

文章编号:1001-9014(2007)01-0038-05

基于免疫克隆算法的协同神经网络参数优化

马秀丽¹, 刘芳², 焦李成¹

(1. 西安电子科技大学 智能信息处理研究所, 陕西 西安 710071;

2. 西安电子科技大学 计算机学院, 陕西 西安 710071)

摘要:针对现有协同神经网络参数优化方法的不足,提出了一种基于免疫克隆算法的参数优化方法.与在平衡注意参数条件下的算法和在不平衡注意参数条件下基于遗传算法和模拟退火算法的优化方法相比,新方法具有全局兼局部寻优能力,不易陷入局部极值,并且迭代步长是自适应调整的.对纹理图像与遥感图像的分类识别结果表明:新方法不仅具有更快的收敛速度而且具有更优的分类识别性能,同时验证了注意参数及所有参数对各原型模式之间竞争态势的影响,从而达到更佳分类识别效果.

关键词:协同神经网络;注意参数;免疫克隆算法;图像分类

中图分类号:TP18 **文献标识码:**A

PARAMETERS OPTIMIZATION OF SYNERGETIC NEURAL NETWORK BASED ON IMMUNITY CLONAL ALGORITHM

MA Xiu-Li¹, LIU Fang², JIAO Li-Cheng¹

(1. Institute of Intelligent Information Processing, Xidian University, Xi'an 710071, China;

2. School of Computer Science, Xidian University, Xi'an 710071, China)

Abstract: Due to the shortages of optimization algorithms available in synergetic neural network (SNN), an algorithm of parameters optimization on immunity clonal algorithm (ICA) was proposed here. Compared with the algorithm under balanced attention parameters and that under unbalanced attention parameters on genetic algorithm (GA) and simulated annealing algorithm (SA), the new algorithm has the global and local searching ability and is not easy to get into local optima. And the iterative step is adjusted adaptively. Experiments on textural images and remote sensing images show that the proposed algorithm has not only faster convergent speed but also better classification performance. Simultaneously, the effect of attention parameters and all parameters on the competition of prototype patterns is verified and then better recognition result can be obtained.

Key words: synergetic neural network; attention parameter; immunity clonal algorithm; image classification

引言

自从 Haken 教授^[1]提出协同神经网络 (Synergetic Neural Network, SNN) 以来,人们对其学习算法进行了广泛的研究,尤其是在原型模式向量的选择和注意参数的设置等方面.

在平衡注意参数条件下,协同神经网络的输出可以从序参量初值直接推断出来.这样在给定原型向量目标模式的前提下,协同神经网络就失去了自学习能力.由于作为神经网络重要特征的自学习能

力被抑制,协同神经网络目标识别的能力也被削弱,从而影响其广泛应用,所以研究不平衡注意参数条件下协同神经网络的学习算法和动力学行为是非常重要的和有意义的.

鉴于现有协同神经网络参数优化方法^[2-5]搜索能力有限的缺点,提出了一种基于免疫克隆算法^[6]的参数优化方法并将其成功用于图像分类.

1 协同神经网络

Haken^[1]指出协同模式识别方法可描述为一个有势的动力学过程,势函数为:

收稿日期:2005-12-21,修回日期:2006-06-18

Received date: 2005-12-21, revised date: 2006-06-18

基金项目:国家“863”计划(2002AA135080),国家自然科学基金(60133010)和“十·五”国防预研资助项目(11307050103)

作者简介:马秀丽(1978-),女,河北南官人,西安电子科技大学在读博士生.主要研究方向:模式识别,图像处理,智能信息处理和数据挖掘等.

$$V = -\frac{1}{2} \sum_{k=1}^M \lambda_k (v_k^+ q)^2 + \frac{1}{4} B \sum_{k \neq k'} (v_k^+ q)^2 (v_{k'}^+ q)^2 + \frac{1}{4} C q (q^+ q)^2, \quad (1)$$

其中,原型模式向量 v_k 和试验模式 q 是具有零均值和单位长度的 N 维列向量,为了满足原型模式向量之间线性无关的条件,要求 $M \leq N$. v_k^+ 是 v_k 的伴随向量. $\lambda_k > 0$ 是注意参数;当所有 λ_k 都相等时,称为平衡注意参数,否则,称为不平衡注意参数.

为了减少系统的维数,引入序参量 ξ_k ,

$$\xi_k = (v_k^+, q) = v_k^+ q, \quad (2)$$

相应的动力学方程为:

$$\dot{\xi}_k = \lambda_k \xi_k - B \sum_{k' \neq k} \xi_k^2 \xi_{k'} - C \left(\sum_{k'=1}^M \xi_{k'}^2 \right) \xi_k. \quad (3)$$

在协同模式识别系统中,注意参数提供了选择性注意机制.对 Haken 模型的研究表明:通过序参量初值和注意参数可以决定序参量的生存性^[1].

2 基于免疫克隆算法的 SNN 参数优化

2.1 协同神经网络参数分析

当 $B = C = \lambda_k > 0$ 时,根据序参量初值可以预测网络的最终输出.因此,动力学方程式(3)失去了通过进一步学习使网络正确识别试验模式的能力^[1].

当动力学方程式(3)中的参数 $B > 0, C > 0, \lambda_k > 0$ 时, (λ_k, B, C) 共同决定了协同神经网络的分类识别性能,同时对它们进行研究才是最合理和有效的^[3].目前,还没有成熟的研究理论来控制参数 (λ_k, B, C) ^[5].

参数 (λ_k, B, C) 的调解是一种全局性的行为.在此,提出了一种基于免疫克隆算法(Immunity Clonal Algorithm, ICA)的参数优化方法,该方法具有全局兼局部寻优能力,可以避免遗传算法易陷入局部极值和模拟退火算法收敛速度慢的缺点,可以自适应调整迭代步长.

2.2 基于免疫克隆算法的参数优化

对于任意的初始状态,免疫克隆算法均以概率 1 收敛到满意种群集^[6].提出的基于免疫克隆算法的参数优化方法的具体迭代步骤描述如下:

Step 1: 初始化抗体种群

随机产生规模为 N 的初始抗体种群 $A(0)$, 每个个体表示一个 (λ_k, B, C) 的集合;

Step 2: 计算初始种群的亲和度

亲和度由对训练样本的正确识别率表示;

Step 3: 克隆

对当前第 k 代父本种群 $A(k)$ 中的个体进行克隆生成相应的子群,从而得到 $A'(k)$.克隆规模根据亲和度大小动态调整,公式为:

$$q_i(k) = \text{Int} \left(N_c * \frac{f(A_i(k))}{\sum_{j=1}^N f(A_j(k))} \right), \quad i = 1, 2, \dots, N \quad (4)$$

其中, $f(\cdot)$ 是亲和度函数. $N_c > N$ 的值与克隆规模有关. $\text{Int}(x)$ 表示大于 x 的最小整数.可以认为,克隆将低维空间(N 维)的问题转化到更高维(N_c 维)的空间中解决,然后将结果投影到低维空间(N 维)中,从而获得对问题更全面的认识.

Step 4: 克隆变异

以概率 p_m 对 $A'(k)$ 进行变异,得到 $A''(k)$;

Step 5: 计算亲和度

计算 $A''(k)$ 中每个个体的亲和度;

Step 6: 克隆选择

在每个子群中,如果最优个体的亲和度大于相应的父本,则选择该个体进入新的父代种群 $A(k+1)$;

Step 7: 计算亲和度

计算新种群 $A(k+1)$ 中每个个体的亲和度;

Step 8: 判断停机条件是否满足

停机条件可设定为具体的迭代次数或训练样本的分类识别率达到指定的精度.如果满足停机条件则停止迭代,迭代过程中保留的最优个体即为搜索到的最优参数,否则保留当前代的最优个体,然后转向 Step 3.

当优化注意参数时,算法步骤跟上述的相同,只不过单个个体不再是所有参数 (λ_k, B, C) 的集合而仅是注意参数 λ_k 的集合.

3 仿真实验及讨论

通过对纹理图像和遥感图像的分类来验证新方法的性能.实际上,免疫克隆算法(ICA)是针对遗传算法(Genetic Algorithm, GA)易早熟的缺点而提出的^[6],因此与基于 GA 的参数优化方法^[3-4]进行了比较.同时与平衡注意参数条件下的算法以及基于模拟退火算法(Simulated Annealing algorithm, SA)的参数优化方法^[5]也进行了比较.其中判断样本正确识别的条件是:将识别结果直接与已知的类属信息进行比较从而判断样本是否被正确识别.

另外,在大多数研究图像分类识别的文献中,大部分都采用 10 次独立实验的统计平均结果.由于新方法比较稳定,10 次、20 次、50 次甚至更高次数的

统计平均结果的精度相差不到 1 个百分点,因此也采用了相同的统计方法.

3.1 纹理图像分类

为了验证所提出方法的性能,从 Brodatz 纹理库中选择 16 幅比较相似的纹理图像进行实验,如图 1 所示. 每幅图像大小 640×640 ,将每幅图像分割成 25 幅互不重叠的子图像作为一类. 所以,共有 16 类,每类有 25 个样本,从每类随机选择三分之一的样本作为训练样本,其余的作为测试样本. 同时,利用 Brushlet^[7]进行特征提取.

(1) 仅优化注意参数

实验中,SCAP 算法^[8]用来求解原型模式向量,迭代步长自适应调整 $1/D^{[9]}$, $B = C = 1$, 迭代次数 60. 参数 λ_k 的搜索空间为 $(0, +\infty)$ 且初值均为 1. 免疫克隆算法的参数设置如下: 初始种群 5, 变异概率 0.2, 克隆规模根据亲和度大小自适应调整. 在遗传算法(GA)中, 初始种群 20, 交叉概率 0.9, 变异概率 0.2, 采用了与实数编码相适应的遗传操作算子: 转盘选择算子, 算术交叉算子和非一致性变异算子. 模拟退火算法(SA)^[5]的参数设置为: $T_0 = 1000$, $\beta = 0.8$, $K = 0.01$, $\alpha = 0.99$, 在搜索过程中, 每隔 10 个点采样一次, 即每迭代一次搜索 10 个点. 三种算法的停机条件为: 在指定的迭代次数内分类识别率大于 99%, 并采用了最优解保存策略. 表 1 给出了 10 次独立实验的统计平均结果, 其中基于 ICA 参数优化方法的最优识别率达到 97.426%. 同时, 图 2 还给出了三种方法前 100 代的收敛曲线对比结果.

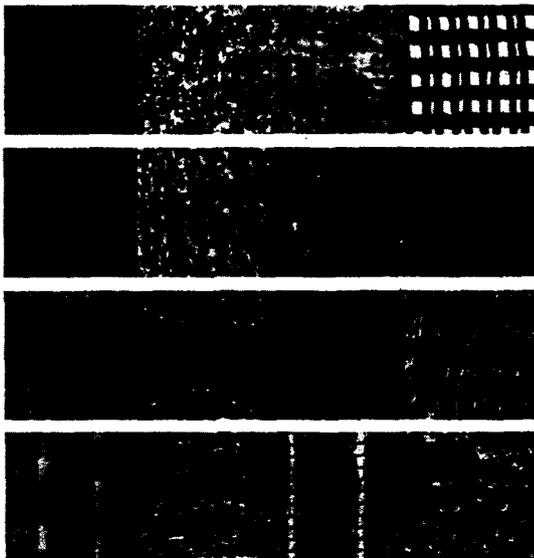


图 1 Brodatz 纹理图像
Fig.1 Brodatz textural images

表 1 注意参数优化方法识别结果比较
Table 1 Comparison of different methods

	$\lambda_k = 1$	apGA	apSA	apICA
trTime(s)	0	135.527	698.394	49.597
teTime(s)	0	0.108	0.099	0.096
trRate(%)	89.844	97.500	95.550	98.516
teRate(%)	93.750	95.682	95.000	96.728

其中, trTime 和 teTime 分别表示训练时间和测试时间, trRate 和 teRate 表示训练样本和测试样本的分类识别率, avRate 表示平均识别率, apGA 和 pGA 分别表示基于 GA 的注意参数及所有参数的优化方法, 同理可得 apSA 和 pSA 以及 apICA 和 pICA.

从表 1 和图 2 可以看出, 当 $B = C = 1$ 仅优化注意参数时, 基于 ICA 参数优化方法的分类识别率比基于 GA 和 SA 方法都略有提高, 而它的收敛速度要快得多. 与在平衡注意参数条件下, 即 $\lambda_k = 1$ 时的分类识别率相比, 在不平衡注意参数条件下, 各种改进算法的分类识别率都有较大的提高.

(2) 所有参数均被优化

为了验证参数 (λ_k, B, C) 对协同神经网络的分类识别性能的影响, 同时对它们进行优化. 它们的搜索空间为 $(0, +\infty)$ 且初值均为 1. 其余参数的设置与上面实验完全相同. 表 2 给出了 10 次独立实验的统计平均识别率, 其中基于 ICA 参数优化方法的最优识别率达到 97.794%. 图 3 给出了 3 种方法前 100 代的收敛曲线对比结果.

表 2 所有参数优化方法识别结果比较
Table 2 Comparison of different methods

	$\lambda_k = B = C = 1$	pGA	pSA	pICA
trTime(s)	0	152.420	763.447	38.271
teTime(s)	0	0.221	0.106	0.099
trRate(%)	89.844	97.600	96.875	98.985
teRate(%)	93.750	95.820	94.000	97.279

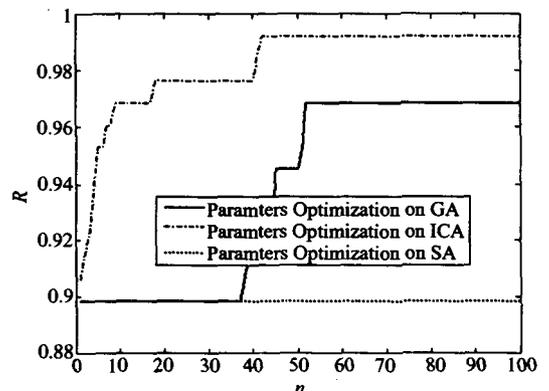


图 2 3 种方法收敛曲线对比
Fig.2 Comparison of methods with convergent curves

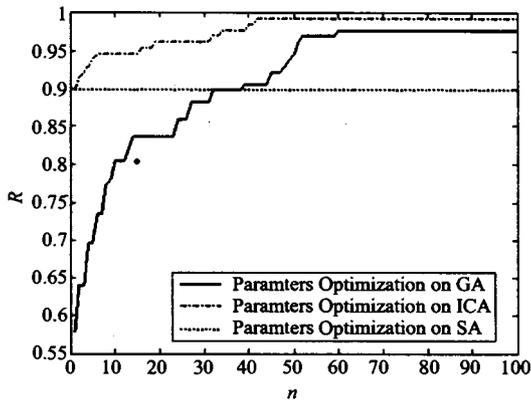


图3 3种方法收敛曲线对比

Fig.3 Comparison of methods with convergent curves

从表2和图3可以看出, (λ_k, B, C) 同时被优化时,与基于GA和SA的参数优化方法相比,基于ICA的参数优化方法具有更快的收敛速度和更高的分类识别率,因为它兼顾全局和局部搜索,不易陷入局部极值。

对比表1和表2可以看出,在完全相同的参数条件下,与仅优化注意参数相比,同时优化 λ_k, B 和 C 时,优化算法的平均识别率都略有提高。所以不同的 (λ_k, B, C) 将影响各原型目标模式之间的竞争,从而获得更优的分类识别结果。

3.2 遥感图像分类

图像集合包含完整的和残缺的二值舰船和飞机图像,以及不同旋转角度的图像共1064幅,其中舰船456幅,飞机608幅,部分图像如图4所示。将图像集分成两组,训练样本集中舰船和飞机图像分别为120和160幅,检验集中分别为336和448幅。采用相对矩^[10]的方法进行特征提取。

(1) 仅优化注意参数

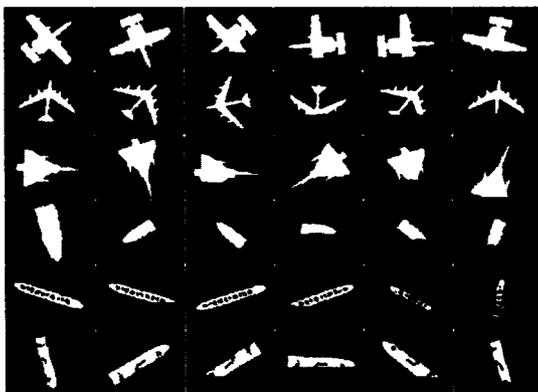


图4 实验所用的部分遥感图像

Fig.4 Part of remote sensing images for experiment

实验中,采用SCAP算法^[8]求解原型模式向量,迭代步长自适应调整 $0.001/D$ ^[9],迭代次数80, $B = C = 1$. 参数 λ_k 的搜索空间为 $(0, +\infty)$ 且初值均为1. ICA的参数设置如下:初始种群5,变异概率0.2,克隆规模根据亲和度大小自适应调整。在GA中,初始种群20,交叉概率0.9,变异概率0.2,采用了与实数编码相适应的遗传操作算子:转盘选择算子,算术交叉算子和非一致性变异算子。SA^[5]的参数设置为: $T_0 = 1000, \beta = 0.8, \alpha = 0.99, K = 0.01$,在搜索过程中,每隔10个点采样一次。3种算法的停机条件相同:在指定的迭代次数内训练样本的分类识别率达到100%。表3给出了10次独立实验的统计平均结果,其中基于ICA参数优化方法的最优识别率达到98.822%。图5给出了3种方法前100代的收敛曲线对比结果。

表3 注意参数优化方法识别结果比较

Table 3 Comparison of different methods

	$\lambda_k = 1$	apGA	apSA	apICA
trTime(s)	0	144.967	60.958	105.058
teTime(s)	0	0.153	0.124	0.157
trRate(%)	96.333	98.000	98.630	98.333
teRate(%)	boat	91.670	96.102	91.700
	plane	100.00	99.020	99.860
avRate(%)	96.335	97.769	96.270	98.109

从表3和图5可以看出,基于ICA的参数优化方法具有更高的分类识别率,并且3种参数优化方法的分类识别率都优于平衡注意参数条件下的分类识别率。实验中观察到,新方法不仅具有较快的收敛速度而且不易陷入局部极值。

(2) 所有参数均被优化

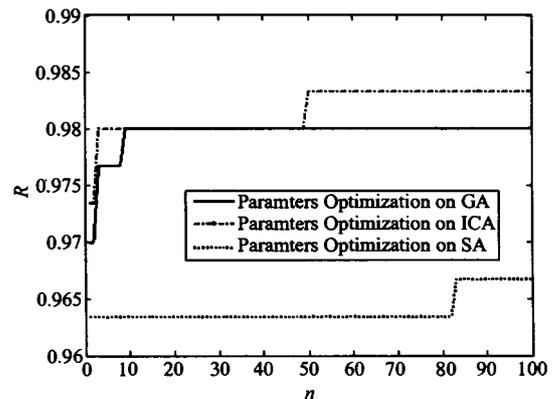


图5 3种方法收敛曲线对比

Fig.5 Comparison of methods with convergent curves

在该实验中, (λ_k, B, C) 同时被优化, 它们的搜索空间为 $(0, +\infty)$ 且初值均为 1, 优化算法其余参数的设置与上面实验完全相同, 停机条件为: 在指定的迭代次数内分类识别率保持 20 次不变. 表 4 给出了 10 次独立实验的统计平均结果, 其中基于 ICA 参数优化方法的最优识别率达到 98.953%. 最后, 图 6 给出了 3 种方法前 100 代的收敛曲线对比结果.

表 4 所有参数优化方法识别结果比较
Table 4 Comparison of different methods

	$\lambda_k = B = C = 1$	pGA	pSA	pICA
trTime(s)	0	26.456	18.651	37.764
teTime(s)	0	0.108	0.117	0.103
trRate(%)	96.333	97.633	97.860	97.967
teRate(%)	boat	91.670	95.654	96.980
	plane	100.00	99.509	99.440
avRate(%)	96.335	97.814	98.360	98.416

从表 4 和图 6 可以看出, 与其它方法相比, 基于 ICA 的参数优化方法具有更优的分类识别率. 由于停机条件为在指定迭代次数内分类识别率保持不变的次数, 所以从训练时间可以得出, 基于 GA 和 SA 的参数优化方法比基于 ICA 的参数优化方法更快的陷入局部极值. 同时, 不平衡注意参数条件下的分类识别率都高于平衡注意参数条件下的分类识别率.

对比表 3 和表 4 可以看出, 当同时优化注意参数 λ_k 、参数 B 和 C 时, 优化算法的识别率都略有提高. 所以不同的 (λ_k, B, C) 将影响到各原型目标模式之间的竞争态势, 从而达到更佳分类识别效果. 因此, 同时对 (λ_k, B, C) 进行研究才是最合理的.

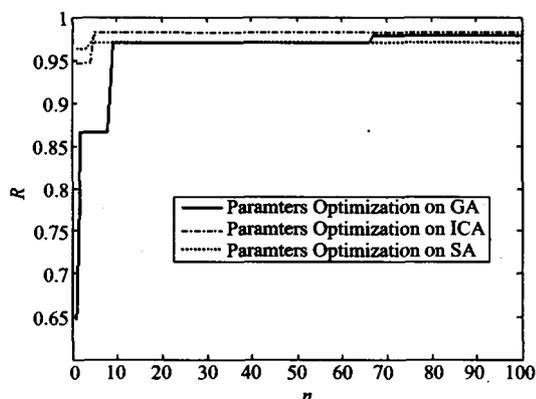


图 6 3 种方法收敛曲线对比

Fig. 6 Comparison of methods with convergent curves

4 结论

提出了一种基于免疫克隆算法的协同神经网络参数优化方法, 与在平衡注意参数条件下的算法和在不平衡注意参数条件下基于遗传算法和模拟退火算法的参数优化方法相比, 新方法具有全局寻优能力, 不易陷入局部极值, 并且迭代步长是自适应调整的. 通过对纹理图像与遥感图像的分类表明: 新方法不仅具有更快的收敛速度而且具有更优的分类识别性能. 同时验证了 (λ_k, B, C) 对各原型模式之间竞争态势的影响, (λ_k, B, C) 共同决定了协同神经网络的分类识别性能, 同时对它们进行研究才是最合理和有效的.

REFERENCES

- [1] Haken H. Synergetic computers and cognition-top-down approach to neural nets [M]. Berlin: Springer-Verlag, 1991.
- [2] Wang F Y, Lever P J A, Pu B. A robotic vision system for object identification and manipulation using synergetic pattern recognition [J]. *Robot. Comput. Integrated Manufacturing*, 1993, 10(6): 445—459.
- [3] HU Dong-Liang, QI Fei-Hu. Study on unbalanced attention parameters in synergetic approach on pattern recognition [J]. *Acta Electronica Sinica* (胡栋梁, 戚飞虎. 协同模式识别中不平衡注意参数的研究. 电子学报), 1999, 27(5): 15—17.
- [4] WANG Hai-Long, QI Fei-Hu, REN Qing-Sheng. Parameters optimization of synergetic neural network [J]. *J. Infra-red Millim. Waves* (王海龙, 戚飞虎, 任庆生. 协同神经网络中参数的优化. 红外与毫米波学报), 2001, 20(3): 215—218.
- [5] GAO Jun, DONG Huo-Ming, SHAO Jing, et al. Parameters optimization of synergetic recognition approach [J]. *Chinese Journal of Electronics*, 2005, 14(2): 192—197.
- [6] LIU Ruo-Chen, DU Hai-Feng, JIAO Li-Cheng. An immune monoclonal strategy algorithm [J]. *Acta Electronica Sinica* (刘若辰, 杜海峰, 焦李成. 一种免疫单克隆策略算法. 电子学报), 2004, 32(11): 1880—1884.
- [7] Meyer F G, Coifman R R. Brushlets: A tool for directional image analysis and image compression [J]. *Applied and Computational Harmonic Analysis*, 1997, 4(1): 147—187.
- [8] Wagner T, Boebel F G. Testing synergetic algorithms with industrial classification problems [J]. *Neural Networks*, 1994, 7(8): 1313—1321.
- [9] ZHAO Tong, QI Fei-Hu, Feng Jiong. Analysis of the recognition performances of synergetic neural network [J]. *Acta Electronica Sinica* (赵同, 戚飞虎, 冯炯. 协同式神经网络的识别性能分析. 电子学报), 2000, 28(1): 74—77.
- [10] WANG Bo-Tao, SUN Jing-Ao, CAI An-Ni. Relative Moments and their applications to geometric shape recognition [J]. *Journal of Image and Graphics* (王波涛, 孙景鳌, 蔡安妮. 相对矩及在几何形状识别中的应用. 中国图像图形学报), 2001, 6(3): 296—300.