技术与方法 文章编号:0253-3626(2011)07-0846-03

基于遗传算法和 BP 神经网络的 CTG 识别研究

周红标[1],张宇林[1],胡金平[2]

(1. 淮阴工学院电子与电气工程学院电气教研室,淮安 223003;2. 兰州理工大学电信学院电子信息教研室,兰州 730050)

【摘 要】目的:针对正常类、非典型类和异常类 3 类胎心宫缩监护图(Cardiotocography,CTG),利用 BP 神经网络建立分类识别模型。方法:根据遗传算法(Genetic algorithm,GA)具有全局寻优的特点,引入遗传算法对神经网络权值和阈值进行优化,形成一种训练神经网络的混合算法(GA-BP 算法),以克服 BP 算法易陷入局部最优解的缺陷,并以 UCI 数据库中的 CTG 数据集为例进行测试。结论:仿真结果表明,对于正常类、非典型类和异常类 3 类样本,BP 算法的分类准确率分别为 97.32%、71.97%和 95.45%,而 GA-BP 算法的分类准确率分别提高到 98.24%、82.67%和 95.65%。可见 GA-BP 分类模型具有较强的学习能力和泛化能力,分类准确度更令人满意,所提出的方法是可行的。

【关键词】胎心宫缩监护;BP 神经网络;遗传算法;优化

【中国图书分类法分类号】TP391.9;R318.04 【文献标志码】A 【收稿日期】2010-11-22

CTG recognition based on genetic algorithm and BP neural network

ZHOU Hong-biao[1], *ZHANG Yu-lin*[1], *HU Jin-ping*[2]

(*1. Department of Electrical, Faculty of Electronic and Electrical Engineering, Huaiyin Institute of Technology;*
2. Department of Electronic Information, College of Telecom, Lanzhou Technology University)

【Abstract】*Objective*: To establish BP neural network (NN) classification model for normal, atypical and abnormal cardiotocography (CTG). *Methods*: Genetic algorithm (GA) was used to optimize the initial weights and threshold value of the neural network based on the inherent local minimum problem of NN and the good global convergence and global search ability of GA. The UCI's CTG set was used to test the algorithm. *Results*: The simulative results indicated that the recognition rates of the normal, atypical and abnormal CTG reached 98.24%, 82.67%, and 95.65% based on GA-BP, more than 97.32%, 71.97%, and 95.45% based on BPNN. The model has good ability both in learning and generalization and the algorithm can also effectively be used in other signal processing fields.

【Key words】cardiotocography; BP neural network; genetic algorithm; optimization

胎心宫缩监护(Cardiotocography,CTG),又称电子胎儿监护(Electronic fetalmonitoring,EFM),是指在待产中对于胎心变化和宫缩情况同时进行的持续描记,主要应用于产程中监测胎儿情况以改善新生儿预后[1]。CTG 既能够持续而精确地评估胎心基线的变异、胎动时胎心率变化,也能精确评估宫缩时胎心率变化。根据胎心率、变异性和是否出现胎心减速,并比较其与宫缩出现时间的关系,可将胎心率分为正常、非典型和异常 3 类[2]。临床上胎心宫缩图一直依靠人工判读,判读结果容易受主观因素影响。不同的专家间常常对同一 CTG 描记图有不同的解释,甚至同一专家在不同时间对同一描记图也有不同的解释。因此利用现代计算机技术和信号处理技术研究胎心宫缩图智能识别方法不仅能够减少主观因素的影响,而且能够更准确地评价胎儿状况,有效改善胎儿预后[3]。

作者介绍:周红标(1980-),男,讲师,硕士,研究方向:生物医学信号处理。

通信作者:张宇林,男,副教授,Email:15189544918@163.com。

基金项目:淮安市 2010 年度科技支撑资助项目(编号:SN1045);淮阴工学院科研基金资助项目(编号:HGG1009)。

台湾大学的 Chen Yu Chen 等人采用中值滤波、峰值检测和谷值检测法获得 FHR 基线、基线变异和胎心率变化信息,并基于 LabVIEW 软件开发了 CTG 类型的分析系统[4]。加拿大麦吉尔大学的 PA Warrick 等人将子宫压力作为激励,胎心率作为响应建立了线性的 CTG 系统辨识模型[5,6]。美国弗吉尼亚大学 H.Cao 等人为了对 CTG 做定量监测,分别采用相关分析和样本熵对子宫收缩期胎心率变异和胎心率减速作了定量检测[7]。上述线性时域分析法和非线性的样本熵分析法的优点是理论较容易被产科医生接受,也便于进行系统集成,但所建模型没有将 CTG 的全部信息体现出来。此外,CTG 本身是一个非线性动力学系统,应寻找更合适的非线性分析方法。神经网络作为一种智能识别算法在各领域中应用极其广泛,由于其具有容错、联想、推测、记忆、自适应、自学习和处理复杂多模式的功能,为 CTG 智能诊断开辟了一条新的途径[8]。神经网络智能诊断运用较多的是 BP 算法,BP 神经网络优点是寻优具有精确性,但同时存在收敛速度慢等缺点[9]。遗传算法[10](Genetic algorithm,GA)是近几年发展起来的基于生物进化原理的一种具有鲁棒性的识别方法。GA 基于样本

适应度函数对初始群体进行遗传操作，来指导学习和确定搜索的方向，从而获得网络结构优化的最优解。本文采用 BP 神经网络建立 CTG 识别模型，利用遗传算法来优化 BP 神经网络的权值和阈值，有效弥补了 BP 神经网络的缺陷，取得了较好的识别准确度。

1 BP 神经网络模型建立

BP 神经网络是一种多层前馈神经网络，该网络的建立过程包括以下几个步骤[11,12]：

步骤 1：网络初始化。根据系统输入输出序列(X,Y)确定网络输入层节点数 n、隐含层节点数 l，输出层节点数 m，初始化输入层、隐含层和输出层神经元之间的连接权值 ω_{ij}，ω_{jk} 初始化隐含层阈值 a，输出层阈值 b，给定学习速率和神经元激励函数。

步骤 2：隐含层输出计算。根据输入向量 X，输入层和隐含层间连接权值 ω_{ij} 以及隐含层阈值 a，计算隐含层输出 H。

$$H_j=f\left(\sum_{i=1}^{n}\omega_{ij}x_j-a_j\right)\quad j=1,2,\cdots,l \tag{1}$$

式(1)中，l 为隐含层节点数；f 为隐含层激励函数，该函数有多种表达形式。

步骤 3：输出层输出计算。根据隐含层输出 H，连接权值 ω_{jk} 和阈值 b，计算 BP 神经网络的预测输出 O。

$$O_k=\sum_{j=1}^{l}H_j\omega_{jk}-b_k\quad k=1,2,\cdots,m \tag{2}$$

步骤 4：误差计算。根据网络预测输出 O 和期望输出 Y，计算网络预测误差 e。

$$e_k=Y_k=O_k\quad k=1,2,\cdots,m \tag{3}$$

步骤 5：权值更新。根据网络预测误差 e 更新网络连接权值 ω_{ij}、ω_{jk}。

$$\omega_{ij}=\omega_{ij}+\eta H_j(1-H_j)x(i)\sum_{k=1}^{m}\omega_{jk}e_k$$

$$i=1,2,\cdots n;j=1,2,\cdots,l \tag{4}$$

$$\omega_{jk}=\omega_{jk}+\eta H_j(e_k)\qquad k=1,2,\cdots,m \tag{5}$$

式(4)和(5)中，η 为学习速率。

步骤 6：阈值更新。根据网络预测误差 e 更新网络节点阈值 a、b。

$$a_j=a_j+\eta H_j(1-H_j)\sum_{k=1}^{m}\omega_{jk}e_k\quad j=1,2,\cdots,l \tag{6}$$

$$b_k=b_k+e_k\quad k=1,2,\cdots,m \tag{7}$$

步骤 7：判断算法迭代是否结束，若没有结束，返回步骤2。

2 GA 优化 BP 神经网络

GA 借鉴了达尔文的进化论和孟德斯鸠的遗传学说[13]。GA 基本的操作分为选择、交叉和变异[14]。利用 GA 优化 BP 神经网络初始权值和阈值的过程如下：

步骤 1：参数设定及种群初始化。设定 GA 中的参数，主要包括个体编码长度 L，群体大小 M，遗传代数 G，交叉概率 P_c，变异概率 P_m 等。种群中的个体编码方法采用实数编码，由输入层与隐含层连接权值、隐含层阈值、隐含层与输出层连接权值以及输出层阈值 4 部分组成。

步骤 2：计算适应度函数。根据个体得到 BP 神经网络的初始权值和阈值，用训练数据训练 BP 神经网络后预测系统输出，把预测输出和期望输出之间的误差绝对值和 E 作为个体适应度值 F，计算公式为

$$F=k\left(\sum_{i=1}^{n}abs(y_i-o_i)\right) \tag{8}$$

式(8)中，n 为网络输出节点数；y_i 为 BP 神经网络第 i 个节点的期望输出；o_i 为第 i 个节点的预测输出；k 为系数。

步骤 3：遗传操作。根据个体的适应度值计算个体的选择概率

$$P_i=f_i/\sum_{j=1}^{n}f_j \tag{9}$$

式(9)中 $f_i=k/F_i$，F_i 为个体 i 的适应度值。采用选择概率 P_i、交叉概率 P_c、变异概率 P_m 等生成新个体，产生下一代种群。

步骤 4：将新个体插入种群中，并计算新个体的适应度函数。

步骤 5：判断算法是否满足结束条件。若找到满意个体，则算法结束；反之，重复步骤 3、4 直至满足适应度函数的预设标准，程序终止条件为 G 代内最高适应度值无显著变化。遗传算法优化 BP 神经网络算法流程如图 1 所示[15]。

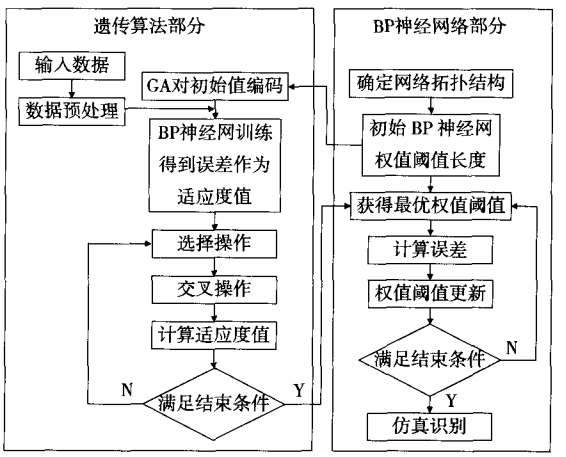


图 1 遗传算法优化 BP 神经网络算法流程图

3 结果与分析

本文的实验研究采用 UCI 的 Cardiotocography 数据集（可从 UCI 网站下载），它包含正常、非典型和异常 3 类样本，共 2 126 组数据。每组数据都由 3 位专业的产科医生进行分

类，他们对每组数据的标签取得了一致意见，在 Ayres de Campos 论文[16]中也提及这些数据是公认的可以用来评价分类指标的数据。其中正常样本 1 655 组，非典型样本 295 组，异常样本 176 组。每组数据经计算机处理，提取出 21 个特征值作为 BP 神经网络输入，输入的指标主要包括胎心率基线（FHR baseline）、胎心率加速（Acc）、胎心率减速（Dec）、长变异（LTV）、短变异（STV）、宫缩和胎动等。3 组数据的神经网络的输出分别标记为 100、010、001。

对于 BP 神经网络进行 CTG 识别，首先从数据集中随机选取 1 926 组数据作为训练集，剩余 200 组作测试集。对训练集和测试集数据分别进行了归一化，采用最大最小法，函数形式如下：

$$x_k=(x_k-x_{min})/(x_{max}-x_{min}) \tag{10}$$

式中，x_{min} 为数据序列中的最小数；x_{max} 为序列中的最大数。归一化是将原始数据规整到[0,1]范围内，可以用 MATLAB 自带的 mapminmax 函数完成。其次用 newff 函数建立BP 神经网络模型，输入层有 21 个节点，隐含层有 18 个节点，输出层有 3 个节点。隐含层和输出层传递函数分别采用tansig 和 purelin，训练函数采用 trainlm。学习步数设为 100步，学习目标设定为 0.000 1，学习率设定为 0.1。最后对网络进行训练和测试，得到的分类误差如图 2 所示。

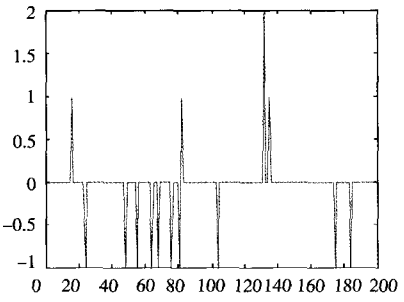


图 2 BP 网络分类误差图

对于 GA-BP 神经网络进行 CTG 识别，种群规模为 50，进化次数为 100 次，交叉概率为 0.4，变异概率为 0.2，遗传算法优化 BP 神经网络时适应度变化曲线如图 3 所示，GA-BP 网络分类误差如图 4 所示。

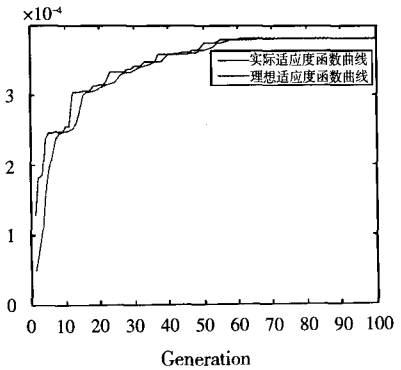


图 3 适应度变化曲线图

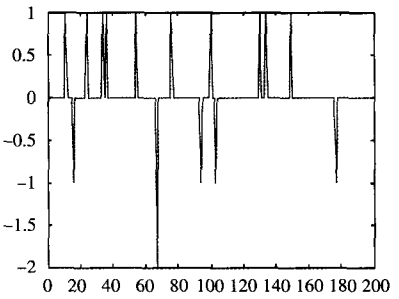


图 4 GA-BP 网络分类误差图

表 1 BP 和 GA-BP 网络的分类准确率

类别	正常类	非典型类	异常类
BP	97.32	71.97	95.45
GA-BP	98.24	82.67	95.65

由于神经网络的初始权值和阈值是随机产生，因此每次测试输出的分类准确率都不相同，故采用 BP 和 GA-BP 网络分别对数据进行 10 次处理，计算分类准确率的平均值来比较两者的分类性能，如表 1 所示。

对比图 2、4 和表 1，对于正常和异常两类样本，GA-BP 和 BP 的识别准确率都比较高，都达到了 95%以上，但是对于非典型类数据的识别，GA-BP 分类准确率为 82.67%，明显高于 BP 的 71.97%。同时，GA-BP 网络的稳定性也大大提高。可以通过 3 类样本 2 个网络的 10 次实验的极差和均方差来说明。BP 网络极差分别为 2.53、15.01 和 1.80，均方差分别为 2.47、38.97 和 1.35。而 GA-BP 网络极差分别为 1.16、4.07 和 1.24，均方差分别为 0.35、1.28 和 0.27。

非典型类样本识别准确率的提高，主要是由于 BP 神经网络训练算法是随机选取网络初始值，很容易陷入局部极小值，而 GA-BP 网络首先利用 GA 在全局范围内确定了较优的网络初值，很大程度上消除了局部极小的可能，获得了良好的网络性能，致使网络具有较强的泛化能力。而正常类和异常类两类样本的识别准确率提高幅度不大，主要是因为 BP 算法下这 2 组的识别率都已达到 95%以上，要想使 3 类样本达到更高的识别准确率，仅仅利用 GA 进行网络的优化是不可能的，只有从样本入手，如增加样本数量，使样本在 3 类中分布均匀等。此外，产科医生对正常类和异常类 CTG 的认同程度较高，而对非典型类 CTG 认同程度较低也是非典型类样本的识别准确率明显低于正常类和异常类的原因之一。

4 结束语

将神经网络引入到医学信号 CTG 的识别当中是目前生物医学信息处理的热点问题，但由于传统的 BP 算法是基于梯度下降，容易达到局部最优解，导致识别准确率较低。本文提出利用 GA 优化 BP 网络权值和阈值，将其按一定规律串接起来作为染色体进行遗传操作，优化后的网络初始值代入 BP 网络中进行 2 次训练，其学习速率大大加快，克服了陷入局部最优解的情况，网络的稳定性也大大提高，从而为 CTG

临床研究 文章编号：0253-3626（2011）07-0849-05

替比夫定治疗 HBeAg 阳性慢乙肝患者的 HBeAg 血清转换预测因素探讨

黄 晶，陈小苹，陈学福，陈文莉，陈 仁，马晓军，罗晓丹
（广东省医学科学院、广东省人民医院感染科，广州 510080）

【摘 要】目的：观察替比夫定治疗 HBeAg 阳性慢性乙型肝炎 52 周的临床疗效并探讨 HBeAg 血清转换预测因素。方法：100 例入组患者，口服替比夫定 600 mg，每天 1 次，治疗 52 周进行疗效评估。以 52 周出现 HBeAg 血清转换有无为应变量，以年龄、性别、基线 ALT、基线 HBeAg、基线 HBV DNA、治疗期间 HBV DNA 变化模式（12、24、36 周 HBV DNA 低于检测下限）、治疗期间 HBeAg 变化模式（12、24、36 周 HBeAg 下降>1 log 及 12、24、36 周 HBeAg 下降>2 log）14 个因素作为自变量进行多因素 Logistic 回归分析。结果：替比夫定治疗 52 周完全病毒学应答率 81%，HBeAg 血清转换率 31%，病毒学突破率 4%。治疗 52 周时 HBeAg 血清转换与基线 HBeAg 值相关（$P<0.01$），而与基线 ALT 值和基线 HBV DNA 无相关性。其中基线 HBeAg≤500 PEIU/ml 与 HBeAg>500 PEIU/ml 的患者 52 周时 HBeAg 血清转换率有统计学差异（$P<0.01$）。多因素 Logistic 回归分析显示 3 个因素进入 Logistic 回归方程，按其作用强弱依次为：24 周时 HBeAg 下降>2 log PEIU/ml；12 周时 HBeAg 下降>1 log PEIU/ml；基线 HBeAg 值，HBV DNA 未能进入。结论：替比夫定治疗 HBeAg 阳性慢性乙型肝炎，52 周 HBeAg 血清转换的预测因素以 24 周 HBeAg 下降>2 log PEIU/ml 的预测效果最强。

【关键词】替比夫定；慢性乙型肝炎；HBeAg 阳性；HBeAg 血清转换；预测因素
【中国图书分类法分类号】R512.6 【文献标志码】A 【收稿日期】2011-06-23

作者介绍：黄 晶（1976-），主治医师，硕士，
研究方向：慢性乙型肝炎的诊断与治疗。
通信作者：陈小苹，女，主任医师，Email：xiaopchen2003@yahoo.com.cn。

智能识别系统的临床应用提供了可能。

参 考 文 献

[1] 胡雅毅，刘兴会.产时胎心宫缩监护的评估及分级处理[J].中国实用妇科与产科杂志，2010，26(2)：95-98.
[2] 刘俊涛，宋亦军.胎儿血氧饱和度监测的临床应用[J].中国实用妇科与产科杂志，2010，26(2)：108-110.
[3] 张 力，胡晓吟，刘兴会.胎儿监护手段应用的循证医学评价[J].中国实用妇科与产科杂志，2010，26(2)：101-105.
[4] Chen C Y，Chen J C，Yu C，et al.A Comparative study of a new cardiotocography analysis program[J].31st Annual International Conference of the IEEE EMBS，2009，31(9)：2567-2570.
[5] Warrick P A，Kearney R E，Precup D，et al. System-identification noise suppression for intra-partum cardiotocography to discriminate normal and hypoxic fetuses[J].Computers in Cardiology，2006，33(5)：937-940.
[6] Warrick P A，Emily F，Hamilton，et al.Classification of normal and hypoxic fetuses from systems modeling of intrapartum cardiotocography [J].IEEE Transactions on biomedical engineering，2010，57(4)：771-779.
[7] Cao H，Lake D E，Chisholm C A，et al.Toward quantitative monitoring of human cardiotocography during labor[J].Proceedings of the 25th Annual Intemational Conference of the IEEE EMBS，2003，1(9)：236-239.
[8] 陆尧胜，卢 岑，李 晓，等.胎心宫缩图计算机分析系统的设计与实现[J].微计算机信息，2010，26(6-3)：26-28.
[9] 雷松林，郑永来.基于遗传算法和 BP 网络的岩爆预测[J].地下空间与工程学报，2009，5(2)：267-271.
[10] 周 燕.遗传算法与 BP 神经网络相结合的说话人识别系统[J].传感器与微系统，2009，28(6)：98-103.
[11] 乔晶晶，潘宏侠.基于遗传算法优化神经网络的齿轮故障诊断[J].水电能源科学，2010，28(6)：106-108.
[12] Liu M，Liang Y C.The application of GA-BP algorithm to intelligent diagnosis of coronary heart disease[J].CMCE，2010，4(8)：367-370.
[13] Tang H C，Wang P，Liu C X.Fault Diagnosis based on hybrid algorithm of an improved adaptive GA and BP[J].ICCASM，2010，14(10)：142-145.
[14] 章 勤，田 晶，孙傲冰，等.基于 BP 神经网络的矽肺病预测组合模型研究[J].计算机科学，2009，36(4)：264-267.
[15] MATLAB 中文论坛编著.MATLAB 神经网络 30 个案例分析[M].北京航空航天大学出版社，2010.1-35.
[16] Ayres de Campos.A program for automated analysis of cardiotocograms[J].J Matern Fetal Med，2000，35(5)：311-318.

（责任编辑：冉明会）