



(12)发明专利申请

(10)申请公布号 CN 111053549 A

(43)申请公布日 2020.04.24

(21)申请号 201911341051.X

(22)申请日 2019.12.23

(71)申请人 威海北洋电气集团股份有限公司

地址 264200 山东省威海市环翠区高技区
火炬路159号

(72)发明人 刘伟 曲媛媛 郑旭东 秦志亮
刘晓炜 谢耘

(74)专利代理机构 威海科星专利事务所 37202
代理人 初姣姣

(51)Int.Cl.

A61B 5/0402(2006.01)

A61B 5/00(2006.01)

权利要求书2页 说明书8页 附图1页

(54)发明名称

一种智能生物信号异常检测方法及系统

(57)摘要

本发明涉及一种能够对生物信号进行高效准确、低成本的自动分析的智能生物信号异常检测方法及系统,包括模型训练阶段和模型应用阶段,所述模型训练阶段包括:获取生物信号;对获取的信号进行预处理;将预处理后的数据送入CNN-BiLSTM-Attention组合模型进行特征的提取,其中首先令信号进入CNN卷积神经网络进行空间特性的提取,而后进入双向长短时记忆网路提取时间特性,再通过各个网络连接层,最后通过SoftMax层计算最终结果,在搭建网络的每个环节,通过设置Attention机制,进行信号不同序列点侧重程度的优化;计算网络模型的损失,直至模型损失小于阈值,则模型训练结束,否则利用Adam对模型参数进行优化;本发明有效提升了生物医学信号的分析效果。

1. 一种智能生物信号异常检测方法,包括模型训练阶段和模型应用阶段,其特征在于所述模型训练阶段包括以下步骤:

步骤1:获取生物信号;

步骤2:对步骤1中获取的信号进行预处理;

步骤3:将步骤2中预处理后的数据送入CNN-BiLSTM-Attention组合模型进行特征的提取,其中首先令信号进入CNN卷积神经网络进行空间特性的提取,而后进入双向长短时记忆网络BiLSTM提取时间特性,再通过各个网络连接层,最后通过SoftMax层计算最终结果,在搭建网络的每个环节,通过设置Attention机制,进行信号不同序列点侧重程度的优化;

步骤4:运用交叉熵损失函数计算步骤3中网络模型的损失,直至模型损失小于阈值,则模型训练结束,否则利用Adam对模型参数进行优化;

步骤5:将训练好的CNN-BiLSTM-Attention组合模型中连接层的中间结果输出,作为机器学习算法的输入端,进行机器学习模型的训练。

2. 根据权利要求1所述的一种智能生物信号异常检测方法,其特征在于所述模型应用阶段包括以下内容:

步骤6:通过步骤1-5所构建的模型对新的信号序列进行片段截取与分类,综合分类结果,将多个片段的分类结果进行统计,设置阈值,若片段中分类结果为异常超过阈值比例,则最终诊断为异常,否则为正常;

步骤7:按照时间顺序采集生物信号,通过循环步骤6,对信号分类结果的长期变化趋势进行有效的监控。

3. 根据权利要求1所述的一种智能生物信号异常检测方法,其特征在于所述步骤2中预处理具体包括:

步骤2-1:数据提取;对于采集的一维信号通道数据进行信号的截取与标注,按序截取预设长度信号片段,提取信号序列局部峰值点周围信号片段,对于信号的分类进行标注,将所有分段信号作为训练网络模型的输入端;

步骤2-2:数据平衡处理;当出现类间样本不平衡的情况,采用模拟生成同类中类似样本的处理方法,增大异常类别的数据集。

4. 根据权利要求2所述的一种智能生物信号异常检测方法,其特征在于步骤2-1中数据提取具体包括以下步骤:

步骤2-1-1:将连续的生物信号通过滑动窗分隔成连续的片段;

步骤2-1-2:对步骤2-1-1所获数据进行归一化处理;

步骤2-1-3:获取步骤2-1-2所获数据的局部最大值;

步骤2-1-4:选取预设长度信号;其中通过设定阈值,来确定数据片段中的R峰候选集,将每个选定的R峰作为生物信号的峰值点,选取峰值点周围预设长度的信号,对长度不一致的信号用0填充至预设长度,获取生物信号的原始样本集。

5. 根据权利要求2所述的一种智能生物信号异常检测方法,其特征在于步骤2-2所述数据平衡处理,具体包括以下内容:

利用同类样本中的欧氏距离,剔除极端的数据,在样本中查找一个能够很好地反映数据集性质的区域,具体指剔除极端值之后,按照样本间距离选取最相近比例样本确定该区域,在该区域进行随机数插值处理,进行样本的再造。

6. 根据权利要求1所述的一种智能生物信号异常检测方法,其特征在于所述步骤3中 CNN-BiLSTM-Attention组合模型采用如下结构:设计卷积池化过程及双向长短时记忆网络结构,最后接入连接层进行分类;首先设置进入模型训练的样本量,然后搭建卷积2-4层网络,每个卷积层后均跟着一个LeakyRelu激活层,相当于一种Attention,神经元激活值设置为0.2-0.5,滤波器个数为8-32,卷积核大小为2-10,从池化层出来以后,完成了第一次的卷积池化;再继续进入卷积层循环上述过程多次,滤波器大小依次增大,卷积和大小不变,池化层设置不变,循环卷积池化层后,设置Self Attention层,采用Dense-Tanh函数,把前面网络的多维特征连入双向长短时记忆网络中进行特征的二次锐化,双向长短时记忆网络输出所有中间结果序列,随后采用Tanh-Softmax函数,对数据进行序列点重要性的衡量,最后,把通过深度网络提取出的序列信息进行拉伸,连接进入全连接层,然后进入两层连接密集层,在这期间,为防止网络出现过拟合的问题,设置Dropout机制,嵌入隐含层中,Dropout随机失活限制值设置为0.3-0.6,若神经元计算出来的值小于0.3-0.6,则此神经元将不再连入下一层密集层中,通过这种方式,对连接密集层各层的神经元进行筛选,最后进入Softmax层进行心电信号的分类。

7. 一种利用如权利要求1所述智能生物信号异常检测方法的系统,设有信号获取单元,数据预处理单元,用于完成模型训练阶段的集成分类器搭建单元和用于完成模型应用阶段的综合判别单元,其特征在于所述集成分类器搭建单元设有特征提取单元和第二层分类训练单元,所述特征提取单元用于对预处理后数据进行特征提取的特征提取单元,并设有采用CNN卷积神经网络进行空间特性的提取的空间特性提取模块、采用双向长短时记忆网络(BiLSTM)提取时间特性的时间特性提取模块、用于输出中间结果的网络模型连接层模块、计算最终结果的SoftMax层模块以及Attention机制模块;

所述第二层分类训练单元利用网络模型连接层模块输出的中间结果进行训练。

8. 根据权利要求7所述的一种利用智能生物信号异常检测方法的系统,其特征在于所述综合判别单元用于将经过预处理的新的生物信号送入集成分类器搭建模块内训练完毕的集成分类器进行分类,并综合连续多个节拍信号的分类结果,确定最终分类结果。

9. 根据权利要求7所述的一种利用智能生物信号异常检测方法的系统,其特征在于所述数据预处理单元设有数据截取模块和数据平衡处理模块,其中数据截取模块依次设有数据分割模块、归一化处理模块、局部最大值获取模块以及预设长度信号获取模块;所述数据平衡处理单元依次设有极端数据剔除模块、随机数插值处理模块。

一种智能生物信号异常检测方法及系统

技术领域：

[0001] 本发明涉及生物信号检测技术领域，具体的说是一种能够对生物信号进行高效准确、低成本的自动分析的智能生物信号异常检测方法及系统。

背景技术：

[0002] 凡是存在于自然界的生命体，无论是动物还是人类，都有属于自己的生命体征，随着经济社会的发展，这些生命体征通过现有的仪器能够被转化为信号序列的形式所采集，这些生物信号是生命体生理功能的部分映射，对生命体的生理状态以及生理功能起到有效的检测。对于人类而言，近些年来，人们的生活节奏加快、工作压力增大，心脑血管疾病的发病率和死亡率越来越高，人们的身体正在受到严重威胁，这些生物信号能够对人的身体起到一个有效的反应，因此，对于生物信号的监测与分析是很有意义的；对于除人类以外的其他生命体而言，生物信号也同样具有预警作用。

[0003] 生命体的生理信号包括，压力信号（动/静脉压、膀胱压、子宫内压等等）、生物电信号（心电、肌电、眼电、脑电等等）、低幅神经信号（SNA）、组织氧含量和体温等等。它能够有效的对生理功能进行记录，由于大部分生物信号都需要专业人员进行分析和诊断，因此对生物信号进行详细分析与准确判别，能够在没有其他辅助的情况之下，及时监控生命体的身体状态，对生命体的身体状态起到预警作用，有效预防突发性疾病、降低突发性疾病的死亡率。

[0004] 人工智能技术利用现有的仪器采集生物序列信号，对其进行自动分析和诊断。有效的机器学习方法，如支持向量机、多层感知机、决策树、随机森林等等，都能够对信号进行分类判别，这些方法在建模前，都要进行信号的预处理，从信号中提取人工特征（具体有小波包分析、功率谱估计、能量熵等等），用于实现最终的分类任务。

[0005] 这些人工提取的信号特征能够展现信号某方面的形态，但却很难利用信号序列的全部特征信息，对于特定的分类任务，人工提取的特征可能不能够满足机器智能学习此分类任务的需求，且对于某些变化较小、不易察觉的异常搏动而言，智能算法可能会失去它的作用。在处理信号分类的问题上，自动的特征提取和代表性方法被证实更加具有可扩展性，且能够做到更加精准的预测，从原始信号序列出发，充分利用信号的全部信息，构建端到端的深度学习框架，它能够允许机器学习最适合其所需特定分类任务的特征。

[0006] 针对生物信号分类任务，近些年来出现了许多的深度学习框架。最经典的方法是搭建卷积神经网络模型，在此基础上，也有自适应的卷积神经网络、导联卷积神经网络等，这些方法证明了使用卷积神经网络能够对生物信号进行高水平的分类。除此之外，还有其他一些深度学习框架也能够进行可靠的信号序列分类，如基于深双向长短时记忆网络模型用于生物信号的分类等。

[0007] 对于一些生命体的生理信号而言，长时间的连续检测显得很重要，对于慢性症状而言，生理状况变化趋势的检测是非常有必要的，因此，对于生物信号的分类不仅仅是单一片段的分类检测，对于连续信号的长时间内的监测分析也是十分重要、十分有发展前景的。

在没有辅助手段、专业人员的情况下,为了方便人们及时对自己以及其他生命体的身体状态进行监控,生物信号的智能化分析显得越来越重要,传统的生物信号诊断是由专业诊断人员完成的,但如果需要长时间检测生命体征,24小时获得的信息全部由人来诊断显得费时费力,且由于专业人员疲劳等因素,诊断的准确率也会有所下滑。因此,对生物信号进行高效准确、低成本的自动分析就显得尤为重要。

[0008] 现阶段此类数据处理方法主要存在以下不足:一、对于生物序列信号的预处理主要存在以下缺点和不足:

[0009] 1、针对不同方向的生物序列信号,有不同的处理方法,提取方式较为细致,不能够具备很好的广泛应用能力,当信号形式改变后可能会不再适用。

[0010] 2、对于信号进行分类研究时,样本不平衡影响着分类算法的效果,而在预处理过程中,样本不平衡的问题却很少被提及、很少采取处理方式避免。

[0011] 3、对于信号的处理分析,通常采取片段式处理分析的方式,但对于生物信号而言,有时长时间的连续监测、对变化趋势进行有效反应会显得更重要,现行方法下针对信号的连续监测与分析做的工作很少,在采集仪器上已经实现的信号连续监测,所得到的信息没有被充分的利用起来。

[0012] 二、对于生物时间序列信号进行识别的算法主要存在以下几条缺点和不足:

[0013] 1、人工提取的信号特征能够展现使用传统机器学习方法,如支持向量机、多层感知机、决策树、随机森林等等,进行信号的识别,这些算法虽然拥有很强的数学可解释性,但对于信号特征的要求十分高,需要有经验的人进行特征的有效提取,才能够使识别准确率达到一个基本的要求,这种方法费时费力,且不能够具备很好的泛化能力。

[0014] 2、信号某方面的形态特征,很难反映信号序列的全部特征信息,对于特定的分类任务,人工提取的特征可能不能够满足机器智能学习此分类任务的需求,例如,对于某些变化较小、不易察觉的心脏异常搏动而言,智能算法可能会失去它的作用。

[0015] 3、针对生物信号分类任务,为弥补传统机器学习方法的不足,近些年来出现了许多的深度学习框架,如CNN、RNN、LSTM、SAE等等,它能够对信号序列进行特征自动提取,但是这些网络结构的深度学习框架算法在识别的准确度上还有待提高,且对于信号分类思想的可解释性较差。

发明内容:

[0016] 本发明针对现有技术中存在的缺点和不足,提出了一种适用于长时间监测,能够有效提升生物信号的分类准确度以及提升生物信号智能化分析的可靠性,且能够做到对生理状况变化趋势的进行检测的智能生物信号异常检测方法及系统。

[0017] 本发明通过以下措施达到:

[0018] 一种智能生物信号异常检测方法,包括模型训练阶段和模型应用阶段,其特征在于所述模型训练阶段包括以下步骤:

[0019] 步骤1:获取生物信号;

[0020] 步骤2:对步骤1中获取的信号进行预处理;

[0021] 步骤3:将步骤2中预处理后的数据送入CNN-BiLSTM-Attention组合模型进行特征的提取,其中首先令信号进入CNN卷积神经网络进行空间特性的提取,而后进入双向长短时

记忆网路 (BiLSTM) 提取时间特性,再通过各个网络连接层,最后通过SoftMax层计算最终结果,在搭建网络的每个环节,通过设置Attention机制,进行信号不同序列点侧重程度的优化;

[0022] 步骤4:运用交叉熵损失函数计算步骤3中网络模型的损失,直至模型损失小于阈值,则模型训练结束,否则利用Adam对模型参数进行优化;

[0023] 步骤5:将训练好的CNN-BiLSTM-Attention组合模型中连接层的中间结果输出,作为机器学习(如随机森林算法)算法的输入端,进行机器学习模型的训练。

[0024] 本发明所述模型应用阶段包括以下内容:

[0025] 步骤6:通过步骤1-5所构建的模型对新的信号序列进行片段截取与分类,综合分类结果,将多个片段的分类结果进行统计,设置阈值,若片段中分类结果为异常超过阈值比例,则最终诊断为异常,否则为正常;

[0026] 步骤7:按照时间顺序采集生物信号,通过循环步骤6,对信号分类结果的长期变化趋势进行有效的监控。

[0027] 本发明所述步骤2中预处理具体包括:

[0028] 步骤2-1:数据提取;对于采集的一维信号通道数据进行信号的截取与标注,按序截取预设长度信号片段,提取信号序列局部峰值点周围信号片段,对于信号的分类进行标注,将所有分段信号作为训练网络模型的输入端;

[0029] 步骤2-2:数据平衡处理;当出现类间样本不平衡的情况,采用模拟生成同类中类似样本的处理方法,增大异常类别的数据集。

[0030] 本发明步骤2-1中数据提取具体包括以下步骤:

[0031] 步骤2-1-1:将连续的生物信号通过滑动窗(可以为10s)分隔成连续的片段;

[0032] 步骤2-1-2:对步骤2-1-1所获数据进行归一化处理;

[0033] 步骤2-1-3:获取步骤2-1-2所获数据的局部最大值;

[0034] 步骤2-1-4:选取预设长度信号;其中通过设定阈值,来确定数据片段中的R峰候选集,将每个选定的R峰作为生物信号的峰值点,选取峰值点周围预设长度的信号,对长度不一致的信号用0填充至预设长度,获取生物信号的原始样本集。

[0035] 本发明步骤2-2所述数据平衡处理,具体包括以下内容:

[0036] 利用同类样本中的欧氏距离,剔除极端的数据,在样本中查找一个区域(具体指剔除极端值之后,按照样本间距离选取最相近比例样本(比例可以是70%,也可以是71%等等),作为一个区域),使其能够很好地反映数据集的性质,在该区域进行随机数插值处理,进行样本的再造,以避免因样本重复造成的模型过拟合问题,也决绝了再造样本边缘化问题;

[0037] 本发明所述步骤3中CNN-BiLSTM-Attention组合模型采用如下结构:设计卷积池化过程及双向长短时记忆网络结构,最后接入连接层进行分类;首先设置进入模型训练的样本量,然后搭建卷积2-4层网络,每个卷积层后均跟着一个LeakyRelu激活层,相当于一种Attention,神经元激活值设置为0.2-0.5,滤波器个数为8-32,卷积核大小为2-10,从池化层出来以后,完成了第一次的卷积池化;再继续进入卷积层循环上述过程多次,滤波器大小依次增大,卷积和大小不变,池化层设置不变,循环卷积池化层后,设置Self Attention层,采用Dense-Tanh函数,把前面网络的多维特征连入双向长短时记忆网络中进行特征的二次

锐化,双向长短时记忆网络输出所有中间结果序列,随后采用Tanh-Softmax函数,对数据进行序列点重要性的衡量,最后,把通过深度网络提取出的序列信息进行拉伸,连接进入全连接层,然后进入两层连接密集层,在这期间,为防止网络出现过拟合的问题,设置Dropout机制,嵌入隐含层中,Dropout随机失活限制值设置为0.3-0.6,若神经元计算出来的值小于0.3-0.6,则此神经元将不再连入下一层密集层中,通过这种方式,对连接密集层各层的神经元进行筛选,最后进入Softmax层进行心电信号的分类。

[0038] 本发明还提出了一种智能生物信号异常检测系统,设有信号获取单元,数据预处理单元,用于完成模型训练阶段的集成分类器搭建单元和用于完成模型应用阶段的综合判别单元,其特征在于所述集成分类器搭建单元设有特征提取单元和第二层分类训练单元,所述特征提取单元用于对预处理后数据进行特征提取的特征提取单元,并设有采用CNN卷积神经网络进行空间特性的提取的空间特性提取模块、采用双向长短时记忆网络(BiLSTM)提取时间特性的时间特性提取模块、用于输出中间结果的网络模型连接层模块、计算最终结果的SoftMax层模块以及Attention机制模块;

[0039] 所述第二层分类训练单元利用网络模型连接层模块输出的中间结果进行训练。

[0040] 本发明所述综合判别单元用于将经过预处理的新的生物信号送入集成分类器搭建模块内训练完毕的集成分类器进行分类,并综合连续多个信号片段的分类结果,确定最终分类结果。

[0041] 本发明所述数据预处理单元设有数据截取模块和数据平衡处理模块,其中数据截取模块依次设有数据分割模块、归一化处理模块、局部最大值获取模块以及预设长度信号获取模块;所述数据平衡处理单元依次设有极端数据剔除模块、随机数插值处理模块。

[0042] 本发明提供了一种特别适用于长时间监测的混合深度学习的生物信号分类算法,能够广泛应用于生物信号的分类任务;本发明与现有技术主要有以下不同:(1)数据分类结果处理方法的不同:对信号序列片段的分类结果进行综合,使得在一段时间内的分类结果较为稳定,有效减少错报。对于信号序列所预示的生理状态起到长时间的监控。训练集样本采用人工创造相似样本的方式扩大训练集,提升实际预测效果。(2)数据模型结构的不同:首先,对于生物信号分类而言,提出新的网络结构,建立CNN-BiLSTM-Attention深度学习网络模型;并在网络模型中应用Attention机制,进行信号不同序列点侧重程度的优化;其次充分结合深度学习方法与机器学习方法各自的优势,提出深度学习与随机森林相结合的分类方法,提出通过深度学习模型充分提取序列的信息特征,再通过随机森林进行分类的方法,并成功应用于实际案例。本发明与其他方法相比,能够有效提升预测准确度,使分类结果更加真实可靠。

附图说明:

[0043] 附图1是本发明的流程图。

[0044] 附图2是本发明的系统结构框图。

具体实施方式:

[0045] 下面结合附图和实施例,对本发明作进一步的说明。

[0046] 本发明提出了一种智能生物信号异常检测方法及系统,其中具体步骤如下:

[0047] 步骤1:前期信号的监测和采集,通过采集终端实现对生物信号的采集;

[0048] 步骤2:对于采集的一维信号通道数据,信号的截取与标注,按序截取预设长度信号片段,提取信号序列局部峰值点周围信号片段,对于信号的分类进行标注,将所有分段信号作为训练网络模型的输入端;

[0049] 步骤3:将经过上两个模块中得到的数据集作为训练测试对象,搭建深度学习网络模型,进行特征的自提取。考虑信号时空特性、建立CNN-BiLSTM-Attention深度学习网络模型,对信号的空间特性和时间特性进行特征的自提取,首先,令信号进入CNN卷积神经网络进行空间特性的提取,而后进入双向长短时记忆网路(BiLSTM)提取时间特性,再通过各个网络连接层,最后通过SoftMax层计算最终结果。在搭建网络的每个环节,通过设置Attention机制,进行信号不同序列点侧重程度的优化;

[0050] 步骤4:运用交叉熵损失函数计算步骤1中网络模型的损失,直至模型损失小于阈值0.1,则模型训练结束,否则,利用Adam对模型参数进行优化;

[0051] 步骤5:输出得到的网络模型连接层的中间结果,作为信号特征进行第二层分类训练;

[0052] 步骤6:步骤5中得到的新的信号特征组,作为随机森林算法(系列机器学习算法如决策树、支持向量机等等)的输入端,进行机器学习模型的训练;

[0053] 步骤7:通过分类模块实现对新的信号序列进行片段截取与分类,综合分类结果,将多个片段的分类结果进行统计,设置阈值,若片段中分类结果为异常超过一定比例,则最终诊断为异常,否则为正常。

[0054] 步骤8:按照时间顺序(具体到秒)进行生物信号的采集工作,通过循环步骤7,对信号分类结果的长期变化趋势进行有效的监控。

[0055] 实施例1:

[0056] 一种利用上述智能生物信号异常检测方法及系统实现的心电图信号识别算法,该方法包括以下内容:

[0057] 步骤1:前端心电图系统信号收集,通过导联式心电图仪对人的心电信号进行监测和采集;

[0058] 步骤2:心电信号的节拍提取,步骤1中所获取的长时间内的电一维信号序列,为非周期性高度重复的序列,因此我们进行信号每个心脏活动周期的节拍信号的提取;

[0059] 步骤3:心电信号的整理与标注,步骤2中所获取的节拍信号,利用专家经验,请心电图领域专家对采集的心电信号进行标注,每个节拍至少由两名心脏病专家进行批注;

[0060] 其中:步骤1所述的信号的采集,心电图信号需要通过心电图仪进行采集或网上获取数据集,作为具体应用例,我们使用PhysioNet MIT-BIH心律失常数据库中的心律失常信号作为标记的心电数据进行分析;

[0061] 步骤2所述的对心电图信号进行节拍提取。不同受试者的心电记录均是一种非周期性高度重复的序列,为此采用一种简单、有效的心电信号预处理手段,进行心电信号的节拍提取,具体为:利用系统中的数据截取模块,将连续的心电信号通过10秒窗分割成连续的片段,通过最大值最小值归一化处理,利用原始信号点减去最小值点,除以最大值与最小值的差值的方式,将振幅控制在0到1的范围内,然后,通过计算片段中大于整个片段90%的点的位置,寻找局部最大值,选取出局部最大值后,通过设置阈值的方式来确定片段中的R峰

候选集,对于每一个选择的R峰,作为一个心电节拍信号的峰值点,选取它周围预设长度的信号,对于长度不一致的信号,用0进行填充,填充到预设长度,得到心电信号的原始样本集。

[0062] 心电信号经过数据截取模块,能够明显提升最后的分类准确度与判别结果,若十多秒的心电信号不经过心拍提取,则专家进行标注过程中就可能会出现个别片段标注混淆,样本混乱,不利于实现后期分类器的训练任务(换句话说,分类器的分类准确度会下降至70%以下),通过本发明中此模块的运算,能够有效的、完整的提取心拍数据,为后期提升分类器的分类效果打下了基础,其他专利中大部分也均有心拍数据提取模块,但使用的具体提取方法与本专利不同。

[0063] 过采样技术进行数据集增强,心拍信号经数据截取模块中所获取的带种类标注的节拍信号,为防止样本不平衡而对智能算法分类的影响,采用过采样方法,进行数据集的增强;

[0064] 其中:对于专家标注好的心电信号样本,进行过采样的处理,以提升分类准确度。生物信号获取成本较高,而所采集的心电图异常信号相比于正常信号而言,样本严重偏少,不同类间样本比例严重失衡,因此,采用人为扩充数据集样本的方法,利用系统中的数据平衡模块,通过人工创造新样本的方式增加样本量,使类与类之间的样本比例达到平衡。本发明人工创造样本的方法是,利用同类样本中的欧氏距离,剔除较为极端的数据,在样本中查找到一个区域,使其能够很好的反应数据集的性质,在该区域进行随机数插值处理,进行样本的再造,避免了因样本重复所造成的模型过拟合等问题,也解决了再造样本边缘化的问题。

[0065] 数据平衡模块在此过程中能够起到以下作用:心电信号经过心拍数据的提取模块后,因异常信号样本数偏少,在模型训练过程中可能会造成阻碍,若不采取一定的样本平衡措施,则测试集样本分类预测准确度仅能达到94%,通过本发明中此模块的处理,能够有效的提升样本分类准确度,使准确度上升至97.5%,以及有效提升后期样本分类结果综合判别的效果,使心电信号识别更加的可靠。

[0066] 系统中集成分类器搭建单元执行以下内容:

[0067] 步骤1:将经过信号提取模块与平衡模块中得到的序列信号作为训练测试对象,搭建深度学习网络模型,进行特征的自提取。考虑信号时空特性、建立CNN-BiLSTM-Attention深度学习网络模型,对信号的空间特性和时间特性进行特征的自提取,首先,令信号进入CNN卷积神经网络进行空间特性的提取,而后进入双向长短时记忆网路(BiLSTM)提取时间特性,再通过各个网络连接层,最后通过SoftMax层计算最终结果。在此期间通过设置Attention机制,进行信号不同序列点侧重程度的优化;

[0068] 步骤2:运用交叉熵损失函数计算步骤1中网络模型的损失,直至模型损失小于阈值0.1,则模型训练结束,否则,利用Adam对模型参数进行优化;

[0069] 步骤3:输出经过步骤1、步骤2后得到的网络模型连接层的中间结果,作为信号特征进行第二层分类训练;

[0070] 步骤4:步骤3中得到的新的信号特征组,作为随机森林算法的输入端,进行机器学习模型的训练与分类任务;

[0071] 其中:步骤1所述的建立CNN-BiLSTM-Attention深度学习网络模型,设计4层卷积2

次池化过程及1层双向长短时记忆网络结构,最后接入连接层进行分类,训练过程中共104989个参数。首先设置进入模型训练的样本量为1000,搭建两层卷积,每个卷积层后均跟着一个LeakyRelu激活层(相当于一种Attention),神经元激活值设置为0.2,滤波器个数为8,卷积核大小为5,随后进入一个池化层,池化层使用尺寸为5,池化方法采用最大池化,步长为2,从池化层出来以后,完成了第一次的卷积池化;再继续进入卷积层循环上述过程3次,其中卷积层滤波器的大小依次为16,32,64,卷积和大小不变,池化层设置不变,循环卷积池化层后,设置Self Attention层,采用Dense-Tanh函数,把出来的64维特征连入双向长短时记忆网络中进行特征的二次锐化,双向长短时记忆网络输出所有中间结果序列,输出维数设置为32,随后采用Tanh-Softmax函数,对数据进行序列点重要性的衡量,最后,把通过深度网络提取出的序列信息进行拉伸,连接进入全连接层,序列长度为512,进入两层连接密集层,每一层的神经元个数依次为64、32,在这期间,为防止网络出现过拟合的问题,设置Dropout机制(随机失活部分神经元),嵌入隐含层中,Dropout随机失活限制值设置为0.5,若神经元计算出来的值小于0.5,则此神经元将不再连入下一层密集层中,通过这种方式,对连接密集层各层的神经元进行筛选,最后进入Softmax层进行心电信号的分类。

[0072] 步骤2所述运用交叉熵损失函数计算步骤1中网络模型的损失。在训练网络的过程中,采用Softmax输出层上的交叉熵作为整个网络的损失函数,为了训练网络,采用Adam优化方法,学习率learning-rate=0.001,beta1=0.9,beta2=0.999,设置epochs为20,batch_size为500,网络的收敛性质较好,训练网络的时间在合理范围内。

[0073] 步骤3所述的提取中间结果作为机器学习算法输入层的方法,将训练好的CNN-BiLSTM-Attention深度学习网络模型中连接层的中间结果提取出来,作为描述心电信号的新特征,作为随机森林算法的输入端,由于深度学习网络模型通过交叉熵损失函数进行了迭代优化,因此中间结果可以作为原始心电信号的映射,是有一定意义的;

[0074] 步骤4所述的利用随机森林算法对信号特征进行第二层分类训练将步骤7中所提取出的新特征作为输入端,进行随机森林分类的训练,进行分类结果的输出;

[0075] 本单元的作用:通过结合卷积神经网络与长短时记忆网络,在其中加入部分注意力机制的网络结构,对心拍数据进行有效的特征提取,采用传统机器学习分类器进行分类的算法,能够使训练样本准确率达到97.5%,相比于其他分类方法,有一定的优势。

[0076] 为进行单个模块的比较,在同样经过本发明中心拍数据提取模块与数据平衡模块处理后,参照相似现有技术中的分类方法,总结以下几种分类器在此应用实施例上的效果:

[0077] 1、采用经典卷积神经网络分类器,卷积、池化层各3层,每层卷积后均跟着一个池化层,参数与本发明中卷积池化层相同,直接对心拍数据进行分类识别,分类准确率达到92.06%;

[0078] 2、采用卷积神经网络与长短时记忆网络相结合的分类器,卷积、池化层参数与层数以及长短时记忆网络参数与本发明相同,对心拍数据进行分类识别,分类准确率达到93.75%;

[0079] 3、采用长短时记忆网络与注意力机制相结合的分类器,长短时记忆网络参数与本发明一致(具体相关专利中没有层数多少的描述),注意力机制采用相关发明中Tanh函数,放在长短时记忆网络层之后,对心拍数据进行分类识别,分类准确率达到93.03%(因训练过程极为缓慢,因此选择迭代次数为10);

[0080] 4、采用卷积神经网络与长短时记忆网络相结合的分类器进行特征提取,进入支持向量机等机器学习算法进行分类输出(采用网格寻优方法,设置可选参数范围, γ 为 0.001-1000, 惩罚系数 C 为 0.001-1000, 进行最优分类结果计算), 卷积层池化层为 2 层(与相关发明中一致), 每一层的网络参数与本发明前两层一致, 对心拍数据进行分类识别, 分类准确度达到 96.14%;

[0081] 其中系统中的综合判别单元执行以下内容:

[0082] 对心电信号进行心拍信号的提取, 然后通过集成分类器搭建模块进行分类, 综合连续多个节拍信号分类结果, 确定分类结果。

[0083] 其中: 综合 10 秒内连续多个节拍信号分类结果, 将多个心电节拍信号的分类结果进行统计, 若节拍信号中分类结果中判别为该类异常的比例超过 30%, 则最终诊断为该类异常。

[0084] 本单元的作用: 通过有效分类器分类出的样本片段, 在实际应用中依旧存在着部分判错的风险(尽管实施例中的样本数据把判错概率在了很小的范围内), 因心电信号短时间内的波动性与易变性, 因此本发明考虑加入综合判别模块, 以时间为轴, 顺时间推移, 综合连续时间内多个片段的分类结果, 做出实时综合分类判别, 进一步降低判错概率, 使心电分析更加具有可靠性, 其他相近文献中均没有此模块的设置描述。

[0085] 经过实验验证, 本发明在心电信号分类方面的技术效果: 对于心电图信号的分类诊断, 本算法通过以上设计能够使得 MITBIH 数据集上心电信号的片段分类准确度达到 97.5%, 结合多片段分类结果综合, 使得心电异常诊断更加具有可靠性。

[0086] 本发明针对现有技术中生物信号分类技术的缺点, 提出一种长时间监测的混合学习生物信号识别算法, 利用了深度学习与机器学习两方面的优势, 有效提升了生物医学信号的分析效果。

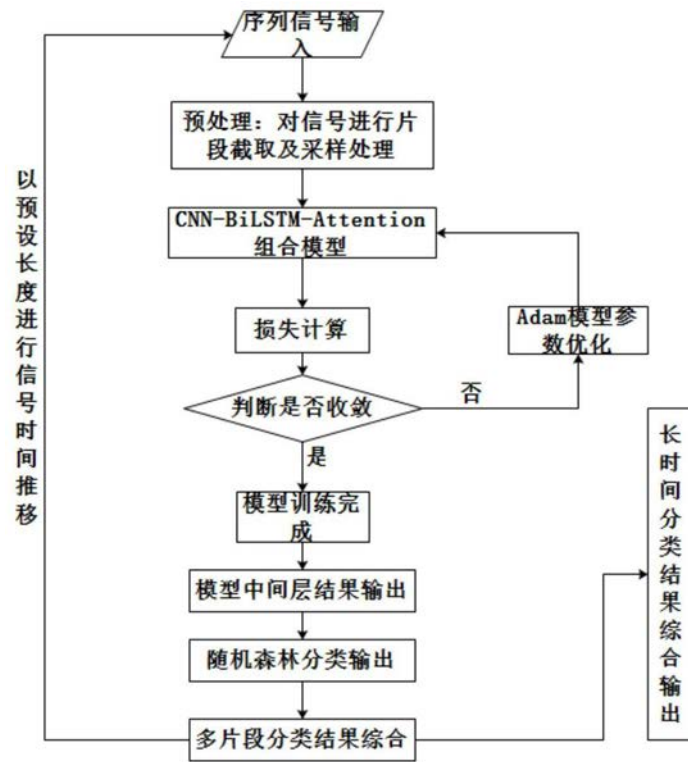


图1

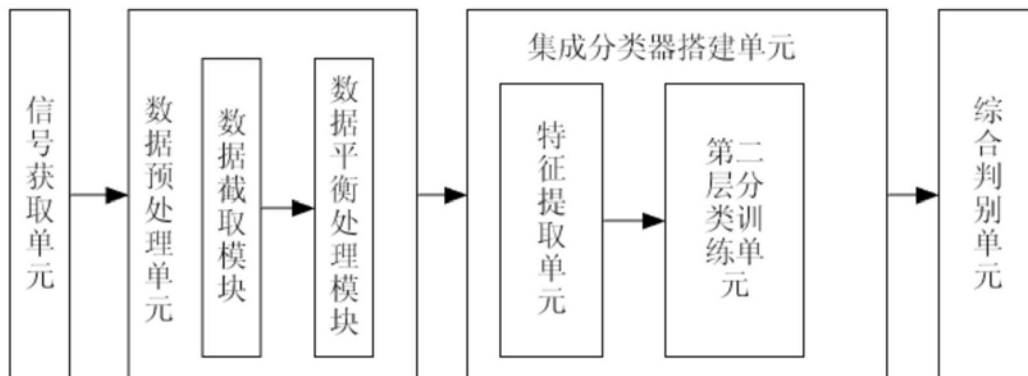


图2

专利名称(译)	一种智能生物信号异常检测方法及系统		
公开(公告)号	CN111053549A	公开(公告)日	2020-04-24
申请号	CN2019111341051.X	申请日	2019-12-23
[标]申请(专利权)人(译)	威海北洋电气集团股份有限公司		
申请(专利权)人(译)	威海北洋电气集团股份有限公司		
当前申请(专利权)人(译)	威海北洋电气集团股份有限公司		
[标]发明人	刘伟 曲媛媛 郑旭东 秦志亮 刘晓炜 谢耘		
发明人	刘伟 曲媛媛 郑旭东 秦志亮 刘晓炜 谢耘		
IPC分类号	A61B5/0402 A61B5/00		
CPC分类号	A61B5/04012 A61B5/0402 A61B5/7235 A61B5/7267		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及一种能够对生物信号进行高效准确、低成本的自动分析的智能生物信号异常检测方法及系统，包括模型训练阶段和模型应用阶段，所述模型训练阶段包括：获取生物信号；对获取的信号进行预处理；将预处理后的数据送入CNN-BiLSTM-Attention组合模型进行特征的提取，其中首先令信号进入CNN卷积神经网络进行空间特性的提取，而后进入双向长短时记忆网络提取时间特性，再通过各个网络连接层，最后通过SoftMax层计算最终结果，在搭建网络的每个环节，通过设置Attention机制，进行信号不同序列点侧重程度的优化；计算网络模型的损失，直至模型损失小于阈值，则模型训练结束，否则利用Adam对模型参数进行优化；本发明有效提升了生物医学信号的分析效果。

