



(12) 发明专利申请

(10) 申请公布号 CN 112200068 A

(43) 申请公布日 2021.01.08

(21) 申请号 202011072931.4

(22) 申请日 2020.10.09

(71) 申请人 武汉科技大学

地址 430081 湖北省武汉市青山区和平大道947号

(72) 发明人 钱龙 王文波

(74) 专利代理机构 北京金智普华知识产权代理有限公司 11401

代理人 杨采良

(51) Int. Cl.

G06K 9/00 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

G06N 3/12 (2006.01)

权利要求书2页 说明书9页 附图3页

(54) 发明名称

基于LSTM神经网络心电提取方法、系统、设备及介质

(57) 摘要

本发明属心电监测技术领域,公开了一种基于LSTM神经网络心电提取方法、系统、设备及介质,采用GA对LSTM网络的超参数作寻优处理,得出学习率、隐层神经元数和训练次数的最佳组合;利用寻优的参数组合构建的GA-LSTM模型作为母体胸部心电信号与腹壁混合信号之间的非线性变换函数;在此基础上应用非线性变换函数计算母体胸部信号的最佳估计;从母体腹壁的混合信号中进行分离,即可分离提取得到最佳的胎儿心电信号的估计。本发明出的心电信号提取模型可以获取更为清晰的胎儿心电信号。



1. 一种基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,应用于计算机终端,所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法包括:

采用GA对LSTM网络的超参数作寻优处理,得出学习率、隐层神经元数和训练次数的最佳组合;

利用寻优的参数组合构建的GA-LSTM模型作为第一待测样本胸部心电信号与腹壁混合信号之间的非线性变换函数;在此基础上应用非线性变换函数计算所述第一待测样本胸部信号的最佳估计;

从所述腹壁的混合信号中进行分离提取得到最佳的第二待测样本心电信号的估计。

2. 如权利要求1所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法包括以下步骤:

步骤一,采集第一待测样本心电信号数据集,并将采集的第一待测样本心电信号数据集的部分样本作为训练数据集;

步骤二,利用GA优化LSTM网络超参数;GA-LSTM模型的输入信号向量和目标输出信号向量输入最优参数组合的GA-LSTM网络,并输出腹壁混合信号中的第一待测样本心电成分;

步骤三,利用误差信号表示将目标信号与输出信号之差;同时GA-LSTM根据均方误差最小化拟合误差,得到第一待测样本胸部心电信号经历非线性变换传输到腹壁的最优拟合函数;

步骤四,将心电数据输入GA-LSTM模型,利用最优拟合函数计算得到腹壁混合信号中的第一待测样本心电;并基于第一待测样本心电计算得到第二待测样本心电信号。

3. 如权利要求2所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,步骤二中,所述利用GA优化LSTM网络超参数包括:

(1) 将LSTM网络模型中学习率、批处理大小、神经网络隐藏层单元数目作为优化对象,执行种群的初始化以及染色体编码与解码处理;

(2) 计算初始种群中各个个体适应度大小;

(3) 对染色体进行选择、交叉和变异处理;

(4) 对染色体解码、计算种群内个体的适应度,保留适应度小的个体,淘汰适应度大的个体;

(5) 判断是否符合遗传终止条件,若不符合遗传终止条件则回到步骤(3);若符合遗传终止条件,则将GA求出的最优参数作为LSTM网络模型的最终参数。

4. 如权利要求3所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,步骤(1)中,所述染色体编码包括:将LSTM网络中的隐藏层神经元数目、学习率和训练次数作为GA的初始化对象,通过实数编码形式进行染色体编码;所述隐藏层神经元的区间范围为 $[5, 40]$,所述学习率的区间范围为 $[0.001, 0.1]$,所述训练次数的区间范围为 $[50, 500]$ 。

5. 如权利要求3所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,步骤(3)中,所述对染色体进行选择、交叉和变异处理包括:

采用tournament selection算法作为GA的选择策略在当前种群中选择适应性较好的个体作为亲本,并将遗传信息传递给子代,进行选择处理;

在交叉之前在父代中利用random.shuffle函数进行洗牌运算,当在 $(0, 1)$ 之间产生的

随机数小于所给的交叉率大小,则进行交叉变换;

当在(0,1)之间产生的随机数小于所给的变异率大小,则进行变异处理。

6.如权利要求5所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,所述变异处理包括:

隐藏层神经元个数、学习率和训练次数变异的规则如下所示:

$$c.ln = \text{abs}(c.ln + \text{random.randint}(-3, 3))$$

$$c.lr = \text{abs}(c.lr + \text{random.uniform}(-0.001, 0.001))$$

$$c.maxe = \text{abs}(c.maxe + \text{random.randint}(-30, 30));$$

其中:c.ln为种群中隐藏层神经元个数;c.lr为种群中的学习率;c.maxe为种群中的训练次数。

7.如权利要求3所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,步骤(4)中,所述对染色体解码、计算种群内个体的适应度包括:

适应度函数如下:

$$\text{fitness} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2};$$

其中,N表示心电数据总量, \hat{y}_n 为母亲腹壁混合信号的网络预测值, y_n 为母亲腹壁混合信号的真实值。

8.如权利要求2所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法,其特征在于,步骤四中,所述基于第一待测样本心电计算得到第二待测样本心电信号包括:

$$\hat{r}_i = u_i - \hat{s}_i;$$

其中, \hat{r}_i 表示胎儿心电信号; u_i 表示腹壁混合信号; \hat{s}_i 表示腹壁混合信号 u_i 中的母体心电成分; $\hat{s}_i = \hat{f}(m_i)$ 。

9.一种计算机设备,其特征在于,所述计算机设备包括存储器和处理器,所述存储器存储有计算机程序,所述计算机程序被所述处理器执行时,使得所述处理器执行权利要求1-7任意一项所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法。

10.一种计算机可读存储介质,储存有指令,当所述指令在计算机上运行时,使得计算机执行如权利要求1-7任意一项所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法。

基于LSTM神经网络心电提取方法、系统、设备及介质

技术领域

[0001] 本发明属于心电监测技术领域,尤其涉及一种基于LSTM神经网络心电提取方法、系统、设备及介质。

背景技术

[0002] 目前,孕妇围产期胎儿监护对了解胎儿在子宫内的生长发育状况十分重要,通常采用多普勒超声波、胎音听诊、胎心率检测以及胎儿心电图的变化来诊断胎儿在母体子宫内的健康状况,后者通过心电图波形分析可以更早期地反映出妊娠期胎儿心脏的各种电生理活动。检测胎儿的心电图中QRS波的变化可及时判断胎儿是否存在缺氧、脐带绕颈等妊娠问题。然而胎儿心电信号通常采用间接法从母亲腹壁获取,因而会包含多种噪声:电极干扰、母体心电成分以及基线漂移等。所以,如何消除采集的胎儿心电所包含的噪声,从而分离出清晰的胎儿心电信号成为一个热门的钻研课题。

[0003] 近年来,关于胎儿心电图提取的方式提出了多种检测方法:独立成分分析法假定胎儿心电、母体心电以及噪声相互统计独立,从而建立ICA模型,在此模型基础上提取胎儿心电信号,但是此方法不适合处理超高斯和亚高斯信号,而且容易陷入局部最优解;盲源提取技术在各个源信号未知的条件下,从腹壁混合信号中分离得到胎儿的心电信号。但是在该模型中,胎儿心电图的相关特征并未得到充分利用,导致提取的胎儿心电准确率较低且不能清楚地反映信号的生理学意义,而且该模型的算法大多需要精确估计胎儿心电的延迟周期,对胎儿时间延迟的周期估计较为敏感,因此该算法的心电数据提取的结果会受到时间延迟周期估计的限制;自适应滤波法计算简单,但对于非平稳性胎儿心电的测量具有局限性;匹配滤波法算法胎儿心电识别率低,容易出现误漏诊;小波分解法难以选择小波基等参数,对于母体心电和胎儿心电重叠的部分很难进行胎儿心电提取;奇异值分解法分解出的分矩阵解释性往往不强且分解矩阵随时间越来越大;神经网络和支持向量机方法取得了较好的胎儿心电提取效果,但存在泛化能力弱、易落入局部最优值、网络结构设计难等问题。

[0004] 近年来,循环神经网络(Recurrent Neural Network,RNN)因处理序列信息功能强大,可预测时间信息的能力得到了发展。Hochreiter和Schmidhuber对RNN网络单元结构进行改进提出了长短时记忆网络(Long Short Term Memory networks,LSTM)模型,通过设计“门”结构解决了梯度消失和梯度爆炸以及RNN信息记忆能力不足的问题,可充分运用远距离的时序信息。LSTM神经网络在语音识别、文本处理等领域已得到成功应用。但LSTM神经网络的关键超参数难以确定,如隐层神经元数、学习率和训练次数对模型的拟合能力起着决定性作用,学习率和训练次数直接影响模型收敛速度和计算时长。LSTM的网络结构参数直接控制模型拓扑结构,因此采用不同的超参数建立的网络模型其预测性能具有较大差异,所以如何选择合适参数对于建立模型来说显得至关重要。目前,往往依赖研究者的经验和多次实验结果去选择网络模型的超参数,具有较大的随机性,降低了模型的预测性能。

[0005] 通过上述分析,现有技术存在的问题及缺陷为:现有胎儿心电图提取方法提取的

胎儿心电准确率较低且不能清楚地反映信号的生理学意义,对于非平稳性胎儿心电的测量具有局限性、容易出现误漏诊;对于母体心电和胎儿心电重叠的部分很难进行胎儿心电提取;神经网络和支持向量机方法存在泛化能力弱、易落入局部最优值、网络结构设计难等问题。

发明内容

[0006] 针对现有技术存在的问题,本发明提供了一种基于LSTM神经网络心电提取方法、系统、设备及介质;具体涉及一种基于遗传算法优化的LSTM神经网络胎儿心电提取方法。

[0007] 本发明是这样实现的,一种基于遗传算法优化的LSTM神经网络胎儿心电提取方法,应用于计算机终端,所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法包括:

[0008] 首先,采用GA对LSTM网络的超参数作寻优处理,得出学习率、隐层神经元数和训练次数的最佳组合;

[0009] 然后,利用寻优的参数组合构建的GA-LSTM模型作为母体胸部心电信号与腹壁混合信号之间的非线性变换函数;在此基础上应用非线性变换函数计算母体胸部信号的最佳估计;

[0010] 最后,从母体(第一待测样本)腹壁的混合信号中进行分离,即可分离提取得到最佳的胎儿(第二待测样本)心电信号的估计。

[0011] 进一步,所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法包括以下步骤:

[0012] 步骤一,采集母体心电信号数据集,并将采集的母体心电信号数据集的部分样本作为训练数据集

[0013] 步骤二,利用GA优化LSTM网络超参数;GA-LSTM模型的输入信号向量和目标输出信号向量输入最优参数组合的GA-LSTM网络,并输出腹壁混合信号中的母体心电成分;

[0014] 步骤三,利用误差信号表示将目标信号与输出信号之差;同时GA-LSTM根据均方误差最小化拟合误差,得到母体胸部心电信号经历非线性变换传输到腹壁的最优拟合函数;

[0015] 步骤四,将心电数据输入GA-LSTM模型,利用最优拟合函数计算得到腹壁混合信号中的母体心电;并基于母体心电计算得到胎儿心电信号。

[0016] 进一步,步骤二中,所述利用GA优化LSTM网络超参数包括:

[0017] (1) 将LSTM网络模型中学习率、批处理大小、神经网络隐藏层单元数目作为优化对象,执行种群的初始化以及染色体编码与解码处理;

[0018] (2) 计算初始种群中各个个体适应度大小;

[0019] (3) 对染色体进行选择、交叉和变异处理;

[0020] (4) 对染色体解码、计算种群内个体的适应度,保留适应度小的个体,淘汰适应度大的个体;

[0021] (5) 判断是否符合遗传终止条件,若不符合遗传终止条件则回到步骤(3);若符合遗传终止条件,则将GA求出的最优参数作为LSTM网络模型的最终参数。

[0022] 进一步,步骤(1)中,所述染色体编码包括;将LSTM网络中的隐藏层神经元数目、学习率和训练次数作为GA的初始化对象,通过实数编码形式进行染色体编码;所述隐藏层神经元的区间范围为[5,40],所述学习率的区间范围为[0.001,0.1],所述训练次数的区间范

围为[50,500]。

[0023] 进一步,步骤(3)中,所述对染色体进行选择、交叉和变异处理包括:

[0024] 采用tournament selection算法作为GA的选择策略在当前种群中选择适应性较好的个体作为亲本,并将遗传信息传递给子代,进行选择处理;

[0025] 在交叉之前在父代中利用random.shuffle函数进行洗牌运算,当在(0,1)之间产生的随机数小于所给的交叉率大小,则进行交叉变换;

[0026] 当在(0,1)之间产生的随机数小于所给的变异率大小,则进行变异处理。

[0027] 进一步,所述变异处理包括:

[0028] 隐藏层神经元个数、学习率和训练次数变异的规则如下所示:

[0029] $c.ln = \text{abs}(c.ln + \text{random.randint}(-3, 3))$

[0030] $c.lr = \text{abs}(c.lr + \text{random.uniform}(-0.001, 0.001))$

[0031] $c.maxe = \text{abs}(c.maxe + \text{random.randint}(-30, 30))$;

[0032] 其中: $c.ln$ 为种群中隐藏层神经元个数; $c.lr$ 为种群中的学习率; $c.maxe$ 为种群中的训练次数。

[0033] 进一步,步骤(4)中,所述对染色体解码、计算种群内个体的适应度包括:

[0034] 适应度函数如下:

[0035]
$$\text{fitness} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2}$$
;

[0036] 其中, N 表示心电数据总量, \hat{y}_n 为母亲腹壁混合信号的网络预测值, y_n 为母亲腹壁混合信号的真实值。

[0037] 进一步,步骤四中,所述基于母体心电计算得到胎儿心电信号包括:

[0038] $\hat{r}_i = u_i - \hat{s}_i$;

[0039] 其中, \hat{r}_i 表示胎儿心电信号; u_i 表示腹壁混合信号; \hat{s}_i 表示腹壁混合信号 u_i 中的母体心电成分; $\hat{s}_i = \hat{f}(m_i)$ 。

[0040] 本发明的另一目的在于提供一种一种计算机设备,所述计算机设备包括存储器和处理器,所述存储器存储有计算机程序,所述计算机程序被所述处理器执行时,使得所述处理器执行所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法。

[0041] 本发明的另一目的在于提供一种计算机可读存储介质,储存有指令,当所述指令在计算机上运行时,使得计算机执行所述基于遗传算法优化的LSTM神经网络胎儿心电提取方法。

[0042] 结合上述的所有技术方案,本发明所具备的优点及积极效果为:本发明以LSTM神经网络模型为技术基础研究建立了胎儿心电信号提取模型,利用遗传算法(genetic algorithm,GA)对LSTM的关键超参数进行优化,利用遗传算法强大的全局寻优能力求得模型的最优参数组合。本发明模型以单导联母体的胸部心电信号作为神经网络的输入信号,以单导联母体腹壁混合信号作为提取目标的输出信号,采用GA优化的LSTM神经网络模型方法来检测和评算母体胸部的心电信号传输到母体腹壁的最佳信号估计;然后将母体胸部信

号的最佳估计从母体腹壁的混合信号中进行分离,即可分离提取得到纯净的胎儿心电信号。实验结果表明,与目前传统的NLMS方法、SVM方法、LSTM双导联神经网络模型方法相比,本发明提出的心电信号提取模型可以获取更为清晰的胎儿心电信号。

附图说明

[0043] 为了更清楚地说明本申请实施例的技术方案,下面将对本申请实施例中所需要使用的附图做简单的介绍,显而易见地,下面所描述的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下还可以根据这些附图获得其他的附图。

[0044] 图1是本发明实施例提供的胎儿心电提取方法原理图。

[0045] 图2是本发明实施例提供的基于遗传算法优化的LSTM神经网络胎儿心电提取方法流程图。

[0046] 图3是本发明实施例提供的LSTM神经网络结构图。

[0047] 图4是本发明实施例提供的4种方法的胎儿心电信号提取结果比较示意图。

[0048] 图5是本发明实施例提供的胎儿心电信号提取实验结果对比(500-700点)示意图。

[0049] 图6是本发明实施例提供的胎儿心电信号提取实验结果对比(2050-2500点)示意图。

具体实施方式

[0050] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0051] 针对现有技术存在的问题,本发明提供了一种基于LSTM神经网络心电提取方法、系统、设备及介质,下面结合附图对本发明作详细的描述。

[0052] 如图1所示,本发明实施例提供的基于遗传算法优化的LSTM神经网络胎儿心电提取方法包括:

[0053] 首先,采用GA对LSTM网络的超参数作寻优处理,得出学习率、隐层神经元数和训练次数的最佳组合;

[0054] 然后,利用寻优的参数组合构建的GA-LSTM模型作为母体胸部心电信号与腹壁混合信号之间的非线性变换函数;在此基础上应用非线性变换函数计算母体胸部信号的最佳估计;

[0055] 最后,从母体腹壁的混合信号中进行分离,即可分离提取得到最佳的胎儿心电信号的估计。

[0056] 如图2所示,本发明实施例提供的基于遗传算法优化的LSTM神经网络胎儿心电提取方法包括以下步骤:

[0057] S101,采集母体心电信号数据集,并将采集的母体心电信号数据集的部分样本作为训练数据集;

[0058] S102,利用GA优化LSTM网络超参数;GA-LSTM模型的输入信号向量和目标输出信号向量输入最优参数组合的GA-LSTM网络,并输出腹壁混合信号中的母体心电成分;

[0059] S103,利用误差信号表示将目标信号与输出信号之差;同时GA-LSTM根据均方误差最小化拟合误差,得到母体胸部心电信号经历非线性变换传输到腹壁的最优拟合函数;

[0060] S104,将心电数据输入GA-LSTM模型,利用最优拟合函数计算得到腹壁混合信号中的母体心电;并基于母体心电计算得到胎儿心电信号。

[0061] 步骤S102中,本发明实施例提供的利用GA优化LSTM网络超参数包括;

[0062] (1)将LSTM网络模型中学习率、批处理大小、神经网络隐藏层单元数目作为优化对象,执行种群的初始化以及染色体编码与解码处理;

[0063] (2)计算初始种群中各个个体适应度大小;

[0064] (3)对染色体进行选择、交叉和变异处理;

[0065] (4)对染色体解码、计算种群内个体的适应度,保留适应度小的个体,淘汰适应度大的个体;

[0066] (5)判断是否符合遗传终止条件,若不符合遗传终止条件则回到步骤(3);若符合遗传终止条件,则将GA求出的最优参数作为LSTM网络模型的最终参数。

[0067] 步骤(1)中,本发明实施例提供的染色体编码包括;将LSTM网络中的隐藏层神经元数目、学习率和训练次数作为GA的初始化对象,通过实数编码形式进行染色体编码;所述隐藏层神经元的区间范围为[5,40],所述学习率的区间范围为[0.001,0.1],所述训练次数的区间范围为[50,500]。

[0068] 步骤(3)中,本发明实施例提供的对染色体进行选择、交叉和变异处理包括:

[0069] 采用tournament selection算法作为GA的选择策略在当前种群中选择适应性较好的个体作为亲本,并将遗传信息传递给子代,进行选择处理;

[0070] 在交叉之前在父代中利用random.shuffle函数进行洗牌运算,当在(0,1)之间产生的随机数小于所给的交叉率大小,则进行交叉变换;

[0071] 当在(0,1)之间产生的随机数小于所给的变异率大小,则进行变异处理。

[0072] 本发明实施例提供的变异处理包括:

[0073] 隐藏层神经元个数、学习率和训练次数变异的规则如下所示:

[0074] $c.ln = \text{abs}(c.ln + \text{random.randint}(-3, 3))$

[0075] $c.lr = \text{abs}(c.lr + \text{random.uniform}(-0.001, 0.001))$

[0076] $c.maxe = \text{abs}(c.maxe + \text{random.randint}(-30, 30))$;

[0077] 其中: $c.ln$ 为种群中隐藏层神经元个数; $c.lr$ 为种群中的学习率; $c.maxe$ 为种群中的训练次数。

[0078] 步骤(4)中,本发明实施例提供的对染色体解码、计算种群内个体的适应度包括:

[0079] 适应度函数如下:

[0080]
$$\text{fitness} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2}$$
;

[0081] 其中, N 表示心电数据总量, \hat{y}_n 为母亲腹壁混合信号的网络预测值, y_n 为母亲腹壁混合信号的真实值。

[0082] 步骤S104中,本发明实施例提供的基于母体心电计算得到胎儿心电信号包括:

[0083]
$$\hat{r}_i = u_i - \hat{s}_i$$
;

[0084] 其中, \hat{r}_i 表示胎儿心电信号; u_i 表示腹壁混合信号; \hat{s}_i 表示腹壁混合信号 u_i 中的母体心电成分; $\hat{s}_i = \hat{f}(m_i)$ 。

[0085] 下面结合具体实施例对本发明的技术方案作进一步说明。

[0086] 实施例1:

[0087] 胎儿心电提取原理

[0088] 目前普遍使用置电极法采集胎儿心电信号:在时刻 i , 由置于孕妇胸部的电极采集孕妇胸部心电信号 m_i , 且同时刻由置于孕妇腹壁的电极采集孕妇腹壁混合信号 u_i 。而在孕妇腹壁心电信号 u_i 中包含母体心电成分 s_i 、胎儿心电 d_i 和噪声 z_i 三类信号, 三类信号函数关系如式 (1) 所示:

$$[0089] \quad u_i = s_i + d_i + z_i, i = 1, 2, \dots, n \quad (1)$$

[0090] 其中, 母体心电成分 s_i 是在时刻 i 由母体胸部心电 m_i 经非线性变换传输到孕妇腹壁所形成的信号, 所以腹壁混合信号中的母体心电成分 s_i 的相位和幅度等相关参数都会发生改变, 非线性变换函数关系如式 (2) 所示:

$$[0091] \quad s_i = f(m_i) \quad (2)$$

[0092] 式 (2) 中, $f(\cdot)$ 表示心电信号的非线性变换函数。假设能求得 $f(\cdot)$ 的最佳估计值 $\hat{f}(\cdot)$, 那么就能够利用 m_i 求得 s_i 的最佳估计值 $\hat{s}_i (\hat{s}_i = \hat{f}(m_i))$, 因此从腹壁混合信号 u_i 中消除母体心电成分的最佳估计值 \hat{s}_i 就得到仅含少量噪声的胎儿心电的最佳估计值 \hat{r}_i , 函数关系如式 (3) 所示:

$$[0093] \quad \hat{r}_i = u_i - \hat{s}_i = u_i - \hat{f}(m_i) \quad (3)$$

[0094] 因此对于胎儿心电信号提取, 首先使用采集的胎儿心电信号数据集的部分样本作为训练数据集 $\{(M_i, u_i), i = 1, 2, \dots, 1\}$, 求得输入信号 M_i 和目标输出信号 u_i 之间的非线性变换函数 $f(\cdot)$ 的最佳估计值 $\hat{f}(\cdot)$, 其中信号 M_i 由母体胸部心电信号 m_i 和它的 J 维时间导数构成。本发明选取GA优化的LSTM神经网络对非线性变换函数 $f(\cdot)$ 来进行拟合, 首先利用训练集数据构建GA-LSTM模型, 然后将测试集数据输入GA-LSTM模型, 最后进行胎儿心电信号的提取。

[0095] LSTM神经网络

[0096] LSTM神经网络解决了网络单元以链式方式链接的传统递归神经网络梯度消失和爆炸的问题, 可有效提高学习时间。在处理有关时间序列的预测和非线性映射问题中, 具备记忆能力的LSTM模型表现出较强的优势。LSTM中添加了一种叫做记忆单元 (Memory Cell) 的结构来记忆过去的信息, 并且增加了输入门、输出门和遗忘门三种门结构来控制历史信息的传递。

[0097] 设LSTM神经网络输入为 (x_1, x_2, \dots, x_T) , 隐层状态为 (h_1, h_2, \dots, h_T) , 网络在时刻 t , 各个单元和门控的计算如式 (4) - (9) 所示:

$$[0098] \quad i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$[0099] \quad f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

$$\tilde{c}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

[0100]

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (7)$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t \circ \tanh(c_t) \quad (9)$$

[0103] 式(4)-(9)中: i_t, f_t, o_t 分别为输入门、遗忘门和输出门的计算;其中,输入门主要用来决定保留多少当前时刻的输入信息到当前时刻的单元状态;遗忘门主要用来决定保留多少上一时刻的单元状态 c_{t-1} 的信息到当前时刻单元状态 c_t 中;输出门主要用来决定当前时刻的单元状态有多少输出。 h_t 为网络最终输出, \tilde{c}_t 为当前输入的单元状态; c_t 为当前时刻的单元状态; w_i, w_f, w_c, w_o 分别为三个门控和单元状态的权重矩阵; b_i, b_f, b_c, b_o 分别为各个门控和单元状态的偏置; $\sigma(\cdot)$ 和 $\tanh(\cdot)$ 为传输函数, \cdot 代表向量内积,符号 \circ 表示按元素相乘。

[0104] 3.GA优化的LSTM网络模型

[0105] GA是计算机模拟达尔文生物进化论的生物科学研究算法。在GA中关于种群的遗传演化的研究发现,以染色体作为种群遗传的主要载体,并且借助多种随机操作:基因选择、交叉和基因变异等,产生和不断演化出一种新的解集种群,据个体适应度和选择函数的取值可以选择最优的种群个体,即为GA中优化问题的最优解。

[0106] 本发明利用GA对LSTM网络的关键超参数进行寻优处理,采用GA强大的全局随机搜索能力,得到LSTM网络中神经元个数、学习率和训练次数的最优组合。基本思路如下:

[0107] 染色体编码

[0108] 将LSTM网络中的隐藏层神经元数目、学习率和训练次数作为GA的初始化对象,通过实数编码形式进行染色体编码。隐藏层神经元的区间范围为 $[5, 40]$,学习率的区间范围为 $[0.001, 0.1]$,训练次数的区间范围为 $[50, 500]$ 。

[0109] 适应度函数

[0110] 适应度函数是用来决定种群中哪些个体可以进行下一代遗传操作,根据个体适用度的不同,采用“优胜劣汰”进行个体筛选。本发明对适应度函数的定义如式(10)所示:

$$\text{fitness} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2} \quad (10)$$

[0112] 其中, N 表示心电图数据总量, \hat{y}_n 为母亲腹壁混合信号的网络预测值, y_n 为母亲腹壁混合信号的真实值。

[0113] 选择算子、交叉算子和变异算子

[0114] 选择算子是在当前种群中选择适应性较好的个体作为亲本,并将遗传信息传递给子代。在这里采用tournament selection算法作为GA的选择策略。该选择算法具有高效的算法执行率和易于实现的特点,算法复杂度远低于其他选择策略且易于并行化,在选择过程中不易陷入局部个体最优点,并且不需要对所有个体的适应度值排序。

[0115] 交叉算子使用洗牌交叉算法,在交叉之前在父代中利用random.shuffle函数进行洗牌运算,然后当在 $(0, 1)$ 之间产生的随机数小于所给的交叉率大小,则进行交叉变换。在

变异算子中,当在(0,1)之间产生的随机数小于所给的变异率大小,则进行变异操作。其中隐藏层神经元个数、学习率和训练次数变异的规则如式(11)所示:

$$\begin{aligned}
 [0116] \quad & c.Ln = \text{abs}(c.Ln + \text{random.randint}(-3,3)) \\
 [0117] \quad & c.lr = \text{abs}(c.lr + \text{random.uniform}(-0.001,0.001)) \\
 [0118] \quad & c.maxe = \text{abs}(c.maxe + \text{random.randint}(-30,30)) \quad (11)
 \end{aligned}$$

[0119] 式中: $c.Ln$ 为种群中隐藏层神经元个数; $c.lr$ 为种群中的学习率; $c.maxe$ 为种群中的训练次数。

[0120] 实施例2:

[0121] 本发明将GA与LSTM神经网络相融合,构建基于GA-LSTM的胎儿心电提取模型。首先采用GA对LSTM网络的超参数作寻优处理,得出学习率、隐层神经元数和训练次数的最佳组合,进一步提高模型的非线性映射能力;然后利用寻优的参数组合构建的GA-LSTM模型作为母体胸部心电信号与腹壁混合信号之间的非线性变换函数 $f(\cdot)$;在此基础上应用非线性变换函数 $f(\cdot)$ 求得母体胸部信号的最佳估计,最后从母体腹壁的混合信号中进行分离,即可分离提取得到最佳的胎儿心电信号的估计。模型具体操作流程如下:

[0122] Step 1:选择训练数据集。为了得到函数 $f(\cdot)$ 的最佳拟合效果,GA-LSTM模型的输入数据由母体心电 m_i 及其J维时间导数构成,本发明取 $J=2$ 。GA-LSTM模型的输入信号用向量 u 表示,目标输出信号用向量 M 表示,如式(12)所示:

$$\begin{aligned}
 [0123] \quad & u = [u_1 \ u_2 \ \dots \ u_L]^T \\
 & M = [m_1 \ m_2 \ \dots \ m_L]^T = \begin{bmatrix} m_1 & \dot{m}_1 & \ddot{m}_1^{(2)} \\ m_2 & \dot{m}_2 & \ddot{m}_2^{(2)} \\ \vdots & \vdots & \vdots \\ m_L & \dot{m}_L & \ddot{m}_L^{(2)} \end{bmatrix} \quad (12)
 \end{aligned}$$

[0124] Step 2:利用GA优化LSTM网络超参数。

[0125] (2.1)将LSTM网络模型中学习率、批处理大小、神经网络隐藏层单元数目作为优化对象,执行种群的初始化以及染色体编码与解码操作。

[0126] (2.2)计算初始种群中各个个体适应度大小;

[0127] (2.3)对染色体进行选择、交叉和变异操作;

[0128] (2.4)对染色体解码、计算种群内个体的适应度,在改算法中适应度越小,越应则保留该个体,否则淘汰该个体;

[0129] (2.5)若不符合遗传终止条件则回到(2.3)步;若符合遗传终止条件,则将GA求出的最优参数作为LSTM网络模型的最优参数;

[0130] Step3:训练GA-LSTM。将 u 和 M 输入最优参数组合的GA-LSTM网络,GA-LSTM网络模型的输出为腹壁混合信号中的母体心电成分 $s=f(M)$ 。将目标信号 u 与输出信号 s 之差用误差信号 e 表示,即 $e=u-s$ 。GA-LSTM根据均方误差最小化拟合误差 $E(eTe)$,最终得到母体胸部心电信号经历非线性变换 $f(\cdot)$ 传输到腹壁的最优拟合函数 $\hat{f}(\cdot)$ 。

[0131] Step4:提取胎儿心电信号。将心电数据 $\{(m_i, u_i), i=1, 2, \dots, N\}$ 送入GA-LSTM模型,利用最优拟合函数 $\hat{f}(\cdot)$ 计算得到腹壁混合信号 u_i 中的母体心电成分 $\hat{s}_i = \hat{f}(m_i)$ 。则胎儿心

电信号 \hat{r}_i 就可以利用下式计算： $\hat{r}_i = u_i - \hat{s}_i$ 。

[0132] 表1四种心电提取模型的统计分析

模型	TP	FP	FN	ACC /%	Se / %	PPV / %	F1 / %
NLMS	96	9	14	80.67	87.27	91.43	89.30
SVM	98	8	12	83.05	89.09	92.45	90.07
LSTM	103	5	7	88.79	93.64	94.50	94.06
GA-LSTM	105	3	5	92.92	95.45	97.22	96.33

[0134] 在本发明的描述中,除非另有说明,“多个”的含义是两个或两个以上;术语“上”、“下”、“左”、“右”、“内”、“外”、“前端”、“后端”、“头部”、“尾部”等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本发明和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。此外,术语“第一”、“第二”、“第三”等仅用于描述目的,而不能理解为指示或暗示相对重要性。

[0135] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,都应涵盖在本发明的保护范围之内。



图1

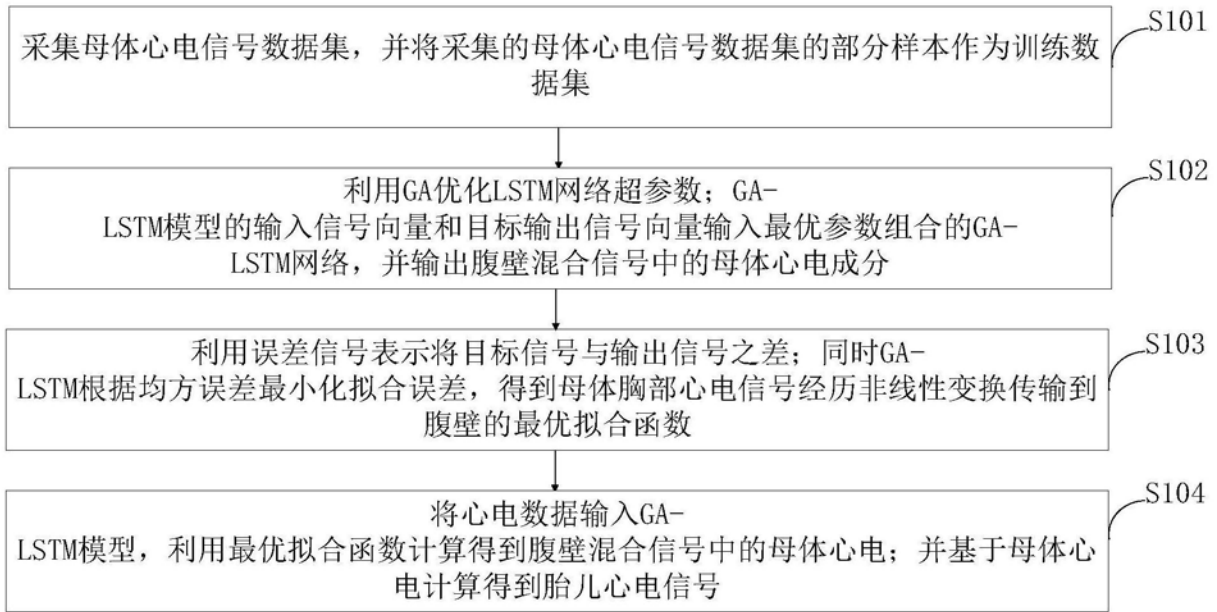


图2

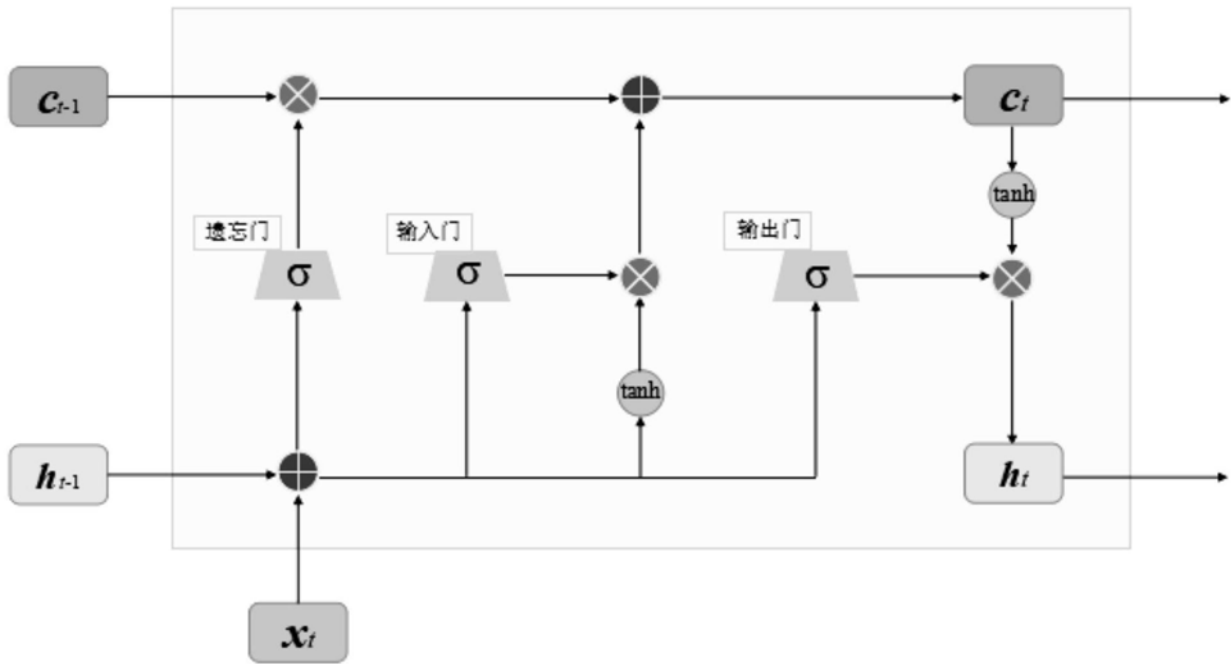


图3

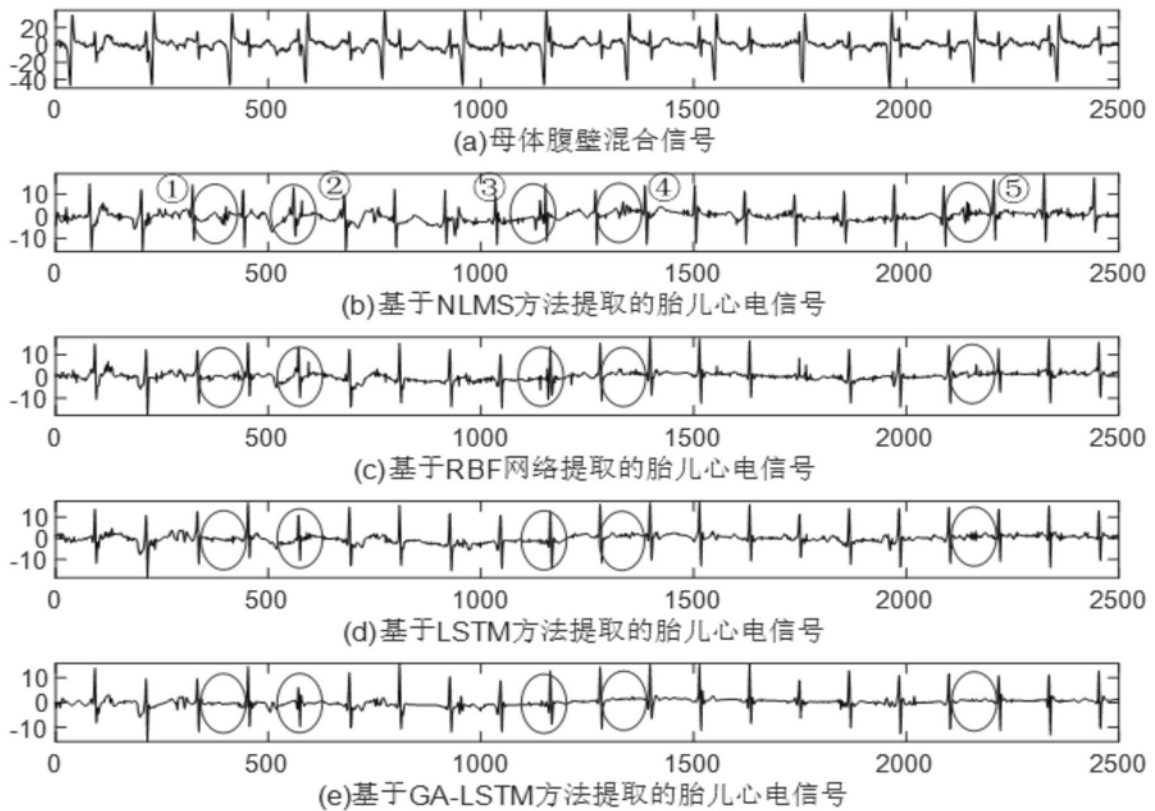


图4

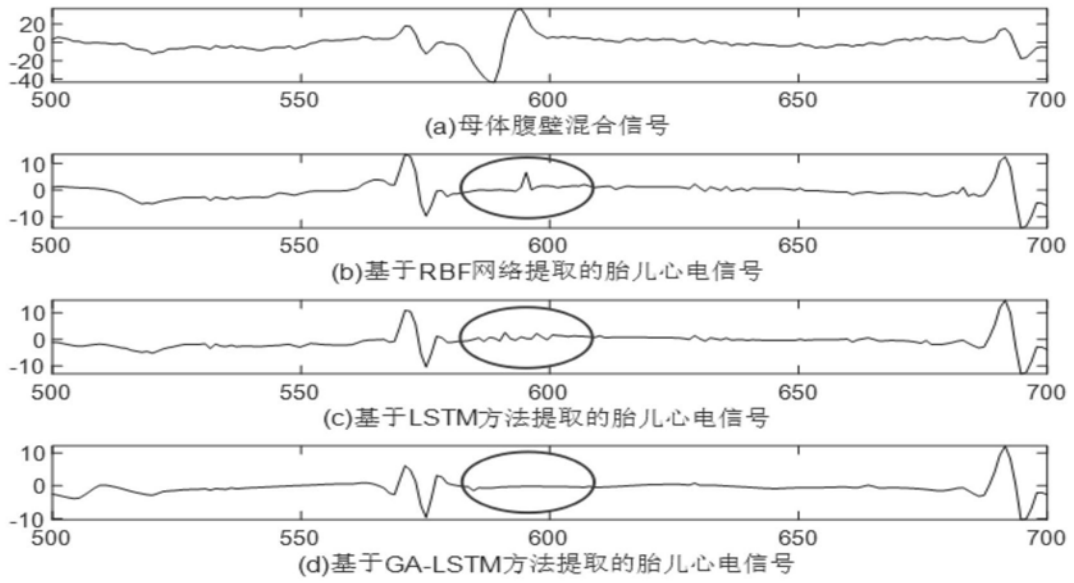


图5

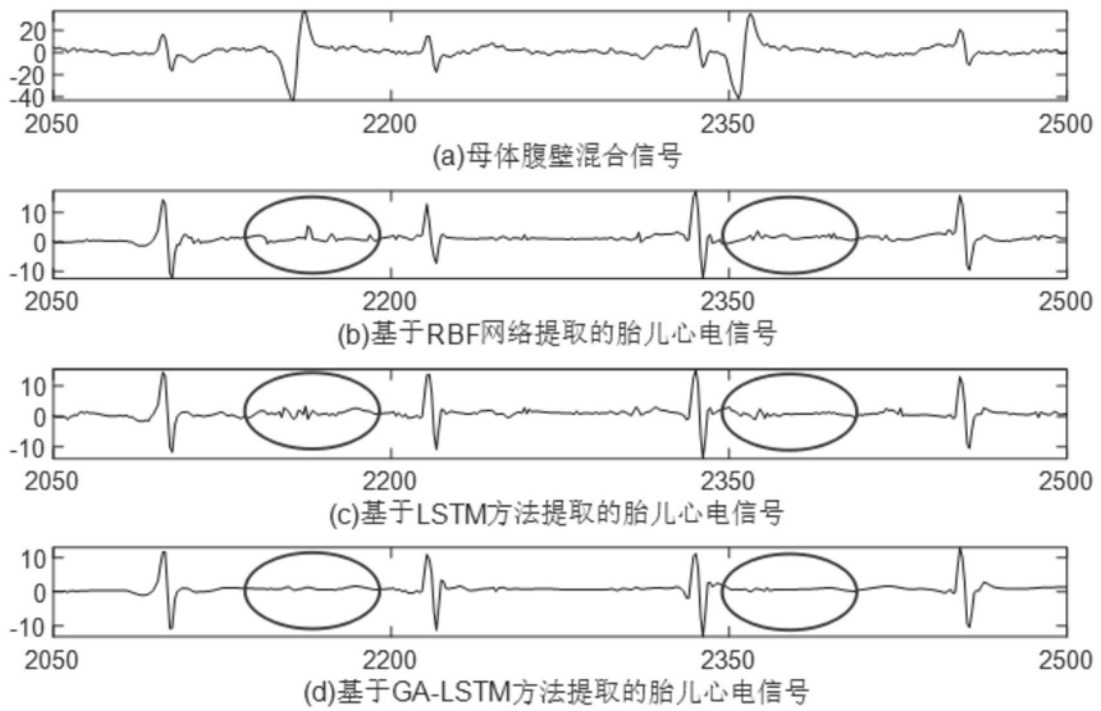


图6