,当某些汉字图像被压缩时,可能会造成某些复杂的汉字信息丢失,如“藏”、“赣”、“鄂”等。因此,模型输入修改为 64×64 大小,从而相应地调整每个层的特征平面神经元数量。

3.3 网络实现和训练

采用 TensorFlow 框架用于实现改进后的 LeNet-5

网络模型。将车牌字符分割后,得到的汉字和数字/字符图像作为训练数据。部分训练样本如图 3 所示。



图3 部分训练数据样本

为了增加数据量,部分车牌数据通过图像处理变形获得。共获得汉字图像 35 460 个,类别 31 个,平均每个类别 1 143 个;数字/字符图像 32 100 个,类别 34 个,平均每个类别 944 个。随机抽取 80% 的数据作为训练样本,其余 20% 作为测试样本。所有数据标准化为 64×64 的二值图像。分别为每个分类字符设置标签,汉字 31 个标签,从 0 到 30,数字/字符 34 个标签,从 0 到 33。在 CPU 为 i5-5200U 上训练模型,进行有监督学习。在训练中设置不同的迭代次数,在 2 000 ~ 10 000 之间训练模型,直到收敛为止。

4 算法测试和性能分析

4.1 性能测试实验

为了测试改进后的 LeNet-5 网络的性能,通过实验进行测试。将数据集的 80% 作为训练模型数据,20% 作为测试数据。测试样本数据由 6 420 个数字/字母图像和 7 092 个汉字图像组成。分别与支持向量机(SVM)和 BP 神经网络进行了对比实验。

支持向量机也分别设计 2 个分类器进行训练,一个用于汉字分类,一个用于数字/字母分类。支持向量机使用的特征包括笔画、密度和网格特征,在相同的实验环境下进行实验,结果如表 1 所示。

表1 改进的 LeNet-5 与 SVM 实验结果

算法	汉字准	字符/数字	100 字符
	准确率/%	准确率/%	耗时/s
改进的 LeNet-5	98.53	98.85	0.375
SVM	91.65	94.58	0.198

从表 1 可以看出,采用卷积神经网络可以显著地提高车牌中的字符识别准确率,汉字分类准确率可以提高约 7%,字符/数字准确率可以提高约 4%。使用支持向量机方法,平均识别 100 个字符的时间为 0.198 s,卷积神经网络的耗时为 0.375 s,由于卷积网络需要对车牌字符图像进行卷积运算,故耗时较多。

在相同的实验环境下,改进的 LeNet-5 网络和 BP

神经网络分别进行对比实验测试,结果如表 2 所示。

表 2 改进的 LeNet-5 和 BP 网络实验结果

算法	汉字准确 率/%	字符/数字 准确率/%	100 字符 耗时/s
改进的 LeNet-5	98.53	98.85	0.375
BP 神经网络	92.77	93.86	0.582

为了测试改进的 LeNet-5 性能,将常用的 BP 神经网络进行对比实验。BP 网络设计为三层结构,隐含层为 450 个节点,学习率设为 0.01。从表 2 可以看出,改进的 LeNet-5 网络识别准确率比 BP 网络高出约 6%。时间性能方面,在相同实验环境配置条件下,改进的 LeNet-5 识别 100 个字符为 0.375 s,BP 神经网络需要 0.582 s,改进的 LeNet-5 网络所需时间也少于 BP 网络。

4.2 识别错误样本分析

在车牌字符识别测试中,出现一些识别错误的情况。分析这些错误识别的字符原因,有如下特点:(1)车牌中某些字符本身就比较模糊,因为光照、倾斜、变形、粘连、分裂等原因所致;(2)车牌分割时,由于噪音或其他原因的干扰引起的字符分割错误,如“京”与“琼”,“1”与“T”等,“0”与“Q”等。对于这些识别错误,部分可以通过加大训练样本来加以改善,部分可以通过提高字符分割正确率等预处理方法来加以改善。

5 结束语

为了改善传统车牌识别方法中过分依赖车牌特征和鲁棒性不强等问题,对卷积神经网络 LeNet-5 的结构进行了改进,并进行了性能测试实验,与三层 BP 神经网络和 SVM 进行了对比。结果表明,改进的 LeNet-5 在车牌字符识别中有良好的识别性能,与三层 BP 神经网络相比,改进的 LeNet-5 在准确率和识别速度方面都优于 BP 神经网络;与 SVM 相比,识别准确率高 于 SVM,识别速度慢于 SVM。由于实验中没有采用 GPU,若在实际应用部署时使用 GPU 加速,运行速度会有很大程度的改善。改进的 LeNet-5 卷积神经网络具有优异的鲁棒性、准确性和较好的实时性。下一步将进一步完善数据预处理及车牌字符分割算法,构建性能优良的车牌识别系统。

参考文献:

[1] BENGIO Y, DELALLEA O. On the expressive power of deep architectures[C]//Proceedings of the 22nd international conference on algorithmic learning theory. Espoo, Finland:Springer-Verlag,2011:18-36.

[2] AREL I,ROSE D C,KARNOWSKI T P. Research frontier: deep learning - a new frontier in artificial intelligence research[J]. IEEE Computational Intelligence Magazine,2010,5(4):13-18.

[3] RUSSAKOVSKY O,DENG Jia,SU Hao,et al. Image net large scale visual recognition challenge [J]. International Journal of Computer Vision,2015,115(3):211-252.

[4] BENGIO Y. Learning deep architectures for AI[J]. Foundations and Trends in Machine Learning,2009,2(1):1-127.

[5] VINCENT P,LAROCHELLE H,LAJOIE I,et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. Journal of Machine Learning Research,2015,11:3371-3408.

[6] 曲之琳,胡晓飞. 基于改进激活函数的卷积神经网络研究 [J]. 计算机技术与发展,2017,27(12):77-80.

[7] 虞力英.《中华人民共和国机动车号牌》标准修订内容理解与实施[J]. 道路交通科学技术,2015(1):47-51.

[8] 中华人民共和国公安部. 中华人民共和国机动车号牌 [S]. 北京:中华人民共和国公安部,2007.

[9] BOLOTOVA Y A,DRUKI A A,SPITSYN V G. License plate recognition with hierarchical temporal memory model [J]. International Forum on Strategic Technology,2014,12(2):136-139.

[10] 周成伟. 基于卷积神经网络的自然场景中数字识别[J]. 计算机技术与发展,2017,27(11):101-105.

[11] JAGANNATHAN J,SHERAJDHEEN A,Deepak M V, et al. License plate character segmentation using horizontal and vertical projection with dynamic thresholding[C]//International conference on emerging trends in computing, communication & nanotechnology. Tirunelveli, India: IEEE,2013: 700-705.

[12] WANG Jianxia,ZHOU Wanzhen,XUE Jingfu,et al. The research and realization of vehicle license plate character segmentation and recognition technology [C]//International conference on wavelet analysis & pattern recognition. [s. l.]:[s. n.],2010:101-104.

[13] SCHERER D,MILLER A,BEHNKE S. Evaluation of pooling operations in convolutional architectures for object recognition[C]//Proceedings of 20th international conference on artificial neural networks. Thessaloniki, Greece: Springer Verlag;2010:92-101.

[14] KRIZHEVSKY A,SUTSKEVER I,HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems,2012,25(2):1097-1105.

[15] DAHL G E,SAINATH T N,HINTON G E. Improving deep neural networks for LVCSR using rectified linear units and dropout[C]//International conference on acoustics, speech& signal processing. Vancouver:[s. n.],2013:8609-8613.

[16] WANG H,CAI Y F,CHEN L. A vehicle detection algorithm based on deep belief network[J]. Scientific World Journal, 2014,2014:647380.