

多特征信息融合 LSTM-RNN 检测 OSA 方法

朱兆坤^{1,2,3} 李金宝¹

¹(齐鲁工业大学(山东省科学院)山东省人工智能研究院 济南 250014)

²(黑龙江大学计算机科学技术学院 哈尔滨 150080)

³(黑龙江大学软件学院 哈尔滨 150080)

(zzklove3344@hotmail.com)

Multi-Feature Information Fusion LSTM-RNN Detection for OSA

Zhu Zhaokun^{1,2,3} and Li Jinbao¹

¹(Shandong Artificial Intelligence Institute, Qilu University of Technology (Shandong Academy of Science), Jinan 250014)

²(School of Computer Science and Technology, Heilongjiang University, Harbin 150080)

³(Software Technology Institute, Heilongjiang University, Harbin 150080)

Abstract Obstructive sleep apnea (OSA) is the most common sleep respiratory disorder, and it is a potential threat to many physiological systems, especially the cardiovascular system. Most of the previous methods for OSA detection extracted the shallow features from electrocardiograph (ECG) which would be used in classifiers, and they failed to achieve excellent performances on the ECG signal with high noise and large datasets. To solve this kind of problem, this paper proposes a long short-term memory recurrent neural network (LSTM-RNN) based on combination of multiple kinds of feature signals. The method fuses multiple kinds of shallow feature signals that are extracted from ECG signals and learns the deep feature from the fused signals. The accuracy of OSA detection model in long ECG segments is increased and the generalization ability on large datasets is improved. An effective preprocessing method is proposed for shallow feature signals to highlight the variation of OSA time sequences. The preprocessing method may improve the convergence of training neural networks, reduce the impact of outlier noise, and further improve the detection accuracy of the model for the ECG segments with high noise. The experimental results indicate that our method is superior to the existing methods in the accuracy of per-segment OSA detection.

Key words obstructive sleep apnea; electrocardiograph; preprocessing; long short-term memory; recurrent neural network

摘要 阻塞性睡眠呼吸暂停(obstructive sleep apnea, OSA)是最常见的睡眠呼吸疾病,它对人体的很多生理系统尤其对心血管系统是一个潜在的威胁.现有使用心电信号(electrocardiograph, ECG)提取浅层特征检测OSA的方法在长片段、高噪声的ECG信号和大数据集上表现较差.针对上述问题,提出一种多特征心电信号融合的长短期记忆循环神经网络,融合从ECG信号中提取的多种浅层特征信号,通过在融合信号上学习深层特征来检测OSA,提升模型在长片段ECG上的检测准确率和大数据集上的

收稿日期:2019-08-23;修回日期:2020-04-30

基金项目:国家自然科学基金项目(61370222);黑龙江省自然科学基金重点项目(ZD2019F003)

This work was supported by the National Natural Science Foundation of China (61370222) and the Key Program of the Natural Science Foundation of Heilongjiang Province of China (ZD2019F003).

通信作者:李金宝(ljinb@sdas.org)

泛化能力.同时还针对浅层特征信号提出一种有效的数据预处理方法,用以突出 OSA 的时序变化,提高神经网络训练的收敛性,并降低由异常值噪声带来的影响,进一步提升模型在高噪声 ECG 片段上的检测准确率.实验证明:提出的方法在片段 OSA 检测准确率上优于已有的方法.

关键词 阻塞性睡眠呼吸暂停;心电信号;预处理;长短时记忆;循环神经网络

中图法分类号 TP391

阻塞性睡眠呼吸暂停(obstructive sleep apnea, OSA)是最常见的睡眠呼吸障碍,被国际睡眠障碍分类划分为第二大类别睡眠障碍^[1],它通常是指睡眠过程中上呼吸道气流周期性的部分或完全缺失^[2].根据 Senaratna 等人^[3]的统计,OSA 在成人中占比在 6%~17% 之间,男性比例明显高于女性且会随着年龄的增长出现的比例会越高,在高龄人群中男性与女性 OSA 的占比分别高达 90% 与 78%.当 OSA 发生时,体内二氧化碳水平升高,促使患者频繁苏醒,影响睡眠质量,导致容易困倦和疲劳^[4].OSA 也是导致高血压、充血性心力衰竭、中风、糖尿病等医学疾病的一个重要因素^[5].因此,早期检测与治疗 OSA 是十分必要的.多导睡眠仪(polysomnography, PSG)是诊断 OSA 的黄金标准.使用 PSG 诊断 OSA 时,需要患者到专业的睡眠实验室,在睡眠过程中佩戴多个传感器,并在专业人员的指导下完成心电信号(electrocardiograph, ECG)、呼吸信号(respiration)和血氧饱和度信号(blood oxygen saturation, SpO2)等多种生理信号的采集.这些信号将会被睡眠专家人工分析与打分,形成最终的诊断结果.使用 PSG 检测 OSA 存在耗时、花费高,以及检测过程中会给患者带来不舒适感等缺点,因此,有超过 85% 的 OSA 患者没有得到及时的诊断与治疗^[5].

OSA 发生时上呼吸道部分或完全阻塞,导致人体吸入氧气量减少,必须加大力度吸气,致使胸腔内负压不断增大,进而破坏睡眠过程中心血管系统的稳定状态,产生心率不齐、心率过缓和心率突变等不良现象. ECG 是一种常用的检测 OSA 的单通道信号.随着科学技术的发展,越来越多的智能设备如智能手环、智能手表等可以通过非侵入式的方式获取到 ECG 信号,让患者能在家庭环境下以较低的成本、较舒适地完成 OSA 的检测.因此,有很多使用 ECG 信号检测 OSA 的方法被提出.

这些方法大致可以分为两大类别:1)传统机器学习方法^[6-10];2)深度学习方法^[11-13].传统机器学习方法通常从 ECG 信号中提取时域、频域和非线性等特征并利用支持向量机(support vector machine,

SVM)等分类器对特征进行分类,在低噪声的 ECG 信号和小数据集上有着较好的检测效果.但传统机器学习方法提取的特征在高噪声的 ECG 信号和不同个体上的差异较大,因此,在高噪声的 ECG 信号和大规模数据集上表现较差.深度学习方法通常以 ECG 信号或从 ECG 信号中提取的浅层特征信号作为输入,对低噪声的短片段 ECG 信号有着较好的检测效果.而长片段 ECG 信号由于含有较多的 QRS 波群(QRS complex),每个波群具有相同的结构,使长片段 ECG 信号之间高度相似,致使长片段中 OSA 表现并不显著.因此深度神经网络直接学习原始 ECG 信号容易出现不收敛和拟合度较低的现象.而现有利用浅层特征信号的方法一般只使用 RR 间隔(RR interval, RRI)作为神经网络的输入,只能学习到心率的变化规律,忽略了 ECG 信号中的其他信息,如 QRS 波群电压变化规律、呼吸气流变化规律等,也没有考虑预处理方法对实验结果的影响,导致 OSA 检测效果较差.

针对以上问题,本文提出一种基于多特征融合心电信号的长短期记忆循环神经网络(long short-term memory recurrent neural network, LSTM-RNN)^[14]检测 OSA.使用从 ECG 信号中提取的浅层特征信号替代原始 ECG 信号作为输入.与原始 ECG 信号相比,浅层特征信号具有更明显的时序变化和更短的信号长度,使其在较长片段和含有较高噪声的 ECG 信号上的表现优于原始 ECG 信号.由于单个特征信号只能从一个侧面反映 ECG 信号的变化.因此,我们将多个浅层特征信号融合,同时从多个侧面捕捉 ECG 信号的变化,并针对浅层特征信号提出了一种预处理方法,在最大化单个浅层特征信号分类效果的基础上进行融合,获得更好的分类效果.最后,利用 LSTM-RNN 在时间序列信号上的优异表现和对较大数据集有较好的泛化能力的特点,在融合后的多个浅层特征信号上学习深层特征.

本文的主要贡献有 2 个方面:

1) 提出一种基于多特征信号信息融合的 LSTM-RNN,通过融合多种从原始 ECG 信号中提取的浅层

特征信号作为神经网络的输入,提高模型在长片段和较高噪声 ECG 信号上的分类准确率;

2) 提出一种针对浅层特征信号的预处理方法,解决神经网络在训练过程中容易不收敛和分类准确率较低的问题。

1 相关工作

传统机器学习方法常用于片段 ECG 信号上的 OSA 分类. Cartwright^[5] 通过提取 QRS 波群主成分和呼吸心率互信息,结合 RRI 的标准差和序列相关系数, 使用最小二乘支持向量机(least squares support vector machine, LS-SVM)对提取的特征进行分类, 在基于片段 ECG 信号的 OSA 分类上得到了 85% 的分类准确率; Varon 等人^[6] 从 ECG 信号中提取 RRI 并从中推导出瞬时心率(instantaneous heart rate, IHR), 计算 IHR 多个频带的能量作为特征并使用二次分类器进行分类, 在基于片段 ECG 信号的 OSA 分类上得到了 84.7% 的准确率; Babaeizadeh 等人^[7] 利用 Hermite 基函数近似 QRS 波群, 使用 Hermite 扩张系数和 RRI 的均值、标准差和总能量作为特征, 并使用 LS-SVM 对特征进行分类, 在基于片段 ECG 信号的 OSA 分类问题上得到了 84% 的分类准确率; Martín-González 等人^[10] 从 ECG 信号中提取 RRI, 并从中提取了多种线性与非线性特征, 包括倒频谱系数、滤波器组和去趋势波动分析, 使用逻辑回归(logistic regression, LR)等对特征进行分类, 在基于片段 ECG 信号的 OSA 分类问题上得到了 84% 的分类准确率. 传统机器学习方法在低噪声的 ECG 片段上取得了较好的效果, 但在含有较高噪声的 ECG 信号和较大数据集上因特征值差异较大, 分类效果并不理想.

深度学习方法使用神经网络从信号中学习特

征, 使用的信号包括原始 ECG 信号以及从 ECG 中提取的浅层特征等. Cheng 等人^[11] 利用 RNN 检测 OSA, 通过从 ECG 中提取的 RRI 进行分类, 在基于整晚 ECG 信号的 OSA 记录筛选上得到了 97.8% 的准确率, 但没有给出 ECG 片段分类的结果. Li 等人^[12] 使用 2 阶段的学习方法: 第 1 个阶段使用 2 层堆叠的稀疏自编码器(sparse autoencoder, SAE)从 RRI 中学习特征表示; 第 2 个阶段使用人工神经网络(artificial neural network, ANN)和 SVM 对从 SAE 得到的特征进行分类, 并利用隐 Markov 模型(hidden Markov model, HMM)对 2 种分类器的结果进行综合, 在基于片段 ECG 信号的 OSA 分类问题上得到了 84.7% 的分类准确率. Urtnasan 等人^[13] 利用卷积神经网络(convolutional neural network, CNN)针对 10 s 的滑动窗口 ECG 信号提取特征, 然后进行分类, 在基于整晚 ECG 信号的 OSA 记录筛选上得到了 90.8% 的分类准确率, 但在长片段的 ECG 信号上因 OSA 发生与不发生之间差异较小表现不佳, CNN 网络容易出现不收敛或拟合度较低的问题. 此外, 使用原始 ECG 信号作为神经网络的输入时对 ECG 信号的质量要求较高, 对含有较高噪声的 ECG 信号表现较差.

本文从长片段 ECG 信号中提取多个浅层特征信号, 并将它们融合作为 LSTM-RNN 的输入, 使 LSTM-RNN 同时学习特征信号的时序变化和各个特征信号之间的关系, 获得更多的深层特征, 提高网络对较高噪声 ECG 信号的 OSA 检测准确率和对较大数据集的泛化能力.

2 数据预处理

2.1 数据去噪

原始 ECG 信号中存在基线漂移问题, 影响 R 峰

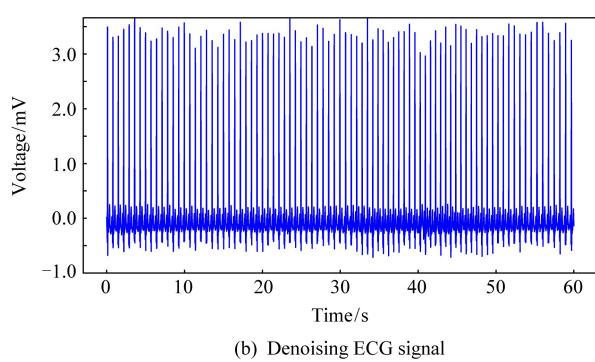
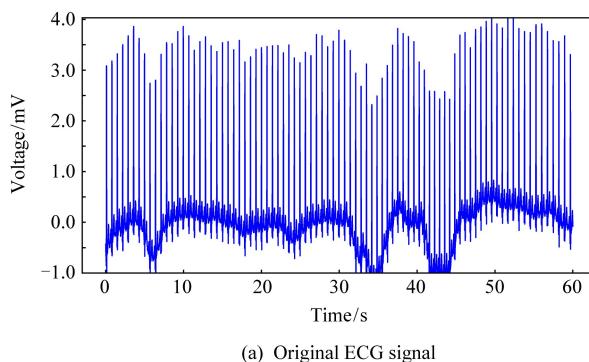


Fig. 1 The comparision of original and denoising ECG signal

图 1 原始 ECG 信号与去噪后的 ECG 信号对比

位置的计算,造成基线漂移的噪声信号频率范围在0~1.5 Hz之间,本文使用小波变换方法去除ECG信号基线漂移。对信号进行小波分解,可以得到细节系数(detail coefficient) C_d 和近似系数(approximate coefficient) C_a 两个部分,分别包含高频部分与低频部分。本文使用的ECG信号的频率为100 Hz,因此需要对该信号的近似系数使用小波迭代分解6次,得到0~1.5 Hz的噪声信号 $C_{a,6}$ 。将 $C_{a,6}$ 中所有系数值置0后再逐层进行小波重构,可以得到去噪后的ECG信号。本文使用的小波为db6小波^[15],对ECG信号的去噪效果如图1所示。

2.2 特征信号提取与预处理

本文从ECG信号中提取RRI, R峰幅值信号(R peaks amplitude, RAMP)和由心电信号导出的呼吸信号(ECG-derived respiration, EDR)三种浅层特征信号。首先对ECG信号使用Pan-Tompkins^[16]算法检测R峰位置并计算RRI, RAMP取R峰位置左右各25个值中的最大值。通过R峰幅值和QRS波群主成分可以重构EDR信号^[17]。RRI体现心率的变化, RAMP体现QRS波群电压的变化, EDR

信号体现呼吸气流的变化。这3种从ECG中提取的浅层特征信号能够从不同侧面反映OSA发生时信号的变化。我们通过融合这3种信号增加浅层特征信息,有利于LSTM-RNN学习到更多的深层特征。通过实验我们发现,对浅层特征信号上使用常规的数据预处理方法如最小-最大标准化和z-score标准化等,神经网络容易出现不收敛的问题。上述的预处理方法将信号数值标准化在较小数值范围内,缩小了信号在时间序列上的变化。为了解决该问题,我们提出一种有效的预处理方法:对特征信号减去均值后放大值域范围,减去均值可以使信号在正负区间内分布均匀,放大值域范围可以扩大信号在时间序列上的变化。对信号的扩大倍数主要取决于该信号的初始值域范围。因为LSTM-RNN中使用tanh作为神经网络的激活函数,tanh在-3~3的范围内变化明显。当我们把信号的初始值域扩大后,信号的时序变化能够在神经网络中有效地传递,使网络更好地学习特征,具体的实验细节见4.2节。对浅层特征信号进行预处理后,还需要将3种浅层特征信号的采样频率统一,使它们能够融合成一个多维的数据

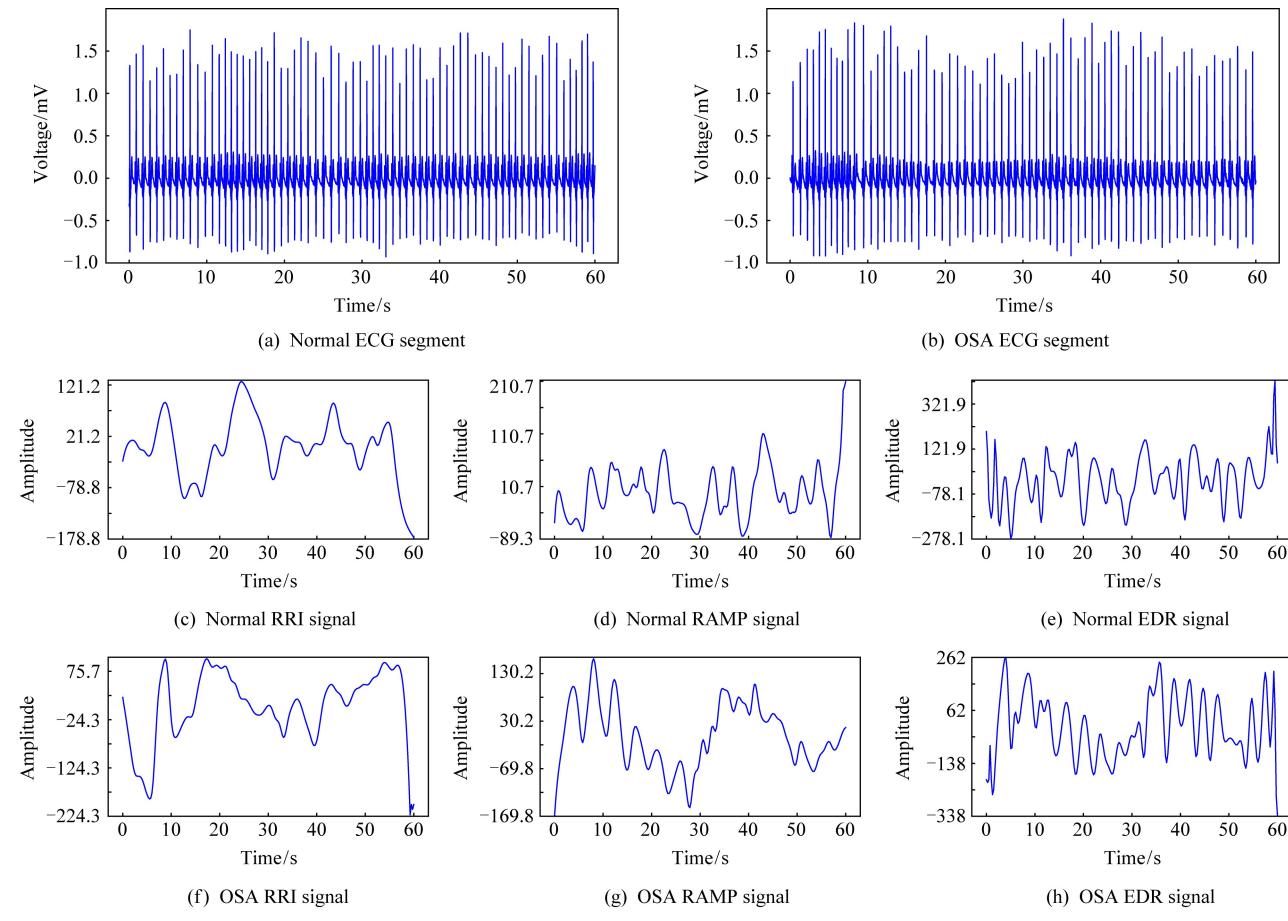


Fig. 2 Comparision on ECG, RRI, RAMP and EDR between normal breathing and OSA

图2 正常呼吸片段与OSA片段在ECG, RRI, RAMP, EDR上的对比

作为神经网络的输入,RRI 信号体现心率的变化,一般人的心跳在每秒 4 次以下;RAMP 使用 R 峰幅值作为采样点,采样频率与 RRI 信号相同;EDR 信号体现呼吸信号的变化,一般呼吸信号频率低于 0.5 Hz.所以,我们将 RRI,RAMP 和 EDR 信号的采样频率统一至 4 Hz,可以最大程度保留浅层特征信号中的信息,并防止因为采样频率过高使信号的细节变化被放大,导致特征信号中保留过多的无用信息.对 RRI 和 RAMP 信号,采用二次样条插值方法将采样频率提高至 4 Hz.对 EDR 信号,采用随机下采样方法将采样频率降至 4 Hz.图 2 给出了 OSA 片段和正常呼吸片段在原始 ECG 信号和从原始 ECG 信号中提取的 RRI,RAMP,EDR 信号的对比.对比图 2(a)与图 2(b),可以看出在长片段 ECG 信号上正常呼吸和 OSA 的 ECG 信号差异不明显.而对比图 2(c)与图 2(f)、图 2(d)与图 2(g)、图 2(e)与图 2(h)这 3 组,可以看出浅层特征信号在正常呼吸和 OSA 的差异较大,更有利于神经网络从中提取深层特征.

3 LSTM-RNN 模型

RNN^[18-20] 网络具有记忆能力,可以高效率对序列的非线性特征进行学习.LSTM-RNN 是一种被广泛使用的 RNN 模型,通常由多个 LSTM 单元串联构成,每个 LSTM 单元中由记忆单元 Cell、遗忘门、输入门和输出门构成,Cell 是 LSTM 单元的主体,门主要通过一个 sigmoid 函数和一个点乘让信息有选择地通过.LSTM-RNN 常被应用于捕捉数据的长期与短期时间依赖,门控单元使其拥有较强的泛化能力,在时间序列数据上有着较好的表现.

本文使用 LSTM-RNN 学习经过融合的多特征浅层信号,同时捕捉特征信号在时间序列上的变化规律和多个特征信号在同一时刻的关系,LSTM-RNN 模型如图 3 所示.我们将 3 个浅层特征信号依据时间串联成一个多通道输入数据,使用 3 个串联的 LSTM 块从多通道输入提取深层特征,设置 LSTM 块的时间步长为 1.

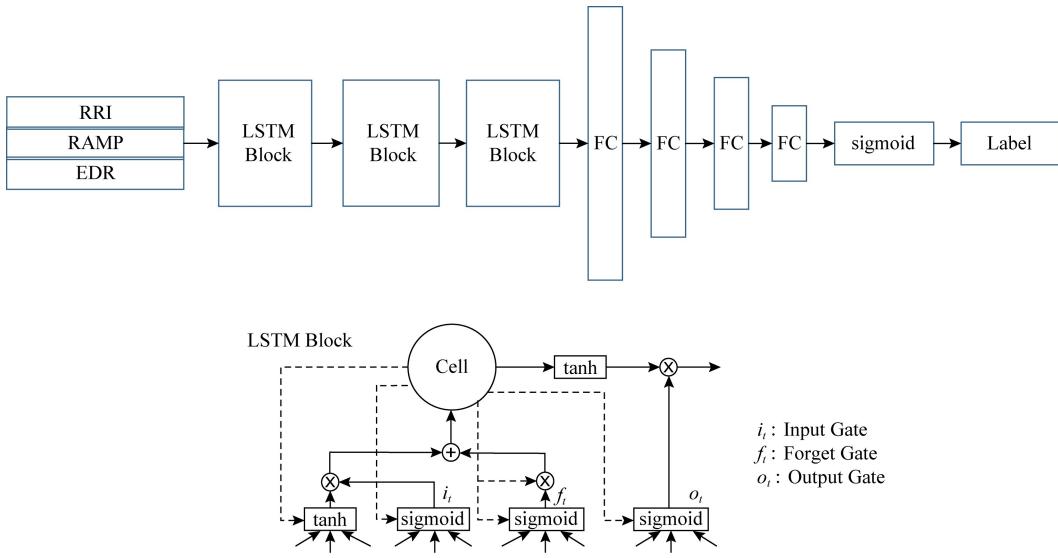


Fig. 3 The proposed LSTM-RNN model

图 3 本文使用的 LSTM-RNN 模型

第 3 个 LSTM 块仅返回最后一个时刻的输出,减少深层特征个数,降低参数量.在 3 个串联的 LSTM 块后,使用 4 层全连接网络对深层特征进行分类.使用二元交叉熵作为损失函数:

$$L(Y, P) = - \sum_x Y_x \lg(P_x), \quad (1)$$

其中, P_x 代表 LSTM-RNN 对 ECG 片段 x 的预测值, Y_x 代表 ECG 片段 x 的真实标签. LSTM-RNN 网络输入维度、输出维度与具体参数详见 4.3 节.

4 实验结果与分析

4.1 实验设置

1) 数据集.本文使用 PhysioNet 的 Apnea-ECG 数据集作为训练数据与测试数据^[21-23],该数据集由菲利普斯大学提供,是一个被广泛使用的基于 ECG 信号检测 OSA 的公开数据集.Apnea-ECG 数据集

中共有 70 条 ECG 记录, PhysioNet 将其划分为 2 个子数据集:训练集与测试集, 其中每个数据集有 35 条记录. 每条 ECG 记录的每分钟都带有一个标签 A 或 N, A 代表非正常呼吸, N 代表正常呼吸, 该标签由睡眠专家标注. 训练集和测试集分别有 17 045 个和 17 268 个被标注的 ECG 片段, 我们去掉数据集中 R 峰无法检测和 $RRI < 30$ 或 $RRI > 180$ 的异常片段, 筛选后训练集和测试集分别保留 15 961 个和 15 938 个片段.

2) 实验环境. 采用 ADAM^[24] 优化方法, 参数 $beta_1$ 和 $beta_2$ 分别设置为 0.9 和 0.999, 学习率为 0.001. 实验环境为基于 Windows 10 操作系统、Keras2.0.1 深度学习框架, 使用 GPU 进行训练, 配置 NVIDIA CUDA10.0 与 cuDNN7.6.0 深度学习库加速 GPU 运算, 用于训练和测试的软件为 Python3.6, 硬件配置为 AMD Ryzen 1800 处理器, Nvidia GeForce GTX1080 GPU, DDR4 32 GB 内存.

3) 评估标准. 准确率(accuracy, Acc)、敏感度(sensitivity, Sen) 和 特异性(specificity, Spe), 计算方法分别为:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \quad (2)$$

$$Sen = \frac{TP}{TP + FN}, \quad (3)$$

$$Spe = \frac{TN}{TN + FP}. \quad (4)$$

其中, 真正例(true positive, TP) 代表正例中判断正确的样本数量; 真负例(true negative, TN) 代表负例中判断正确的样本数量; 假正例(false positive, FP) 代表正例中判断错误的样本数量; 假负例(false negative, FN) 代表负例中判断错误的样本数量. 准确率越高则说明网络对样本的分类效果越好, 敏感度越高则说明网络对正样本的识别效果越好, 特异性越高则说明网络对负样本的识别效果越好.

4.2 原始 ECG 信号作为输入时的分类效果

使用长片段原始 ECG 信号时, LSTM-RNN 在训练中损失基本无下降, 训练结果如图 4 所示.

在前 26 轮训练过程损失无明显下降, 造成这一结果的原因是 ECG 信号中的 QRS 波群具备一定的结构, 导致传统的 LSTM-RNN 无法从结构型的信号中学习有效的特征. 本文提出的基于多特征信号融合的方法使用从 ECG 信号中提取的特征信号作为输入, 这些信号没有特定的结构, 不需要针对信号本身设计相应的网络模型.

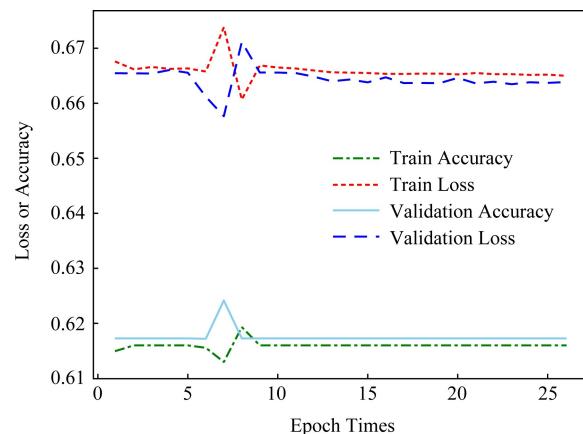


Fig. 4 Training result of denoising ECG by LSTM-RNN

图 4 去噪 ECG 信号使用 LSTM-RNN 的训练结果

4.3 不同预处理方法对分类结果的影响

在神经网络中, 最常用的数据标准化方法主要有最小-最大标准化和 z-score 标准化, 计算方法分别为:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (5)$$

$$x^* = \frac{x - x_{\mu}}{x_{\sigma}}, \quad (6)$$

其中, x_{\min} 为浅层特征信号的最小值, x_{\max} 为浅层特征信号的最大值, x_{μ} 为原始浅层特征信号平均值, x_{σ} 为原始浅层特征信号的标准差, x 为原始浅层特征信号, x^* 为初始化后的浅层特征信号.

通过实验发现, 这 2 种预处理方法因将信号处理至较小数值, 使浅层特征信号在时间序列上变化规律并不明显, 会使 LSTM-RNN 容易出现不收敛的问题. 所以, 我们通过减去均值并扩大值域范围, 使浅层特征信号在时间序列上变化规律更加明显. 具体结果如表 1 所示. 当我们对 3 种特征信号不做任何预处理、做最小-最大标准化和 z-score 标准化时, LSTM-RNN 都无法收敛, 说明常用的最小-最大标准化和 z-score 标准化不适用于 RRI, RAMP, EDR 信号. 在不做任何预处理时, 信号在正负区间内分布不均匀, 使 LSTM-RNN 无法捕捉到有效的时序变化. 做最小-最大标准化时和 z-score 标准化后, 信号在时间序列上的变化被缩小, 导致 LSTM-RNN 无法从微小的变化中学习到有效的变化规律. 我们将所有的特征信号减去均值, 使其在正负区间内分布均匀, tanh 激活后有更好的数据分布. 我们对 3 种特征信号进行不同程度的放大, 从 10~100 000 倍, 探究每个特征信号的最优放大倍数. RRI 因为初始

信号值域范围较大,在不经过任何放大时获得了较好的分类效果.RAMP 和 EDR 初始信号值域范围较小,在不放大时易出现不收敛现象.RAMP 和 EDR 信号分别放大 100 倍和 10 000 倍,得到了最好的分类效果.但对 RRI, RAMP, EDR 信号继续放大后,分类准确率则会有所降低.因为当信号值域范围

过大时,在神经网络中逐层传递值域范围也会较大,但在经过 tanh 函数激活后的变化较小.另外,我们注意到 z-score 标准化后扩大 100 倍的分类结果低于我们提出的预处理方法,造成这一结果的原因是 RRI, RAMP, EDR 在 z-score 标准化之后缩小了特征信号在事件序列上的差异.

Table 1 Performance of Different Preprocessing on Feature Signals Based on Per-segment OSA Classification

表 1 不同预处理方法在片段 OSA 问题上的分类结果

Preprocessing Methods	Accuracy		
	RRI	RAMP	EDR
None			
Subtracting mean	0.796		
min-max			
z-score			
z-score and magnifying 100 times	0.769	0.780	0.805
Subtracting mean and magnifying 10 times	0.791		
Subtracting mean and magnifying 100 times	0.789	0.787	
Subtracting mean and magnifying 1 000 times	0.780	0.774	0.815
Subtracting mean and magnifying 10 000 times	0.756	0.761	0.817
Subtracting mean and magnifying 100 000 times	0.735	0.725	0.805

4.4 多特征信号融合对分类结果的影响

我们比较了浅层特征信号单个或多个组合的分类结果,如表 2 所示.EDR 信号提供了单通道特征的最好的分类结果,在测试集上的分类准确率达到了 81.7%,比 RRI 和 RAMP 的分类准确率提高了约 2 个百分点.当我们融合 2 个特征信号时,都获得

了比单个信号更好的分类准确率,其中 RRI 与 EDR 的组合提供了双通道特征的最好分类准确率,达到了 83.7%.当我们融合 RRI, RAMP, EDR 信号时,分类准确率达到了 85.4%,好于单个特征信号和 2 个特征信号组合的分类准确率.实验结果表明,融合多个特征信号可以提高对 OSA 分类的准确率.

Table 2 Performance of Different Combination of Feature Signals Based on Per-segment OSA Classification

表 2 不同特征信号组合在片段 OSA 问题上的分类结果

Signals	TP	TN	FP	FN	Acc	Sen	Spe
RRI	4 394	8 301	1 537	1 706	0.797	0.720	0.844
RAMP	4 635	7 930	1 908	1 476	0.788	0.758	0.806
EDR	4 361	8 667	1 171	1 739	0.817	0.715	0.881
RRI+RAMP	4 681	8 534	1 304	1 419	0.829	0.767	0.867
RRI+EDR	4 843	8 506	1 332	1 257	0.837	0.794	0.865
RAMP+EDR	4 163	9 012	826	1 937	0.826	0.682	0.916
RRI+RAMP+EDR	4 924	8 663	1 175	1 176	0.854	0.809	0.881

表 3 给出了实验中 LSTM-RNN 获得最好分类准确率时的模型参数.参数 dropout 和 Recurrent dropout 的设置通过有选择地激活神经元提高模型的泛化能力,但其数值过大时会导致欠拟合.在 LSTM-RNN 中不宜使用 ReLU 激活函数,因为 ReLU 激

活函数会忽略负值,缩小信号的时序变化,导致网络训练时不收敛.

我们还验证了本文方法在 OSA 记录筛选问题上的准确率,即从所有 ECG 记录中筛选出患有 OSA 的记录.在 Apnea-ECG 数据库的 35 条测试记录中,

采用 Apnea-ECG 推荐的计算方法,即如果一条 ECG 记录中存在 100 min 以上的 OSA 片段,则该条 ECG 记录为重度 OSA 患者;存在 10~100 min 的 OSA 片段则为轻度 OSA 患者;存在 5 min 以下 OSA 片段则为健康人。

本文方法能够 100% 区分健康人与重度患者,但在 5 个中度患者中,把一名中度患者错判为重度患者,这是因为本文方法倾向于把正常片段判断为 OSA 片段,我们将在后续工作中引入混合决策机制解决该问题。

Table 3 Parameters Used in LSTM-RNN

表 3 LSTM-RNN 参数

LayerName	Input Shape	Layer Type	Dropout	Recurrent Dropout	Activation	Output Shape	Parameters
LSTM1	(240,3)	LSTM	0.1	0.4	tanh	(240,384)	595 968
LSTM2	(240,384)	LSTM	0.2	0.5	tanh	(240,384)	1 181 184
LSTM3	(240,384)	LSTM	0.3	0.6	tanh	(384,)	1 181 184
FC1	(384,)	FC				(128,)	49 280
FC2	(128,)	FC				(64,)	8 256
FC3	(64,)	FC				(32,)	2 080
FC4	(32,)	FC			sigmoid	(1,)	33

Notes: FC(fully-connected layer).

4.5 对比分析

我们从已发表的文献中选择了一些与本文使用相同数据库的文献作对比,如表 4 所示。已有方法大多使用传统机器学习方法,对于部分片段拥有很高的准确率,但将其应用于较大数据集时,通常不能获

得较好的分类准确率。本文使用 LSTM-RNN 学习经过融合的从 ECG 信号中提取的多特征浅层特征信号,在分类准确率上达到了 85.4%,高于 Li 等人^[12]使用 SAE 学习 RRI 的方法,也高于传统机器学习方法。

Table 4 Comparision for Per-segment Classification Between Our Method and Others

表 4 本文方法与其他方法在片段分类问题上的比较

Reference	Year	Population	Signal	Methods	Classifier	Accuracy
Carolina, et al. ^[5]	2015	70	RRI&EDR	Feature Engineering	LS-SVM	0.847
Saeed, et al. ^[6]	2010	70	RRI	Feature Engineering	Quadric	0.847
Hemant, et al. ^[7]	2016	70	QRS Complex	Feature Engineering	LS-SVM	0.838
Nguyen, et al. ^[9]	2014	35	HRV	RQA	Decision Fusion	0.850
Li, et al. ^[12]	2018	70	RRI	SAE	Decision Fusion	0.847
Our method		70	RRI&RAMP&EDR	LSTM-RNN		0.854

Notes: TQWT(tunable-Q factor wavelet transform), HRV(heart rate variability), RQA(recurrence quantification analysis)。

5 总 结

本文提出多特征信号信息融合的 LSTM-RNN 和一种针对浅层特征信号的预处理方法,通过融合多个突出时序变化后的浅层特征信号作为 LSTM-RNN 的输入,使 LSTM-RNN 能够学习到更多更有效的深层特征,提升对含有较高噪声的 OSA 片段的分类准确率与较大数据集的泛化能力。在 Apnea-ECG 数据集基于片段的 OSA 检测问题上得到了 85.4% 的分类准确率。

参 考 文 献

- [1] Sateia M. International classification of sleep disorders third edition: High lights and modifications [J]. Chest, 2014, 146(5): 1387-1394
- [2] Mendonça F, Mostafa S S, Ravelo-García A G, et al. A review of obstructive sleep apnea detection approaches [J]. IEEE Journal of Biomedical and Health Informatics, 2019, 23(2): 825-837
- [3] Senaratna C V, Perret J L, Lodge C J, et al. Prevalence of obstructive sleep apnea in the general population: A systematic review [J]. Sleep Medicine Reviews, 2017, 34: 70-81

- [4] Choia S H, Yoona H, Kima H S, et al. Real-time apnea-hypopnea event detection during sleep by convolutional neural networks [J]. *Computers in Biology and Medicine*, 2018, 100(1): 123–131
- [5] Cartwright R. Obstructive sleep apnea: A sleep disorder with major effects on health [J]. *Disease a Month*, 2001, 47(4): 109–147
- [6] Varon C, Caicedo A, Testelmans D, et al. A novel algorithm for the automatic detection of sleep apnea from single-lead ecg [J]. *IEEE Transactions on Biomedical Engineering*, 2015, 62(9): 2278–296
- [7] Babaeizadeh S, White D P, Pittman S D, et al. Automatic detection and quantification of sleep apnea using heart rate variability [J]. *Journal of Electrocardiology*, 2010, 43(6): 535–541
- [8] Sharma H, Sharma K K. An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions [J]. *Computers in Biology and Medicine*, 2016, 77(2016): 116–124
- [9] Hassan A. Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting [J]. *IEEE Journal of Biomedical and Health Informatics*, 2016, 29(1): 22–30
- [10] Martín-González S, Navarro-Mesa J L, Juliá-Serdá G, et al. Heart rate variability feature selection in the presence of sleep apnea: An expert system for the characterization and detection of the disorder [J]. *Computers in Biology and Medicine*, 2017, 91(1): 47–58
- [11] Cheng Maowei, Sori W J, Jiang Feng, et al. Recurrent neural network based classification of ECG signal features for obstruction of sleep apnea detection [C] //Proc of 2017 IEEE Int Conf on Computational Science and Engineering and IEEE/IFIP Int Conf on Embedded and Ubiquitous Computing, CSE and EUC. Piscataway, NJ: IEEE, 2017: 199–202
- [12] Li Kunyang, Pan Weifeng, Li Yifan, et al. A method to detect sleep apnea based on deep neural network and hidden Markov model using single-lead ECG signal [J]. *Neurocomputing*, 2018, 294(1): 94–101
- [13] Urtnasan Er, Park J, Lee K. Multiclass classification of obstructive sleep apnea/hypopnea based on a convolutional neural network from a single-lead electrocardiogram [OL]. [2019-04-06]. <https://doi.org/10.1088/1361-6579/aac7b7>
- [14] Hochreiter S, Schmidhuber J. Long short-term memory [J]. *Neurocomputing*, 1997, 9: 1735–1780
- [15] Mikhled A, Khaled D. ECG signal denoising by wavelet transform thresholding [J]. *American Journal of Applied Sciences*, 2008, 5(5): 276–281
- [16] Pan J, Tompkins W J. A real-time QRS detection algorithm [J]. *IEEE Transactions on Biomedical Engineering*, 1985, 32(3): 230–236
- [17] Han D, Rao Y N, Principe J C, et al. Real-time PCA (principal component analysis) implementation on DSP [C] //Proc of IEEE Int Joint Conf on Neural Networks. Piscataway, NJ: IEEE, 2004: 2159–2162
- [18] You Yang, Lu Cewu, Wang Weiming, et al. Relative CNN-RNN: Learning relative atmospheric visibility from images [J]. *IEEE Transactions on Image Processing*, 2019, 28(1): 45–55
- [19] Chen Dehua, Yin Suna, Le Jiajin, et al. A link prediction model for clinical temporal knowledge graph [J]. *Journal of Computer Research and Development*, 2017, 54(12): 2687–2697 (in Chinese)
(陈德华, 殷苏娜, 乐嘉锦, 等. 一种面向临床领域时知识图谱的链接预测模型 [J]. *计算机研究与发展*, 2017, 54(12): 2687–2697)
- [20] Liang Bin, Liu Quan, Xu Jin, et al. Aspect-based sentiment analysis based on multi-attention CNN [J]. *Journal of Computer Research and Development*, 2017, 54(8): 1724–1735 (in Chinese)
(梁斌, 刘全, 徐进, 等. 基于多注意力卷积神经网络的特定目标情感分析 [J]. *计算机研究与发展*, 2017, 54(8): 1724–1735)
- [21] Penzel T, Mc Names J, Murray A, et al. Systemantic compratition of different algorithms for apnoea detection based on electrocardiogram recordings [J]. *Medical and Biological Engineering and Computing*, 2002, 40(4): 402–407
- [22] Penzel T, Moody G B, Mark R G, et al. The Apnea-ECG database [J]. *Computers in Cardiology*, 2000, 27(1): 255–258
- [23] Penzel T, McNames J, Murray A, el al. Systemantic compratition of different algorithms for apnoea detection based on electrocardiogram recordings [J]. *Medical and Biological Engineering and Computing*, 2002, 40(4): 402–407
- [24] Kingma D P, Ba J. ADAM: A method for stochastic optimization [C/OL] //Proc of 2015 Int Conf on Learning Representations (ICLR). [2019-05-02]. <https://arxiv.org/pdf/1412.6980v8.pdf>



Zhu Zhaokun, born in 1994. Master. His main research interests include biological signal processing, temporal neural network.



Li Jinbao, born in 1969. PhD, professor of Shandong Artificial Intelligence Institute, Qilu University of Technology (Shandong Academy of Science). His main research interests include computer vision, wireless sensor network.