

引用格式:宋晓茹,吴雪,高嵩,等.基于深度神经网络的手写数字识别模拟研究[J].科学技术与工程,2019,19(5):193-196

Song Xiaoru, Wu Xue, Gao Song, et al. Simulation Study on Handwritten numeral recognition based on deep neural network [J]. Science Technology and Engineering, 2019, 19(5): 193-196

基于深度神经网络的手写数字识别模拟研究

宋晓茹 吴雪 高嵩* 陈超波

(西安工业大学电子信息工程学院,西安 710021)

摘要 当前图像识别大多采用基于特征提取的传统机器学习方法与卷积神经网络的方法,但传统图像识别技术需要手动提取图片特征,而卷积神经网络对硬件要求高,训练时间长等。针对以上问题,提出基于深度神经网络模型的手写体图像识别方法,让机器自动学习特征,并在此基础上,通过改进成本函数,加入 dropout 防止过拟合,来提高手写数字识别的识别率。仿真实验对比结果表明,基于深度神经网络模型的方法比当前传统算法的识别率提高了 3.41%,有效解决了人工识别费力耗时问题,对手写数字的研究具有重要意义。

关键词 图像识别 特征提取 深度神经网络 dropout

中图法分类号 TP391.41; **文献标志码** A

随着科学技术与发展,计算机视觉已经广泛运用于图像理解、地图导航、医疗制药、无人机和无人驾驶汽车等领域。而这些应用的核心技术就是图像处理^[1]、图像定位^[2] 和图像分类^[3] 等视觉识别任务。同时由于图像本身的复杂性和不确定性,图像识别任务也一直是研究的一个挑战。传统的图像识别分为两步,首先是进行特征的选择,其次构造分类器。

特征选择是图像识别的关键一步,特征选择的好坏直接影响到分类效果的好坏,文献[4]对图像的颜色特征进行提取,然后使用粒子群算法来对 SVM 的参数进行优化。到目前为止,国外的特征提取方法已经日趋成熟并有了很多新的方法。例如文献[5]对于特征提取,使用了质心特征、对角特征,分区特征和基于峰值程度的特征,文献[6]介绍了像素惯性矩(PMI)和距离编码(DDC)两种独特的特征提取方法。然而,这些方法的普遍缺点是需要人为手动的选取特征。用于图像分类的方法也有很多,文献[7]提出采用支持向量机对结构损伤进行识别,但支持向量机缺点是选取核函数的方法难以确定。文献[8]采用基于类 Haar 特征和 adabost 的分类方法,对车辆进行识别

研究,但此方法的缺点是对离群点非常敏感。此外,近几年来,深度学习在图像识别与检测方面也取得了瞩目的成绩,文献[9,10]提出了卷积神经网络在图像识别中的应用,文献[11]提出了一种基于递归神经网络(RNN)的汉字识别方法,这两种方法的优点是不需要人为手动提取特征,但是缺点是所需硬件配置高,若在一般的 CPU 上进行训练,花费的时间是很长的。

基于此,为了避免手动提取特征的复杂性和盲目性^[12],使用普通的 CPU 就可以进行训练的情况下,借助深度神经网络学习的方法,设计了一种深度神经网络模型,将传统的三层的 BP 神经网络扩展到 5 层,让机器能够自动学习出深层次的更有利于表达图片本身的更健硕的特征。并在此基础上为了使网络能够加快学习的速度,将均方差成本函数改进为交叉熵成本函数,针对因网络结构层数多带来的过拟合问题,采用 dropout 的方法来减少过拟合。在 Mnist 图像数据集上的实验结果表明,该方法相比于传统神经网络的方法具有更好的准确率。

1 深度神经网络图像识别模型的设计

深度学习的概念由 Hinton 等于 2006 年提出,已被成功应用到特征提取,图片分类等图像识别任务中。深度神经网络能够逐层提取图片的特征,一级一级学习出更加复杂的特征,因此很自然地想到可以用 DNN 模型来进行手写体识别的任务。基于手写数字库,采用的识别模型如图 1 所示。其中深

2018 年 10 月 9 日收到 国家重点研发计划(2016YFE0111900)和

陕西省重点研发计划(2018KW-022,2017KW-009)资助
第一作者简介:宋晓茹,女,汉族,博士,副教授。E-mail:545313193@qq.com。

*通信作者简介:高嵩,男,汉族,博士,教授。E-mail:602507619@qq.com。

度神经网络第一层 784 个神经元, 隐含层总共有三层, 第一个隐含层的单元数 500, 第二个隐层的单元数为 300, 第三个隐藏层的单元数为 300, 输出层 10 个神经元, 神经元的激励函数选为 sigmoid 函数, 输出层采取 softmax。并在深度神经网络中将成本函数改为交叉熵函数和加入 dropout 来防止过拟合。使用深度神经网络进行数字识别分为两步: 训练阶段和测试阶段。

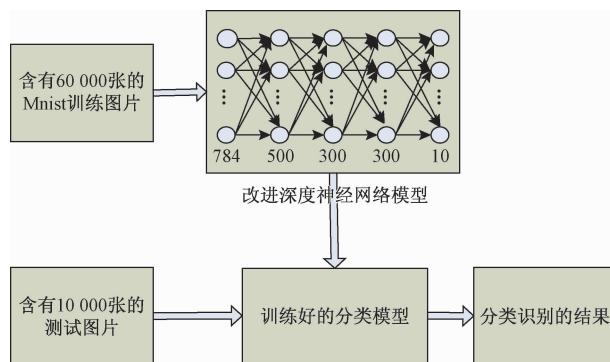


图 1 深度神经网络识别模型

Fig. 1 Deep neural network identification model

1.1 传统神经网络模式识别的方法

神经网络结构是一种前馈网络, 首先按照前向传播逐层计算, 每一层神经元的输入为上一层输出的加权和, 直到计算到输出层为止。然后将输出值与期望值进行比较, 衡量网络性能的好坏可用均方差误差成本函数, 其定义为

$$E_p = \frac{1}{2n} \sum_{i=1}^n e_i^2 = \frac{1}{2n} \sum_{i=1}^n \|y - a\|^2 \quad (1)$$

式(1)中, n 为总的训练样本个数, a 为网络输出向量, y 为训练的标签向量。

然后进入网络的反向传播, 根据微积分的链式求导原则, 采用梯度下降法, 用误差调整网络的各个层的权值, 调整权值的学习算法为

$$w_{ij(k+1)} = w_{ij(k)} + \Delta w_{ij} = w_{ij(k)} - \eta g_k \quad (2)$$

式(2)中, $w_{ij(k+1)}$ 是更新后的权重, $w_{ij(k)}$ 是当前的权重值, η 是学习率, g_k 是当前梯度。

1.2 深度神经网络代价函数的改进

为了让神经网络能够学习得更快, 使用交叉熵成本函数来代替均方差成本函数。分析之前的深度神经网络均方差误差成本函数公式 [式(1)], 假设只有一个训练样本 ($n = 1$), 且只有一个神经元, σ 代表 sigmoid 函数, 则成本函数变为

$$c(w, b) = \frac{1}{2}(y - a)^2 \quad (3)$$

式(3)中, $a = \sigma(z)$, $z = wx + b$, 由于采用的是梯度下降法, 因此分别对 w 和 b 求偏导数, 可得:

$$\frac{\partial c}{\partial w} = (a - y)\sigma'(z)x \quad (4)$$

$$\frac{\partial c}{\partial b} = (a - y)\sigma'(z) \quad (5)$$

对于 sigmoid 函数, 曲线图如 2 所示, 当神经元输出接近于 1 或者接近于 0 时, 曲线很平缓, 导致导数 $\sigma'(z)$ 接近于 0, 使得对 w 和 b 的偏导数接近于 0, 这样会使得网络更新很慢。

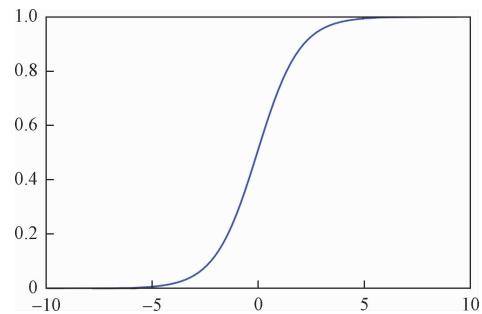


图 2 sigmoid 函数曲线

Fig. 2 Sigmoid function curve

定义交叉熵成本函数(假设为两分类问题):

$$c = -\frac{1}{n} [y \ln a + (1 - y) \ln(1 - a)] \quad (6)$$

由式(6)得: 当期望值 y 与神经网络输出相等时, $c = 0$ 。计算其对权重的偏导数:

$$\frac{\partial c}{\partial w} = \frac{1}{n} x_j \sum_x [\sigma(z) - y] \quad (7)$$

由式(7)可以看出权重的学习主要取决于 $\sigma(z) - y$, 也就是输出值与期望值的误差。此成本函数优点是: 当偏差大的时候权重更新得快, 偏差小时权重更新的小。因此, 代价函数选取为交叉熵代价函数。

1.3 防止过拟合方法的实现

在训练集上的表现好, 但是不能泛化到测试集上, 这是深度神经网络经常出现的过拟合问题。基于此, 在深度神经网络中加入 dropout 的方法来防止过拟合。

dropout 工作原理的本质是对神经网络模型结构的改变, 然后计算其多个模型的平均值。dropout 工作时以一定的概率 p 使得一些神经元暂时不输出, 然后在剩下的概率为 $1 - p$ 的神经元网络中进行学习训练。之后, 恢复所有的神经元, 再进行下一次训练时, 又以概率 p 随机选择一些神经元暂时不输出, 一直重复此过程。使用 dropout 方法相当于每次训练都选择了不一样的网络结构, 这样减少了对神经元相互间的适应性和依赖性。加入 dropout 方法后会使得网络学习到更加健硕和具有鲁棒性的特征, 现选取经验值 0.7。

2 实验结果与分析

2.1 实验数据

实验采用的是目前国际上常用到的 Yann LeCun 等建立的 Mnist 数据集库,它总共包含 70 000 张图片,其中 60 000 张训练图片和 10 000 张测试图片,训练集与测试集图片不重复。每张图片都是 28×28 像素,都有对应的标签。本次实验的 CPU 参数为: Intel(R) Core(TM) i5-5200 CPU @ 2.20 GHz 2.20, 内存为 4.00 GB, 操作系统为 windows10 64 bit, 使用 tensorflow 作为实验仿真平台进行仿真训练。

2.2 各模型的测试对比分析

基于手写数字库,选取第一层 784 个神经元,隐含层总共有三层,第一个隐含层的神经元数 500,第二个隐层的神经元数为 300,第三个隐藏层的神经元数为 300,神经元的激励函数采用 sigmoid 函数,输出层有 10 个神经元,采用 softmax 函数概率输出。实验中的学习率取为 0.001,误差函数选为交叉熵函数,加入 dropout 项,keep_prob 值为 0.7。批量的个数为 100,采用截断正态分布来初始化网络的各项参数,总共迭代 30 次并对对该网络性能进行测试。其中的 loss 曲线如图 3 所示,相应的分类正确率如图 4 所示,可以看到,在迭代了 30 步之后,正确率在 0.979 左右。

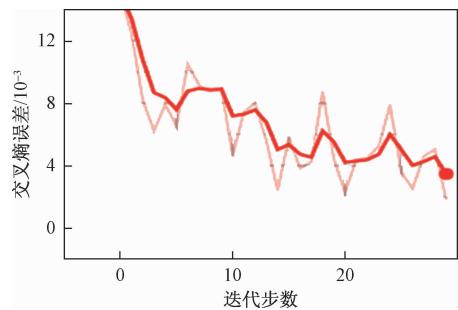


图 3 加入 dropout 和改进成本函数的 loss 曲线

Fig. 3 Add dropout and improve the loss curve of cost function

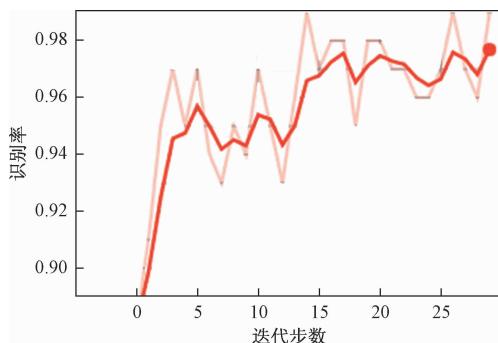


图 4 加入 dropout 和改进成本函数的识别率曲线

Fig. 4 Add dropout and improve cost function recognition rate curve

为了测试比较改进深度神经网络的性能,使用传统的神经网络、未改进深度神经网络进行仿真实验对比。在传统神经网络模型中,隐含层只有一层,使用的是均方差误差函数,未加入 dropout 方法。仿真结果如图 5 所示。使用未改进深度神经网络(即只是将传统三层的神经网络的深度增加到五层)的仿真结果如图 6 所示。从两幅图可以看出,传统神经网络的识别率大概在 0.94 左右,而在不改变成本函数和没有未 dropout 方法的五层深度网络中的识别率大概在 0.97 左右,提高了 0.03 左右。

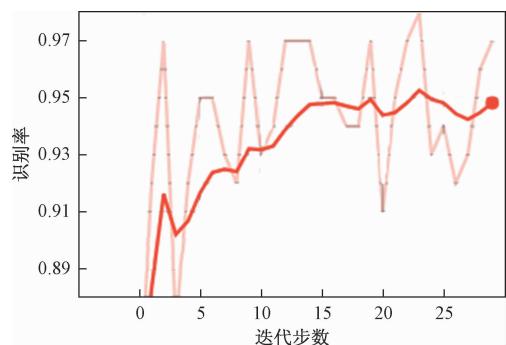


图 5 传统神经网络的识别率

Fig. 5 Recognition rate of traditional neural network

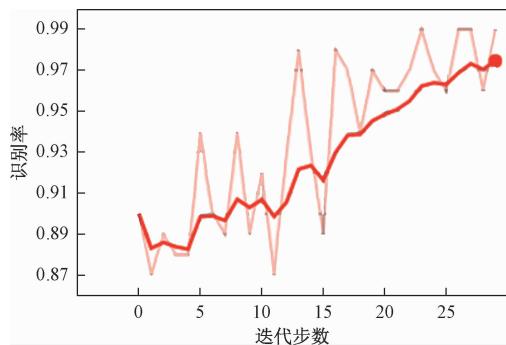


图 6 改进之前的深度网络的识别率

Fig. 6 Improve the recognition rate of the previous deep network

实验仿真的对比结果如表 1 所示。

表 1 Mnist 数据集上不同方法的实验结果对比

Table 1 Comparison of experimental results of different methods on Mnist data set

模型方法	正确率/%
传统神经网络	94.49
改进之前的深度网络	97.52
改进成本函数和加入 dropout 方法	97.90

3 结论

基于深度神经网络的方法对手写体数字进行识别研究,该方法既避免了传统神经网络手工提取特征的过程和对先验知识的依赖性,又避免了卷积神

经网络或者 LSTM 网络需要硬件设备高和训练时间长的缺点。同时在代价函数,防止过拟合两个方面对模型进行了优化,并将其与传统的神经网络进行比较,通过实验对测试数据进行仿真,结果表明,深度学习及其各种改进方法,对手写体数字的识别具有比传统神经网络更加快速准确的识别效果,在识别率和实时性方面得到了有效的提高。

参 考 文 献

- 1 Waris M A, Iosifidis A, Gabbouj M. Object proposals using CNN-based edge filtering[C]// International Conference on Pattern Recognition. New York:IEEE, 2017:627-632
- 2 Zhang P, Zhuo T, Huang W, et al. Online object tracking based on CNN with spatial-temporal saliency guided sampling[J]. Neurocomputing, 2017, 257:115-127
- 3 Smirnov E A, Timoshenko D M, Andrianov S N. Comparison of regularization methods for image net classification with deep convolutional neural networks[J]. Aasri Procedia, 2014, 6(1): 89-94
- 4 韩晓艳,赵东. 基于粒子群的支持向量机图像识别[J]. 液晶与显示, 2017, 32(1): 69-75
Han Xiaoyan, Zhao Dong. Image recognition based on particle swarm support vector machine[J]. Liquid Crystal and Display, 2017, 32 (1): 69-75
- 5 Kumar M, Jindal M K, Sharma R K, et al. Offline handwritten numeral recognition using combination of different feature extraction techniques[J]. National Academy Science Letters, 2018, 41 (1): 29-33
- 6 Prasad B K, Sanyal G. Multiple hidden markov model post processed with support vector machine to recognize english handwritten numerals [J]. International Journal on Artificial Intelligence Tools, 2018, 27 (5):1-19
- 7 张颖,戚世雷,方有亮. 基于支持向量机的钢框架结构损伤定位试验研究[J]. 科学技术与工程, 2018, 18(7): 25-30
Zhang Ying, Qi Shilei, Fang Youliang. Experimental study on damage location of steel frame structure based on support vector machine [J]. Science Technology and Engineering, 2018, 18(7): 25-30
- 8 张雪芹,方婷,李志前,等. 基于类 Haar 特征和 AdaBoost 的车辆识别技术[J]. 华东理工大学学报(自然科学版), 2016, 42 (2): 260-265
Zhang Xueqin, Fang Ting, Li Zhiqian, et al. Vehicle identification technology based on similar Haar features and AdaBoost[J]. Journal of East China university of Science and Technology (Natural Science Edition), 2016, 42(2): 260-265
- 9 段萌,王功鹏,牛常勇. 基于卷积神经网络的小样本图像识别方法[J]. 计算机工程与设计, 2018, 39(1): 224-229
Duan Meng, Wang Gongpeng, Niu Changyong. A small sample image recognition method based on convolution neural network [J]. Computer Engineering and Design, 2018, 39(1): 224-229
- 10 常亮,邓小明,周明全,等. 图像理解中的卷积神经网络[J]. 自动化学报, 2016, 42(9): 1300-1312
Chang Liang, Deng Xiaoming, Zhou Mingquan, et al. Convolution neural network in image understanding[J]. Journal of Automation, 2016, 42(9): 1300-1312
- 11 Zhang X Y, Yin F, Zhang Y M, et al. Drawing and recognizing Chinese characters with recurrent neural network[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 99: 849-862
- 12 傅天驹,郑嫦娥,田野,等. 复杂背景下基于深度卷积神经网络的森林火灾识别[J]. 计算机与现代化, 2016(3): 52-57
Fu Tianju, Zheng Chang'e, Tian Ye, et al. Identification of forest fire based on deep convolutional neural network in complex background[J]. Computer and Modernization, 2016(3): 52-57

Simulation Study on Handwritten Numeral Recognition Based on Deep Neural Network

SONG Xiao-ru, WU Xue, GAO Song*, CHEN Chao-bo

(School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China)

[Abstract] The traditional machine learning method based on feature extraction and the convolutional neural network method are mostly used in image recognition at present. However, the traditional image recognition technology requires manual extraction of image features, while convolutional neural network has high requirements on hardware and long training time. In view of the above problems, a handwritten image recognition method was proposed based on the deep neural network model, which enables the machine to learn the features automatically. On this basis, it improves the recognition rate of handwritten numeral recognition by improving the cost function and adding dropout to prevent overfitting. The results of simulation experiment show that the method based on the deep neural network model has improved the recognition rate by 3.41% compared with the current traditional algorithm, which effectively solves the problem of labor and time consuming in manual recognition.

[Key words] image recognition feature extraction deep neural network dropout