



(12)发明专利申请

(10)申请公布号 CN 110801221 A

(43)申请公布日 2020.02.18

(21)申请号 201911263286.1

(22)申请日 2019.12.09

(71)申请人 中山大学

地址 510275 广东省广州市海珠区新港西路135号

(72)发明人 贺奥迪 刘官正

(74)专利代理机构 广州市智远创达专利代理有限公司 44619

代理人 王会龙

(51)Int.Cl.

A61B 5/0402(2006.01)

A61B 5/00(2006.01)

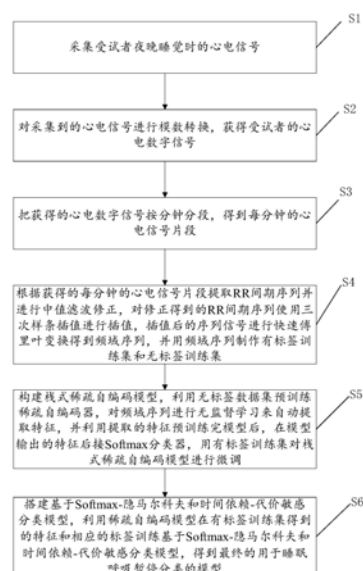
权利要求书2页 说明书7页 附图3页

(54)发明名称

基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备

(57)摘要

本发明公开了一种基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备,方法包括:采集睡觉时的心电信号;对心电信号进行模数转换,获得心电数字信号;把心电数字信号按分钟分段得到心电信号片段;根据心电信号片段提取RR间期序列进行修正,基于三次样条插值及快速傅里叶变换得到频域序列,制作训练集;构建栈式稀疏自编码模型,利用无标签数据集预训练稀疏自编码器,对频域序列进行无监督学习提取特征,用有标签训练集对栈式稀疏自编码模型进行微调;搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,利用稀疏自编码模型在有标签训练集得到的特征和相应的标签训练基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,得到睡眠呼吸暂停分类模型。



1. 一种基于无监督特征学习的睡眠呼吸暂停片段检测方法, 其特征在于, 包括以下步骤:

S1、采集受试者夜晚睡觉时的心电信号;

S2、对采集到的心电信号进行模数转换, 获得受试者的心电数字信号;

S3、把获得的心电数字信号按分钟分段, 得到每分钟的心电信号片段;

S4、根据获得的每分钟的心电信号片段提取RR间期序列并进行中值滤波修正, 对修正得到的RR间期序列使用三次样条插值进行插值, 插值后的序列信号进行快速傅里叶变换得到频域序列, 并用频域序列制作有标签训练集和无标签训练集;

S5、构建栈式稀疏自编码模型, 利用无标签数据集预训练稀疏自编码器, 对频域序列进行无监督学习来自动提取特征, 并利用提取的特征预训练完模型后, 在模型输出的特征后接Softmax分类器, 用有标签训练集对栈式稀疏自编码模型进行微调;

S6、搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型, 利用稀疏自编码模型在有标签训练集得到的特征和相应的标签训练基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型, 得到最终的用于睡眠呼吸暂停分类的模型。

2. 根据权利要求1所述的基于无监督特征学习的睡眠呼吸暂停片段检测方法, 其特征在于, 步骤S4具体包括:

S41、将所述心电数字信号记为 Q , 并对心电数字信号按分钟分段, 第 n 分钟的心电信号片段记作 Q_n , 即 $Q = \{Q_1, Q_2, \dots, Q_n\}$;

S42、对每分钟的心电信号片段进行滤波, 然后采用Pan-Tompki算法提取出R波位置, 相邻的后一个R波减去前一个R波得到RR间期序列;

S43、对RR间期序列进行中值滤波修正, 将超出 $[0.4, 2]$ 范围以外的数据点进行修正, 去除奇异点, 得到修正后的RR间期序列;

S44、对修正的RR间期序列使用三次样条插值进行插值, 统一每个RR间期序列的维数;

S46、对插值后的RR间期序列进行快速傅里叶变换得到频域序列;

S47、用频域序列制作有标签训练集和无标签训练集; 其中, 将数据库中没有标签的数据直接作为无标签数据集; 将经过利用多导睡眠监测仪的监测结果对每分钟心电信号片段进行标注的数据作为有标签训练集, 其中正常片段标注为0, 睡眠呼吸暂停片段标注为1。

3. 根据权利要求1所述的基于无监督特征学习的睡眠呼吸暂停片段检测方法, 其特征在于, 步骤S5具体包括:

S51、构建自编码器, 对于每个输入向量 $X = \{x_1, x_2, \dots, x_n\}$, 通过编码器对输入向量进行编码, 编码结果为 $H = f(W^1 X + b)$, 其中编码后得到的特征向量 $H = \{h_1, h_2, \dots, h_m\}$, 以及通过解码器对特征向量 H 进行解码, 解码过程为 $\hat{X} = f(W^2 H + b)$, 其中解码后得到重构向量 $\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n\}$, n, m 分别为输入层和隐藏层的维数, W^1 和 W^2 均为权重矩阵, 分别位于输入层和隐藏层之间以及隐藏层和输出层之间, b 为偏置项, 其代价函数 $J(W, b) = \left[\frac{1}{M} \sum_{i=1}^M L(X, \hat{X}) \right] + \frac{\lambda}{2} \sum_{l=1}^{N-1} (W^{(l)})^2$, 其中 $L(X, \hat{X}) = \frac{1}{2} \|\hat{X} - X\|^2$;

S52、在自编码器的基础上构建稀疏自编码器, 隐藏层中加入稀疏惩罚项使得网络能学习到更加抽象的特征, 其代价函数 $J_{\text{sparse}}(W, b) = J(W, b) + \beta P_{\text{penalty}}$, 其中,

$P_{\text{penalty}} = \sum_{j=1}^m KL(\rho || \rho_j)$, β 为稀疏惩罚项系数;

S53、对稀疏自编码器进行堆栈操作,通过把前一层稀疏自编码器的隐藏层输出向量作为后一层稀疏自编码器的输入向量,构建得到栈式稀疏自编码器模型;

S54、用无标签训练集对栈式稀疏自编码器模型应用逐层贪婪预训练的策略进行预训练,在此过程中神经网络以最小化重构误差作为优化目标,对每个稀疏自编码器依次进行学习,利用前向传播和反向传播算法进行多次迭代调整编码器和解码器的权值矩阵完成一个稀疏自编码器的训练,最终在编码器的输出端输出特征向量;再将前面一个稀疏自编码器的特征向量作为后面一个稀疏自编码器的输入,继续进行同样的学习过程直到所有堆栈起来的稀疏自编码器都学习完毕;

S55,完成所有稀疏自编码器权重矩阵的初始化并训练好栈式稀疏自编码器模型后,在栈式稀疏自编码器模型后面接入Softmax分类器,利用有标签训练集对栈式稀疏自编码器模型进行微调。

4. 根据权利要求1所述的基于无监督特征学习的睡眠呼吸暂停片段检测方法,其特征在于,步骤S6具体包括:

S61、搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型;

S62、把栈式稀疏自编码器模型在有标签训练集提取到的特征结果以及对应片段的标签输入到Softmax-隐马尔科夫和时间依赖-代价敏感模型,利用MetaCost通过对训练集样本进行多次取样训练Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,得到最终的睡眠呼吸暂停片段分类模型。

5. 根据权利要求4所述的基于无监督特征学习的睡眠呼吸暂停片段检测方法,其特征在于,步骤S62中,MetaCost实现代价敏感的方法如下:

在训练集 $\{(x_1, y_1), (x_2, y_2), \dots\}$ 中进行T次有放回的取样,每次取得n个样本;

记这些子训练集为 $S = \{S_1, S_2, \dots, S_T\}$, 选择一个分类模型, 分别对 $\{S_1, S_2, \dots, S_T\}$ 进行建模, 得到模型集 $M = \{M_1, M_2, \dots, M_T\}$;

使用 $M = \{M_1, M_2, \dots, M_T\}$ 分别对训练集 $\{(x_1, y_1), (x_2, y_2), \dots\}$ 的每个样本进行分类, 得到在对应模型 M_T 下, 对应样本 x 属于第 j 类的概率 $P(j|x; M_T)$, 设定代价矩阵 $C(i, j)$, 表示第 j 类样本被错分为第 i 类的代价, 定义条件风险作为代价函数 $R(i|x) = \sum_j P(j|x; M_T) C(i, j)$, 最小化代价函数 $R(i|x)$, 并在此过程中修改训练集中数据的标签 $y = \arg\min_i R(i|x)$, 使用修改标签后的训练集进行建模, 得到最终的睡眠呼吸暂停片段分类模型。

6. 一种基于无监督特征学习的睡眠呼吸暂停片段检测设备, 其特征在于, 包括存储器以及处理器, 所述存储器内存储有计算机程序, 所述计算机程序能够被所述处理器执行, 以实现如权利要求1至5任意一项所述的基于无监督特征学习的睡眠呼吸暂停片段检测方法。

基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备

技术领域

[0001] 本发明涉及医疗检测技术领域,尤其涉及一种基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备。

背景技术

[0002] 睡眠呼吸暂停(sleep apnea)是一种在睡眠期间呼吸气流较小或停止的常见疾病,主要是由于睡眠期间,喉咙后部的舌头或者其它软组织松弛,阻塞了呼吸道,或者由于呼吸中枢受抑制、中枢神经对呼吸反馈控制不稳定导致的。全世界睡眠呼吸暂停的发病率约2%到4%,目前全球约有9.36亿人患有睡眠呼吸暂停(OSA),我国约有6000万人患有该类疾病,睡眠呼吸暂停发生时,由于机体的缺氧会导致在睡眠期间大脑反复发生微觉醒,严重影响患者的睡眠质量,因此睡眠呼吸暂停患者常常白天嗜睡。此外,睡眠呼吸暂停发生时胸腔和肺动脉压力的波动会导致血压的变化,长期的呼吸暂停会导致高血压,该病患者的常见症状为白天困倦、性格暴躁、注意力不集中,另外,长期的睡眠呼吸暂停容易引发高血压、冠心病、脑血栓等一系列心脑血管疾病。因此,睡眠呼吸暂停的及时诊断具有很重要的意义。

[0003] 目前,睡眠呼吸暂停诊断的金标准是多导睡眠监测(PSG),多导睡眠监测仪连续记录心电图(ECG)、脑电图(EEG)、眼电图(EOG)、肌电图(EMG)、呼吸气流、血氧饱和度、胸腹式呼吸运动、鼾声等生理信号,然后,受过专业训练的技术人员参考美国睡眠医学会(AASM)制定的标准对多通道信号进行分析。然而,这种诊断方法具有一定的局限性,由于需要将大量的传感器粘贴到身上,很多测试者在睡觉时会感到不舒服从而影响睡眠质量,多导睡眠监测的测试费用昂贵,这种方法并不适用于广大人群的日常生活中的监测。因此,实现简单、快速、低负荷的睡眠呼吸暂停监测,及时筛查出睡眠呼吸暂停患者,具有重要的现实意义。

[0004] 目前,很多研究人员已经在通过多导睡眠监测中的单通道或少数通道诊断睡眠呼吸暂停做出了大量的探索。为了更加便捷快速的完成睡眠呼吸暂停的初步筛查,大量的研究者探索仅采用心电图、口鼻气流、鼾声中的一个信号或几个信号完成对睡眠呼吸暂停的初步筛查。1984年,Guilleminaultd等人观察了400个睡眠呼吸暂停综合症患者的连续24个小时的心电图和PSG图,发现当患者出现呼吸暂停症状时,心率减慢,当恢复正常时,心率加快,提出了心率变异性可以用来检测睡眠呼吸暂停综合症。

[0005] 在目前的利用单导数据进行睡眠呼吸暂停检测方法中,大部分方法都是利用从心电信号提取的RR间期信号进行分析,基于心电信号的方法已被证实可用于睡眠呼吸暂停的检测,基于深度学习的特征提取方法也被广泛研究。然而,目前所有的特征提取方法均是基于有监督学习建立的,表现出对数据标签的高度依赖,没有标签的数据无法用到网络训练中。由于在实际使用中,带标签的数据获取难度是非常大的,通常需要付出大量的时间或者金钱用于采集数据。

发明内容

[0006] 针对上述问题,本发明的目的在于提供一种基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备,成功将无监督学习方法引入睡眠呼吸暂停特征提取的任务中,解决了前人研究中对训练数据标签以及先验知识高度依赖的问题。

[0007] 本发明实施例提供了一种基于无监督特征学习的睡眠呼吸暂停片段检测方法,包括以下步骤:

[0008] S1、采集受试者夜晚睡觉时的心电信号;

[0009] S2、对采集到的心电信号进行模数转换,获得受试者的心电数字信号;

[0010] S3、把获得的心电数字信号按分钟分段,得到每分钟的心电信号片段;

[0011] S4、根据获得的每分钟的心电信号片段提取RR间期序列并进行中值滤波修正,对修正得到的RR间期序列使用三次样条插值进行插值,插值后的序列信号进行快速傅里叶变换得到频域序列,并用频域序列制作有标签训练集和无标签训练集;

[0012] S5、构建栈式稀疏自编码模型,利用无标签数据集预训练稀疏自编码器,对频域序列进行无监督学习来自动提取特征,并利用提取的特征预训练完模型后,在模型输出的特征后接Softmax分类器,用有标签训练集对栈式稀疏自编码模型进行微调;

[0013] S6、搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,利用稀疏自编码模型在有标签训练集得到的特征和相应的标签训练基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,得到最终的用于睡眠呼吸暂停分类的模型。

[0014] 优选地,步骤S4具体包括:

[0015] S41、将所述心电数字信号记为 Q ,并对心电数字信号按分钟分段,第 n 分钟的心电信号片段记作 Q_n ,即 $Q = \{Q_1, Q_2, \dots, Q_n\}$;

[0016] S42、对每分钟的心电信号片段进行滤波,然后采用Pan-Tompki算法提取出R波位置,相邻的后一个R波减去前一个R波得到RR间期序列;

[0017] S43、对RR间期序列进行中值滤波修正,将超出 $[0.4, 2]$ 范围以外的数据点进行修正,去除奇异点,得到修正后的RR间期序列;

[0018] S44、对修正的RR间期序列使用三次样条插值进行插值,统一每个RR间期序列的维数;

[0019] S46、对插值后的RR间期序列进行快速傅里叶变换得到频域序列;

[0020] S47、用频域序列制作有标签训练集和无标签训练集;其中,将数据库中没有标签的数据直接作为无标签数据集;将经过利用多导睡眠监测仪的监测结果对每分钟心电信号片段进行标注的数据作为有标签训练集,其中正常片段标注为0,睡眠呼吸暂停片段标注为1。

[0021] 优选地,步骤S5具体包括:

[0022] S51、构建自编码器,对于每个输入向量 $X = \{x_1, x_2, \dots, x_n\}$,通过编码器对输入向量进行编码,编码结果为 $H = f(W^1X + b)$,其中编码后得到的特征向量 $H = \{h_1, h_2, \dots, h_m\}$,以及通过解码器对特征向量 H 进行解码,解码过程为 $\hat{X} = f(W^2H + b)$,其中解码后得到重构向量 $\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n\}$, n, m 分别为输入层和隐藏层的维数, W^1 和 W^2 均为权重矩阵,分别位于输入层和隐藏层之间以及隐藏层和输出层之间, b 为偏置项,其代价函数

$J(W, b) = \left[\frac{1}{M} \sum_1^M L(X, \hat{X}) \right] + \frac{\lambda}{2} \sum_{l=1}^{N-1} (W^{(l)})^2$, 其中 $L(X, \hat{X}) = \frac{1}{2} \|\hat{X} - X\|^2$;

[0023] S52、在自编码器的基础上构建稀疏自编码器,隐藏层中加入稀疏惩罚项使得网络能学习到更加抽象的特征,其代价函数 $J_{\text{sparse}}(W, b) = J(W, b) + \beta P_{\text{penalty}}$, 其中, $P_{\text{penalty}} = \sum_{j=1}^m KL(\rho || \rho_j)$, β 为稀疏惩罚项系数;

[0024] S53、对稀疏自编码器进行堆栈操作,通过把前一层稀疏自编码器的隐藏层输出向量作为后一层稀疏自编码器的输入向量,构建得到栈式稀疏自编码器模型;

[0025] S54、用无标签训练集对栈式稀疏自编码器模型应用逐层贪婪预训练的策略进行预训练,在此过程中神经网络以最小化重构误差作为优化目标,对每个稀疏自编码器依次进行学习,利用前向传播和反向传播算法进行多次迭代调整编码器和解码器的权值矩阵完成一个稀疏自编码器的训练,最终在编码器的输出端输出特征向量;再将前面一个稀疏自编码器的特征向量作为后面一个稀疏自编码器的输入,继续进行同样的学习过程直到所有堆栈起来的稀疏自编码器都学习完毕;

[0026] S55、完成所有稀疏自编码器权重矩阵的初始化并训练好栈式稀疏自编码器模型后,在栈式稀疏自编码器模型后面接入Softmax分类器,利用有标签训练集对栈式稀疏自编码器模型进行微调。

[0027] 优选地,步骤S6具体包括:

[0028] S61、搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型;

[0029] S62、把栈式稀疏自编码器模型在有标签训练集提取到的特征结果以及对应片段的标签输入到Softmax-隐马尔科夫和时间依赖-代价敏感模型,利用MetaCost通过对训练集样本进行多次取样训练Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,得到最终的睡眠呼吸暂停片段分类模型。

[0030] 优选地,步骤S62中,MetaCost实现代价敏感的方法如下:

[0031] 在训练集 $\{(x_1, y_1), (x_2, y_2), \dots\}$ 中进行T次有放回的取样,每次取得n个样本;

[0032] 记这些子训练集为 $S = \{S_1, S_2, \dots, S_T\}$, 选择一个分类模型,分别对 $\{S_1, S_2, \dots, S_T\}$ 进行建模,得到模型集 $M = \{M_1, M_2, \dots, M_T\}$;

[0033] 使用 $M = \{M_1, M_2, \dots, M_T\}$ 分别对训练集 $\{(x_1, y_1), (x_2, y_2), \dots\}$ 的每个样本进行分类,得到在对应模型 M_T 下,对应样本 x 属于第 j 类的概率 $P(j|x; M_T)$, 设定代价矩阵 $C(i, j)$, 表示第 j 类样本被错分为第 i 类的代价,定义条件风险作为代价函数 $R(i|x) = \sum_i P(j|x; M_T) C(i, j)$, 最小化代价函数 $R(i|x)$, 并在此过程中修改训练集中数据的标签 $y = \arg \min_i R(i|x)$, 使用修改标签后的训练集进行建模,得到最终的睡眠呼吸暂停片段分类模型。

[0034] 本发明实施例还提供了一种基于无监督特征学习的睡眠呼吸暂停片段检测设备,包括存储器以及处理器,所述存储器内存储有计算机程序,所述计算机程序能够被所述处理器执行,以实现如上述的基于无监督特征学习的睡眠呼吸暂停片段检测方法。

[0035] 本发明与现有技术相比,具有如下有益效果:

[0036] 本发明在处理心电信号时仅仅通过心电信号提取的RR间期序列,经过快速傅里叶变换得到频域序列,然后通过一种频域栈式稀疏自编码器的无监督特征学习深度网络模型,成功将无监督学习方法引入睡眠呼吸暂停特征提取的任务中,可以利用大量的无标签数据预先训练自编码器,再利用少量的有标签数据对网络模型进行微调,使用无监督的方

法以降低特征学习过程中对带标签数据的依赖。另外基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型充分利用分类的心电信号片段间的时序性和疾病片段分类的代价敏感问题来提高分类模型的识别性能。

[0037] 同时,本发明在训练好模型之后不需要再额外采集呼吸信号、血氧信号以及脑电信号等额外信号,只需要采集心电信号即可。同时,由于本发明采用的方法能够利用大量的无标签数据,大大降低了训练成本,准确性也得到了进一步提升,因此在实现方面较为可行。

附图说明

[0038] 为了更清楚地说明本发明的技术方案,下面将对实施方式中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施方式,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。

[0039] 图1是本发明第一实施例提供的基于无监督特征学习的睡眠呼吸暂停片段检测方法的一种流程示意图。

[0040] 图2是本发明第一实施例提供的基于无监督特征学习的睡眠呼吸暂停片段检测方法的另一种流程示意图。

[0041] 图3是本发明实施例提供的自编码器的结构示意图。

[0042] 图4是本发明实施例提供的栈式稀疏自编码器的示意图。

[0043] 图5是对图4的栈式稀疏自编码器进行微调的示意图。

[0044] 图6是本发明实施例提供的代价敏感训练过程示意图。

具体实施方式

[0045] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0046] 请参阅图1及图2,本发明第一实施例提供了一种基于无监督特征学习的睡眠呼吸暂停片段检测方法,包括:

[0047] S1、采集受试者夜晚睡觉时的心电信号。

[0048] S2、对采集到的心电信号进行模数转换,获得受试者的心电数字信号。

[0049] S3、把获得的心电数字信号按分钟分段,得到每分钟的心电信号片段。

[0050] S4、根据获得的每分钟的心电信号片段提取RR间期序列并进行中值滤波修正,对修正得到的RR间期序列使用三次样条插值进行插值,插值后的序列信号进行快速傅里叶变换得到频域序列,并用频域序列制作有标签训练集和无标签训练集。

[0051] 其中,具体地,步骤S4包括:

[0052] S41、将所述心电数字信号记为 Q ,并对心电数字信号按分钟分段,第 n 分钟的心电信号片段记作 Q_n ,即 $Q = \{Q_1, Q_2, \dots, Q_n\}$;

[0053] S42、对每分钟的心电信号片段进行滤波,然后采用Pan-Tompki算法提取出R波位置,相邻的后一个R波减去前一个R波得到RR间期序列;

[0054] S43、对RR间期序列进行中值滤波修正,将超出 $[0.4, 2]$ 范围以外的数据点进行修正,去除奇异点,得到修正后的RR间期序列;

[0055] S44、对修正的RR间期序列使用三次样条插值进行插值,统一每个RR间期序列的维数;

[0056] S46、对插值后的RR间期序列进行快速傅里叶变换得到频域序列;

[0057] S47、用频域序列制作有标签训练集和无标签训练集;其中,将数据库中没有标签的数据直接作为无标签数据集;将经过利用多导睡眠监测仪的监测结果对每分钟心电信号片段进行标注的数据作为有标签训练集,其中正常片段标注为0,睡眠呼吸暂停片段标注为1。

[0058] 具体实施时,制作训练集和测试集所用的心电信号来自于多导睡眠监测仪(PSG)采集的心电信号,把得到的心电数字信号按分钟进行分段,然后从心电信号片段中提取RR间期序列并对异常信号进行修正,把修正的RR间期序列利用快速傅里叶变换(FFT)得到频域序列,然后把没有标签的频域数据直接作为无标签数据集,由医生参照PSG的其它通道信号对每分钟心电信号片段进行标注的频域信号作为有标签训练集,其中0表示正常,1表示呼吸暂停。

[0059] S5、构建栈式稀疏自编码模型,利用无标签数据集预训练稀疏自编码器,对频域序列进行无监督学习来自动提取特征,并利用提取的特征预训练完模型后,在模型输出的特征后接Softmax分类器,用有标签训练集对栈式稀疏自编码模型进行微调。

[0060] 其中,具体地,S5包括:

[0061] S51、构建如图3所示的自编码器,对于每个输入向量 $X = \{x_1, x_2, \dots, x_n\}$,通过编码器对输入向量进行编码,编码结果为 $H = f(W^1 X + b)$,其中编码后得到的特征向量 $H = \{h_1, h_2, \dots, h_m\}$,以及通过解码器对特征向量H进行解码,解码过程为 $\hat{X} = f(W^2 H + b)$,其中解码后得到重构向量 $\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n\}$,n,m分别为输入层和隐藏层的维数, W^1 和 W^2 均为权重矩阵,分别位于输入层和隐藏层之间以及隐藏层和输出层之间,b为偏置项,其代价函数 $J(W, b) = \left[\frac{1}{M} \sum_1^M L(X, \hat{X}) \right] + \frac{\lambda}{2} \sum_{l=1}^{N-1} (W^{(l)})^2$,其中 $L(X, \hat{X}) = \frac{1}{2} \|\hat{X} - X\|^2$;

[0062] S52、在自编码器的基础上构建如图4所示的稀疏自编码器,隐藏层中加入稀疏惩罚项使得网络能学习到更加抽象的特征,其代价函数 $J_{\text{sparse}}(W, b) = J(W, b) + \beta P_{\text{penalty}}$,其中,

$P_{\text{penalty}} = \sum_{j=1}^m KL(\rho || \rho_j)$, β 为稀疏惩罚项系数;

[0063] S53、对稀疏自编码器进行堆栈操作,通过把前一层稀疏自编码器的隐藏层输出向量作为后一层稀疏自编码器的输入向量,构建得到栈式稀疏自编码器模型;

[0064] S54、用无标签训练集对栈式稀疏自编码器模型应用逐层贪婪预训练的策略进行预训练,在此过程中神经网络以最小化重构误差作为优化目标,对每个稀疏自编码器依次进行学习,利用前向传播和反向传播算法进行多次迭代调整编码器和解码器的权值矩阵完成一个稀疏自编码器的训练,最终在编码器的输出端输出特征向量;再将前面一个稀疏自编码器的特征向量作为后面一个稀疏自编码器的输入,继续进行同样的学习过程直到所有堆栈起来的稀疏自编码器都学习完毕;

[0065] S55,完成所有稀疏自编码器权重矩阵的初始化并训练好栈式稀疏自编码器模型

后,在栈式稀疏自编码器模型后面接入Softmax分类器,利用有标签训练集对栈式稀疏自编码器模型进行微调(如图5所示)。

[0066] S6、搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,利用稀疏自编码器模型在有标签训练集得到的特征和相应的标签训练基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型(如图6所示),得到最终的用于睡眠呼吸暂停分类的模型。

[0067] 其中,步骤S6具体包括:

[0068] S61、搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型;

[0069] S62、把栈式稀疏自编码器模型在有标签训练集提取到的特征结果以及对应片段的标签输入到Softmax-隐马尔科夫和时间依赖-代价敏感模型,利用MetaCost通过对训练集样本进行多次取样训练Softmax-隐马尔科夫和时间依赖-代价敏感分类模型,得到最终的睡眠呼吸暂停片段分类模型。

[0070] 其中,MetaCost实现代价敏感的方法如下:

[0071] 在训练集 $\{(x_1, y_1), (x_2, y_2), \dots\}$ 中进行T次有放回的取样,每次取得n个样本;

[0072] 记这些子训练集为 $S = \{S_1, S_2, \dots, S_T\}$,选择一个分类模型,分别对 $\{S_1, S_2, \dots, S_T\}$ 进行建模,得到模型集 $M = \{M_1, M_2, \dots, M_T\}$;

[0073] 使用 $M = \{M_1, M_2, \dots, M_T\}$ 分别对训练集 $\{(x_1, y_1), (x_2, y_2), \dots\}$ 的每个样本进行分类,得到在对应模型 M_T 下,对应样本 x 属于第 j 类的概率 $P(j|x; M_T)$,设定代价矩阵 $C(i, j)$,表示第 j 类样本被错分为第 i 类的代价,定义条件风险作为代价函数 $R(i|x) = \sum_i P(j|x; M_T) C(i, j)$,最小化代价函数 $R(i|x)$,并在此过程中修改训练集中数据的标签 $y = \arg\min_i R(i|x)$,使用修改标签后的训练集进行建模,得到最终的睡眠呼吸暂停片段分类模型。

[0074] 具体实施时,本发明把从心电信号片段中提取的RR间期序列经过快速傅里叶变换得到频域序列,利用无监督栈式稀疏自编码器模型提取特征,然后用基于Softmax-隐马尔科夫和时间依赖-代价敏感模型分类。在网络训练完成后只需采集心电信号即可快速完成睡眠呼吸暂停片段的识别。本发明提供的方法简单易行,抗干扰效果好,操作简单,能够快速准确地测量受试者的每分钟睡眠情况,进而实现对睡眠呼吸暂停的早期检测。

[0075] 本发明第二实施例还提供了一种基于无监督特征学习的睡眠呼吸暂停片段检测设备,包括存储器以及处理器,所述存储器内存储有计算机程序,所述计算机程序能够被所述处理器执行,以实现如上述的基于无监督特征学习的睡眠呼吸暂停片段检测方法。

[0076] 示例性的,所述计算机程序可以被分割成一个或多个模块,所述一个或者多个模块被存储在所述存储器中,并由所述处理器执行,以完成本发明。所述一个或多个模块可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述所述计算机程序在所述基于无监督特征学习的睡眠呼吸暂停片段检测设备中的执行过程。

[0077] 所述基于无监督特征学习的睡眠呼吸暂停片段检测设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述基于无监督特征学习的睡眠呼吸暂停片段检测设备可包括,但不仅限于,处理器、存储器、显示器。本领域技术人员可以理解,上述部件仅仅是基于无监督特征学习的睡眠呼吸暂停片段检测设备的示例,并不构成对基于无监督特征学习的睡眠呼吸暂停片段检测设备的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件,例如所述基于无监督特征学习的睡眠呼吸暂停片段检测设备还可以包括输入输出设备、网络接入设备、总线等。

[0078] 所称处理器可以是中央处理单元 (Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor,DSP)、专用集成电路 (Application Specific Integrated Circuit,ASIC)、现成可编程门阵列 (Field-Programmable Gate Array,FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述基于无监督特征学习的睡眠呼吸暂停片段检测设备的控制中心,利用各种接口和线路连接整个所述基于无监督特征学习的睡眠呼吸暂停片段检测设备的各个部分。

[0079] 所述存储器可用于存储所述计算机程序和/或模块,所述处理器通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器内的数据,实现所述基于无监督特征学习的睡眠呼吸暂停片段检测的设备的功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序 (比如声音播放功能、文字转换功能等) 等;存储数据区可存储根据手机的使用所创建的数据 (比如音频数据、文字消息数据等) 等。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡 (Smart Media Card, SMC),安全数字 (Secure Digital,SD) 卡,闪存卡 (Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0080] 其中,所述基于无监督特征学习的睡眠呼吸暂停片段检测设备集成的模块如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读存储介质中。基于这样的理解,本发明实现上述实施例方法中的全部或部分流程,也可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一个计算机可读存储介质中,该计算机程序在被处理器执行时,可实现上述各个方法实施例的步骤。其中,所述计算机程序包括计算机程序代码,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器 (ROM,Read-Only Memory)、随机存取存储器 (RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是,所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减,例如在某些司法管辖区,根据立法和专利实践,计算机可读介质不包括电载波信号和电信信号。

[0081] 需说明的是,以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外,本发明提供的装置实施例附图中,模块之间的连接关系表示它们之间具有通信连接,具体可以实现为一条或多条通信总线或信号线。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0082] 以上所述是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也视为本发明的保护范围。

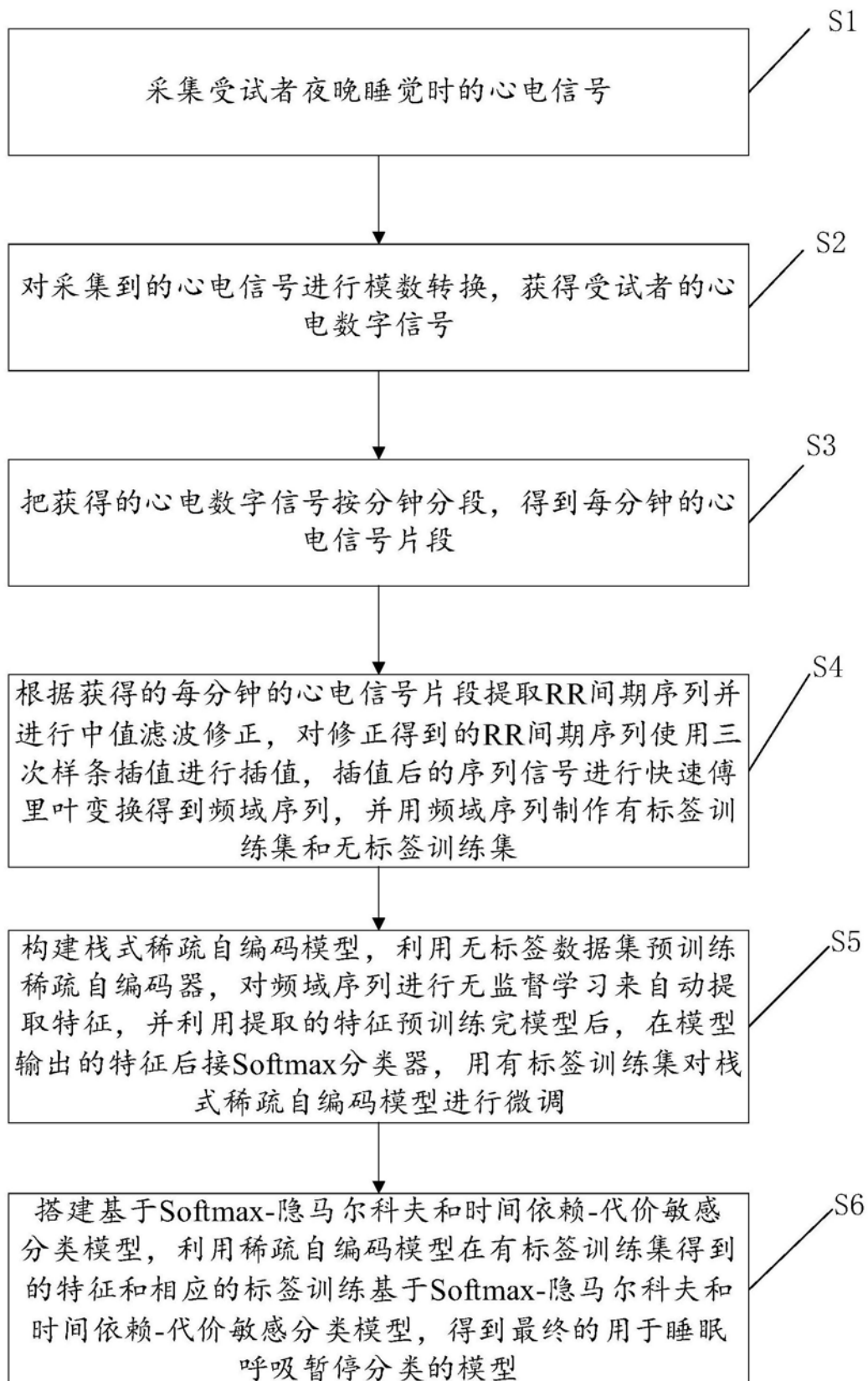


图1

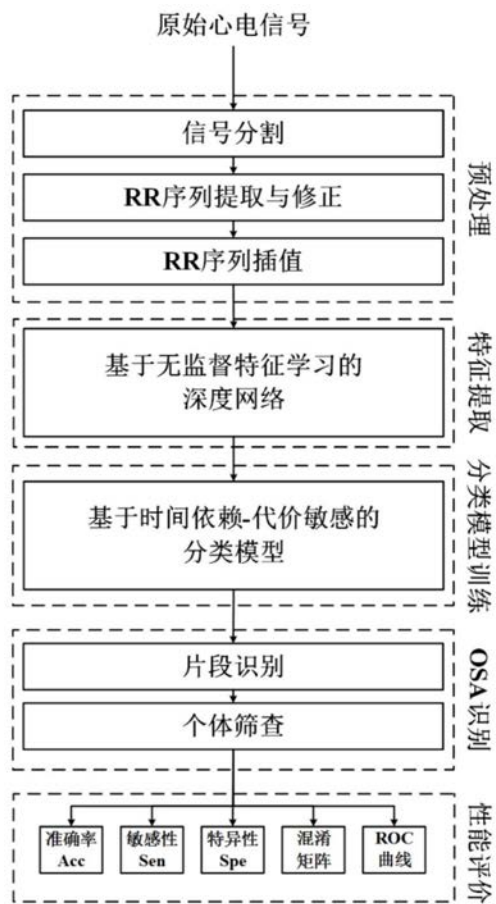


图2

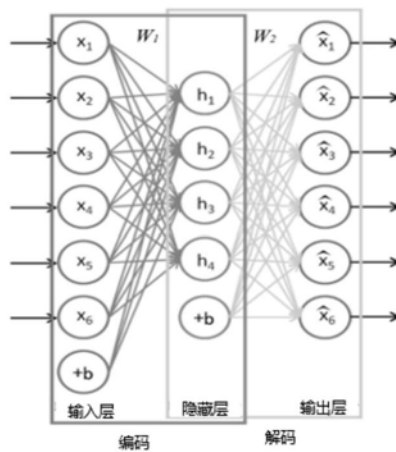


图3

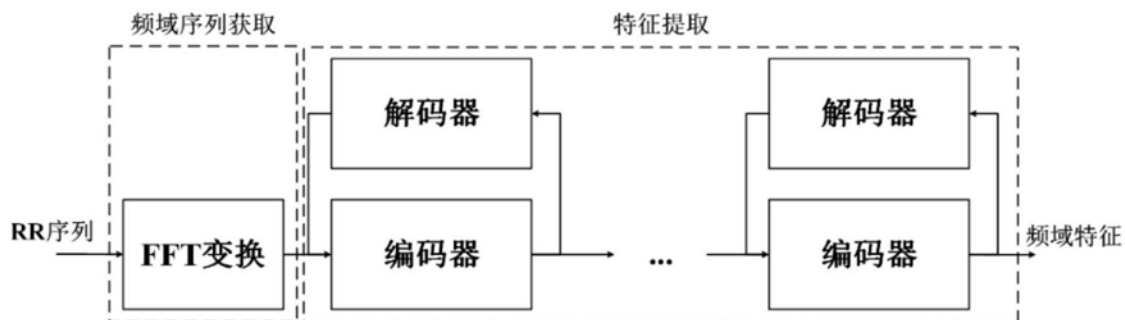


图4

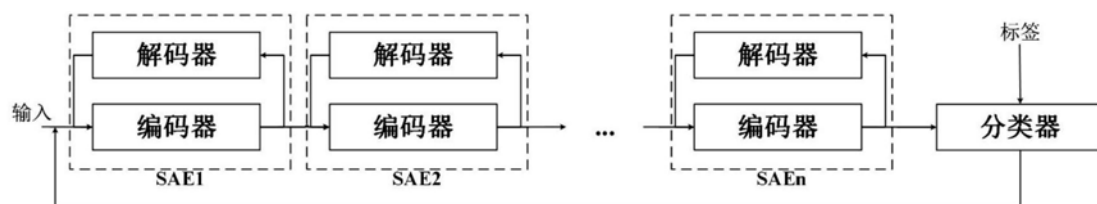


图5

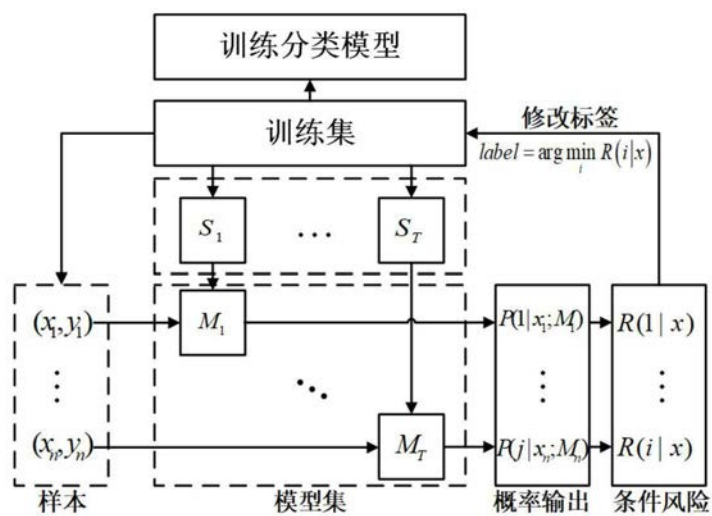


图6

专利名称(译)	基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备		
公开(公告)号	CN110801221A	公开(公告)日	2020-02-18
申请号	CN201911263286.1	申请日	2019-12-09
[标]申请(专利权)人(译)	中山大学		
申请(专利权)人(译)	中山大学		
当前申请(专利权)人(译)	中山大学		
[标]发明人	刘官正		
发明人	贺奥迪 刘官正		
IPC分类号	A61B5/0402 A61B5/00		
CPC分类号	A61B5/0402 A61B5/4812 A61B5/4815 A61B5/4818 A61B5/725 A61B5/7267		
代理人(译)	王会龙		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种基于无监督特征学习的睡眠呼吸暂停片段检测方法及设备，方法包括：采集睡觉时的心电信号；对心电信号进行模数转换，获得心电数字信号；把心电数字信号按分钟分段得到心电信号片段；根据心电信号片段提取RR间期序列进行修正，基于三次样条插值及快速傅里叶变换得到频域序列，制作训练集；构建栈式稀疏自编码模型，利用无标签数据集预训练稀疏自编码器，对频域序列进行无监督学习提取特征，用有标签训练集对栈式稀疏自编码模型进行微调；搭建基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型，利用稀疏自编码模型在有标签训练集得到的特征和相应的标签训练基于Softmax-隐马尔科夫和时间依赖-代价敏感分类模型，得到睡眠呼吸暂停分类模型。

