

文章编号:1001-9081(2008)07-1662-03

基于小样本的神经网络参数优化选择方法

吴贵芳¹, 林青松¹, 孙秀明¹, 徐金梧², 徐科²

(1. 河南科技大学 电子信息工程学院, 河南 洛阳 471003; 2. 北京科技大学 机械工程学院, 北京 100083)

(easyfancy@126.com)

摘要:提出了一种基于小样本的神经网络参数优化选择方法,并应用于冷轧带钢表面缺陷在线检测系统中。该方法利用小样本对神经网络的每一种可能的组合参数进行模拟训练及测试,通过直方图求取最优化的一组神经网络参数。实验表明,利用该方法选择出来的神经网络参数,可以实现对冷轧带钢表面缺陷的最优识别。

关键词:小样本;最优参数;神经网络;冷轧带钢;表面缺陷

中图分类号: TP391.4 **文献标志码:** A

An optimizing parameter selection method of neural network based on small sample

WU Gui-fang¹, LIN Qing-song¹, SUN Xiu-ming¹, XU Jin-wu², XU Ke²

(1. College of Electronic Information Engineering, Henan University of Science and Technology, Luoyang Henan 471003, China;

2. College of Mechanical Engineering, University of Science and Technology Beijing, Beijing, 100083, China)

Abstract: An optimized parameter selection method of neural network based on small sample was presented, and it was applied to online surface defect inspection system of cold rolled strips. Small sample of the method was utilized to train and test neural network by every possible combination of parameters to get a group of neural network parameters by plotting histogram of recognition rate under different parameter combinations. Experiments show that optimum recognition of surface defect of cold rolled strips can be accomplished by the parameters that are selected by the optimizing parameter selection method.

Key words: small sample; optimum parameter; neural network; cold rolled strip; surface defects

0 引言

在应用神经网络解决实际问题时,神经网络参数的选择往往是影响分类器精度的一个重要因素。许多研究者靠实际经验或者从当前实验数据中对可能的网络参数一个一个地测试,然后找出一组实验范围内得到效果较好的参数^[1-2]。如果完全采用经验值作为网络参数,那就没有将具体问题的独特个性考虑进去,一旦个性表现得比较突出,利用神经网络得出的识别结果将与理想的结果会产生很大的偏差。如果采用实际的实验数据所得的结果进行一组一组地对比,那么在工作量上就显得非常大;另外训练网络一般比较复杂,则每实验一次所花的时间就会很长,并且要人工遍历参数的各种可能性,效率也相当低,实现起来也不太可能。

在表面检测领域,针对不同的对象,有各种不同的自动检测系统^[3-4]。冷轧带钢表面质量在线监测系统是集图像处理、模式识别和自动控制于一体的一套实时监测系统,能在现场恶劣条件下,24 小时不间断地实时监测冷轧带钢表面的质量。图 1 是一套安装在生产线上的冷轧带钢表面缺陷在线检测系统的总体结构图^[2]。在冷轧带钢表面质量在线监测系统中,带钢表面缺陷的识别是一个特别关键的环节,在很大程度上决定着系统的整体性能。系统能否有效识别已经检测出来的缺陷的类型,直接关系到该系统的智能性及实用性,并且

直接影响后续工作中对带钢的质量定级。

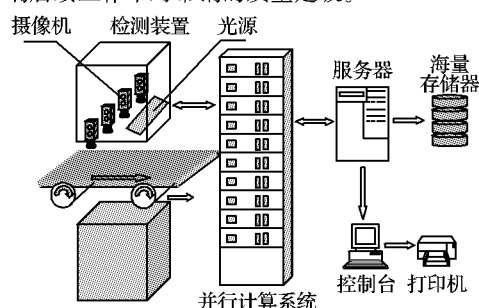


图 1 冷轧带钢表面质量在线系统的总体结构

但是在利用神经网络对系统进行缺陷识别的时候,最优参数的选择是其中最关键的问题之一^[5]。为了解决该参数选择问题,本文提出一种基于小样本^[6-7]的神经网络参数优化选择方法。该方法从具体需要解决的实际问题出发,既要考虑到样本太大造成的时间上的代价,又要考虑选择之后的结果能否真实反映出这一实际问题所需要的网络参数。

1 参数优化原理

基于小样本的神经网络参数优化选择,基本原理是从小样本的角度来寻找描述实际问题的内在参数,使找出的参数能真实地反映整个问题的解,为避免盲目参数的选择提供了

收稿日期:2008-01-25;修回日期:2008-04-06。 基金项目:国家自然科学基金资助项目(50074010);国家 863 计划项目(2003AA331080);河南科技大学人才科学研究基金资助项目(09001121)。

作者简介:吴贵芳(1978-),男,江西黎川人,副教授,博士,主要研究方向:数字图像处理、模式识别;林青松(1965-),男,河南洛阳人,副教授,硕士,主要研究方向:检测技术、自动化装置;孙秀明(1982-),女,河南开封人,硕士研究生,主要研究方向:数字图像处理;徐金梧(1949-),男,浙江宁波人,教授,博士,主要研究方向:故障诊断、图像处理;徐科(1972-),男,浙江宁波人,研究员,博士,主要研究方向:表面检测、模式识别。

可能。图2所示为基于小样本神经网络参数优化选择方法的流程图。

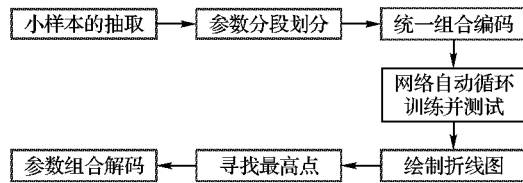


图2 算法流程

1.1 小样本的抽取

为了能真实地反映出实际问题的特性,将以一定的比例 η ,从样本库中抽取出一个样本集合。假设有一个原始样本 $P_0 = \{X_1, X_2, \dots, X_n\}$,其中每个样本的个数分别为 $p_0^1, p_0^2, \dots, p_0^n$,同时假定需要抽取的样本 $P_s = \{X'_1, X'_2, \dots, X'_n\}$ 的样本总数为 N ,则定义抽取规则如下。

按样本库中每一种类型样本的比例进行随机抽取,则每样本抽取的个数 p_s^i 表示为:

$$p_s^i = \frac{p_0^i}{\sum_{j=1}^n p_0^j} N; i = 1, 2, \dots, n \quad (1)$$

因此这样抽取出来的样本 P_s 具有与原始样本 P_0 一样的分布规律,即这个样本在一定范围内可以代替原始样本,所以用这个样本得出来的神经网络参数,应当在一定程度上具备了整个原始样本所适应的神经网络的参数。因此定义通过该方法得到的样本为小样本,也可以称为比例样本。在实际使用中可以抽取一组训练样本 p_{s1} 与一组测试样本 p_{s2} 。

1.2 参数分段划分与统一组合编码的确定

使用神经网络进行实际工程应用的时候,需要充分了解神经网络使用过程中的相关参数,在所有参数中,进一步选择其中需要确定的参数个数,并定义为 p ,理论上这些参数应当尽量与样本无关,即这些参数是用来代表实际问题的本质特性的。然后将每一个参数在一定经验值或理论值范围内进行若干细分。假设对这 p 个参数分别分为 C_1, C_2, \dots, C_p 份,然后以不重复组合的形式得到 M 个组合,可以用式(2)来表示:

$$M = \prod_{i=1}^p C_i^1 \quad (2)$$

接下来用 $1, 2, \dots, M$ 分别对每一参数组合进行编码。编码时首先需要确定所有参数的排列顺序,从前到后固定每一个参数的排列位置,比如顺序定为 (P_1, P_2, \dots, P_M) ,再分别对其进行编号。当前面的参数被给定一个编号后,后面的参数以由后到前的顺序遍历所有可能的值,每一次取不同的值时,相应的编号也增加1,直到出现 M 组为止,即代表编码结束。

1.3 神经网络自动循环测试及折线图绘制

利用程序,分别用 M 组参数中的具体参数作为神经网络的参数,并对每一组参数用小样本集 P_{s1} 进行训练。训练结束后,利用该组参数下已经训练好的神经网络,对测试样本集 p_{s2} 进行识别,同时统计相应的识别率。自动完成所有训练及测试之后,将不同参数下的识别率,以折线图的形式绘制出来,并做进一步的曲线分析。根据折线图,从中选择出识别率最高的一组参数作为神经网络的参数。

2 神经网络参数的优化求取过程

利用冷轧带钢表面质量在线监测系统,从现场采集相关的缺陷样本,分布如表1所示。

然后根据实际需要,从表1中按照小样本的抽取规则,抽

取一组训练小样本集 P_{s1} 和测试小样本集 P_{s2} ,并确定样本的抽取比例为8%和4%,于是得到如表2所示一个小样本分布情况。

表1 系统样本分布表

项目名称	训练样本数量	测试样本数量
划伤	1 223	612
粘结横纹	1 398	699
点粘结	1 203	602
羽纹	1 153	577
白斑	463	232
辊印	295	147
折边	193	96
锈斑	121	61
乳化液斑	89	45
橘皮纹	80	40
边裂	70	35
其他	71	35
总计	6 360	3 180

表2 小样本的分布情况

样本类别	小样本 p_{s1} 数量	小样本 p_{s2} 数量
划伤	98	49
粘结横纹	112	56
点粘结	96	48
羽纹	92	46
白斑	37	19
辊印	24	12
折边	15	8
锈斑	10	5
乳化液斑	7	4
橘皮纹	6	3
边裂	6	3
其他	6	3
总计	509	254

对于LVQ3神经网络来说,先根据小样本确定神经网络的竞争层神经元个数为106个。另外最需要确定的与描述系统本质有关的参数有三个,分别是:初始学习率 $\eta(0)$ 、相对学习率 ε 和窗口宽 m 。根据LVQ3神经网络的使用经验^[8],这三个参数的经验值取值范围一般如式(3)所示。

$$\begin{cases} 0.01 \leq \eta(0) \leq 0.1 \\ 0.05 \leq \varepsilon \leq 0.3 \\ 0.1 \leq m \leq 0.5 \end{cases} \quad (3)$$

因此可以把这三个参数分别在其取值范围内进行分割,将 $\eta(0)$ 分割为0.01、0.03、0.05、0.08以及0.1这五组;将 ε 分为0.05、0.1、0.2以及0.3这四组;将 m 分为0.1、0.2、0.3、0.4以及0.5这五组。因此总共可以组合成 $M = C_5^1 \times C_4^1 \times C_5^1 = 100$ 组,然后分别以1、2、 \dots 、100的方式进行编码,每一个编码对应一组不同的参数值。在这里的编码方法是对参数 $\eta(0)$ 、 ε 和 m 分别进行取值,当前面的取定一个值后,后面的参数以由后到前的顺序遍历所有可能的值,每一次取不同的值时,编号增加1。因此编号为5的参数组对应的各个参数值分别为(0.01, 0.05, 0.5),同理,第16组参数值为(0.01, 0.3, 0.1),以此方法可以计算出每一组参数值。

对于BP神经网络来说,主要是确定初始学习率 $\eta(0)$ 以及惯性项系数 $\alpha^{[1-2]}$,根据经验值, $\eta(0)$ 一般取0.9~1.0之间,而(一般是取0.7~0.9。于是可以将 $\eta(0)$ 分为0.9、0.91、

0.92、...、0.99 共 10 组,将(也分为 0.7、0.72、0.74、...、0.9 共 11 组,总共可以将两个参数组合成 110 组值,分别以 1、2、...、110 作为相应参数组合的编号,编码的方法与前面一样。因此第 5 组参数值为 (0.9, 0.78), 第 16 组参数值为 (0.91, 0.78), 以同样的方法可以得出每一组编号的具体参数值。

接下来分别利用小样本测试集 P_{s1} 对不同参数组合编号下的 LVQ3 神经网络和 BP 神经网络进行训练,同时统计出每一组参数下的测试集 P_{s2} 识别率,并绘制出折线图,分别得到如图 3 以及图 4 的测试曲线。

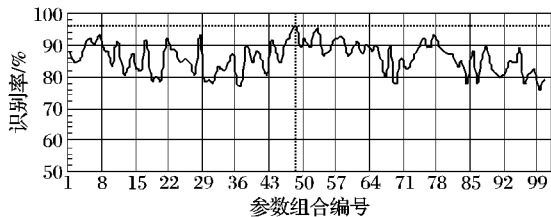


图 3 LVQ3 神经网络参数优化选择测试曲线

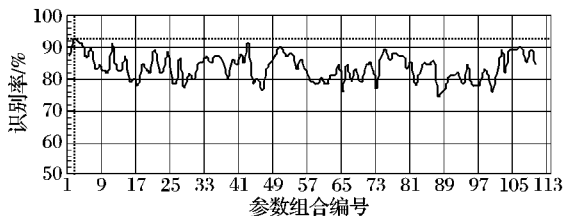


图 4 BP 神经网络参数优化选择测试曲线

从图 3 中可以看出,坐标 (48,96.2%) 为 LVQ3 神经网络识别率中的最高点,即第 48 组参数值下,识别率最高为 96.2%,因此对其进行解码可知它对应的参数组合值是 $\eta(0) = 0.05, \varepsilon = 0.1, m = 0.3$ 。在图 4 中可以看到,坐标 (2,

92.32%) 为 BP 神经网络识别率中的最高点,即在第 2 组参数值下,识别率最高为 92.3%,解码后得到它所对应的参数组合值是 $\eta(0) = 0.9, \alpha = 0.72$ 。于是就认为这两组参数在冷轧带钢表面缺陷的识别过程中分别代表两种神经网络的最佳参数。

3 实验结果

根据前面的神经网络参数优化选择方法得出的参数,确定如表 3 所示的 LVQ3 神经网络参数表,用它来构造冷轧带钢表面缺陷识别系统。同时利用几组不同的初始学习率 $\eta(0)$ 、相对学习率 ε 以及窗口宽度 m ,并分别构造出不同的冷轧带钢表面缺陷识别系统,并利用表 2 所示的训练样本以及测试样本来对所有不同参数下的 LVQ3 神经网络进行训练与测试,得到如表 3 的训练结果以及表 4 的测试结果。

表 3 LVQ3 神经网络参数表

项目名称	参数值
输入层神经元个数	22
竞争层神经元个数	1276
输出层神经元个数	12
初始学习率	0.05
相对学习率	0.1
窗口宽度 m	0.3
迭代次数	51040

从表 4 及表 5 的测试条件中可以看出,序号为 1 的测试条件中的神经网络参数均是由基于小样本的神经网络参数优化选择方法得到的,而从第 2 组到第 8 组的参数,都是由三个参数的随机组合而形成的。相应的识别率直方图如图 5 所示。

表 4 不同网络参数下的训练结果

组号	网络类型	测试条件			测试结果		
		$\eta(0)$	ε	m	正确识别个数	识别准确率/%	与第 1 组的差值/%
1	LVQ3	0.05	0.10	0.3	5996	94.28	0.00
2	LVQ3	0.01	0.05	0.1	5192	81.64	12.64
3	LVQ3	0.10	0.30	0.5	5404	84.97	9.31
4	LVQ3	0.08	0.20	0.4	5381	84.61	9.67
5	LVQ3	0.03	0.30	0.2	5488	86.29	7.99
6	LVQ3	0.06	0.08	0.3	5684	89.37	4.91
7	LVQ3	0.01	0.30	0.1	5475	86.08	8.20
8	LVQ3	0.05	0.10	0.5	5736	90.19	4.09

表 5 不同网络参数下的测试结果

组号	网络类型	测试条件			测试结果		
		$\eta(0)$	ε	m	正确识别个数	识别准确率/%	与第 1 组的差值/%
1	LVQ3	0.05	0.10	0.3	2852	89.69	0.00
2	LVQ3	0.01	0.05	0.1	2493	78.40	11.29
3	LVQ3	0.10	0.30	0.5	2598	81.70	7.99
4	LVQ3	0.08	0.20	0.4	2530	79.56	10.13
5	LVQ3	0.03	0.30	0.2	2645	83.18	6.51
6	LVQ3	0.06	0.08	0.3	2791	87.77	1.92
7	LVQ3	0.01	0.30	0.1	2609	82.04	7.65
8	LVQ3	0.05	0.10	0.5	2755	86.64	3.05

从表 4 及表 5 的测试结果中可以看出,在序号为 1 的参数下,所得到的训练集的识别准确率与测试集的识别准确率分别为 94.28% 与 89.69%,而在其他的 7 组参数下, LVQ3 神经网络对训练集的最高识别准确率是在第 8 组参数下得到,为 90.19%,最高测试集的识别准确率是在第 6 组参数下得

到,为 87.77%,这两个数据均低于第 1 组参数下的识别率。同时在这 7 组参数下,训练集的最低识别率出现在第 2 组参数下,其识别率为 81.64%,而测试样本的最低识别率也出现在第 2 组参数下,识别率为 78.40%,因此,它们分别比第一

(下转第 1671 页)

树是否有重复,并引入了优先级的概念,对分析产生的树有一定程度上的约束,但对树局部的重复却仍无能为力。而 CYK 图算法在分析表格中保存着分析过程中产生的每条边,实现了分析边的共享,避免了重复计算的问题,因此不论是在产生的边数还是花费的时间上都比算符优先计算具有优势。而引入优先级后的图算法 Prio-CYK 在产生的边数和耗时上则表现出更大的优势。

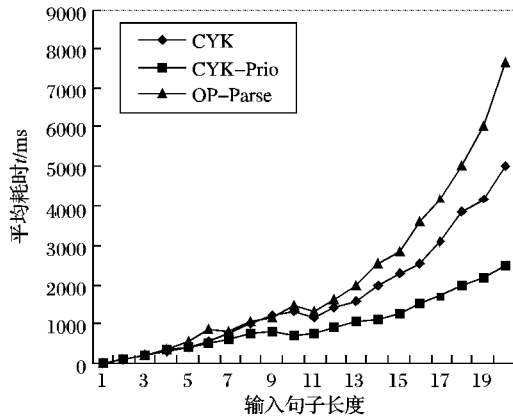


图2 分析耗时比较

5 结语

本文结合二元组合文法 (BCG) 的特点,提出了一种适合

(上接第1664页)

组参数下的识别准确率低 12.64% 与 11.29%。同时对后面 7 组参数下分别求它们的平均识别率,则对训练集与测试集的平均识别率分别为 86.16% 和 82.75%,分别比第 1 组参数下的识别准确率低 8.12% 和 6.94%。

因此本实验充分证明了经过基于小样本的神经网络参数优化选择方法,可以得出一组较优的神经网络参数,它能比较准确地体现当前问题的分类信息,利用这组参数构造的神经网络进行冷轧带钢表面缺陷识别,可以得到较高的识别准确率。

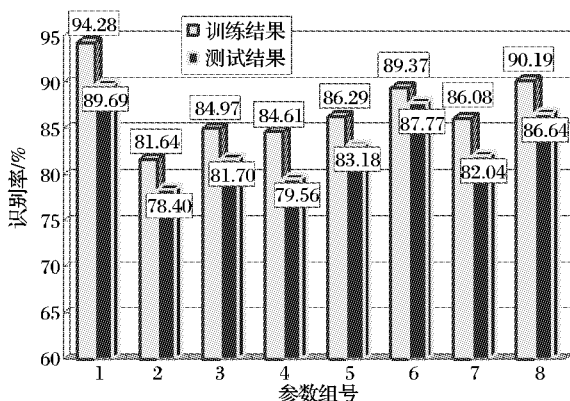


图5 不同网络参数下识别率的直方图

4 结语

由于神经网络在使用的时候,参数的选择在很大程度上影响着神经网络使用效果的好坏。由于神经网络涉及到多样本的训练与测试,所以在选择参数的时候,都只能靠经验进行选择。然而被选择的参数能否真正反映需要解决问题的本质,一直没有得到有效的解决。基于小样本的神经网络参数优化选择方法,从需要解决的问题的本质出发,研究了样本的

BCG 文法分析的图算法,该算法引入优先级的概念进行剪枝和部分歧义的消解,实现了树分析空间的共享,降低了空间和时间花费。相对于传统的图算法,该算法融合了 BCG 文法的优先级特性,可以解决部分剪枝和歧义消解问题,减轻概率型选择最优树操作的工作量。相对于算符优先的分析算法,该算法实现了分析树的共享,降低了分析空间的花费,提高了分析的效率。文中通过理论上的性能分析和汉语句法分析的试验,证明了该图算法在 BCG 文法分析中的可行性。

参考文献:

- [1] 曹雁峰, 万建成, 卢雷. 基于二元运算关系的汉语计算语法模型 [J]. 山东大学学报: 工学版, 2005, 35(1): 88-93.
- [2] YANG XIAO, WAN JIAN-CHENG, QIAO YONG-BO. A binary combinatorial grammar for Chinese and its parsing algorithm [C] // Proceedings of Sixth Intelligent Systems Design and Applications. Jinan: IEEE Computer Society, 2006: 761-766.
- [3] YANG XIAO, WAN JIAN-CHENG. A parsing algorithm of natural language based on operator precedence [C] // Proceedings of IEEE International Conference on Natural Language Processing and Knowledge Engineering. Wuhan: Beijing University of Posts and Telecommunications Press, 2005: 73-78.
- [4] 周强, 黄昌宁. 基于局部优先的汉语句法分析方法 [J]. 软件学报, 1999, 10(1): 1-6.

选择规律,在充分考虑了神经网络参数的经验值范围的基础上,确定出具体使用情况下的最优参数。实验证明,利用该方法选择得出的神经网络参数,在经验值范围内的使用效果是最佳的,能有效提高样本的识别率。

参考文献:

- [1] CAI K, XIA J T, LI L T, *et al.* Analysis of the electrical properties of PZT by a BP artificial neural network [J]. Computational Materials Science, 2005, 34(2): 166-172.
- [2] 徐科, 徐金梧, 班晓娟. 冷轧带钢表面质量自动监测系统的模式识别方法研究 [J]. 钢铁, 2002, 37(6): 28-31.
- [3] MORIYAMA I, TANABE Y, ISHIMARU I, *et al.* Surface inspection system with automatic discrimination of microscratches and particles on chemical mechanical polished wafers [C] // Proceedings of the 1999 IEEE International Symposium on Semiconductor Manufacturing. [S. l.]: IEEE Press, 1999: 139-142.
- [4] HAYATI S, BALARAM J, SERAJI H, *et al.* Remote surface inspection system [C] // Proceedings of 1993 IEEE International Conference on Robotics and Automation. [S. l.]: IEEE Press, 1993, 3: 875-882.
- [5] PEDRYCZ W, KWAK K C. Linguistic models as a framework of user-centric system modeling [J]. IEEE Transactions on Systems, Man and Cybernetics, 2006, 36(4): 727-745.
- [6] TANG BIN, LUO SI-WEI, HUANG HUA. High performance face recognition system by creating virtual sample [C] // Proceedings of the 2003 International Conference on Neural Networks and Signal Processing. Washington, DC: IEEE Press, 2003, 2: 972-975.
- [7] MOSZYNSKI M, BALCERZYK M, KAPUSTA M, *et al.* CdWO/sub 4/ crystal in gammaRay spectrometry [J]. IEEE Transactions on Nuclear Science, 2005, 52(6): 3124-3128.
- [8] KOHONEN T. New developments of learning vector quantization and the self-organizing map [C] // SYNAPSE'92. Osaka, Japan: [s. n.], 1992.