

# 经验模态分解的单通道呼吸信号自动睡眠分期

白雨欣, 令狐荣乾

北方工业大学理学院, 北京

收稿日期: 2023年5月16日; 录用日期: 2023年6月9日; 发布日期: 2023年6月16日

## 摘要

睡眠是人体基本的生理需求, 可以保证机体的生长发育、为机体储蓄能量、维持机体免疫等。对睡眠质量的准确评估是认识睡眠障碍并采取有效干预措施的关键。如果用经验丰富的睡眠专家进行人工睡眠分期是比较耗时并且主观的。目前, 研究人员提出了许多准确、有效、有针对性的睡眠分期方法。比如, 基于深度学习以及经验模态分解算法的单通道呼吸信号自动睡眠分期方法, 它被成功地用于呼吸信号 (RESP) 的睡眠分期, 该方法为呼吸信号分解和睡眠阶段自动识别提供了新途径。本文采用的呼吸信号数据集来自SHHS, 它是一个中心队列研究, 用来确定睡眠与呼吸障碍的心血管和其他病症的数据库。首先, 我们对SHHS数据库中的单通道呼吸信号进行了分析, 以便更好地了解人类睡眠情况。其次, 利用经验模态分解算法(EMD)对预处理后的呼吸信号进行分解, 从原始呼吸信号和分解出的6个简单信号中提取时域、非线性动力学、统计学等方面的9个特征。最后, 使用长短期记忆网络(LSTM)构建分类模型, 将提取的呼吸信号特征进行分类识别, 实现自动睡眠分期。实验结果表明, 在4类和5类睡眠分期任务中, SHHS数据库的呼吸信号自动睡眠分期准确率分别为89.22%和88.43%。实验结果表明, 本文提出的自动睡眠分期模型具有较高的分类精度和效率, 具有较强的适用性和稳定性。

## 关键词

经验模态分解算法, 长短期记忆网络LSTM, 呼吸信号, 特征提取, 睡眠阶段分类

# Empirical Modal Decomposition of Single-Channel Respiratory Signals for Automatic Sleep Staging

Yuxin Bai, Rongqian Linghu

College of Science, North China University of Technology, Beijing

Received: May 16<sup>th</sup>, 2023; accepted: Jun. 9<sup>th</sup>, 2023; published: Jun. 16<sup>th</sup>, 2023

## Abstract

Sleep is a basic physiological need of the body to ensure growth and development, save energy for

the body, and maintain immunity of the body. Accurate assessment of sleep quality is the key to recognizing sleep disorders and taking effective interventions. Manual sleep staging is time consuming and subjective when performed by experienced sleep specialists. Currently, researchers have proposed a number of accurate, effective, and targeted sleep staging methods. For example, a single-channel computer signal automatic sleep staging method based on deep learning and empirical modal decomposition algorithms has been successfully used for respiratory signal (RESP) sleep staging, which provides a new way to decompose respiratory signals and identify sleep stages automatically. The data set used in this paper is from SHHS, which is a central cohort study to identify sleep and breathing disorders in a database of cardiovascular and other conditions. First, we analyzed the single-channel respiratory signals from the SHHS database to better understand human sleep. Second, the pre-processed respiratory signals were decomposed using an empirical modal decomposition algorithm (EMD) to extract nine features in the time domain, nonlinear dynamics, and statistics from the original respiratory signals and the six simple signals that were decomposed. Finally, a classification model was constructed using a long short-term memory network (LSTM) to classify and identify the extracted respiratory signal features for automatic sleep staging. The experimental results show that the accuracy of automatic sleep staging of respiratory signals from SHHS database is 89.22% and 88.43% in 4 and 5 categories of sleep staging tasks, respectively. The experimental results show that the automatic sleep staging model proposed in this paper has high classification accuracy and efficiency, and has strong applicability and stability.

## Keywords

Empirical Modal Decomposition Algorithm, Long Short-Term Memory Network LSTM, Respiratory Signal, Feature Extraction, Sleep Stage Classification.

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 介绍

睡眠是评价人类生活质量和身体健康的标准之一, 并且了解睡眠质量和结构对人类的健康至关重要。在人类的生命长河中, 有 1/3 的时光都在休息。睡眠障碍(失眠、昼夜节律紊乱、嗜睡和阻塞性睡眠呼吸暂停)是导致人类白天嗜睡和夜间失眠、认知和情绪障碍、皮肤衰老加速和死亡率增加的危害性因素之一 [1]。因此, 对睡眠质量的准确评估是认识睡眠障碍并采取有效干预措施的关键。

迄今为止, 对睡眠质量的评估主要是基于脑电信号的分析, 但是标准的睡眠阶段测定需要在昂贵的和劳动密集型的多导睡眠图(PSG)中使用脑电图(EEG)信号, 这就需要专用设备和实验室以及临床专家的专业知识。最近的研究人员没有依赖 PSG 的 EEG 信号, 而是专注于呼吸信号, 因为他们既可以显示不同的睡眠阶段, 也可以在大量低成本的简化设备中捕获 [2]。这种方法很适用于移动和物联网(IoT)医疗保健、智能家居中的睡眠健康监测 [3]。目前, 医生在诊断睡眠障碍相关疾病时的首要工作是对患者进行睡眠阶段分期 [4] [5], 也称为睡眠评分。虽然有很多标准来划分睡眠阶段, 但最常见的是四类睡眠阶段标准, 它包括醒觉(Wake)、浅睡眠(Light Sleep)、深睡眠(Deep Sleep)和快速眼动(Rapid Eye Movement) [6] 对于一些更精细的测量, 把睡眠阶段分成五类也是我们研究的重点。2007年, 根据 AASM 规则(American Academy of Sleep Medicine)人们可以把一个人的生活划分成三个不同的时期: 分为清醒期(Wake)、非快速眼动期(N1, N2, N3)和快速眼动期(Rapid Eye Movement) [7] [8], 其中 N1、N2 和 N3 分别代表清醒期到睡眠阶段

的过渡、浅睡和深度睡眠。为了更好地了解志愿者的睡眠状况, 我们应该对他们每个睡眠阶段的多导睡眠图(PSG)进行全面的监测, 其中的数据包括了, 脑电图(EEG)、心电图(ECG)、眼电图(EOG)、肌电图(EMG)、呼吸信号等。在基于 AASM 规则下[8], 为 PSG 中的每 30 秒数据给出睡眠分类标签(W, N1, N2, N3, R), 不同标签代表不同的睡眠状态[9] [10]。然而, PSG 的收集和整理需志愿者到睡眠实验室并由睡眠专家分析。此外, 传统的人工评分操作复杂、耗时、且评分准确性易受睡眠专家主观因素影响。

为了解决上述问题, 近年来许多科研工作者致力于研发睡眠自动评分模型。研究者们利用 Sleep-EDF 数据库、Dreams Subjects 数据库、SHHS 数据库做出了睡眠自动分期模型, 具体准确率如表 1 所示:

**Table 1.** Accuracy of automatic sleep staging models  
**表 1.** 睡眠自动分期模型的准确率

数据库	时间	人物	四分类 准确率	五分类 准确率
Sleep-EDF	2014 年	Zhu	88.9%	89.3%
Sleep-EDF	2016 年	Hassan	94.36%	93.6%
Dreams Subjects	2016 年	Hassan	83.78%	78.95%
Sleep-EDF	2020 年	Liu 等	93.1%	91.9%
Dreams Subjects	2020 年	Liu 等	86.4%	83.4%
SHHS	2020 年	Liu 等	87.5%	85.8%

1998 年, N.E. Huang、Z. Wu 等学者首次将 EMD (经验模式分解)引入美国国家宇航局, 并以此为基础, 开发了一种全新的、具备高度灵活性的自适应信息时频数据处理方法, 可以有效地实现对复杂的非线性、非平稳的数据分析[11], 经验模态分解法(EMD)是将信号分解成一些列特征时间尺度(IMF)分量, 使得各 IMF 分量是窄带信号, 简单的说就是将一个复杂信号分解成多个简单信号的过程。与小波变换不同, EMD 的分解模式基于信息的实际情况, 能够实现对不断变化的环境的快速响应。这种方法用于解决复杂的、不均匀、非线性的信号, 特别适用于时频分析。自 1998 年由 Huang 首次提出以来, EMD(模拟数据集)一直受到广泛的重视, 并且已经有所突出发现和进展。这种方法的特色在于无需依赖预先设计的函数, 就能够通过对特定信息进行加工, 从而产生固有模态的函数。这种方法能够解决复杂、变化快速、波形不规则等问题, 并且能够提供较高的信噪比和较佳的时频聚焦能力[12]。

与 EMD 不同, EEMD 拥有更强大的自我调节功能, 它不仅能够捕捉到复杂的、不断发展的信息, 而且还能够准确地预测和识别出复杂的、不断发展的、不断改进的信息, 因此, EEMD 已成为一种极为普遍的、高效的、基础的、多维的、复杂的信息处理技术。

其次, 从原始呼吸信号和分解出来的 6 个简单信号中提取关于统计、时域和非线性等方面的 9 个特征。利用长短期记忆算法(LSTM)构建分类模型。经过测试我们发现, 对 SHHS 数据库中呼吸信号进行 4 类和 5 类睡眠分期的任务中, 我们的模型准确率分别达到 89.22%, 88.43%。我们的模型的准确度分别达到 89.22%和 88.43%。这种自动化的分期模型既能够克服人工分期的延迟, 又能够抵抗外界环境的变化, 并保持良好的分类精度。

本文工作流程如图 1 所示。首先, 我们把 SHHS 数据库中的 6000 多个呼吸信号分为两组, 第一组是训练组, 也就是我们在每个睡眠阶段随机选择 85%的呼吸信号特征来生成训练集, 其余 15%作为测试组。其次, 利用四阶巴特沃斯滤波器对呼吸信号进行过滤的预处理, 再用经验模态分解算法(EMD)分解出 6 个简单信号和 9 个特征。特征向量导入 LSTM 分类模型中进行交叉验证训练和测试。特征提取后, 使用

5 倍交叉验证对 LSTM 分类器进行训练和测试。

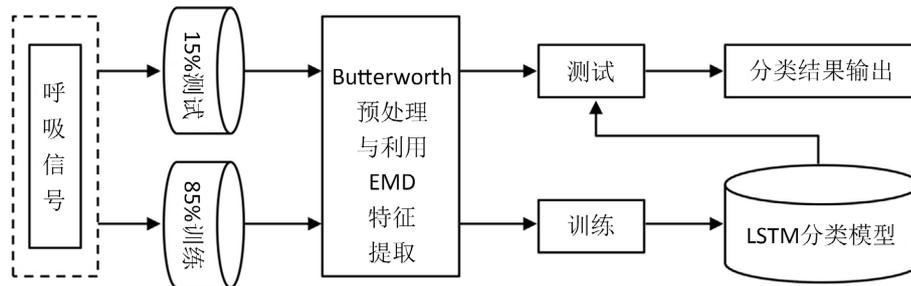


Figure 1. Workflow diagram of this paper

图 1. 本文的工作流程图

## 2. 实验数据与方法

### 2.1. 实验数据

本文研究的 SHHS (睡眠心脏健康研究)数据库是一个关于睡眠障碍和心血管疾病之间关系的长期研究的数据集合。该数据库包含了 6000 多名参与者的睡眠研究结果、病史、人口统计信息和心血管评估等信息。它是研究睡眠障碍对心血管健康影响的研究人员的宝贵资源,并为有关睡眠呼吸暂停、高血压和其他心血管疾病的研究提供了全面的数据收集。SHHS 数据库对合格的研究人员是公开的。这些数据包括 EEG、EOG、EMG、ECG、鼻腔气流和呼吸信号、SaO<sub>2</sub> 和心率测量,以及注释的睡眠阶段、呼吸事件、EEG 唤醒等。

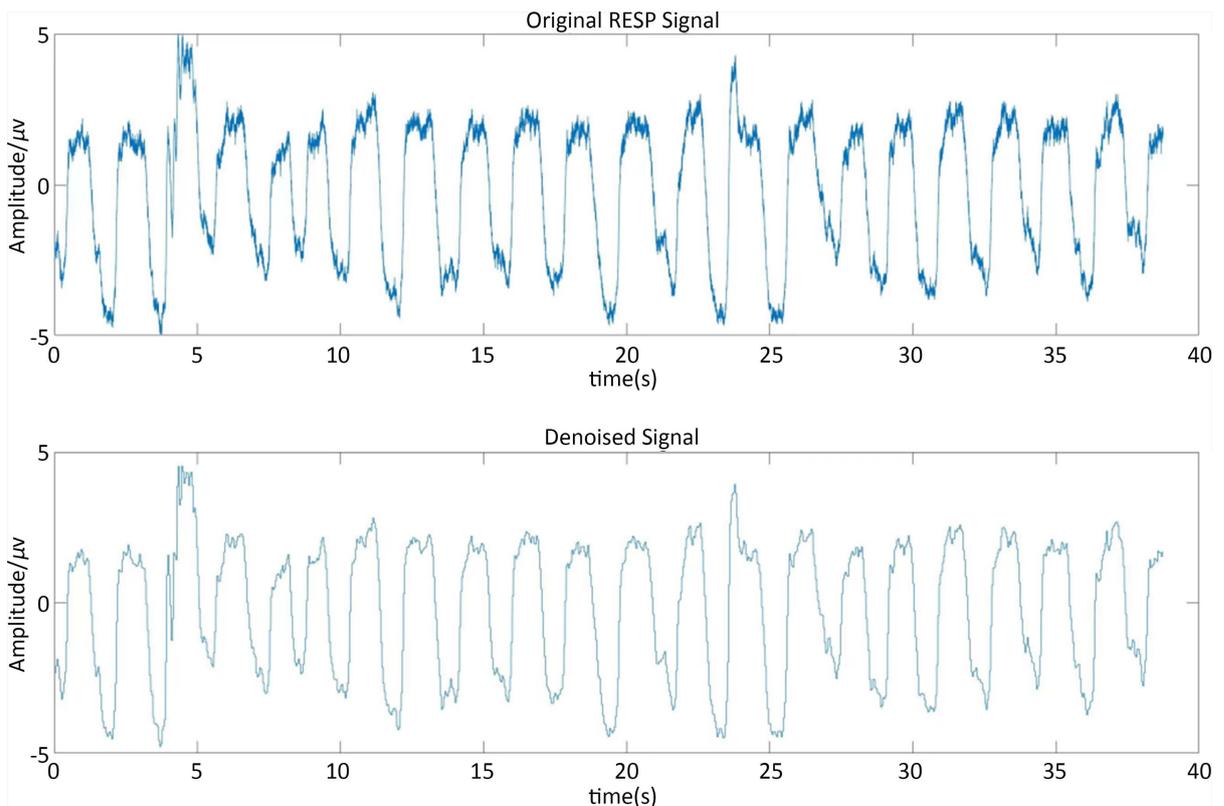


Figure 2. Raw respiratory signal and filtered respiratory signal

图 2. 原始呼吸信号和滤波处理后的呼吸信号

在这项工作中, 我们使用呼吸信号来划分睡眠阶段。呼吸信号是指呼吸时由呼吸肌产生的电信号。这些信号可以用放置在胸部或腹部的传感器测量, 并提供有关呼吸模式和效率的宝贵信息。

呼吸信号反映了参与吸气的肌肉的活动。以及呼气活动, 它反映了参与呼气的肌肉的活动。通过分析这些信号, 临床医生和研究人员可以深入了解呼吸肌的力量和协调, 以及呼吸系统的阻力和顺应性。

## 2.2. 数据预处理

首先, 利用 4 阶巴特沃斯高通滤波器(Butterworth)对分割好的各 30 秒原始呼吸信号进行滤波, 只允许频率范围在 0.5~50 Hz 的信号成分通过。其次, 由于小波分析在时域和频域上具有很好的去噪性能, 本文中, 将巴特沃斯高通滤波后的信号经过小波变换去噪。实验表明, 使用 db4 小波基和分解层数为 6 时效果比较好[13](图 2)。

## 2.3. 经验模态分解算法(EMD)

经验模式分解概念通俗易懂得讲是确定适当的时间尺度来揭示信号的物理特性。换句话讲我们在试图用 EMD 的方法来提取某些功能即 IMF (具体如下)。

经验模态分解法(EMD)是将我们的呼吸信号分解成一系列特征时间尺度(IMF)分量, 使得各 IMF 分量都具有较低的频率范围, 并且有较低的频率阈值。即 IMF 分量必须满足下面两个条件: 在整个信号长度上, 极值点和过零点的数目必须相等或者至多只相差一个; 在任意时刻, 由极大值点定义的上包络线和由极小值点定义的下包络线的平均值为零, 即信号的上下包络线关于时间轴对称。简单的说就是将一个复杂信号分解成多个简单信号的过程。同小波变换相比, EMD 方法是完全根据信号数据本身来确定需要分解出多少个 IMF, 因此更加的具有自适应性。

EMD 的分解过程其实是如下的筛选过程, EMD 方法的实现是要满足两个基本条件的过程。

- 1、信号至少要存在两个极值点, 要找出局部最大值和最小值。
- 2、在两个极端点之间, 时间尺度的变化会影响到他们的特征。

由于 EMD 的方法是自适应的, 分解的基础是基于数据并从数据中得出。在 EMD 方法中, 数据  $X(t)$  被分解成 IMFs,  $c_j$  [12]

$$x(t) = \sum_{j=1}^n c_j + r_n \quad (1)$$

即  $r_n$  代表着  $x(t)$  的残余值, 在提取了  $n$  个 IMF 以后。而  $n$  个 IMF 则表示一种不同振幅、不同频率的简单震荡函数。

EMD 分解步骤如下:

1、通过分析原始信号的极值点, 并利用曲线拟合技术将其组合成一条上下包络线, 从而将其完整地包裹起来。

2、通过上下包络线, 我们可以计算出均值曲线  $m(t)$  的均值, 然后再用原始信号  $f(t)$  减去  $m(t)$ , 最终得到  $H(t)$  即为 IMF。

3、理论上讲, 由于第一二步得到的 IMF 通常不会满足 IMF 的两个条件, 因此必须反复执行第一二步, 直至 SD 达到一个特定阈值, 一般来说, 这个阈值应该在 0.2~0.3 之间, 当 SD 低于这个阈值时, 才可以停止对  $H_1(t)$  进行第一、第二个步骤。这样得到的第一个满足条件的是  $H_1(t)$  就是 IMF [14]。

$$SD = \frac{\sum_{t=0}^r |H_{k-1}(t) - H_k(t)|^2}{\sum_{t=0}^T H_{k-1}^2(t)} \quad (2)$$

4、通过执行 1、2、3 步骤, 直至残差  $r(t)$  达到预先设定的条件, 即  $f(t)-H(t)$ ,

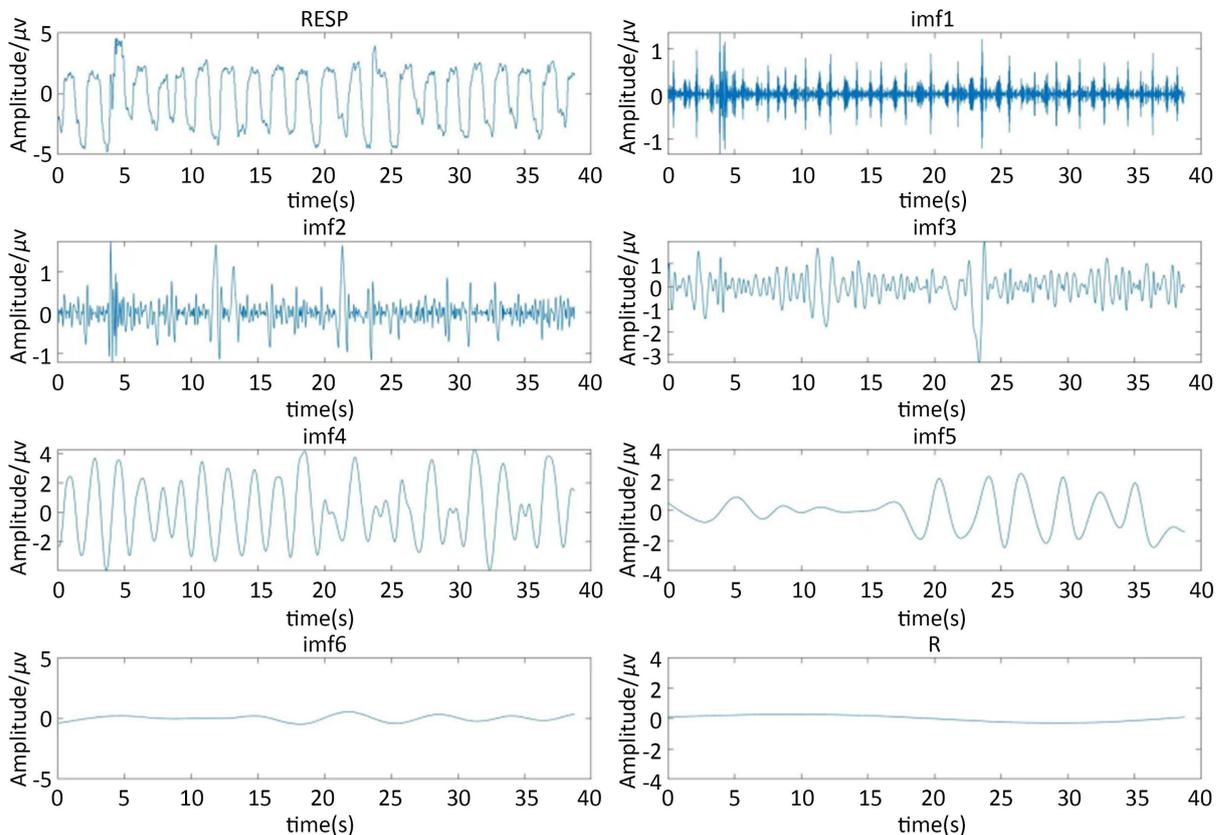


Figure 3. EMD decomposition of respiratory signals

图 3. 呼吸信号的 EMD 分解

经过我们用 EMD 方法分解之后, 我们可以分析得出从 IMF1~IMF6 频率以此降低。R 为残差信号保持在 0 的周围, 这说明 EMD 对呼吸信号的分解性较强(图 3)。

## 2.4. 特征提取

在本文中, 我们提取了原始呼吸信号, 以及经过预处理并由经验模态分解算法分解出的 6 个简单信号的 9 个特征: 均值(ME)、偏度(SK)、峰度(KU)、过零率(Zcr)、排列熵(PE)、样本熵(SE)、灵活性(HA)、复杂度(HC)和流动性(HM)。设定  $S(n)$  ( $n=1,2,3,\dots,M$ ) 为一时间序列, 下面给出上述 9 个特征的数学定义。

### 2.4.1. 均值(ME)、偏度(SK)、峰度(KU)

均值(ME)和方差( $\sigma^2$ )是常用的统计量, 可充分的使用数据, 适应性强, 但容易受到极端值的影响。偏度(SK)可以用来衡量实值变量概率分布的不对称程度, 若偏度介于-0.5 至 0.5 之间, 则表明数据具有较高的对称性。峰度(KU)是对实值变量的概率分布峰态的度量[15]。KU 可以描述数据分布形态的陡缓程度, 因此可以用 KU 来检验数据分布的正态性。

$$\text{ME} = \frac{1}{M} \sum_{n=1}^M S(n) \quad (3)$$

$$\sigma^2 = \frac{1}{M} \sum_{n=1}^M (S(n) - \text{ME})^2 \quad (4)$$

$$SK = \frac{1}{\sigma^3} \sum_{n=1}^M (S(n) - ME)^3 \tag{5}$$

$$KU = \frac{1}{\sigma^4} \sum_{n=1}^M (S(n) - ME)^4 \tag{6}$$

**2.4.2. 过零率(Zcr)**

对于离散信号, 图像单位时间内穿过时间轴的次数称为过零率(Zcr), 它是信号采样点符号的变化率。这种变化响应于信号频谱的变化[16], Zcr 特性已被广泛使用, Zcr 越大, 相应的频率就越高。Zcr 的计算可以用等式(7)~(8)来表示。

$$Zcr = \frac{1}{M-1} \sum_{n=2}^M \varphi_{R<0}(S(n)S(n-1)) \tag{7}$$

其中  $\varphi_{R<0}$  为指标函数, 定义为:

$$\varphi_{R<0}(s) = \begin{cases} 1, & s < 0 \\ 0, & s \geq 0 \end{cases} \tag{8}$$

**2.4.3. 样本熵(SE)**

SE 由 Richman、Moorman 等学者在 2000 年提出, 它可以更准确地反映出时间序列的复杂性, 且计算不依赖数据的长度。参数  $m$  和  $r$  的变化对样本熵的影响程度相同, 因此, SE 具有很好的一致性。SE 被广泛应用于脑电信号和呼吸信号等时间序列的分析[17]。

为了计算 SE, 首先将输入信号  $S(n)$  向量重构为  $M - m + 1$  个  $m$  维向量  $D_m(n)$ , 如(9)所示。

$$D_m(n) = \{S(n), S(n+1), \dots, S(n+m-1)\} (n=1, 2, \dots, M-m+1) \tag{9}$$

距离  $d[D_m(i), D_m(k)]$  定义为向量  $D_m(i)$  和  $D_m(k)$  之间相应位置元素差的绝对值的最大值, 公式(10)中  $h = 0, 1, 2, \dots, M-1$ 。

$$d[D_m(i), D_m(k)] = \max_{h=0,1,2,\dots,M-1} |S(i+h) - S(k+h)| (1 \leq i, k \leq M-m+1, i \neq k) \tag{10}$$

$B_m(r)$  和  $A_{m+1}(r)$  分别定义为  $m$  和  $m+1$  维重构向量中距离不超过预定参数  $r$  的向量对的数量。

$$SE(S, m, r) = -\ln(A_{m+1}(r)/B_m(r)) \tag{11}$$

其中  $m$  是重构维数, 本文中,  $r$  设置为标准差的 0.15 倍,  $m$  设置为 2。

**2.4.4. 排列熵(PE)**

PE 是一种对信号微弱变化具有放大效应的动力学突变检测方法, 它可以快速和精准地对信号的突变时刻做出响应, 是衡量信号复杂度的一个标准[18]。首先, 需要将输入信号  $S(n)$  重构为  $K$  行  $P$  列的矩阵  $C_{K \times P}$ 。其次, 对矩阵  $C_{K \times P}$  的每一行向量  $c$  都按照升序排列, 记录升序排列后的行向量  $v$  中元素在向量  $c$  中对应的位置序列  $ID$ , 例如向量  $c = (8, 7, 10, 3)$ , 升序排列后  $v = (3, 7, 8, 10)$ , 得到位置序列为  $ID = (4, 2, 1, 3)$ 。最后, 计算出每一种位置序列出现次数与  $P$  的全排列的比值  $\xi_t (1 \leq t \leq L)$ 。本文中,  $P$  被设置为 4, 则  $P$  的全排列为 24。

$$PE = -\sum_{t=1}^L \xi_t \ln(\xi_t) \tag{12}$$

**2.4.5. Hjorth 参数(HA, HC, HM)**

Hjorth [19]参数包括活动性(HA)、复杂性(HC)和移动性(HM)。通过直观的观察, 我们可以清楚地感

受到呼吸信号的波动强度、斜率以及斜率变化的幅度。

$$HA = \sigma^2 \tag{13}$$

$$HM = \sqrt{\frac{\sum_{n=1}^{M-1} (S'(n) - \overline{S'})^2 / (M-1)}{\sum_{n=1}^M (S(n) - ME)^2 / (M-1)}} \tag{14}$$

$$HC = \frac{\sqrt{\frac{\sum_{n=1}^{M-2} (S''(n) - \overline{S''})^2 / (M-2)}{\sum_{n=1}^{M-1} (S'(n) - \overline{S'})^2 / (M-1)}}}{\sqrt{\frac{\sum_{n=1}^{M-1} (S'(n) - \overline{S'})^2 / (M-1)}{\sum_{n=1}^M (S(n) - \overline{S})^2 / M}}} \tag{15}$$

其中,  $S'(n)$  表示  $S(n+1)(n=1,2,\dots,M-1)$  和  $S(n)(n=1,2,\dots,M-1)$  对应位置元素的差。 $\overline{S'}$  是  $S'(n)$  的平均值,  $ME$  是  $S(n)$  的平均值。同理,  $S''(n)$  表示  $S'(n+1)$  和  $S'(n)$  对应位置元素的差,  $\overline{S''}$  是  $S''(n)$  的平均值。 $HA$  实际上是输入信号的  $\sigma^2$ ,  $HM$  代表主导频率,  $HC$  代表信号带宽。因此,  $Hjorth$  参数既能反应输入信号在时域特征和也能反应频域特征[18] [19]。

### 2.5. 分类算法

本文中, 把睡眠阶段分为 5 类(AWA, S1, S2, DS, REM)和 4 类(S1 和 S2 合并为 LS)。我们将分割好的每 30 秒原始呼吸和分解得到的 6 个简单信号分别提取 9 个特征, 把特征组合成一行维数为 63 的向量, 并将特征向量导入 LSTM 分类模型中进行交叉验证训练和测试。特征提取后, 使用 5 倍交叉验证对 LSTM 分类器进行训练和测试。我们使用了两种输出方案, 一种是包含 AWA、S1、S2、DS 和 REM5 个类, 另一种是将 S1 和 S2 组合为 LS, 得到 4 个类。注意对于每个数据库, 本文中, 我们在每个睡眠阶段随机选择 85% 的呼吸信号特征来生成训练集, 其余 15% 作为测试集。

### 长短期记忆网络 LSTM

长短期记忆网络(Long Short-Term Memory networks, LSTM)是一种特殊的 RNN。在人工智能领域, 当我们要处理语言文字的时候, 我们会经常使用循环神经网络 RNN。它是一种拥有短期记忆的神经网络, 常用于处理序列数据[20]。RNN 可以用上下文搜索有用的信息, 然后做出判断。但通常, RNN 会出现比较严重的问题, 例如: 小明出生在中国, 并且在那里长大, 他的家乡在黄土高坡, 那里的风景很美丽, 所以小明说哪国语? 如果用 RNN 的方法, RNN 就要一直往前推, 一直推到最开头的部分, 发现小明是中国人。所以推断出他讲中文。这个现象就是“长依赖问题”。如果这个句子很长, RNN 可能把有用的信息舍弃掉, 从而判断失误。幸运的是科学家将 RNN 进行了改造升级, 提出了长短期记忆网络, 他们由

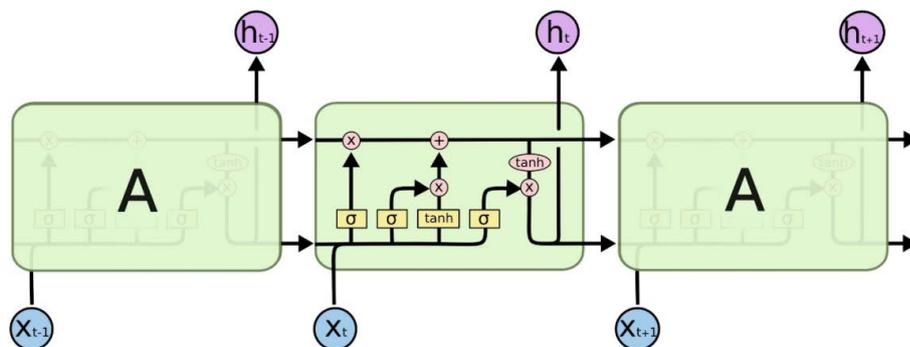


Figure 4. The repetition module in the LSTM contains four interacting layers  
 图 4. LSTM 中的重复模块包含四个相互作用的层

Hochreiter & Schmidhuber (1997)引入。LSTM 提出的目的就是为了解决长期依赖问题, 换句话说, 它可以记住更长时间段的重要信息。与 RNN 结构不同, LSTM 的神经网络层较为复杂。LSTM 有三个门结构, 分别为遗忘门, 输入门和输出门[21] (图 4)。

第一步:

LSTM 的第一步是决定我们要从这个过程中丢弃那些信息。这个决定是由一个“遗忘门作用层”的 sigmoid 层做出的, 这时起到了 LSTM 的第一步是决定我们要从细胞状态中丢弃哪些信息。它查看前一个  $h_{t-1}$  和当前输出  $x_t$ , 并为状态  $C_{t-1}$  中的每个数字中输出一个介于 0 和 1 之间的数字。1 表示“完全保留”, 而 0 表示“完全丢弃”这个公示表示了遗忘门的工作原理。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (16)$$

第二步:

LSTM 的第二步是决定我们要从这个过程中储存那些新信息。tanh 函数层可以从当前的向量中提取出有价值的信息, 而 sigmoid 函数层则可以根据需要调整这些信息, 使其达到最佳的状态。用一下公示来解释一下此过程。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (17)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (18)$$

第三步:

其实第三步是将第一步和第二步结合起来, 进行一个数据和信息更新的过程。我们上一步的状态  $C_{t-1}$  更新为新的状态  $C_t$ 。我们将上一步的状态乘以  $f_t$ , 忘记了我们之前决定忘记的信息。

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (19)$$

第四步:

在 LSTM 的第三步中, 我们需要确定我们的输出内容。为此, 我们需要创建一个 sigmoid 层, 它会根据我们的需求来确定哪些内容需要输出。接着, 我们可以使用 tanh 函数来计算单元格的状态, 并将其与 sigmoid 门的输出相比较, 从而确定最终的输出内容。

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (20)$$

$$h_t = o_t * \tanh(C_t) \quad (21)$$

## 2.6. 模型评价

LSTM 分类器的性能可以通过计算分类的准确率、特异性和灵敏度来评估。准确率是混淆矩阵(Q)主对角线元素的和与样本总数(SUM)的比值。

$$\text{accuracy} = \frac{\sum_{i=1}^A Q_{ii}}{\text{SUM}} \quad (22)$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (23)$$

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (24)$$

其中, True Positive (TP), 样本的真实类别是正类, 分类模型将其预测为正类。False Negative (FN)为样本的真实类别是正类, 分类模型将其预测为负类。False Positive (FP), 样本的真实类别是负类, 模型将其预

测为正类。True Negative (TN), 样本的真实类别是负类, 分类模型将其预测为负类。kappa 系数通常被用来衡量自动评分和传统手动评分之间的一致性。Landis 和 Koch 等人认为, kappa 数大于 0.80 的分类器具有接近完美的分类精度[22]。图 5 为 SHHS 数据库 4 分类的混淆矩阵, 对角线的值表示分类器对该类别预测正确的个数, 每一列纵轴表示这个类别真实的样本数, 例如从第一列可以得 W 期样本数为 2455。因此, 总样本数(SUM)为混淆矩阵 4 列样本数之和, SUM = 4556。

The SHHS database 4 classification confusion matrix

Confusion matrix		True value			
		W	LS	N3	R
Predictive value	W	2342	13	3	5
	LS	82	1064	57	109
	N3	4	86	295	1
	R	27	103	1	364

**Figure 5.** The SHHS database 4 classification confusion matrix  
**图 5.** SHHS 数据库 4 分类的混淆矩阵

以 W 期为例, 从混淆矩阵第一列可得 W 期共有 1000 个样本, 其中有 2342 个样本被分类模型预测为 W 期, 而有 82 个样本被预测为 LS 期, 4 个样本被预测为 N3 期, 27 个样本被预测为 R 期。因此, 对于 W 期这个类别 TP = 2342 (正确预测为 W 期的样本数), FN = 113 (W 期被预测为其它类别的样本数)。

从混淆矩阵第一行得出分类模型将一共将 2363 个样本预测为 W 期, 其中预测正确的有 2342 个样本, 而错误地将 LS 期、N3 期和 R 期预测为 W 期的共有 21 个样本。因此, 对于 W 期这个类别 FP = 21 (其它类别被错误预测为 W 期的样本数)。TN = 总样本数 - (TP + FN + FP), 对于 W 期这个类别 TN = 2078。通过公式(21)和(22)计算出 W 期的 specificity = 99.00%, sensitivity = 95.40%。本文中, 4 类和 5 类睡眠分期中每个类别的 specificity 和 sensitivity 计算结果在表 2 和表 3 中。

### 3. 结果

#### 3.1. 分类结果分析

为了提高模型的分类准确性和防止过拟合, 不仅需要对训练数据进行高效的处理, 还需要调试很多的模型参数(gamma、树深度、学习率、叶子数、列样本率、子样本率、训练次数)。本文中, 对 SHHS 数据库呼吸信号进行 5 类和 4 类睡眠阶段分类的准确性和 kappa 系数如表 4 所示。

从表 4 数据可以看出, 4 类分类器模型比 5 类分类器模型具有更好分类准确率。SHHS 数据库对应的 4 分类的准确率为 89.22%, kappa 系数为 0.8412, 且 kappa 系数都大于 0.8, 睡眠专家认为 kappa 数大于 0.80 的分类器具有接近完美的分类精度。

**Table 2.** Sensitivity and specificity of the four classifications  
**表 2.** 四分类的灵敏度和特异性

数据库		四分类			
		W	LS	N3	R
SHHS	特异性	99.00%	92.46%	97.83%	96.79%
	灵敏度	95.40%	84.04%	82.87%	75.99%

**Table 3.** Sensitivity and specificity of the five classifications  
**表 3.** 五分类的灵敏度和特异性

数据库		五分类				
		AWA	S1	S2	DS	REM
SHHS	特异性	99.01%	96.27%	96.06%	97.27%	97.84%
	灵敏度	94.50%	45.45%	84.43%	91.18%	69.55%

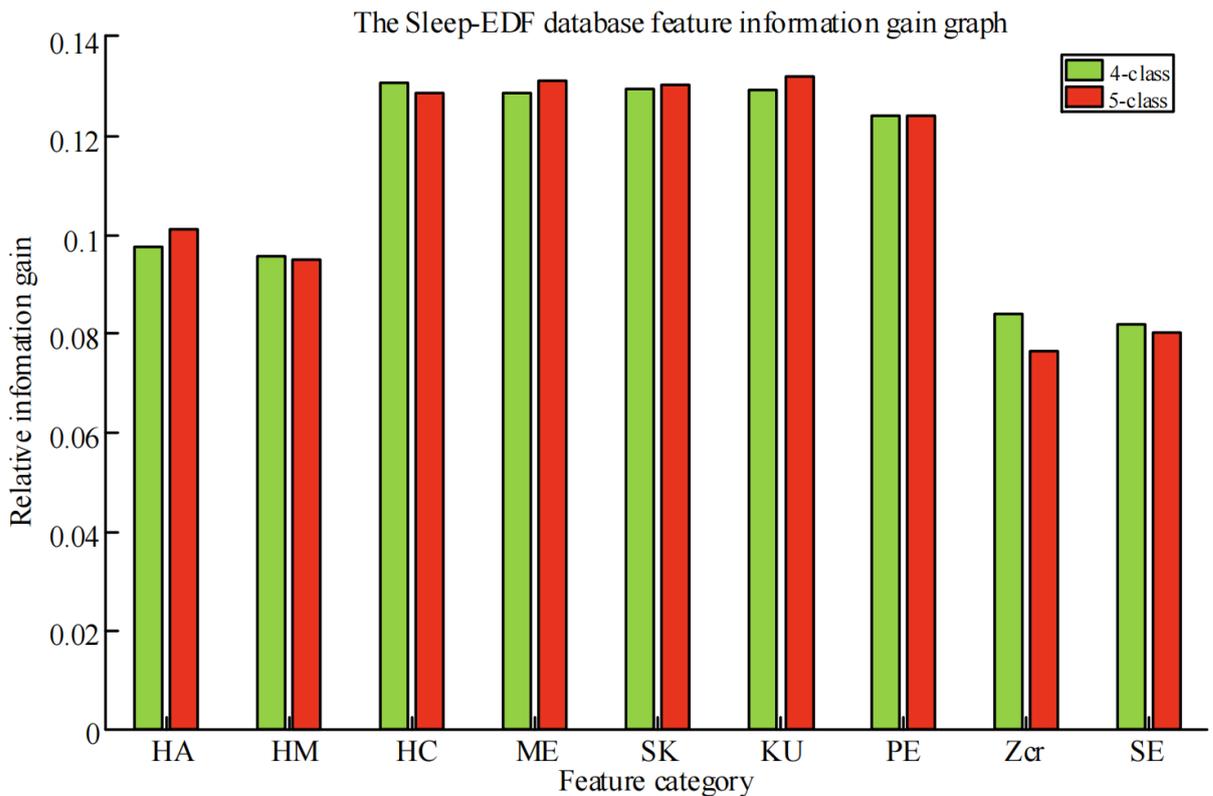
**Table 4.** SHHS database sleep classification model accuracy and kappa coefficient  
**表 4.** SHHS 数据库睡眠分类模型准确率和 kappa 系数

数据库	四分类		五分类	
	准确率	Kappa 系数	准确率	Kappa 系数
SHHS	89.22%	0.8412	88.43%	0.8238

通过表 2 和表 3 得出, 在进行 4 分类和 5 分类的实验中, 分类模型对每个睡眠阶段的识别和分类都是相当精准的, 且分类模型对清醒期的灵敏度和特异性较强。然而, 通过表 3 得出分类模型对 S1 睡眠阶段的分类灵敏度最低, 但对 AWA, S2, DS 睡眠阶段分类整体表现处于 80% 以上。因此, 实验结果进一步验证了本文分类模型的普遍适用性和稳健性。

### 3.2. 特征重要性分析结果

特征对睡眠阶段的分类是至关重要的。因此, 需要对特征的重要性进行计算和评估[23]。信息增益是



**Figure 6.** Relative information gain of Sleep-EDF database features in the LSTM  
**图 6.** LSTM 中 Sleep-EDF 数据库特征的相对信息增益

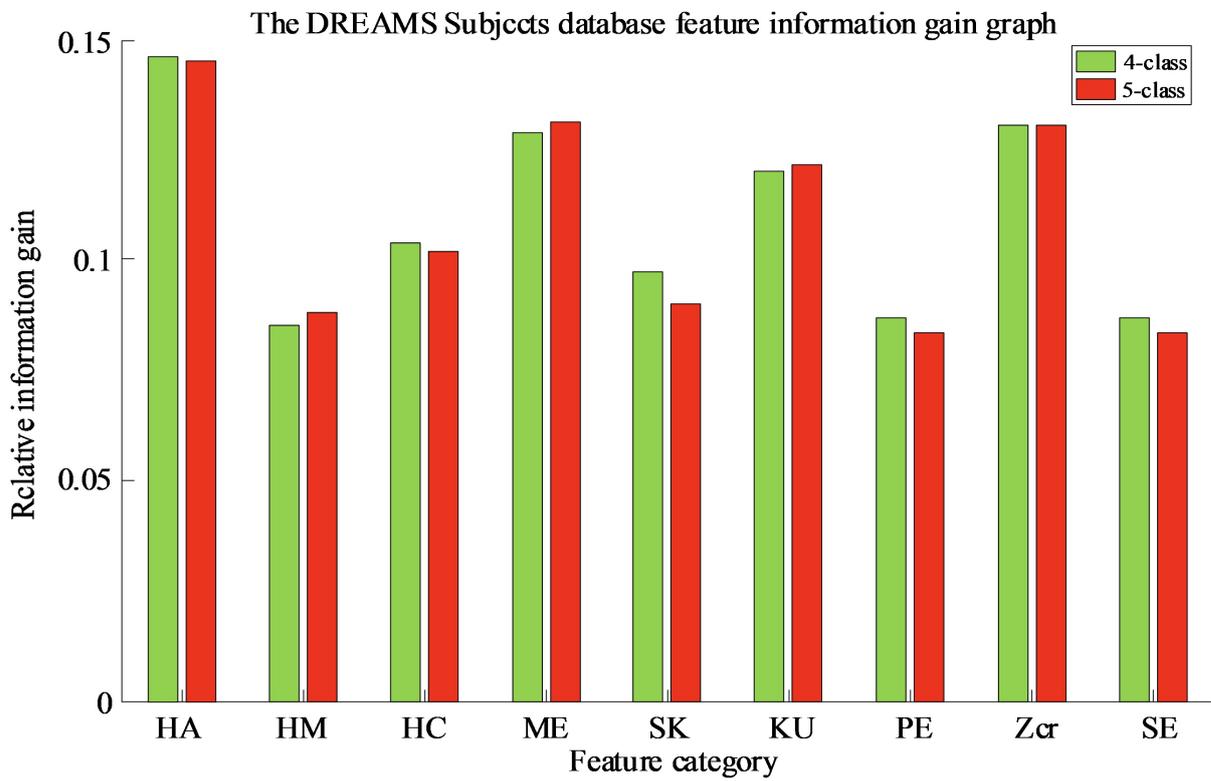


Figure 7. Relative information gain of DREAMS database features in the LSTM

图 7. LSTM 中 DREAMS 数据库特征的相对信息增益

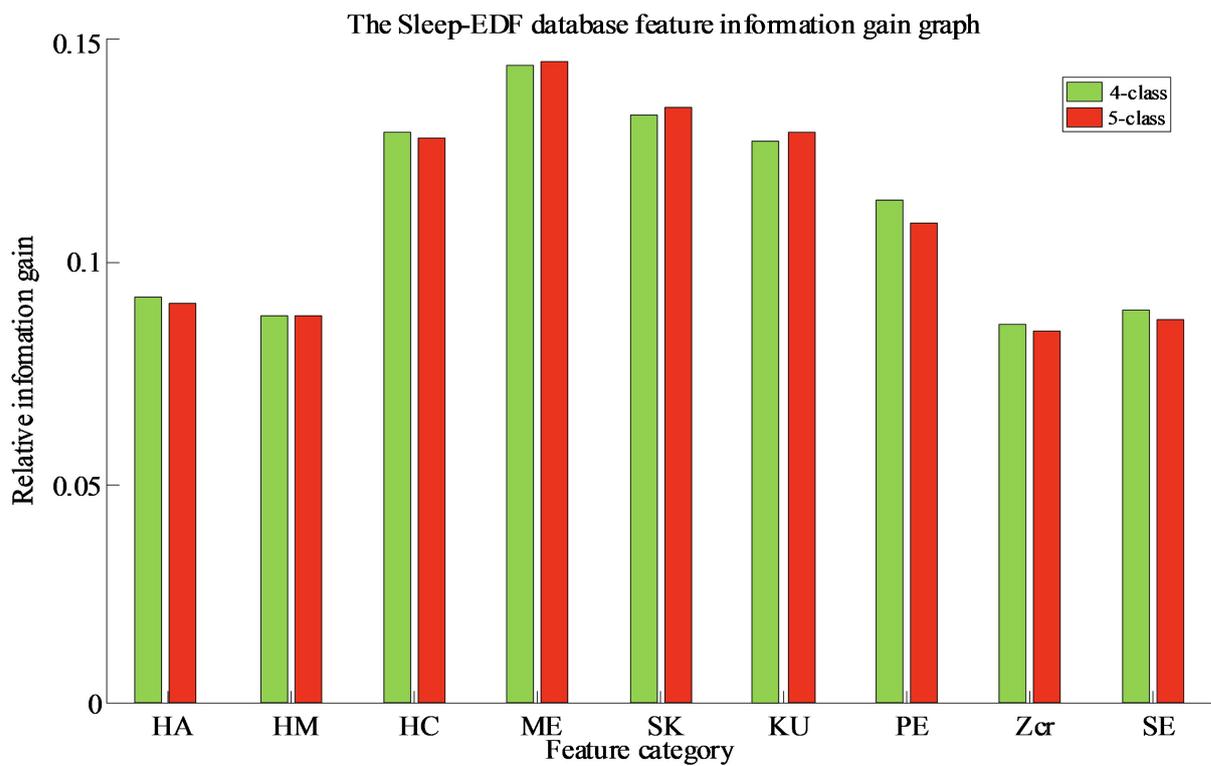


Figure 8. Relative information gain of SHHS database features in the LSTM

图 8. LSTM 中 SHHS 数据库特征的相对信息增益

评价特征重要性的最佳选择, 它以相对于总信息增益的比例来衡量, 比例越大, 特征在分类模型中的作用就越重要。接下来, 本文计算了 Sleep-EDF 数据库、DREAMS Subjects 数据库和 SHHS 数据库的呼吸信号特征(HA、HM、HC、Zcr、ME、SK、KU、SE、PE)的相对信息增益, 如图 6~8 所示。

从图 6~8 可以看出, 特征(HC, ME, SK, KU, PE)在 Sleep-EDF 数据库中具有较高的相对信息增益, 因此, 在分类模型中更为活跃。特征(HA, ME, Zcr, KU)在 DREAMS Subjects 数据库的分类模型中更为活跃, 而特征 Zcr 的相对信息增益最低, 适应能力较差。特征(HC, ME, SK, KU, PE)在 SHHS 数据库的分类模型中更为活跃。上述分析得出在三个数据库中表现最活跃的特征是(ME, KU), 其次是特性(SK, PE)。

#### 4. 讨论

呼吸信号的分解和特征提取是获得准确睡眠评分的关键。在本文中, 我们把 EMD 作为非线性和非平稳据的分析处理的一种选择, 它是一种新的处理非平稳信号的方法。大多数研究人员在没有对呼吸进行预处理(滤波、降噪)的情况下, 对 DREAMS Subjects 数据库中呼吸信号进行 4 类和 5 类睡眠分期的平均准确率分别为 83.39%和 80.06%, 相比于 Sleep-EDF 数据库和 SHHS 数据库, 准确率是最低的。本文采用 4 阶巴特沃斯高通滤波器和小波去噪技术对上述三个数据库中呼吸信号进行了预处理。我们对 DREAMS Subjects 数据库中呼吸信号进行 4 类和 5 类睡眠分期中, 准确率分别为 89.25%和 84.25%。实验表明, 呼吸信号的预处理至关重要。

ME 和 KU 主要反映信号的平均值和峰度, 图 4~6 显示 ME 和 KU 与睡眠分期高度相关。此外, SK、PE 和 HA 对模型分类有显著贡献。在本文中, 我们探索使用新组合的九个特征(KU, SK, HA, HM, HC, Zcr, SE 和 PE)进行分类, 这在很大程度上提高了模型的评分性能。在 Hassan 等人之前的工作中, 分类器中只使用了脑电信号的 4 个统计特征[24], 即 ME、HA、SK 和 KU, 然而分类器的分类准确率平均下降了 4.24%。多数研究方法在随着数据库变化和实验数据量的增加, 准确率没有升高, 反而有明显降低。然而, 我们的研究方法随之改变, 平均准确率保持在 90%左右。更加验证了本文研究方法的可靠性和普遍性。我们还了解到, 利用 xgboost, catboost 或者 cfc 算法有可能对利用呼吸信号进行睡眠分类的准确率更高, 希望以后在算法上更加优化。

本文中, 实验的处理器为 Intel(R) Core(TM) i7-8650U CPU @ 1.90 GHz 2.11 GHz CPU 16 GB, 并使用了 Matlab 和 Python 两个工具来完成实验。我们对一个 30 秒呼吸信号的分解和分类进行了耗时计算, 实验过程中 NSP 算法分解耗时 0.101 秒, 睡眠阶段分类耗时 0.1 秒。因此, 对呼吸信号分解的高计算效率和睡眠评分的高稳定性为可穿戴设备的开发带来了良好的前景。

#### 参考文献

- [1] Younes, M. (2017) The Case for Using Digital EEG Analysis in Clinical Sleep Medicine. *Sleep Science and Practice*, 1, Article No. 2. <https://doi.org/10.1186/s41606-016-0005-0>
- [2] Zhang, Y., et al. (2019) Sleep Stage Classification Using Bidirectional LSTM in Wearable Multi-Sensor Systems. *IEEE Conference on Computer Communications Workshops*, Paris, 29 April-2 May 2019, 443-448. <https://doi.org/10.1109/INFCOMW.2019.8845115>
- [3] Yang, Z., Pathak, P.H., Zeng, Y., Liran, X. and Mohapatra, P. (2017) Vital Sign and Sleep Monitoring Using Millimeter Wave. *MobiHoc '16: Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, Chennai, 10-14 July 2017, 211-220. <https://doi.org/10.1145/2942358.2942381>
- [4] Zhao, D., et al. (2019) Comparative Analysis of Different Characteristics of Automatic Sleep Stages. *Computer Methods and Programs in Biomedicine*, 175, 53-72. <https://doi.org/10.1016/j.cmpb.2019.04.004>
- [5] Berry, R.B., Brooks, R., Gamaldo, C.E., et al. (2012) The AASM Manual for the Scoring of Sleep and Associated Events. Rules, Terminology and Technical Specification. American Academy of Sleep Medicine, Darien.
- [6] Wolpert, E.A. (1969) A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of

- Human Subjects. *Archives of General Psychiatry*, **20**, 246-247.  
<https://doi.org/10.1001/archpsyc.1969.01740140118016>
- [7] Rechtschaffen, A. (1968) A Manual of Standardized Terminology, Technique and Scoring Systems for Sleep Stages of Human Subjects. *Journal of the National Institute of Health*, **1**, 246-247.
- [8] Alickovic, E. and Subasi, A. (2018) Ensemble SVM Method for Automatic Sleep Stage Classification. *IEEE Transactions on Instrumentation and Measurement*, **67**, 1258-1265. <https://doi.org/10.1109/TIM.2018.2799059>
- [9] Punjabi, N.M., et al. (2015) Computer-Assisted Automated Scoring of Polysomnograms Using the Somnolyzer System. *Sleep*, **38**, 1555-1566. <https://doi.org/10.5665/sleep.5046>
- [10] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, E.H., Zheng, Q., Tung, C.C. and Liu, H.H. (1998) The Empirical Mode Decomposition Method and the Hilbert Spectrum for Non-Stationary Time Series Analysis. *Proceedings of the Royal Society of London. Series A*, **454**, 903-995. <https://doi.org/10.1098/rspa.1998.0193>
- [11] Wu, Z.H. and Huang, N.E. (2009) Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Advances in Data Science and Adaptive Analysis*, **1**, 1-41. <https://doi.org/10.1142/S1793536909000047>
- [12] Shen, Y., Zhang, Y.C. and Wang, Z.H. (2011) Satellite Fault Diagnosis Method Based on Predictive Filter and Empirical Mode Decomposition. *Journal of System Engineering and Electronics*, **11**, 84-85.
- [13] Quan, P. and Lei, Z. (1999) Two Denoising Methods by Wavelet Transform. *IEEE Transactions on Signal Processing*, **47**, 3401-3406. <https://doi.org/10.1109/78.806084>
- [14] Caesarendra, W., Kosasih, P.B., Tieu, A.K., et al. (2013) Condition Monitoring of Naturally Damaged Slow Speed Slewing Bearing Based on Ensemble Empirical Mode Decomposition. *Journal of Mechanical Science and Technology*, **27**, 2254-2255. <https://doi.org/10.1007/s12206-013-0608-7>
- [15] Kai, F., Qu, J., Yi, C., et al. (2014) Classification of Seizure Based on the Time Frequency Image of EEG Signals Using HHT and SVM. *Biomedical Signal Processing & Control*, **13**, 15-22. <https://doi.org/10.1016/j.bspc.2014.03.007>
- [16] Skowronek, J. and McKinney, M. (2006) Features for Audio Classification: Percussiveness of Sounds. *Intelligent Algorithms in Ambient and Biomedical Computing*, **7**, 103-118. [https://doi.org/10.1007/1-4020-4995-1\\_7](https://doi.org/10.1007/1-4020-4995-1_7)
- [17] Richman, J.S. and Randall, M.J. (2000) Physiological Time-Series Analysis Using Approximate Entropy and Sample Entropy. *The American Journal of Physiology-Heart and Circulatory Physiology*, **278**, H2039-H2049. <https://doi.org/10.1152/ajpheart.2000.278.6.H2039>
- [18] Kim, D.J., et al. (2013) Disturbed Resting State EEG Synchronization in Bipolar Disorder: A Graph-Theoretic Analysis. *NeuroImage: Clinical*, **2**, 414-423. <https://doi.org/10.1016/j.nicl.2013.03.007>
- [19] Cecchin, T., et al. (2010) Seizure Lateralization in Scalp EEG Using Hjorth Parameters. *Clinical Neurophysiology*, **121**, 290-300. <https://doi.org/10.1016/j.clinph.2009.10.033>
- [20] Schmidhuber, J. (2015) Deep Learning in Neural Networks: An Overview. *Neural Networks*, **61**, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [21] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [22] Koch, L.G.G. (1977) A One-Way Components of Variance Model for Categorical Data. *Biometrics*, **33**, 671-679. <https://doi.org/10.2307/2529465>
- [23] Koprinska, I. (2007) Feature Selection for Brain-Computer Interfaces. Springer, Berlin.
- [24] Quan, S.F., et al. (1997) The Sleep Heart Health Study: Design, Rationale, and Methods. *Sleep*, **20**, 1077-1085.