

混合PSO优化卷积神经网络结构和参数

唐贤伦, 刘庆, 张娜, 周家林

(重庆邮电大学工业物联网与网络化控制教育部重点实验室 重庆 南岸区 400065)

【摘要】为了使卷积神经网络在非经验指导下自动寻得最优连接,并提高其参数优化效率,提出用粒子群优化卷积网络参数,并用离散粒子群优化卷积网络特征图之间连接结构的新方法。先使用粒子群优化所有权值,再采用离散粒子群优化降采样层和卷积层之间特征图连接结构。将该方法用于MNIST数据集和CIFAR-10数据集,实验结果表明,相比其他连接结构的卷积神经网络和其他识别方法,该方法可以有效实现网络结构及参数的优化,加速网络收敛并提高识别准确率。

关键词 卷积神经网络; 离散粒子群优化; 手写字符识别; 粒子群优化; 结构优化
中图分类号 TP183 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2018.02.011

Optimizing Structure and Parameters of Convolutional Neural Networks Using Hybrid PSO

TANG Xian-lun, LIU Qing, ZHANG Na, and ZHOU Jia-lin

(Key Laboratory of Industrial Internet of Things & Networked Control, Ministry of Education,
Chongqing University of Posts and Telecommunications Nan'an Chongqing 400065)

Abstract In order to make convolutional neural network get optimal connection automatically without experienced guidance and improve the optimizing effectiveness for parameters of convolutional neural network, a new method using both particle swarm optimization algorithm and discrete particle swarm optimization algorithm is proposed to optimize parameters and feature maps connecting structure of convolutional neural network. The particle swarm optimization is applied to optimize the weights of convolutional neural network at first, and then the discrete particle swarm optimization is applied to optimize feature maps connections between sub-sampling layer and convolutional layer. The method is applied to MNIST database and CIFAR-10 database, compared to convolutional neural networks of other connecting structures and other recognition methods, results shown that this method can optimize the parameters and structure of the network effectively, accelerate network convergence and improve the recognition accuracy.

Key words convolutional neural network; discrete particle swarm optimization; handwritten character recognition; particle swarm optimization; structural optimization

深度学习可以自动提取特征并将特征进行分类,因此得以广泛应用^[1-3]。作为深度学习方法之一的卷积神经网络(convolutional neural networks, CNNs),在视频人体动作识别^[4]等领域已得到成功应用。

文献[5]把卷积神经网络第一次成功应用在手写字符识别领域。在特征图连接问题上,传统CNNs第一降采样层与第二卷积层之间特征图全连接,对于经典的LeNet^[6]结构,第一降采样层和第二卷积层之间特征图连接也是由人为经验决定的特定连接。虽然利用粒子群优化算法^[7]和遗传算法等智能算法^[8]训练相对简单的神经网络已经取得了很好的结果,把粒子群应用到光学字符识别应用中也取得了显著

成效^[9],但是这些算法训练的网络特征图之间都是经验指导的固定连接结构,而全连接结构使得连接数量过多,降低网络的运算速率,同时全连接结构使得网络结构对称,不利于提取不同的特征;而特定的连接结构又受到经验的影响,不能普遍适用。

本文提出一种将粒子群优化卷积网络参数和离散粒子群优化^[10]卷积神经网络特征图连接结构相结合的新方法,使用粒子群预训练参数,并使用离散粒子群优化第一采样层和第二卷积层特征图之间结构连接,使得网络在非经验指导下自动寻得最优连接。网络结构在传统结构基础上增加一个全连接层,非线性函数采用指数激活单元(exponential linear

收稿日期: 2016-09-21; 修改日期: 2017-03-25

基金项目: 国家自然科学基金(60905066); 重庆市教委科学技术研究项目(KJ1500401)

作者简介: 唐贤伦(1977-),男,博士,教授,主要从事计算机智能方面的研究。

unit, ELU)^[11], 实验证明, 将经过参数和结构优化后的网络应用到手写数字识别应用中可以达到较理想的识别效果。

1 卷积神经网络

权值共享、局域感受野和降采样是卷积神经网络的三大特征, 可以实现识别图像的缩放、位移和扭曲不变性。图1为本文构建网络。

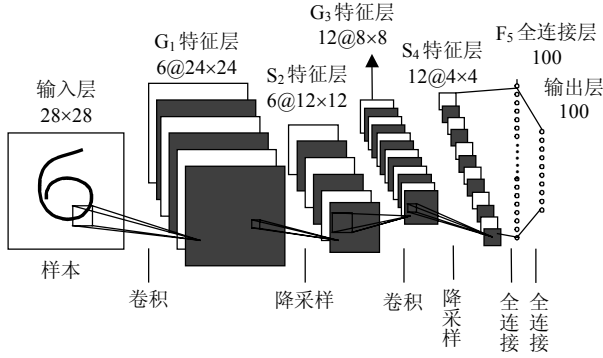


图1 卷积神经网络结构图

本文采用非线性激活函数指数激活单元(exponential linear unit, ELU)^[11]来加速深度网络学习并提高分类识别率。跟修正线性单元(rectified linear units, ReLUs)^[12]相似, ELU激活函数同样可以取得正值来避免梯度消失。与ReLUs激活函数不同的是, ELU激活函数可以取得负值使得平均单元激活值接近于零。ELU非线性函数的表达式为:

$$f(x) = \begin{cases} x & x \geq 0 \\ \alpha(\exp(x) - 1) & x < 0 \end{cases} \quad (1)$$

$$f'(x) = \begin{cases} 1 & x \geq 0 \\ f(x) + \alpha & x < 0 \end{cases} \quad (2)$$

式中, 超参数 α 控制值使当网络输入为负值时输出也为负值, $\alpha > 0$ 。

2 混合PSO优化卷积神经网络

本文采用粒子群优化卷积神经网络参数, 和离散粒子群优化卷积神经网络特征图之间的连接结构, 用BP对该网络微调参数。

2.1 粒子群优化算法

粒子群算法(particle swarm optimization, PSO)通过随机初始化一群粒子, 并根据法则更新其速度和位置, 最终找到最优解。首先对粒子群进行初始化^[13]:

$$U_i(0) = U_{\max} - r(U_{\max} - U_{\min}) \quad (3)$$

$$X_i(0) = 0.1\mathbf{RS} + \mathbf{G} \quad (4)$$

式中, $U_i(0)$ 、 U_{\max} 以及 U_{\min} 分别表示粒子向量初速度、

速度上限及下限; $X_i(0)$ 表示粒子的初始位置; \mathbf{RS} 表示元素为(-1,1)的随机向量, 向量的维数与 $X_i(0)$ 相同; \mathbf{G} 是待优化参数之前经验值组成的向量; r 是(0,1)范围内随机值。速度及位置更新公式为:

$$U_i(t+1) = wU_i(t) + c_1r_1(P_i - X_i(t)) + c_2r_2(P_g - X_i(t)) \quad (5)$$

$$X_i(t+1) = X_i(t) + U_i(t+1) \quad (6)$$

式中, c_1 和 c_2 为加速因子, 为正常数; r_1 和 r_2 为(0,1)范围内的随机数; w 为惯性因子。

惯性因子是粒子保持飞行速度的系数, 本文采用非线性化调整策略对惯性因子进行改进, 即:

$$w(t) = w_{\max} - (w_{\max} - w_{\min}) \sin^2\left(\frac{\pi i}{2\text{Maxgin}}\right)$$

式中, w_{\max} 、 w_{\min} 分别为最大、最小惯性因子; i 、 Maxgin 是当前的及最大的迭代次数。

2.2 离散粒子群优化算法

离散粒子群优化算法(discrete particle swarm optimization, D-PSO)和PSO几乎相似, 有以下不同。位置的初始化方式为:

$$X_i(0) = F(0.1\mathbf{RS} + \mathbf{G}) \quad (7)$$

式中, 函数 $Y = F(X)$ 定义为: $y_i = \begin{cases} 0 & x_i < 0.5 \\ 1 & x_i \geq 0.5 \end{cases}$; \mathbf{RS}

表示元素为(-1,1)的随机向量, 向量的维数与 $X_i(0)$ 相同; \mathbf{G} 表示 S_2 与 C_3 层之间的特征图连接结构为全连接的向量。

位置更新公式为:

$$x_{ij}(t+1) = \begin{cases} 0 & r_i(t) \geq f(u_{ij}(t)) \\ 1 & r_i(t) < f(u_{ij}(t)) \end{cases} \quad (8)$$

其中非线性函数为sigmoid函数:

$$f(u_{ij}(t)) = \frac{1}{1 + e^{-u_{ij}(t)}}$$

式中, $x_{ij}(t)$ 表示粒子 i 在时间 t 第 j 个位置; $u_{ij}(t)$ 为对应速度; $r_{ij}(t)$ 为(0,1)之间随机数。其中 $u_{ij}(t)$ 要设定一个上下限, 保证 $f(u_{ij}(t))$ 值不能太靠近0或1。

2.3 算法流程

算法的基本流程如下。

1) 在 S_2 层和 C_3 层间用全连接方式的情况下, 用PSO优化参数。

① 初始粒子群相关参数; 根据式(3)、式(4)初始化各粒子参数变量速度及位置, 其中式(4)中的 \mathbf{G} 表示按照经验值初始化卷积神经网络时的网络权值向量, 由卷积核、连接层权值及偏置等参数组成; 把位置转换成网络参数形式, 并把均方差作为适应度值。

② 对所有粒子执行如下操作:

- a. 根据式(5)和式(6)更新粒子的速度和位置;
 - b. 计算粒子适应度,若得到的适应度值优于个体极值,则将个体极值位置向量设成当前位置向量;
 - c. 若粒子适应度优于全局最优,就把全局最优位置设成当前位置;
 - d. 如果满足停止条件,全局极值位置即为所求变量值,并停止搜索;否则,返回步骤②继续搜索;
- 2) 用D-PSO优化特征图连接结构。

① 初始化D-PSO相关参数;根据式(3)随机初始化速度和由式(7)初始化位置;对每个粒子,将粒子的位置变量转换成特征图连接结构的形式,计算粒子的适应度值。

② 对所有粒子执行如下操作:

- a. 根据式(5)和式(8)更新粒子的速度和位置;
- b. 计算适应度,若得到的适应度优于个体最优值,就将个体最优值位置设成当前位置;
- c. 若粒子适应度优于全局最优值,就将全局最优值位置设成当前位置;

d. 如果达到停止要求,立即停止循环,全局极值位置为所求变量值;否则,返回步骤②继续搜索。

3) 先将PSO寻得的全局最优位置转换为网络权值的形式初始化卷积神经网络参数,在此基础上再用D-PSO寻得的全局最优位置转换为特征图的连接结构,最后用BP算法微调混合PSO训练后的网络参数。

3 实验结果及分析

在MNIST及CIFAR-10数据集上进行了实验分析。为了避免随机性和不稳定性,实验均进行了10次重复,最后取均值作为最终结果。实验用的电脑处理器是Intel(R)Core(TM)i5-3470 CPU @ 3.20 GHz,安装内存(RAM)为8.00 GB,64位的Windows 7操作系统,没有采用GPU进行并行计算。

3.1 MNIST实验

MNIST手写数字数据集包含60 000个训练样本,10 000个测试样本,共10类,图像为灰度图像。

3.1.1 不同激活函数的网络速度及识别率对比

3种网络结构与图1类似,但缺少 F_3 全连接层,其中 S_2 与 C_3 层之间特征图为全连接方式,使用不同的非线性函数。CNN-S使用常用的Sigmoid函数,CNN-ReL使用ReLU, CNN-ELU使用ELU。网络样本分批训练,每个批次50个样本。

图2及表1表明,相同的结构和训练样本时,CNN-ELU网络收敛速度及识别率都优于其余二者。

网络为Sigmoid函数时,训练速度缓慢,易陷入过拟合;而ReLU函数相比Sigmoid函数具有单侧抑制性及稀疏激活性;与ReLU函数相似,ELU函数同样可以取得正值来避免梯度消失并能取负值以使平均单元激活值接近零,提高学习速度。本文方法采用ELU非线性激活函数,加速了网络收敛。

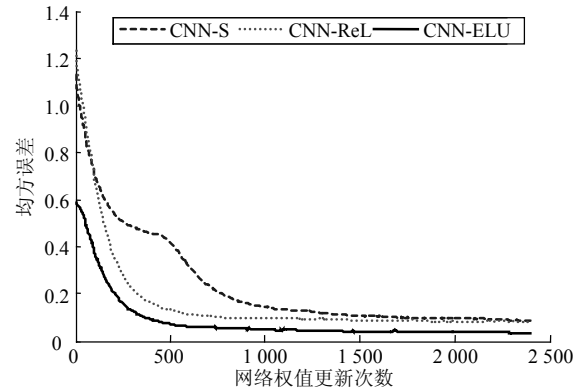


图2 采用不同非线性函数各网络结构误差收敛状况对比

表1 采用不同非线性函数各网络结构错误率对比 %

不同激活函数的CNN	训练次数(epochs)				
	1	5	10	50	100
CNN-S	11.11	4.22	2.68	1.36	1.15
CNN-ReL	3.18	1.72	1.34	0.96	0.94
CNN-ELU	2.12	1.36	1.12	0.94	0.93

3.1.2 不同连接结构的网络识别率对比

实验中CNN-1网络为 S_2 与 C_3 层全连接;CNN-2网络的 S_2 层与 C_3 层特征图连接结构为经验连接;CNN-3网络的 S_2 与 C_3 层结构为随意连接;CNN-DPSO网络 S_2 层与 C_3 层特征图之间连接结构由离散粒子群优化过后得到。离散粒子群的粒子数为20个, $c_1=c_2=1.494$, $w_{max}=0.9$, $w_{min}=0.2$,迭代20次。本文采用ELU函数,由表1可以看出采用ELU的网络,迭代次数超过10次,图像识别率的提高并不明显,故后面所有实验均最多迭代10次。10次迭代后,CNN-3、CNN-2、CNN-1、CNN-DPSO这4种结构的网络识别率分别为1.37%、1.36%、1.28%、1.20%。

由于 S_2 与 C_3 层特征图之间为全连接时,使得连接数量过多,降低网络的运算速率;而固定连接结构受到经验的影响,不能普遍适用。离散粒子群算法在无经验指导的情况下智能优化 S_2 与 C_3 层特征图之间的连接,可以消除冗余连接,从而提高了效率。因此采用DPSO优化后的网络使识别率高于其他连接结构的网络。证明了将DPSO用来优化特征图连接能够使卷积神经网络在无经验指导下获得较好的识别效果。

3.1.3 不同训练方法的识别率对比

实验中NN表示神经网络; DBN表示深度信念网络; CNN-ELU网络结构与3.1.1节中描述相同; CNN-F网络如图1, 用反向传播算法训练; CNN-F-PSO网络结构与CNN-F相同, 用PSO训练网络。CNN-F-HPSO网络采用混合PSO训练网络。

由表2和表3可以看出, 卷积神经网络效果普遍优于NN和DBN网络, 说明CNN结构更适合用于提取二维图像特征。本文采用混合PSO优化参数及特征图间的连接防止网络过早陷入局部最优, 并有效利用特征图间的信息, 使得在相同的迭代次数时, 本文算法识别率略高于其他算法。在达到几乎相同甚至更高的识别率时, 本文方法较传统CNN方法需要更少运行时间, 从而表明本文方法的可行性。

表2 不同训练方法不同训练次数的误识别率对比 %

训练方法	训练次数(epochs)		
	1	5	10
NN	8.12	4.73	3.58
DBN	6.30	3.68	3.15
CNN-F	1.93	0.98	0.89
CNN-F-PSO	1.84	0.93	0.78
CNN-F-HPSO	1.75	0.89	0.71

表3 传统CNN方法与本文方法识别错误率与消耗时间对比

方法	错误率/%	寻优时间/s	训练测试时间/s	总时间/s
CNN-ELU	1.14	-	2 233.5	2 233.5
本文方法	0.71	935.5	919.8	1 855.3

3.1.4 与其他方法识别结果对比

表4 各种方法误识率对比

识别方法	预处理	误识率/%
linear classifier (1-layer NN)	去斜	8.4
K-nearest-neighbors, Euclidean (L2)	无	5.0
40 PCA + quadratic classifier	无	3.3
SVM, Gaussian Kernel	无	1.4
Trainable feature extractor + SVMs [no distortions]	无	0.83
本文方法	无	0.71
committee of 35 conv. net, 1-20-P-40-P-150-10	宽度归一化	0.23

表4中其他方法是国内外学者在MNIST数据集上的研究成果^[14]。由表4可知, 本文方法优于上述传统方法, 但也不是最优方法, 因为相对于其他卷积神经网络, 本文方法网络较浅, 特征图数较少, 且没有对图像进行移动、缩放、倾斜、压缩等人为的扭曲变形预处理, 也没有进行去斜处理和宽度归一化处理。但本文方法加速了网络收敛, 相对上述传统方法, 提高了识别率, 证明该方法的可行性。

3.2 CIFAR-10实验

CIFAR-10数据集包含10类彩色图像, 其中有50 000个训练样本, 10 000个测试样本。简单地对CIFAR-10进行灰度化等预处理操作后, 称之为CIFAR-10- gray。数据的预处理方式与文献[15]的相同, 即先灰度化, 再均值方差归一化, 然后白化处理; 网络迭代训练次数同样是50次。网络结构设置8C-8S-16C-16S-240F也与文献[15]中的3种结构相同。本文将文献[15]中的3种网络结构称为CNN-1、CNN-2、CNN-3网络, 它们的激活函数均用ReLU-Softplus^[15], 池化方式分别为最大池化、随机池化、概率加权池化。结果对比如表5所示。

表5 不同方法的误识率对比 %

数据集	算法			
	CNN-1	CNN-2	CNN-3	本文算法
CIFAR-10-gray	33.34	36.56	34.93	30.10

由于只对数据集进行灰度化预处理, 并且灰度化也会损失一些重要特征, 从表5可知文中方法并没有达到非常好的效果, 但仍略微优于文献[15]中的CNN-1、CNN-2、CNN-3网络的识别效果, 证明了本文方法不仅在MNIST数据集上有较好的识别效果, 应用于CIFAR-10数据集上也是有效可行的。

4 结束语

基于混合PSO的卷积神经网络训练方法与其他方法相比提高了训练效率, 并在训练时间接近时, 网络识别率得到提高; 达到同样的识别效果, 本文方法收敛需要的训练次数较少。因此本文方法加速了误差收敛, 只需少量次数迭代训练网络就能达到甚至优于其他传统方法的识别效果。虽然优化结构和参数时增加了训练复杂度, 但经优化后的网络提高了信息处理效率, 因此比较适合识别大规模数据。

参 考 文 献

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 215: 436-444.
- [2] 戴晓爱, 郭守恒, 任涓, 等. 基于堆栈式稀疏自编码器的高光谱影像分类[J]. 电子科技大学学报, 2016, 45(3): 382-386.
- [3] DAI Xiao-ai, GUO Shou-heng, REN Yu, et al. Hyperspectral remote sensing image classification using the stacked sparse autoencoder[J]. Journal of University of Electronic Science and Technology of China, 2016, 45(3): 382-386.
- [3] ARORA C, SABETZADEH M, BRIAND L, et al. Automated checking of conformance to requirements templates using natural language processing[J]. IEEE

- Transactions on Software Engineering, 2015, 41(10): 944-968.
- [4] ZHANG Zhong, WANG Chun-heng, XIAO Bai-hua, et al. Cross-view action recognition using contextual maximum margin clustering[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(10): 1663-1668.
- [5] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.
- [6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [7] YUAN Q, YIN G. Analyzing convergence and rates of convergence of particle swarm optimization algorithms using stochastic approximation methods[J]. IEEE Transactions on Automatic Control, 2015, 60(7): 1760-1773.
- [8] OULLRTTE R, BROWNE M, HIRASAWA K. Genetic algorithm optimization of a convolutional neural network for autonomous crack detection[C]//Congress on Evolutionary Computation. Portland: IEEE, 2004, 1: 516-521.
- [9] FEDOROVICI L O, PRECUP R E, DRAGAN F, et al. Evolutionary optimization-based training of convolutional neural networks for OCR applications[C]//2013 17th International Conference on System Theory, Control and Computing (ICSTCC). Sinaia: IEEE, 2013: 207-212.
- [10] LIU Dong, JIANG Qi-long, CHEN J X. Binary inheritance learning particle swarm optimisation and its application in thinned antenna array synthesis with the minimum sidelobe level[J]. IET Microwaves, Antennas & Propagation, 2015, 9(13): 1386-1391.
- [11] CLEVERT D A, UNTERTHINER T, HOCHREITER S. Fast and accurate deep network learning by exponential linear units (ELUs)[C]//International Conference on Learning Representations. San Juan: Computer Science, 2016: arXiv:1511.07289.
- [12] XIONG Shi-fu, WU Guo, LIU Di-yuan. The Vietnamese speech recognition based on rectified linear units deep neural network and spoken term detection system combination[C]//9th International Symposium on Chinese Spoken Language Processing (ISCSLP). Singapore: IEEE, 2014: 183-186.
- [13] LI Y H, ZHAN Z H, LIN S J, et al. Competitive and cooperative particle swarm optimization with information sharing mechanism for global optimization problems[J]. Information Sciences, 2015, 293: 370-382.
- [14] LECUN Y, CORTES C, BURGESS J C. The MNIST database of handwritten digits[EB/OL]. [2016-07-12]. <http://yann.lecun.com/exdb/mnist/>.
- [15] 余萍, 赵继生, 张浩. 基于非线性修正函数的卷积神经网络图像识别研究[J]. 科学技术与工程, 2015, 15(34): 221-225.
- YU Ping, ZHAO Ji-sheng, ZHANG Jie. Image recognition of convolutional neural networks based on rectified nonlinear units function[J]. Science Technology and Engineering, 2015, 15(34): 221-225.

编辑 蒋晓