

文章编号: 1001-0920(2013)11-1702-05

一种基于大脑皮层结构的侧抑制神经网络

杨刚¹, 乔俊飞¹, 薄迎春^{1,2}, 韩红桂¹

(1. 北京工业大学 电子信息与控制工程学院, 北京 100124;
2. 中国石油大学 信息与控制工程学院, 山东 青岛 266555)

摘要: 借鉴仿生学原理, 基于大脑皮层结构提出一种新型侧抑制神经网络(S-LINN)模型. 通过模拟大脑皮层内锥体神经元和抑制神经元的连接特点, 在多层结构的S-LINN的不同层神经元之间引入跨越连接, 同时在隐含层内神经元之间进行信息的侧向抑制传输. 引入的两种连接机制有效地提高了网络处理问题的能力, 与其他网络相比能够以更精简的结构较好地解决实际问题. 通过对乳腺癌诊断数据集和异或问题的求解, 表明了S-LINN网络不但能够获得较高的训练精度, 而且具有更强的泛化能力.

关键词: 仿生学; 大脑皮层; 侧抑制; 人工神经网络

中图分类号: TP273

文献标志码: A

A lateral inhibition neural network based on neocortex topology

YANG Gang¹, QIAO Jun-fei¹, BO Ying-chun^{1,2}, HAN Hong-gui¹

(1. College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China; 2. College of Information and Control Engineering, China University of Petroleum, Qingdao 266555, China.
Correspondent: QIAO Jun-fei, E-mail: isi.bjut@gmail.com)

Abstract: On the basis of the connective mode and information transmit mechanism of cerebral cortex, a novel lateral inhibition neural network model with span connection(S-LINN) is proposed. Combining the liminar organization of cerebral cortex, and fully considering the lateral inhibition connection between interneurons and the pyramidal neurons, the proposed S-LINN transforms information to other neurons in different layers, which is used to enhance response contrast and advance the network representation, respectively. The effectiveness and superiority of the proposed network is compared with other popular approaches on two benchmark problems in the areas of real-world regression and classification problems. Simulation results show that the proposed S-LINN can achieve the higher accuracy of approximation and generalization with the comparably compact network structure.

Key words: bionics; cerebral cortex; lateral inhibition; artificial neural network

0 引言

人工神经网络(ANN)是一个模仿人类大脑神经结构和功能的计算模型, 引入生物神经网络(BNN)中神经元连接和信息传输机制, 构造具有生物学结构和功能的复杂网络, 已成为ANN研究热点之一. 其中, 基于抑制神经元的侧向抑制提出的侧抑制网络^[1]、基于small-world特性提出的WS模型^[2]和基于scale-free特性提出BA模型^[3]等都是基于生物神经网络连接特性构建的特殊网络结构.

在进行ANN的设计时, 除了可以从工程应用的角度出发外, 也应该同时遵循结构决定功能的原则. 因此, 可以从结构仿生的角度出发, 根据BNN神经元

连接模态和相关信息传输机制设计ANN模型. 生物解剖学研究表明^[4], 在具有分层结构的大脑皮层内锥体神经元能够跨越不同的层级进行信息的传递, 同时抑制内部神经元的存在, 使得处于兴奋状态的神经元能对其邻近的神经元产生抑制作用. “侧抑制”^[5]是广泛存在于感知系统的一种重要的信息处理机制. 文献[6]在ESN(echo state network)中引入侧抑制连接构造新的网络结构DESN(decoupled ESN), 可以降低传统ESN的计算复杂度. 文献[7]设计了一种部分重设的侧抑制神经网络用于进行图像分割, 取得了较好的效果. 文献[8]构造了一个结构为2-2-1的侧抑制网络求解异或(XOR)问题, 表明了侧抑制机制能够在解

收稿日期: 2012-08-13; 修回日期: 2012-10-22.

基金项目: 国家自然科学基金重点项目(61034008); 教育部博士点基金项目(20080005004).

作者简介: 杨刚(1983-), 男, 博士生, 从事神经网络分析与设计、系统建模的研究; 乔俊飞(1968-), 男, 教授, 博士生导师, 从事智能系统建模及控制、神经网络分析与设计等研究.

决此类模式识别问题时发挥一定的作用.

从结构仿生和信息仿生的角度出发, 本文基于大脑皮层内锥体神经元在不同层神经元之间的跨越连接和中间神经元在同层内的抑制连接提出了一种新型的侧抑制神经网络模型 (S-LINN). S-LINN 在分层结构的基础上在隐含层内引入侧抑制机制实现增大网络输出的对比, 同时引入信息在层与层之间的跨越传输提高网络的表达能力, 并为避免网络陷入局部极小作出贡献. 通过对 UCI 的乳腺癌诊断和异或 (XOR) 问题的测试验证了 S-LINN 处理模式识别问题的能力, 结果表明该网络在获得较高的训练精度的同时具有更好的泛化能力, 在网络规模较小的情况下也能达到满意的效果, 且有效地避免网络陷入局部极小.

1 S-LINN 网络模型

具有层级结构的 S-LINN 网络模拟了大脑皮层内神经元在不同层之间的跨越连接结构和同层神经元之间信息的侧向抑制传递. 多层 S-LINN 网络的结构如图 1 所示.

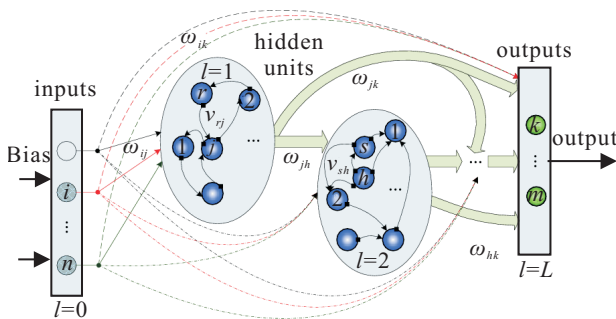


图 1 S-LINN 网络结构框架 ($L + 1$ 层)

1.1 S-LINN 网络的结构

首先给出文中所用符号及其含义. $l, q = 0, 1, \dots, L$ 为网络内各个层的标示, n_l 为第 l 层神经元的个数, 有 $n_0 = n, n_L = m; x_i = [x_1, x_2, \dots, x_n]^T \in R^n$ 为网络输入, $t_i = [t_1, t_2, \dots, t_m]^T \in R^m$ 为网络的期望输出, $y_i = [y_1, y_2, \dots, y_m]^T \in R^m$ 为网络的实际输出; x_i^l 和 y_i^l 分别为第 l 层内第 i 个神经元的输入和输出, $f^l(\cdot)$ 为第 l 层神经元的激活函数, P 为样本个数; ω_{ik} 为不同层内第 i 个神经元与第 k 个神经元之间的连接权值, v_{rj} 为同一个隐含层内第 r 个神经元与第 j 个神经元之间的抑制连接权值; μ 为学习率参数, t 为时间步, e 为网络期望输出与实际输出之间的误差.

根据上述网络定义和符号说明, 对于 S-LINN 有:

1) 输入层. $l = 0$ 层内应有 n 个输入神经元对应网络的 n 维输入, 则第 i 个输入神经元的输出为

$$y_i^0 = f^0(x_i^0), i = 1, 2, \dots, n. \quad (1)$$

2) 隐含层. S-LINN 网络存在不同层之间的信息传输, 存在隐含层内神经元之间的侧抑制连接机制实

现的信息传输和竞争, 因此, 隐含层 l 内的神经元有两类输入: 一是 $[0, l - 1]$ 层内神经元的传入输入, 二是同一层内相邻神经元的抑制输入. 隐含层 l 内神经元 j 的输入连接结构如图 2 所示.

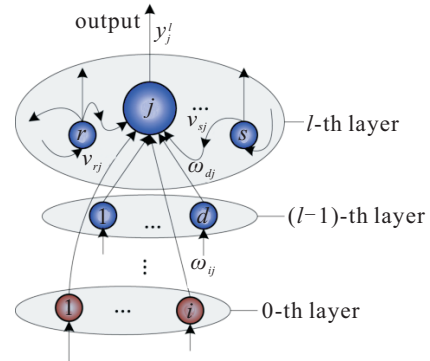


图 2 隐含层神经元的输入连接结构

图 2 中, v_{rj} 和 v_{sj} 分别为第 l 层内第 r 个神经元和第 s 个神经元作用于当前神经元 j 的抑制系数, ω_{ij} 和 ω_{dj} 分别为输入层内第 i 个神经元和第 $l - 1$ 层内第 d 个神经元与当前神经元 j 之间的连接权值. 神经元 j 的输出为

$$y_j^l = y_j'^l - \sum_{r=1, r \neq j}^{n_l} v_{rj} (y_r'^l - \theta_{rj}), \quad (2)$$

$$y_j'^l = f^l \left(\sum_{q=0}^{l-1} \left(\sum_{i=1}^{n_l} (\omega_{ij} y_i^q) \right) \right). \quad (3)$$

其中: $y_j'^l$ 为第 j 个神经元未受周围神经元侧抑制作用时的输出, v_{rj} 为神经元 r 抑制当前神经元 j 的系数, θ_{rj} 为神经元 r 能够抑制神经元 j 时需要达到的阈值.

3) 输出层. 在 S-LINN 中, 输出神经元可以接收来自输入层和隐含层内神经元的输出信号, 因此第 k 个输出神经元的输出为

$$y_k^L = f^L \left[\sum_{l=0}^{L-1} \left(\sum_{i=1}^{n_l} (\omega_{ik} y_i^l) \right) \right]. \quad (4)$$

不失一般性, 可得到 $L + 1$ 层 S-LINN 网络的输出表达式为

$$y = f^L \left[\sum_{k=1}^m \sum_{l=0}^{L-1} \left(\sum_{i=1}^{n_l} (\omega_{ik} y_i^l) \right) \right] = f^L \left[\sum_{k=1}^m \left(\sum_{i=1}^n (\omega_{ik} y_i^0) + \sum_{l=1}^{L-1} \left(\sum_{j=1}^{n_l} (\omega_{jk} y_j^l) \right) \right) \right] = f^L \left[\sum_{k=1}^m \left(\sum_{i=1}^n (\omega_{ik} f^0(x_i^0)) + \sum_{l=1}^{L-1} \left(\sum_{j=1}^{n_l} (\omega_{jk} y_j^l) \right) \right) \right]. \quad (5)$$

由于 y_j^l 是关于 x 的函数, y 也是 x 的函数. 显然, S-LINN 网络实现了 n 维空间到 m 维空间的映射. 但与传统的前馈网络相比, S-LINN 网络的输出不仅与输入神经元有直接关系, 也包含了隐含层内神经元之

间的侧向抑制作用,网络的表述能力明显提高.

1.2 S-LINN 网络中侧抑制作用分析

当仅有神经元 j 和 r 互相抑制时,由式 (2) 可知

$$\begin{aligned} y_j^l &= y_j^l - v_{rj}(y_r^l - \theta_{rj}), \\ y_r^l &= y_r^l - v_{jr}(y_j^l - \theta_{jr}). \end{aligned} \quad (6)$$

设定两个神经元互相抑制时,抑制阈值与抑制系数大小相等,即 $v_{jr} = v_{rj}$, $\theta_{rj} = \theta_{jr}$,则由式 (6) 可得

$$\begin{aligned} y_j^l - y_r^l &= \\ y_j^l - v_{rj}(y_r^l - \theta_{rj}) - [y_r^l - v_{jr}(y_j^l - \theta_{jr})] &= \\ (y_j^l - y_r^l)(1 + v_{rj}), \end{aligned} \quad (7)$$

即

$$\Delta y = \Delta y'(1 + v_{rj}).$$

其中: y 为神经元的输出, y' 为神经元的输入, Δ 为误差. 由于 $0 < v_{jr} < 1$, $\Delta y > \Delta y'$, 即在神经元之间引入的侧抑制连接,对网络输入信号之间的差别具有放大作用,这一特点有助于提高神经网络解决模式识别问题时区分类别的能力.

1.3 网络的权值调整过程

本节从单个神经元对应单个输入样本的角度出发,以单隐含层 ($L = 2$) 的 S-LINN 网络为例,推导基于反向传播原理的学习算法,所得结论也适用于多个隐含层的情况.

S-LINN 网络中隐含层神经元之间侧向抑制系数根据高斯分布选取,学习过程中暂不调整,因此对于 S-LINN 的训练实际上是对参数 ω_{ij} , ω_{jk} 和 ω_{ik} 的确定. 定义 S-LINN 的输入和输出神经元均为线性神经元,当 $l = 0, 2$ 时有 $f^l(x) = x$; 隐含层神经元为非线性神经元,其激活函数为 Sigmoid 函数,当 $l = 1$ 时有 $f^l(x) = (1 + e^{-x})^{-1}$.

对于第 p 个输入样本定义训练的目标函数为

$$J = \frac{1}{2P} \sum_{p=1}^P (t_p - y_p)^2. \quad (8)$$

其中: P 为样本总量, y_p 和 t_p 分别为第 p 个输入样本所对应的网络输出和期望输出.

基于最速下降梯度方法,定义 S-LINN 中网络权值学习规则为

$$\Delta \omega_{ik} = -\mu \frac{\partial J}{\partial \omega_{ik}}. \quad (9)$$

根据网络模型可得第 j 个隐含神经元和第 k 个输出神经元对应于第 p 个输入样本的输出为

$$\begin{aligned} y_{pj}^1 &= y_{pj}^1 - \sum_{r=1, r \neq j}^{n_1} v_{rj}(y_{pr}^1 - \theta_{rj}) = \\ f^1 \left(\sum_{i=1}^{n_0} (\omega_{ij} x_{pi}^0) \right) - \sum_{r=1, r \neq j}^{n_1} v_{rj}(y_{pr}^1 - \theta_{rj}) &= \end{aligned}$$

$$\begin{aligned} f^1 \left(\sum_{i=1}^n (\omega_{ij} x_{pi}^0) \right) - \sum_{r=1, r \neq j}^{n_1} v_{rj}(y_{pr}^1 - \theta_{rj}) &= \\ f^1 \left(\sum_{i=1}^n (\omega_{ij} x_{pi}) \right) - \sum_{r=1, r \neq j}^{n_1} v_{rj}(y_{pr}^1 - \theta_{rj}), \end{aligned} \quad (10)$$

$$\begin{aligned} y_{pk} &= y_{pk}^2 = f^2 \left(\sum_{i=1}^n (\omega_{ik} x_{pi}) + \sum_{j=1}^{n_1} (\omega_{jk} y_{pj}^1) \right) = \\ \sum_{i=1}^n (\omega_{ik} x_{pi}) + \sum_{j=1}^{n_1} (\omega_{jk} y_{pj}^1) &= \\ \sum_{i=1}^n (\omega_{ik} x_{pi}) + \sum_{j=1}^{n_1} \left[\omega_{jk} \left(f^1 \left(\sum_{i=1}^n (\omega_{ij} x_{pi}) \right) - \sum_{r=1, r \neq j}^{n_1} v_{rj}(y_{pr}^1 - \theta_{rj}) \right) \right]. \end{aligned} \quad (11)$$

根据式 (8), (10) 和 (11), 求取 S-LINN 网络内各参数的梯度方向为

$$\begin{cases} \frac{\partial J}{\partial \omega_{jk}} = \frac{\partial J}{\partial y_{pk}} \cdot \frac{\partial y_{pk}}{\partial \omega_{jk}}, \\ \frac{\partial J}{\partial \omega_{ik}} = \frac{\partial J}{\partial y_{pk}} \cdot \frac{\partial y_{pk}}{\partial \omega_{ik}}, \\ \frac{\partial J}{\partial \omega_{ij}} = \frac{\partial J}{\partial y_{pk}} \cdot \frac{\partial y_{pk}}{\partial y_{pj}^1} \cdot \frac{\partial y_{pj}^1}{\partial \omega_{ij}}. \end{cases} \quad (12)$$

其中

$$\begin{aligned} \frac{\partial J}{\partial y_{pk}} &= y_{pk} - t_{pk}, \quad \frac{\partial y_{pk}}{\partial \omega_{jk}} = y_{pk}(1 - y_{pk})y_{pj}^1, \\ \frac{\partial y_{pk}}{\partial \omega_{ik}} &= y_{pk}(1 - y_{pk})x_{pi}, \quad \frac{\partial y_{pk}}{\partial y_{pj}^1} = y_{pk}(1 - y_{pk})\omega_{jk}, \\ \frac{\partial y_{pj}^1}{\partial \omega_{ij}} &= y_{pj}^1(1 - y_{pj}^1)x_{pi}. \end{aligned}$$

由此可得各个权值的修正量分别为

$$\begin{cases} \Delta \omega_{jk} = \mu \sum_{p=1}^P \delta_{pk} y_{pj}^1, \\ \Delta \omega_{ik} = \mu \sum_{p=1}^P \delta_{pk} x_{pi}, \\ \Delta \omega_{ij} = \mu \sum_{p=1}^P \delta_{pj} x_{pi}. \end{cases} \quad (13)$$

其中

$$\begin{aligned} \delta_{pk} &= (y_{pk} - t_{pk})y_{pk}(1 - y_{pk}), \\ \delta_{pj} &= y_{pj}^1(1 - y_{pj}^1) \sum_{k=1}^m \delta_{pk} \omega_{jk}. \end{aligned}$$

2 仿真实验

S-LINN 引入了不同层神经元之间的跨越连接和同一层内神经元之间的侧抑制连接,丰富了网络的表达能力,增大了神经元对输入产生响应的对比. 本节利用 S-LINN 对模式识别问题进行处理,证明网络的有效性;与一些主流方法的处理结果进行比较,证明

网络的高效性. 为了最大限度地消除随机因素对实验结果的影响, 对每个仿真对象均独立运行 50 次后取其平均值进行比较.

2.1 乳腺癌诊断

乳腺癌诊断是通过对乳房肿瘤细胞特征的分析判断肿瘤性质的非线性模式识别问题, 常用于检验神经网络的性能. 本文选用 UCI 机器学习实验室的 Wisconsin breast cancer dataset^[9]进行网络性能分析, 该数据集由 Wolberg 收集, 共有 699 组记录, 每条记录包含 9 个属性, 其中诊断结果为良性和恶性的样本分别为 458 组和 241 组. 数据集内有 16 组样本存在属性缺失, 故有效样本为 683 组. 根据训练/测试样本数量比 (r_{sam}) 分别随机选择数据组成训练样本集和测试样本集, 网络的处理结果如图 3 和表 1 所示.

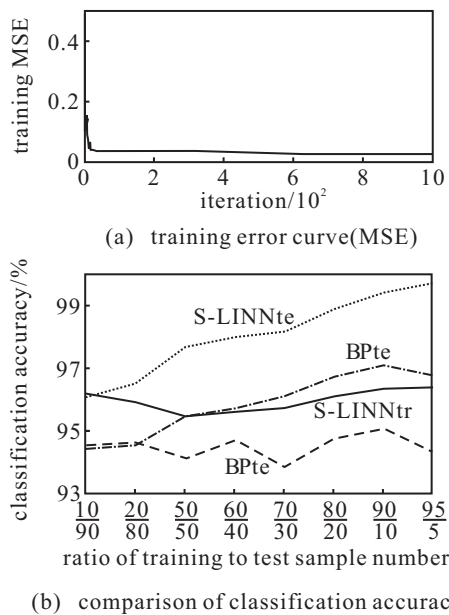


图 3 S-LINN 的学习曲线及正确分类率

图 3(a) 为含有 5 个隐含神经元的 3 层 S-LINN 在处理 Breast Cancer 诊断问题时某次运行的学习误差曲线, $r_{sam}=70/30$, training MSE=0.0242, test MSE=0.0027, training accuracy=97.23%, test accuracy=100%. 由实验结果和图示可知, S-LINN 网络能够很好地收敛于较小的稳定值(0.0242), 且达到了 100% 的测试精度, 表明 S-LINN 能够有效地处理该问题. 图 3(b) 为含有 3 个隐含神经元的 3 层 S-LINN 网络和含有 7 个隐含神经元的 3 层 BP 网络在不同的 r_{sam} 情况下对 Breast Cancer 数据集的学习和测试精度对比曲线, 其中“tr”和“ts”分别表示训练和测试阶段获得的诊断精度. 设定最大学习次数 $T_{max}=1000$, 容许误差 $\epsilon=1e-5$. 此外, 采用二进制数 (0 1) 和 (1 0) 的组合分别表示良性和恶性诊断结果, 因此 S-LINN 网络具有两个输出神经元.

表 1 Breast Cancer 数据集的学习性能比较

methods	network size	test accuracy/%
NN ^[10]	9-11-1	95.2
AR1+NN ^[10]	8-11-1	97.4
RS_SVM ^[11]	Subset5	96.87
LLWNN-RL ^[11]	/	97.2
S-LINN	9-5-2	99.76
S-LINN	9-3-2	99.70

表 1 为 S-LINN 与其他方法在 Breast Cancer 诊断预测中的性能对比. 其中: methods 为求解此问题所采用的各种方法; network size 列为采用的网络规模, 即 9-11-1 表示共使用 9 个输入神经元, 11 个隐含层神经元和 1 个输出神经元; test accuracy% 列为该方法在对应网络规模下求解的测试精度, 用百分数表示. 由图 3(b) 可知, 在含有 3 个隐含神经元(规模远小于含有 7 个隐含神经元的 BP 网络)的情况下, 无论是学习精度还是测试精度方面, S-LINN 都明显优于 BP 网络, 表明在处理此线性不可分问题时 S-LINN 不但具有较高的学习能力和泛化能力, 同时具有明显的优势. 由表 1 可知: S-LINN 的平均测试精度能够达到 99.70% (最高精度为 100% 且达到最高精度的比例为 46/50), 高于其他方法的最高精度 97.4%, 表明 S-LINN 网络具有较强的学习能力; 在仅有 3~5 个隐含神经元和 $r_{sam}=10/90$ 的情况下, S-LINN 仍有大于 96% 的学习和测试精度, 表明在网络规模和训练样本均较少的情况下, S-LINN 仍能保证相对理想的网络性能.

2.2 XOR 问题

XOR 是一个标准的线性不可分问题, 常用于检验网络处理非线性问题的能力, 其真值如表 2 中 #1~#4 所示. 研究表明, 仅含有 1 个隐层神经元的神经网络无法解决该问题, 增加一组样本 (#5) 的 XOR 问题只存在一个全局极小点^[13]. 因此, 为了验证 S-LINN 网络在处理此类非线性问题时的能力, 同时检验 S-LINN 网络能否有效避免陷入局部极小点, 将 S-LINN 网络用于解决改进的 XOR (简记为 XOR*) 问题, 设定 $T_{max}=20000$, $\epsilon=1e-15$. 表 2 后 3 列为对应网络结构处理 XOR* 问题时的实际输出和学习误差.

由表 2 可知, 对于结构为 2-1-1 的前馈 (BP) 网络无法解决的 XOR 问题, S-LINN 能够很好地解决, 并且能够达到或超越结构为 2-2-1 的前馈网络所取得的效果 ($6.96e-10 < 1.86e-9$). 另一方面, 与结构同为 2-2-1 的前馈网络相比, S-LINN 在学习精度方面具有显著的提高 ($9.99e-16 \ll 1.86e-9$). 图 4 为 S-LINN 与 BP 学习误差曲线对比, 与 BP 网络相比, 无论结构为 2-1-1 还是为 2-2-1 的 S-LINN 网络都能够更快、更好地收敛到一个满意的结果. 此外, 对于前馈网络容易陷入局部极小点的问题, S-LINN 网络都能有效地避

免,如图4内圆形所示. BP网络在陷入局部极小点时,不能很快地跳出,而结构为2-1-1的S-LINN网络能够在一段时间的学习之后跳出该点,结构为2-2-1的S-LINN网络能够在学习阶段很快地跳出局部极小点,这对于传统方法是明显的提升.

表 2 异或问题真值表和神经网络输出

Index	Inputs		Output	S-LINN		BP
	x_1	x_2		2-1-1	2-2-1	2-2-1
#1	0.01	0.01	0.01	0.0100	0.0100	0.0100
#2	0.01	0.99	0.99	0.9900	0.9900	0.9899
#3	0.99	0.01	0.99	0.9900	0.9900	0.9899
#4	0.99	0.99	0.01	0.0100	0.0100	0.0100
#5	0.50	0.50	0.99	0.9899	0.9900	0.9901
MSE	/	/	/	6.96e-10	9.99e-16	1.86e-9

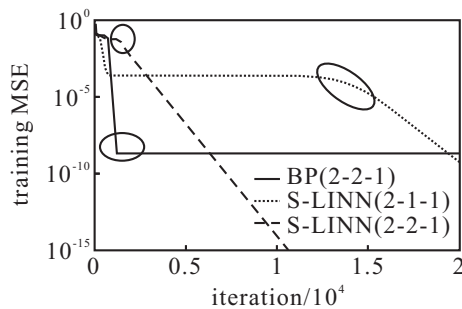


图 4 S-LINN 与 BP 学习误差曲线对比

3 结 论

本文根据生物仿生学原则,通过模仿大脑皮层内锥体神经元的跨层连接和抑制神经元的侧向抑制机制设计了一个多层神经网络模型S-LINN.该网络具有以下特点:

1)通过在隐含层神经元内引入侧抑制连接,增强了神经元输出之间的对比,提高了网络在处理聚类和辨识问题时的性能.

2)尽管S-LINN网络的设计连接较为复杂,但网络具有较强的处理问题能力,同时能够以更简单的网络规模实现或超越其他网络方法相当水平的性能,具有更好的学习能力和泛化能力.

3)S-LINN网络性能优越,尤其是在解决线性不可分问题方面,不但能够快速高效地获得理想的结果,而且能够有效地避免网络陷入局部极小而获得全局最优,为神经网络在复杂系统中的应用和自适应网络的设计提供了技术支持.

参考文献(References)

- [1] Kim S, Wang J. Lateral inhibition and concentration-invariant odor perception[J]. *J of Biology*, 2009, 8(1): 1-3.
- [2] Watts D J, Strogatz S H. Collective dynamics of "small-world" networks[J]. *Nature*, 1998, 393(6684): 440-442.
- [3] Barabasi A, Albert R. Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509-512.
- [4] Cruikshank S J, Lewis T J, Connors B W. Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex[J]. *Nature Neuroscience*, 2007, 10(4): 462-468.
- [5] Hartline H K, Ratliff F. Inhibitory interaction of receptor units in the eye of limulus[J]. *J of General Physiology*, 1957, 40(3): 357-376.
- [6] Xue Y B, Yang L, Haykin S. Decoupled echo state networks with lateral inhibition[J]. *Neural Networks*, 2007, 20(3): 365-376.
- [7] 张莉, 冯大政. 基于部分重设的侧抑制神经网络及其在图像分割中的应用[J]. *控制与决策*, 2012, 27(3): 464-467.
(Zhang L, Feng D Z. Lateral inhibition neural network with partial reset and its application in image segmentation[J]. *Control and Decision*, 2012, 27(3): 464-467.)
- [8] Smith L S. Adding lateral inhibition to a simple feedforward network enables it to perform exclusive-or[J]. *Neural Computation*, 1998, 10(2): 277-277.
- [9] Wolberg W. UCI repository of machine learning databases [DB/OL]. [1992-07-15][2012-01-20]. <http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wis/consin+%28Original%29>.
- [10] Karabatak M, Ince M C. An expert system for detection of breast cancer based on association rules and neural network[J]. *Expert Systems with Applications*, 2009, 36(2): 3465-3469.
- [11] Chen H, Yang B, Liu J, et al. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis[J]. *Expert Systems with Applications*, 2011, 38(7): 9014-9022.
- [12] Senapati M, Mohanty A, Dash S, et al. Local linear wavelet neural network for breast cancer recognition[J]. *Neural Computing & Applications*, 2013, 22(1): 125-131.
- [13] Gori M, Tesi A. On the problem of local minima in backpropagation[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1992, 14(1): 76-86.