

文章编号: 2095-2163(2019)02-0054-04

中图分类号: TP391

文献标志码: A

基于卷积神经网络的手写数字识别系统的设计

吕红

(徐州工业职业技术学院 信息与电气工程学院, 江苏 徐州 221000)

摘要: 当前卷积神经网络应用于手写数字的识别已成为研究的热点之一。本文在 Matlab 环境下输入手写数字图片, 然后对图片进行灰度化、二值化、反色、去噪、分割和大小归一化预处理, 通过卷积神经网络经典模型 LeNet-5, 对比 3 种数据集: MNIST 数据集、MNIST 数据集训练+自建数据集调精和自建数据集训练卷积神经网络的实际识别效果, 选择自建的数据集进行卷积神经网络训练, 在训练好的卷积神经网络中手写体数字图片取得了较好的识别效果。

关键词: 手写体数字; 深度学习; 卷积神经网络

The design of handwritten digital recognition system based on Convolutional Neural Network

LV Hong

(School of Information and Electrical Engineering, Xuzhou College of Industrial Technology, Xuzhou Jiangsu 221000, China)

[Abstract] Now, the application of Convolutional Neural Network in handwritten number recognition has become one of the research hotspots. In this paper, handwritten digital images are inputted in Matlab environment, and the images are preprocessed by graying, binarization, inverse color, de-noising, segmentation and size normalization. The classical model of Convolutional Neural Network Lenet - 5 is used, and the comparison is provided of the actual recognition results of three kinds of data sets: MNIST training set, MNIST training set + own training and own training set. The proposed training set is selected for Convolutional Neural Network training, and for handwritten number pictures, good recognition results have been achieved in the trained Convolutional Neural Network.

[Key words] handwritten number; deep learning; Convolutional Neural Network

0 引言

自 2006 年随着单隐层神经网络到深度神经网络模型的发展, 世界人工智能迎来了新一轮的研究热潮。人工智能化的应用已在逐步改变人们的日常生活。在互联网大数据异常活跃的时代, 人们需要进行很多关于数据类的工作, 比如数据统计、发票税单、银行支票、快递分拣、电脑阅卷等, 如何利用设备自动化、智能化, 高效地识别数字和字符, 提高工作效率则已成为当前亟待解决的研究问题。卷积神经网络(Convolutional Neural Networks, CNN) 作为一类包含卷积计算且具有深度结构的前馈神经网络^[1], 其经典模型 LeNet-5 在识别手写数字方面表现优异。手写数字分为 2 种: 实时手写数字和脱机手写数字(即数字图片)。本文研究的是脱机手写数字的识别, 考虑到阿拉伯数字本身字形信息量小、不同的人写法千差万别, 再加上输入的只是一张图片, 没有上下文的联系, 因此对其进行快速、精确的

识别将具有更高的挑战性。

1 卷积神经网络 LeNet-5 模型

经典的 LeNet-5 网络模型, 是最早的应用于手写数字识别的卷积神经网络^[2], 有着最广泛的用途和区别于其它网络的独特优势。该模型包括 1 个输入层、2 个卷积层、2 个池化层(子采样)和全连接以及输出层, 模型结构如图 1 所示。

在本文数字图像识别中用到的卷积是二维卷积核与二维图像做卷积操作^[3], 就是卷积核滑动到二维图像上所有位置, 并在每个位置上与对应的像素点做内积。一般包括 Full 卷积、Same 卷积和 Valid 卷积三种。其核心是可以减少不必要的权值连接, 引入稀疏或局部连接带来的权值共享策略大大地减少参数量, 从而可以避免过拟合现象的发生; 此外, 由于卷积操作具有平移不变性, 使得学到的特征具有拓扑对应性、鲁棒性的特征。本文的 LeNet-5 中采用的是 Valid 卷积。对应数学公式可表示为:

作者简介: 吕红(1980-), 女, 副教授, 主要研究方向: 软件开发、智能算法。

收稿日期: 2018-12-21

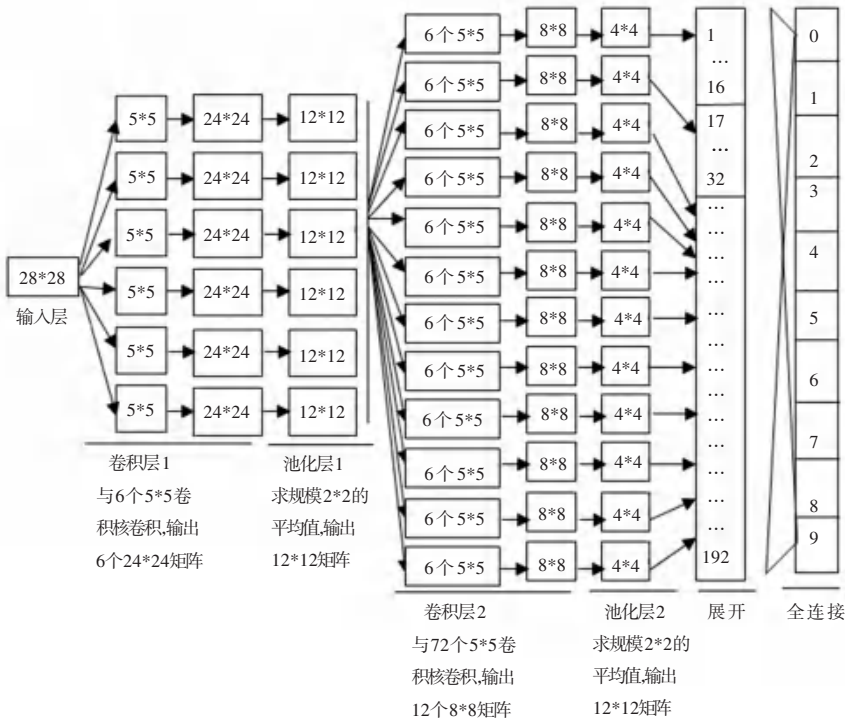


图 1 LeNet-5 模型

Fig. 1 LeNet-5 model

$$\begin{aligned}
 \mathbf{y} &= \text{conv}(\mathbf{x}, \mathbf{w}', \text{valid}') = (y(1), \dots, y(t), \dots, \\
 & \quad y(n - m + 1)) \in R^{n-m+1}, \\
 \mathbf{y}(t) &= \sum_{i=1}^m \mathbf{x}(t + i - 1) \mathbf{w}(i).
 \end{aligned}
 \tag{1}$$

其中, \mathbf{y} 是输出矩阵, 是 $(n - m + 1) * (n - m + 1)$ 的, 同时也是输入 $n * n$ 矩阵 \mathbf{x} 与卷积核 $m * m$ 矩阵 \mathbf{w} 做 valid 卷积的结果, 且 $n > m$ 。每一个元素 $y(t)$ 等于 $\mathbf{x}(t + i - 1)$ 和 $\mathbf{w}(i)$ 相乘 ($1 \leq i \leq m$), 然后相加的和。由图 2 可以更直观地看到 Valid 卷积的操作过程。

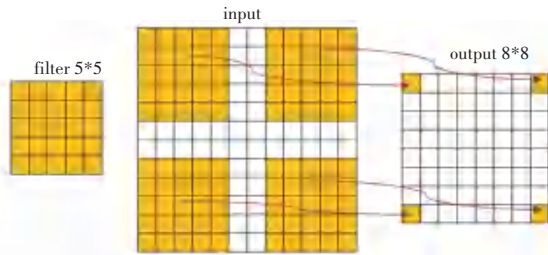


图 2 Valid 卷积操作

Fig. 2 Valid convolution operation

池化操作^[4]利用数字图像各局部相关性, 在保留有用信息前提下, 大幅度减少下一层的输入维度, 有效控制过拟合风险。池化操作有多种形式, 例如

最大池化、平均池化、范数池化和对数概率池化等, 常用的池化方式为最大池化和平均池化, 本项目设计中使用的是平均池化。平均池化是不用重叠的 $2 * 2$ 矩形框将输入矩阵分成不同的区域, 对每个矩形框的数取平均值作为输出矩阵的一个元素。平均池化的特点和优势在于提取均值进行数据压缩。图 3 即演示了 $8 * 8$ 的图像特征矩阵通过 $2 * 2$ 池化层后得到 $4 * 4$ 矩阵。

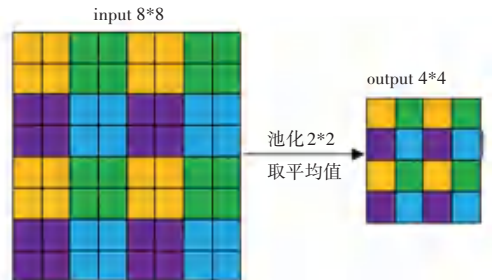


图 3 池化操作

Fig. 3 Pooling operation

2 手写数字识别系统的设计

该识别系统首先通过 LeNet-5 模型训练卷积神经网络, 然后通过已经训练好的卷积神经网络进行

特征提取^[5],最后判断特征值输出识别的结果。手写数字图片识别系统的总体设计框架如图4所示。

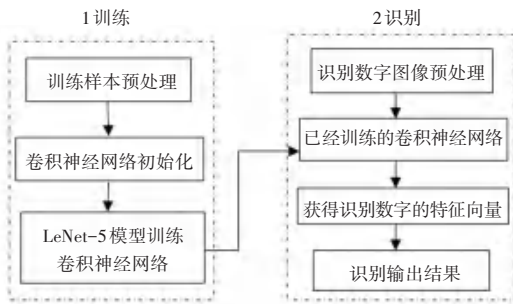


图4 系统框架

Fig. 4 System framework

在系统里需要处理生成训练样本和要识别的手写数字图片,这都需要对图片进行灰度化、二值化、反色、去噪、分割和大小归一化预处理^[6],这里的图片可以是一个数字,也可以是多个数字。

首先,打开需要识别的手写体数字图片,获得对应的二进制图片数据,并将图像灰度化和二值化,如图5所示。

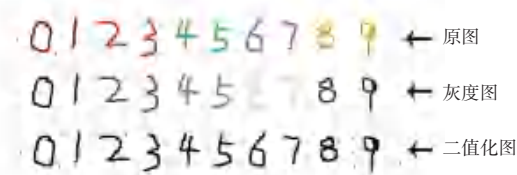


图5 原图、灰度图和二值化图

Fig. 5 Original, grayscale and binary images

其次,对图像进行反色和去噪处理,如图6所示。反色是为了识别图像方便,而去噪则是去掉图片中较大的噪声,从而提高识别的正确率。

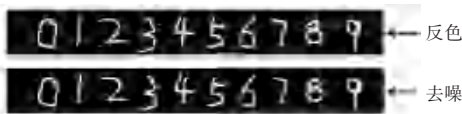


图6 反色和去噪

Fig. 6 Anti-color and de-noising

图像中如果包含多个手写数字,那么就需要对图像进行分割,把每个数字独立地分割出来。这里采用先从上而下、再从下而上扫描图片,找到第一个白色像素点,这样就可以确定手写数字的高度范围;然后在这个范围内从左向右逐列扫描,遇到第一个白色的像素点时认为是一个字符分割的起始位置,直至遇到某一列中没有白色像素点,则认为这个字符的分割结束位置,在此过程中要保存下起始和结束的位置。如此反复,直至扫描至图像的最右端。同理,按照逐行扫描的方法获得每个数字的高度范

围。图7中,则用红色矩形框显示查找到的每个数字的精确位置。



图7 分割

Fig. 7 Division

最后,将得到的数字进行大小归一化处理,也就是将数字图像统一处理成 28×28 的大小,这样可以提高识别率,大小归一化之后的图像效果如图8所示。

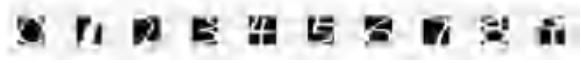


图8 大小归一化

Fig. 8 Size normalization

3 识别结果与分析

在识别的设计过程中,对卷积神经网络的训练采用了3种方式:采用MNIST数据集^[7](共10类,训练集60 000个、测试集10 000个)训练;采用MNIST数据集预训练,并利用自己创建的数据集(共10类,训练集100个,测试集100个)来对已经训练好的卷积神经网络进行调整;采用自己创建的数据集训练。训练次数和识别率见表1。

表1 不同训练集的识别结果

Tab. 1 Recognition results of different training sets %

训练次数	识别结果		
	MNIST 数据集训练	添加自建数据集训练	自建数据集训练
1 200	11	18	97
2 400	13	13	98
4 800	10	14	97
9 600	5	16	97
19 200	7	28	97

从表1中可以看出在实际应用中,使用MNIST训练集训练的卷积神经网络在实际识别中出现了过拟合,而添加自己创建的训练集调整训练出来的卷积神经网络可以得到更佳识别效果,但是效果也并未臻至理想,所以这里采用自己的数据集独立完成卷积神经网络的训练,训练在1 200次时就可以达到比较稳定的97%的识别效果。

4 结束语

本文通过图像去噪预处理,有效地滤除图片的噪声信息,然后通过大小归一化,调整图片信息与

(下转第62页)