计 算 机 研 究 与 发 展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202440389
61(11): 2739−2753, 2024

基于多任务注意力网络的非接触式睡眠监测

李思恒[1,2,3] 金蓓弘[1,2,3] 张扶桑[1,2,3] 王 志[1,2,3] 马俊麒[1,2,3] 苏 畅[1,2,3] 任晓勇[4] 刘海琴[4]

[1](中国科学院软件研究所软件工程技术中心 北京 100190)

[2](基础软件与系统重点实验室(中国科学院);计算机科学国家重点实验室(中国科学院软件研究所) 北京 100190)

[3](中国科学院大学 北京 100049)

[4](西安交通大学第二附属医院 西安 710000)

(lisiheng19@otcaix.iscas.ac.cn)

Multi-Task Attention Network for Contactless Sleep Monitoring

Li Siheng[1,2,3], Jin Beihong[1,2,3], Zhang Fusang[1,2,3], Wang Zhi[1,2,3], Ma Junqi[1,2,3], Su Chang[1,2,3], Ren Xiaoyong[4], and Liu Haiqin[4]

[1](*Technology Center of Software Engineering*, *Institute of Software*, *Chinese Academy of Sciences*, *Beijing* 100190)

[2](*Key Laboratory of System Software*(*Chinese Academy of Sciences*)*and State Key Laboratory of Computer Science*(*Institute of Software*, *Chinese Academy of Sciences*), *Beijing* 100190)

[3](*University of Chinese Academy of Sciences*, *Beijing* 100049)

[4](*The Second Affiliated Hospital of Xi'an Jiaotong University*, *Xi'an* 710000)

Abstract Sleep takes up nearly one-third of a person's day and is closely related to human health. Since the durations and transitions over different sleep stages during sleep directly affect a person's sleep quality, identifying sleep stages has become the most basic and important task in sleep monitoring. However, sleep disorders occurring in sleep can lead to complex sleep structures, thereby increasing the difficulty of classifying sleep stages. Most of the existing contactless solutions for sleep stage classification lack a sufficient understanding of the complexity in sleep structure, ignoring the relationship between sleep stage and sleep disorder. Therefore, these solutions fail to achieve great performance in patients with sleep disorders. In this paper, we propose a sleep monitoring system that focuses on predicting sleep stages from ultra-wideband (UWB) signals. We design a sequence prediction model that combines an attention-based sequence encoder and a contrastive learning module to extract the temporal progression of sleep and improve the generalizability of the encoder. Particularly, the key to our approach is a multi-task fine-tuning strategy that incorporates sleep disorder information into sleep staging to reduce the interference of sleep disorders with sleep stage prediction. We conduct extensive experiments on 110 subjects, including healthy individuals and patients with different severities of sleep disorders. The experimental results demonstrate that the performance of our model is superior to the baseline methods.

Key words contactless sensing; ultra-wideband (UWB); deep neural networks; multi-task learning; sleep stage classification; sleep disorder

摘 要 睡眠几乎占据了一个人每天三分之一的时间,它与人体的健康状况紧密相关.由于睡眠过程中各睡眠阶段的持续时间和转换情况直接影响人的睡眠质量,因此识别睡眠阶段是睡眠监测最基本和最重要

收稿日期:2024-05-31;**修回日期**:2024-08-21

基金项目:国家自然科学基金项目(62072450, 62172394);中国科学院软件研究所基础研究项目(ISCAS-JCMS-202305)

This work was supported by the National Natural Science Foundation of China (62072450, 62172394) and the Basic Research Program of ISCAS (ISCAS-JCMS-202305).

通信作者:金蓓弘(beihong@iscas.ac.cn)

的任务. 然而, 睡眠中出现的睡眠障碍会导致睡眠结构变得复杂, 这增加了睡眠阶段分类的难度. 已有的非接触式睡眠阶段分类工作大多对睡眠结构的复杂性认识不足, 忽视了睡眠阶段和睡眠障碍之间的联系. 因此, 这些工作难以在睡眠障碍患者上取得较好的性能. 提出一种非接触式睡眠监测系统, 利用超宽带(ultra-wideband, UWB)信号来识别人体睡眠阶段的变化情况. 该系统包含了一个序列预测模型, 使用一个基于注意力机制的序列编码器挖掘不同睡眠阶段之间的时序转换关系, 并通过一个对比学习模块提高编码器的泛化性. 值得一提的是, 该序列预测模型采用了一个基于多任务学习的两阶段训练框架, 并在模型的微调阶段通过多专家学习模块将睡眠障碍信息融入模型中, 从而降低了睡眠障碍对睡眠阶段预测造成的干扰. 在 110 名受试者(包括健康个体和不同程度睡眠障碍患者)中进行实验评估, 实验结果表明所提出的模型的性能优于基线方法.

关键词 非接触式感知; 超宽带; 深度神经网络; 多任务学习; 睡眠分期; 睡眠障碍

中图法分类号 TP391

睡眠是人体的一个基本生理过程, 它对人体的身心健康有重要的影响. 高质量的睡眠不仅有助于缓解疲劳, 促进身体的生长和恢复, 还与情绪的稳定性、免疫系统的功能以及心血管健康等方面密切相关[1-2]. 近年来, 越来越多的人受到如慢性失眠症、嗜睡症、睡眠呼吸暂停综合征等睡眠障碍的困扰, 这些睡眠障碍使得人们在夜间无法获得充足的睡眠, 并可能进一步成为如高血压、脑卒中、肥胖症和一些心脏疾病的诱因[3-4]. 根据《2024 中国居民睡眠健康白皮书》[5] 显示, 约 28% 的人夜间睡眠时长不超过 6 h, 而约 64% 的人睡眠质量欠佳. 因此, 居家睡眠监测对了解和改善人体健康有着重要的意义. 监测人体整夜的睡眠结构, 不仅能够对人体一些潜在的睡眠问题进行提前预警, 提醒用户及时就医, 而且能对患者进行诊后随访, 获得患者诊后的睡眠情况.

美国国家卫生研究院的睡眠专家 Marishka Brown 将一晚高质量的睡眠定义为: 充足的总睡眠时长和合理的睡眠阶段分布[6]. 因此, 对人体整夜睡眠进行阶段分类是评估睡眠质量的主要手段. 美国睡眠医学学会根据人体不同的生理特性(包括呼吸、心跳、体动等)和功能, 将人体整夜的睡眠划分为了 5 个阶段: 清醒期(wake)、非快速眼动 1 期(non-rapid eye movement 1, N1)、2 期(N2)、3 期(N3)和快速眼动期(rapid eye movement, REM)[7].

在医疗领域, 多导睡眠图(polysomnography, PSG)监测被称为睡眠监测的“金标准”. 通过在被测者身上佩戴多个专用设备, 可以对其脑电图、肌电图、胸式和腹式呼吸张力图、鼻及口通气量、血氧饱和度等信息进行测量. 然后, 这些 PSG 数据被划分为多个 30 s 长的睡眠时段(sleep epoch), 由专业医疗人员根据 PSG 数据的变化情况, 为每个睡眠时段标注一个睡眠阶段(sleep stage). 使用 PSG 进行睡眠监测可以获得高精度的与睡眠相关的生理参数, 然而, 这一过程需要专业的医护人员进行操作, 费时费力. 并且, 一次 PSG 睡眠监测的花费往往较为昂贵. 因此, 基于 PSG 的睡眠监测目前仅局限于医院场景中.

当前, 一些可穿戴设备已被用于居家环境中的睡眠监测[8-9]. 然而, 这些设备同样要求被测者整夜佩戴, 而长时间的设备佩戴会给被测者带来接触不适等问题, 并进一步影响睡眠质量. 同时, 这些方法也面临着设备佩戴不规范带来的采集到的数据不可用的问题. 此外, 已有工作大多使用公开数据集(这些数据集中的数据往往质量较好)进行性能测试, 而非将其应用于真实的睡眠场景中. 因此, 这些方法在实用性方面有待考究.

为了解决接触式感知带来的困扰, 非接触式感知技术逐渐走进了人们的视野. 非接触式感知技术通过分析人体反射的无线信号来实现对人体的感知, 它无需人体佩戴任何设备, 具有低成本、非侵扰和感知范围广等优势, 有望成为居家日常睡眠监测的有效解决方案. 然而, 利用非接触式感知进行睡眠监测面临 3 个挑战:

1)如何从整夜收集到的超长信号序列中挖掘出信号间的时序依赖关系, 并映射到不同的睡眠阶段. 已有工作或依赖基于手工特征的传统机器学习方法, 或采用包括卷积神经网络(convolutional neural network, CNN)和循环神经网络(recurrent neural network, RNN)在内的深度学习模型. 其中, CNN 通常只能捕获局部特征, 难以建立全局的特征依赖关系[10]; 而 RNN 可以通过逐时间步计算隐状态的方式来对序列元素间的长距离依赖进行建模, 但这使其面临梯度消失和梯度爆炸的问题[11], 并且其递推计算的方式导致计算

效率低下. 因此, 这些方法都无法很好地捕获信号间的长距离依赖.

2)如何同时提取健康人群和睡眠障碍人群的睡眠特征. 与健康个体相比, 睡眠障碍患者的睡眠结构更加复杂和不规律, 并且在不同睡眠阶段展现出与健康个体不同的生理特性. 例如, 在睡眠呼吸暂停低通气综合症(sleep apnea-hypopnea syndrome, SAHS)患者的整夜睡眠中, 频繁出现的呼吸暂停事件使其生命体征(如呼吸)变得紊乱, 并扰乱正常的睡眠阶段转换模式, 使得人体频繁转醒[12]. 因此, 这类人群往往对睡眠监测有着更大的需求, 也是医院中进行睡眠监测的主要对象. 然而, 已有工作忽视了睡眠阶段和睡眠障碍之间的关系, 使得这些关系没有得到挖掘和利用, 而且这些工作多数采用仅包含健康个体的数据集进行训练[13-14], 所以往往能够在健康个体上取得不错的性能, 但在睡眠障碍患者上的性能则不如人意[15].

3)如何增强睡眠阶段分类模型的泛化性. 睡眠阶段分类工作的目标是该工作能应用于未知的睡眠场地中(如住宅卧室), 从而实现日常监测. 然而, 已有工作多在受控的环境中进行实验, 使得训练出的模型对于环境的变化较为敏感. 此外, 这些工作大多仅在规模较小的数据集上进行性能评估[16-17], 其泛化性有待考究.

针对上述3种挑战, 本文结合超宽带(ultra-wideband, UWB)信号处理和深度学习模型, 提出了一个非接触式睡眠监测系统. 首先, 采用基于注意力机制的序列到序列(sequence to sequence, seq2seq)编码器来学习各睡眠阶段的特征和不同睡眠阶段之间的转换模式, 并对不同睡眠阶段之间的长、短距离依赖关系进行建模. 其次, 提出一种基于多任务学习的两阶段训练框架, 用于提取健康个体和睡眠障碍患者的不同睡眠特性. 具体地, 先将模型在包含健康个体的公共数据集上进行预训练, 建模正常睡眠的结构和转换模式. 然后, 结合多任务学习技术, 将睡眠呼吸暂停事件识别作为辅助任务, 将预训练好的模型在包含不同程度的SAHS患者的数据集上进行微调. 这样的做法将睡眠呼吸暂停信息融入了模型中, 可对基于健康数据训练出的模型做出的睡眠阶段预测进行调整, 从而学习到因睡眠呼吸暂停事件频繁出现所导致的睡眠阶段异常转换. 最后, 为了提升模型的泛化性, 采取以下措施: 利用UWB信号距离分辨率高的特征, 设计床定位的算法聚焦床内信号的变化, 从而排除床以外环境噪声的干扰. 同时, 采用对比学习策略, 通过主动引入噪声的方式来增强模型的降噪能力.

本文的主要贡献包括4个方面:

1)对原始的UWB信号进行分析, 提出一系列信号处理方法, 去除环境噪声的干扰, 实现床定位, 并提取到来自被测个体的信息.

2)设计了一个seq2seq的深度学习模型来进行睡眠阶段分类. 该模型使用CNN来提取信号的特征, 使用基于注意力机制的序列编码器来提取信号间的时序依赖关系, 并结合对比学习策略来提高模型的泛化能力.

3)提出一种基于多任务学习的两阶段训练框架, 通过多任务联合训练的方式将睡眠呼吸暂停信息融入模型, 建模睡眠阶段和睡眠障碍间的关系, 从而减轻睡眠障碍对睡眠阶段分类性能的影响.

4)本文研发的系统在110名受试者(包括健康个体和不同程度的SAHS患者)中进行了实验评估, 结果表明该系统在4项分类任务(即2~5个阶段的睡眠分期)中取得了出色的性能, 优于其他基准方法.

1 相关工作

已有的睡眠阶段分类工作可以被分为3类: 基于PSG、基于穿戴式设备和基于非接触式感知的工作.

1.1 基于PSG的睡眠分期

PSG监测是睡眠监测领域公认的金标准. 通过多种专业仪器对受试者身体各部位整夜的活动情况进行记录, 再经由专业医护人员根据脑电图、眼动图和肌电图等信号的波动情况对人体整夜的睡眠阶段进行标注. 由于这一过程费时费力, 已有工作利用机器学习模型实现对PSG信号的睡眠阶段自动分类. 例如, Jia等人[18]借助图卷积网络来学习不同通道的脑电图信号之间的内在联系, 从而实现了5个阶段睡眠分类; Biswal等人[19]利用脑电图、眼电图和肌电图等信号, 通过融合CNN和RNN实现了5阶段睡眠分类; Jia等人[20]利用脉冲神经网络来捕捉脑电图信号中的动态波动特征, 在2个大型脑电图数据集上实现了睡眠阶段的5分类; Chen等人[21]提出了一种基于多源数据融合的神经网络, 从包括脑电图、眼电图、肌电图和心电图在内的多通道PSG数据中, 利用时序卷积和图卷积提取时空特征, 并使用门控循环单元(gated recurrent unit, GRU)提取频域特征, 最后实现睡眠阶段的5分类. 当前, 已有基于脑电图的商

用设备被投入使用. 例如, Sleep Profiler[22] 是一款头戴式设备, 它通过收集脑电图、肌电图等信息来监测用户的睡眠情况. 然而, PSG 设备繁杂、昂贵等特点使其无法被应用于居家环境中, 而已有的商用设备(如 Sleep Profiler)与人体的接触程度尽管比 PSG 设备少, 但仍然会给人体带来较大的不适感, 并且会扰乱人体正常的睡眠习惯.

1.2 基于穿戴式设备的睡眠分期

包括智能手环、手表在内的可穿戴设备通过监测人的身体活动来实现对睡眠的监测. 例如, Chang 等人[23] 利用加速度计、麦克风、方向传感器和光传感器等智能手表内置的传感器收集数据, 从中识别出多种睡眠事件(包括睡姿、体动)、声音事件和照明条件, 并借助隐马尔可夫模型将人体的整夜睡眠分为 REM 睡眠、浅睡眠(N1+N2)和深睡眠(N3); Han 等人[24] 通过在睡眠耳塞中内置麦克风来感知人体睡眠过程中各种睡眠事件发出的声音, 从中提取声音特征, 并结合 CNN 和双向长短期记忆(long short-term memory, LSTM)网络进行睡眠分期(REM 睡眠、浅睡眠和深睡眠). Zhai 等人[25] 从公开数据集的信号数据(腕动计采集的体动和心跳数据)中提取时域、频域、非线性域、几何域等多类特征, 并使用 CNN 和 LSTM 实现了睡眠分期; Zhai 等人[8] 采用简单拼接、注意力加权、张量法等多种数据融合手段, 对从公开(腕动计)数据集上计算的体动和心跳特征进行融合, 最终实现睡眠阶段 3 分类; Huang 等人[9] 利用注意力机制从加速度计采集的 3 维数据中自动提取特征, 在 2 个公共数据集上取得了不错的结果. 而当前市面上也存在众多包含睡眠监测功能的商用设备, 例如智能手环和智能手表等, 但其准确性还有待提高[26]. 长期佩戴穿戴式设备入睡同样给人体带来不适, 且文献 [27] 表明部分人对佩戴设备入睡存有抵触情绪. 同时, 接触式感知存在着因设备佩戴不规范带来的数据不可用问题. 此外, 已有工作大多仅在公开数据集上进行性能评估, 而非应用于真实的睡眠场景中, 其实用性有待考究.

1.3 基于非接触式感知的睡眠分期

已有的非接触式睡眠分期方法可以被划分为 2 类: 基于传统机器学习的方法和基于深度学习的方法. 例如, Gu 等人[28] 利用智能手机内置的传感器来捕捉人睡眠中的体动和发出的声音, 并结合年龄等个人特征, 通过条件随机场(conditional random field, CRF)模型实现了 3 阶段睡眠分期; Zhang 等人[16] 首先从无线信号中提取运动相关的特征来识别清醒/睡眠阶段, 然后在被判定为睡眠的时段中提取呼吸相关的特征, 利用支持向量机(support vector machine, SVM)进行 REM/NREM 阶段的分类; Rahman 等人[17] 使用多普勒连续波雷达来监测人体整夜的呼吸、心跳和体动, 并基于这 3 类手工特征, 结合随机森林分类器实现了睡眠阶段 3 分类. 部分工作[16,28] 为了提高准确率而采用了粗粒度的睡眠时段(如 5 min), 这与金标准要求的睡眠时段(30 s)不一致. 总体而言, 基于传统机器学习的方法需要人工设计特征, 而特征的设计和选择往往依赖于该领域专家的知识和经验. 同时, 手工设计的特征不够通用和灵活, 使得模型的性能也相对受限.

相比而言, 基于深度学习的方法使用端到端神经网络自动从数据中学习特征, 从而实现睡眠阶段分类. 例如, Yu 等人[29] 提出了一个基于 Wi-Fi 信号的睡眠分类系统, 他们结合 CNN-LSTM 模型和关于睡眠阶段的先验医学知识, 在睡眠阶段 4 分类上取得了不错的结果; Zhao 等人[13] 利用 CNN 和 RNN 从无线信号中提取特征从而实现睡眠阶段 4 分类, 同时, 他们结合对抗训练的机制来去除数据中与个人和环境相关的信息, 从而增强模型的泛化性; Zhai 等人[14] 首先在公共数据集上预训练他们的 CNN-RNN 模型, 然后在用生物医学雷达收集的数据集上进行微调, 从而实现睡眠分期. 总体而言, 基于深度学习的方法取得了比基于传统机器学习方法更为优秀的性能. 然而, 已有工作大多仅在规模较小的健康人数据集上进行实验, 并且多数实验在受控环境中(如睡眠实验室)而非日常场景下进行. 同时, 一些与睡眠阶段相关的睡眠信息(如睡眠呼吸暂停)也没有得到挖掘和利用.

与现有工作相比, 本文提出的系统结合了 UWB 信号处理技术和深度学习模型, 并将睡眠呼吸暂停信息融入模型中, 从而实现准确的睡眠分期.

2 系统设计

本节首先展示系统的概览, 然后对 UWB 信号的感知方式进行介绍, 最后对组成系统的各个模块进行详细地描述.

2.1 系统概览

系统的整体框架如图 1 所示, 由 3 个主要的模块组成: 信号抽取、信号预处理和睡眠阶段分类.

1)信号抽取模块. 该模块负责背景消减、床定位和人体存在性检测. 首先对从 UWB 设备中接收到的

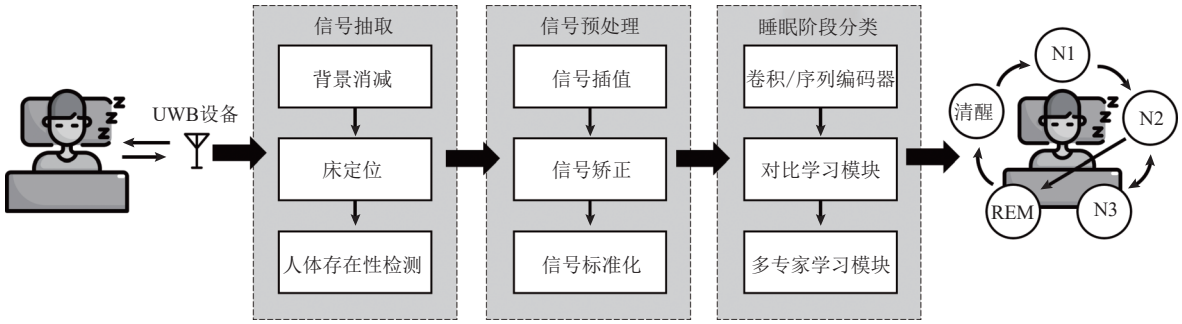


Fig. 1 System framework

图 1 系统框架

信号数据使用背景消减方法，去除环境中静态物体的反射对信号造成的影响. 然后，基于信号的振幅信息对环境中的人体进行定位，并结合人体的历史位置信息和床的宽度信息对床的范围进行定位. 最后，判断该人体是否在床的范围内，并对在床范围内的人体进行信号抽取.

2)信号预处理模块. 该模块对抽取出的代表人体的信号数据进行插值、矫正和标准化等信号预处理操作，从而减少噪声的干扰，提高信号的质量.

3)睡眠阶段分类模块. 该模块构建了一个包含 CNN 和序列编码器的深度学习模型来从 UWB 信号中挖掘特征，最终实现序列预测(即整夜睡眠阶段预测). 该深度学习模型还使用了一个对比学习策略来提升模型的降噪能力. 在此基础上，设计了一个基于多任务学习的两阶段训练框架来将与睡眠阶段相关的睡眠呼吸暂停事件识别作为辅助任务融入模型，使得模型能够学习到更加有效的嵌入表示.

2.2 UWB 信号感知方式

UWB 信号从发射到接收的基本流程如图 2 所示. UWB 设备是收发一体的(即发射端和接收端集成在同一块芯片上)，其发射端生成高斯脉冲信号 $p(t)$，并使用频率为 f_c 的载波对其进行调制，得到发射信号 $x(t)=p(t)\cos(2\pi f_c t)$. UWB 设备以一定的间隔(即脉冲重复间隔)连续发射脉冲信号，并收集从环境中人体或物体反射回来的信号，接收到的信号的时延为 $t=2r/c$，其中 r 代表设备与反射物之间的距离，c 代表信号在空气中传播的速度. 换言之，接收到的信号中包含了被测对象的信息，包括其生命体征和身体活动. 在设备的接收端，对接收到的信号 $y(t)$ 进行解调制(包括同相/正交采样和低通滤波)，得到基带信号.

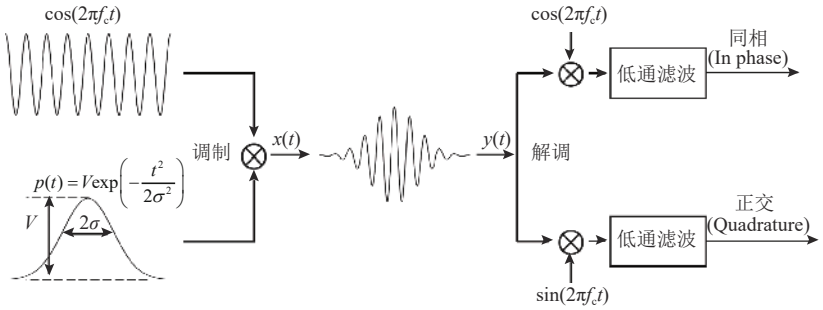


Fig. 2 Transmission and reception of UWB signals

图 2 UWB 信号的发射与接收

该基带信号(一个二维复数矩阵)可表示为如图 3 所示的形式，其中，每一行被称为快时间维度，它包含了在一个脉冲重复间隔期间所接收到的反射回来的脉冲响应，不同脉冲响应的时间延迟对应着设备与反射脉冲信号的对象之间的距离，即横轴表示距离；而每一列被称为慢时间维度，它表示每个脉冲发射的时间，即纵轴表示时间. 因此，定位到被测个体所在的快时间坐标后，在该快时间坐标上，沿着慢时间坐标轴取出对应的一维向量，该向量即可表示被测个体的活动信息(包括生命体征、体动情况等).

2.3 信号抽取

通常，UWB 设备被放置在床头柜上，它发射并接收 UWB 信号. 接收到的信号被转换成同相/正交域下的 I/Q 信号，输出到一个复数矩阵. 该二维复数矩阵(一维代表距离，一维代表时间)包含了监测范围

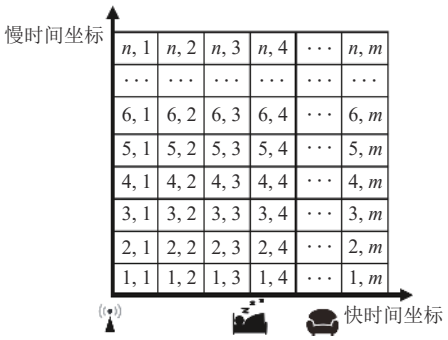


Fig. 3 The matrix of received baseband signals

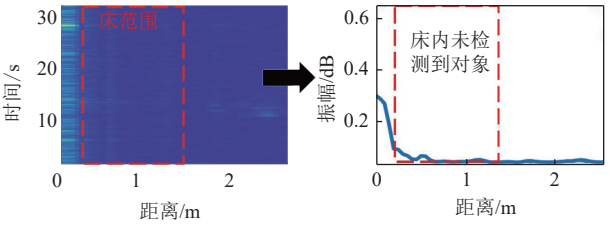
图 3 接收的基带信号矩阵

内所有物体的信息. 因此, 需要从接收到的包含环境信息的 UWB 信号矩阵中抽取出代表被测个体的信号数据. 信号抽取主要分为 3 步: 背景消减、床定位和人体存在性检测.

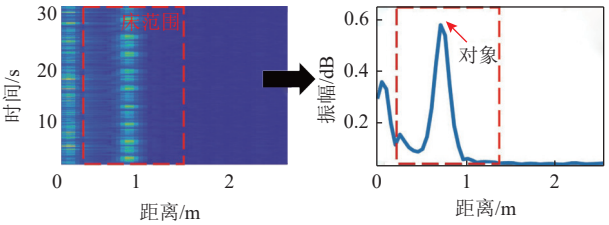
1)背景消减. 首先, 采用 30 s 的滑动窗口, 以 30 s 的滑动步长对整夜的信号数据进行分割; 然后对于每个窗口的信号, 使用背景消减算法[30]来消除原始接收信号中由静态物体反射造成的杂波的影响.

2)床定位. 对被测个体所处的床的范围进行定位, 从而隔绝床以外信号的影响. 在整夜的监测过程中, 人体在大多数时间内都应该在床范围内并处于静止的状态. 因此, 可以结合人体处于静止状态时的历史位置信息和床的宽度信息对床的范围进行定位. 在大部分情况下, 当人体处在监测范围内时, 其所在位置对应的信号强度(即信号的振幅)会明显大于周围环境中物体的信号强度. 因此, 将每个窗口信号的振幅信息按时间进行叠加, 最大的信号峰值(peak)所在的位置(即 30 s 内信号强度最大的位置)就是被测个体所处的位置, 该窗口对应位置上的一维信号数据就是被测个体反射回来的信号. 通常来说, 当人体处于静止时, 其反射的信号以呼吸为主导. 因此, 对所抽取的信号使用自相关函数来检验该信号的周期性. 若信号的周期超过特定的阈值且其频率在正常的人体呼吸频率范围内(6~30 次/min), 则判定人体处于静止呼吸状态, 而其所处位置则成为床位置范围的一个候选位置. 在处理完整夜所有的信号窗口之后, 可以获得一个代表床范围的候选位置序列, 取其中位数 m, 则床的范围为 $[m-w/2, m+w/2]$, 其中 w 是床的宽度.

3)人体存在性检测. 在获得床的范围后, 将监测范围限制在床的范围内并检测被测个体是否在床上. 具体而言, 重新分析每个窗口中的信号, 将其振幅沿按时间进行叠加, 得到一条信号强度曲线, 曲线上的每一个点代表某一距离在 30 s 内的反射信号总强度. 从该曲线上截取与床对应的范围内的部分曲线, 当满足以下情况之一时, 可判定当前人体不在床范围内: ①该部分曲线上不存在峰值; ②该部分曲线上的最大峰值小于特定阈值. 否则, 判定当前人体在床范围内, 并抽取对应最大峰值位置上的信号数据. 图 4(a)和图 4(b)分别给出了人体不在床范围内和在床范围内时对应的 30 s 信号窗口.



(a) 人体不在床范围内



(b) 人体在床范围内

Fig. 4 The strength of reflected signals at different distances

图 4 不同距离上的反射信号强度

2.4 信号预处理

在获得代表被测个体的信号数据后, 通过取模和取相位角的方式得到其对应的振幅序列和相位序列, 并进行信号预处理操作, 从而去除信号中包含的噪声, 获得后续模型的输入. 信号预处理主要分为 3 步: 插值、矫正和标准化.

1)插值. 人体的呼吸、心跳和体动信息都有助于睡眠阶段分类. 因此, 为了保证信号能够捕捉到更多的生命体征细节信息, 给 UWB 设备选择了一个相对较高的帧率. 然而, 较高的帧率和较长的采集时间(如每晚采集 8 h 以上)导致信号丢帧的情况频繁出现. 为了解决这一问题, 提出了一种混合插值的方法来进行信号的插值操作. 设输入的信号序列中某一片段连续丢失的采样点数量为 n, 2 个阈值分别为 N_1 和 N_2, 则该混合插值方法的具体操作为: ①若 $n<N_1$, 对该片段进行样条插值; ②若 $N_1 \leqslant n \leqslant N_2$, 对该片段进行多项式插值; ③若 $n>N_2$, 该片段丢帧严重, 直接丢弃整个序列.

2)矫正. 多径效应和环境噪声等因素导致部分 UWB 信号出现特定的信号趋势, 这些趋势破坏了信

号数据的原本波形，并且可能进一步影响后续的模型性能. 因此，首先采用 ADF(augmented Dickey-Fuller)测试[31]来判定信号是否平稳，若非平稳，则信号存在趋势性，需要对这些非平稳序列进行去趋势(detrend)操作.

3)标准化. 对输入的信号数据采用 Z-score 方法进行标准化操作，从而将数据缩放到相同的范围内.

本文分别对被测个体的振幅序列和相位序列进行如上的信号预处理操作，得到预处理后的振幅和相位序列，这些包含 30 s 信号数据的序列将用于后续的睡眠阶段分类.

2.5 睡眠阶段分类

考虑到人体整夜的睡眠阶段之间具有一定的时序关系与转换模式，融合当前窗口前后一定时段的信息能够更好地帮助模型进行分类任务. 因此，设计了一个 seq2seq 的深度学习模型来实现从 UWB 信号序列到人体睡眠阶段序列的映射. 所设计的模型结构如图 5 所示，该模型由 4 个部分组成：CNN 编码器、序列编码器、对比学习模块和多专家学习模块. 同时，使用一种两阶段训练框架来对模型进行训练.

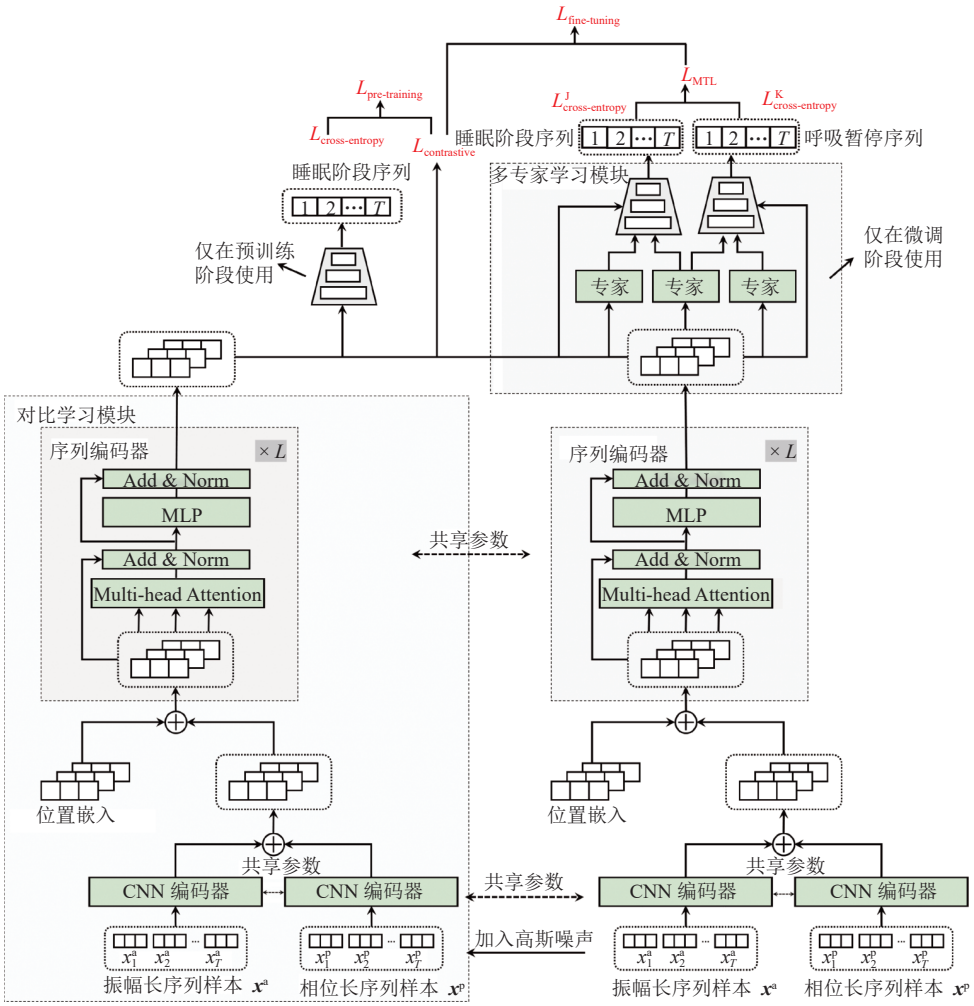


Fig. 5 Model architecture

图 5 模型架构

由上文可知，模型的目标是将每个 30 s 人体信号数据映射成一个睡眠阶段，许多已有工作常将一个 30 s 的数据作为一个样本输入模型进行训练. 本文中，为了引入上下文(context)信息，将 T 个连续的 30 s 人体信号数据(即预处理后的振幅序列和相位序列)拼接在一起(且滑动步长为 R 个 30 s)，分别得到代表振幅和相位的 2 个长序列数据 $\boldsymbol{x}^{\text{a}} \in \mathbb{R}^k$ 和 $\boldsymbol{x}^{\text{p}} \in \mathbb{R}^k$ (其中 $k=f_{\text{s}}\cdot 30\cdot T$，$f_{\text{s}}$ 代表 UWB 设备的采样频率/帧率)，将拼接后的长序列样本送入 seq2seq 模型中，让模型每次同时预测 T 个连续的睡眠阶段.

2.5.1 CNN 编码器

如图 6 所示，CNN 编码器由 b 个块(block)构成，每个块包含 2 个组成部分：一部分由一维卷积层、批量归一化(batch normalization)层和 Leaky ReLU 组成；

另一部分由一维卷积层、Leaky ReLU 和最大池化层组成. 每块之间存在一个跳跃连接. 其中, 一维卷积层用于从信号中提取局部特征: 通过在信号的不同时间窗口上滑动, 卷积层能够捕获信号中的局部模式和特征, 用以反映信号中的趋势、周期性等信息; Leaky ReLU 作为激活函数, 可以缓解深层神经网络中常出现的梯度消失问题, 并防止使用 ReLU 时出现的神经元死亡问题; 最大池化层用于对卷积层输出的特征图(feature map)进行下采样, 在减小其空间维度的同时保留最显著的特征; 批量归一化层对每个批次(batch)中的数据进行归一化, 确保模型在训练过程中能够更加稳定、快速地收敛; 跳跃连接用于缓解梯度消失问题以及加速模型训练过程. 最后, 将 b 层卷积的输出结果输入一个全局平均池化层, 使得输出结果的时间维度(即序列长度)与一开始输入的序列样本对应的睡眠阶段数量 T 相同. 用 CNN 编码器进行特征抽取的操作可以被表示为以下形式:

$$\boldsymbol{E}=CNNs(\boldsymbol{x}^{\mathrm{a}})+CNNs(\boldsymbol{x}^{\mathrm{p}}), \tag{1}$$

其中, $\boldsymbol{x}^{\mathrm{a}}$和$\boldsymbol{x}^{\mathrm{p}}$是被测个体的振幅长序列数据和相位长序列数据, $\boldsymbol{E}\in\mathbb{R}^{T\times d}$是 CNN 编码器输出的序列嵌入(embedding), d 是序列嵌入的特征维度. 输出的序列嵌入中包含 T 个特征向量, 与长序列样本对应的睡眠阶段标签数量相同. 接着, 将其送入序列编码器, 挖掘不同睡眠阶段的信号数据之间的依赖关系.

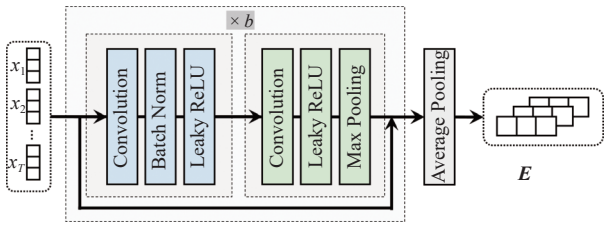


Fig. 6 CNN-based encoder

图 6 CNN 编码器

2.5.2 序列编码器

序列编码器使用多头自注意力机制[32]来挖掘不同睡眠阶段之间的长、短距离依赖, 捕获输入序列中的时序关系和转换模式, 最后得到序列的嵌入, 其中包含序列元素的上下文信息.

为了强化序列中各元素的时序关系, 向 CNN 编码器输出的序列嵌入中注入一个位置嵌入(positional embedding) $\boldsymbol{P}\in\mathbb{R}^{T\times d}$(根据输入序列中元素的位置顺序, 通过正弦/余弦函数生成编码作为 $\boldsymbol{P}$ 的初始值, 且 $\boldsymbol{P}$ 会在模型训练的过程中被不断地更新), 从而得到序列编码器的输入 $\boldsymbol{H}$, 即 $\boldsymbol{H}=\boldsymbol{E}+\boldsymbol{P}$. 此处使用可学习的位置编码[33]而非固定的位置编码, 这样的方式允许模型自适应地优化位置编码, 从而更好地适应不同的任务和数据. 接着, 将 $\boldsymbol{H}$ 输入一个 L 层的多头自注意力模块中, 其中第 l 层 $(1\leqslant l\leqslant L)$ 的具体操作为:

$$\boldsymbol{A}_l^i=softmax\left(\frac{\boldsymbol{X}_{l-1}\boldsymbol{W}_l^{\mathrm{Q}i}(\boldsymbol{X}_{l-1}\boldsymbol{W}_l^{\mathrm{K}i})^{\mathrm{T}}}{\sqrt{d/h}}\right)\boldsymbol{X}_{l-1}\boldsymbol{W}_l^{\mathrm{V}i}, \tag{2}$$

$$\boldsymbol{A}_l=concat(\boldsymbol{A}_l^1,\boldsymbol{A}_l^2,\cdots,\boldsymbol{A}_l^h), \tag{3}$$

$$\boldsymbol{B}_l=LayerNorm(dropout(MLP(\boldsymbol{A}_l))+\boldsymbol{X}_{l-1}), \tag{4}$$

$$\boldsymbol{C}_l=dropout(ReLU(MLP(\boldsymbol{B}_l))), \tag{5}$$

$$\boldsymbol{X}_l=LayerNorm(dropout(MLP(\boldsymbol{C}_l))+\boldsymbol{B}_l), \tag{6}$$

其中, $\boldsymbol{X}_0=\boldsymbol{H}$, h 是多头注意力机制中注意力头(attention head)的数量, $\boldsymbol{W}_l^{\mathrm{Q}i}$、$\boldsymbol{W}_l^{\mathrm{K}i}$和$\boldsymbol{W}_l^{\mathrm{V}i}$是模型中可训练的参数, 其中 $i\in[1,h]$. 序列编码器的最终输出为 $\boldsymbol{X}_L\in\mathbb{R}^{T\times d}$, 在下文中将其表示为 $\boldsymbol{S}$.

2.5.3 对比学习模块

对比学习模块通过数据增强的方式来帮助模型学习到更具有区分性和表征能力的特征表示, 从而提高模型的性能和泛化能力. 如图 5 所示, 该模块同样由一个 CNN 编码器和一个序列编码器构成. 由 2.5.1 节和 2.5.2 节可知, 信号数据输入 CNN 和序列编码器后得到 $\boldsymbol{S}$, 往相同的信号数据中加入高斯噪声, 将其再次输入 CNN 和序列编码器, 得到 $\boldsymbol{S}^+$. $\boldsymbol{S}$ 和 $\boldsymbol{S}^+$来自同源的输入且在语义上是等价的, 被当作一个“正样本对”. 而 $\boldsymbol{S}$ 和同一批次内其他信号数据得到的嵌入 $\boldsymbol{S}^-$(即两者来自不同源的输入)则组成“负样本对”. 对比损失函数表示为:

$$L_{\mathrm{contrastive}}(\boldsymbol{s},\boldsymbol{s}^+,\boldsymbol{s}^-)=-\log\left(\frac{\exp(sim(\boldsymbol{s},\boldsymbol{s}^+)/\tau)}{\exp(sim(\boldsymbol{s},\boldsymbol{s}^+)/\tau)+\sum\limits_{\boldsymbol{s}^-\neq\boldsymbol{s}^+}^{2N}\exp(sim(\boldsymbol{s},\boldsymbol{s}^-)/\tau)}\right), \tag{7}$$

其中, $\boldsymbol{s}$、$\boldsymbol{s}^+$和 $\boldsymbol{s}^-$是 $\boldsymbol{S}$、$\boldsymbol{S}^+$和 $\boldsymbol{S}^-$进行平均池化后的结果, τ是温度参数(用于控制对“难负样本”的惩罚力度[34-35], “难负样本”是与原始样本相似度较高的负样本), N 是批大小(batch size), $sim(\boldsymbol{a},\boldsymbol{b})=\dfrac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}$, 即使用余弦相似度来计算 2 个嵌入之间的相似度. 可以看出, 式(7)的训练目标是最小化“正样本对”之间的距离, 并且最大化“负样本对”之间的距离. 通过这样的方法, 能够使得模型学习到更有区分性的特征, 拥有更强的抗噪能力.

2.5.4 多专家学习模块

在人体睡眠的过程中, 呼吸暂停事件与睡眠阶

段之间有着密切的联系. 例如, 文献 [36] 指出, 相比于 N3 期, 阻塞型呼吸暂停事件通常在 N1 期、N2 期和 REM 期更常见; 中枢型呼吸暂停事件通常出现在 N1 期和 N2 期, 在 N3 期和 REM 期趋于消失; 此外, 所有呼吸暂停事件都不会完整地出现在清醒期等. 文献 [12] 则表明呼吸暂停事件会干扰正常的睡眠阶段转换模式, 例如频繁的呼吸暂停事件会使得人体从睡眠中醒来. 因此, 将睡眠阶段分类设置为主任务, 睡眠呼吸暂停事件识别设置为辅助任务, 通过多任务学习的方式共同训练模型, 从而将睡眠呼吸暂停信息融入模型中, 利用两者之间的关系来提升模型性能.

实现多任务学习的一个简单方法是将序列编码器输出的嵌入 $\boldsymbol{S}$ 分别输入 2 个全连接网络中, 分别完成睡眠阶段分类和睡眠呼吸暂停事件识别. 但由于 2 个任务共用同一个输入, 这样的方式会导致任务间的特有模式和共享模式发生冲突, 进而对模型性能造成负面影响. 为了解决这个问题, 本文使用定制门控网络[37] 来将 $\boldsymbol{S}$ 分别映射为睡眠呼吸暂停相关的嵌入和睡眠阶段相关的嵌入. 该网络的结构如图 7 所示, 包括 3 个组成部分: 专家集群、门控网络和塔网络.

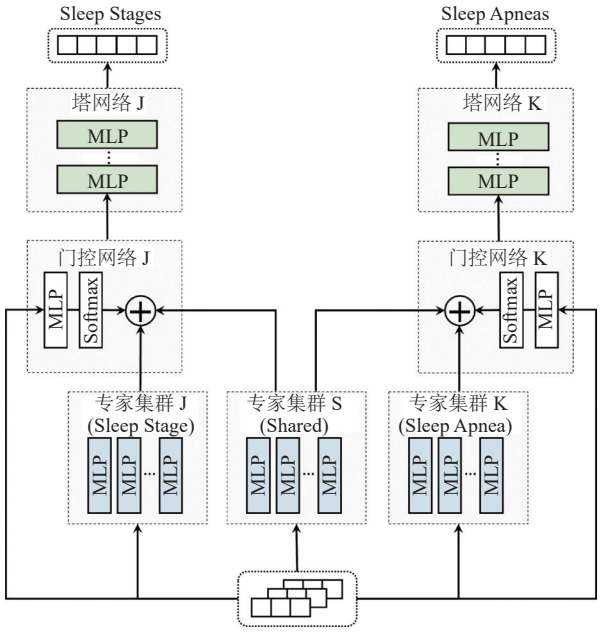


Fig. 7 Customized gate control network

图 7 定制门控网络

1)专家集群. 通过设计多个专家集群, 可以从同一数据中抽取出不同方面的信息. 本文设计了 3 个专家集群, 包括 2 个特定任务专家集群和 1 个共享任务专家集群. 其中, 2 个特定任务专家集群分别用于进行睡眠阶段分类(专家集群 J)和睡眠呼吸暂停事件识别(专家集群 K), 负责从数据中分别学习出 2 类任务各自特有的模式; 而共享任务专家集群(专家集群 S)则负责学习 2 类任务之间的共享模式. 通过这样的方式, 任务间的专有参数和共享参数在模型训练的过程中被显示地分离开来, 从而减轻因任务间的特有模式和共享模式冲突对模型性能造成的影响. 其中, 每个专家集群中包含 e 个多层全连接网络.

2)门控网络. 门控网络负责融合特定任务专家和共享任务专家输出的结果. 每个任务都有一个门控网络, 它为每个专家网络提供对应的权重, 通过加权求和的方式为对应任务输出最后融合的结果. 门控网络是一个以 softmax 为激活函数的单层全连接网络.

3)塔网络. 塔网络根据门控网络融合的结果, 完成具体的预测任务. 每个任务都有一个塔网络, 其结构为一个多层全连接网络.

多专家学习模块的损失函数表示为:

$$L_{\mathrm{MTL}}\left(\boldsymbol{y}^{\mathrm{J}},\hat{\boldsymbol{y}}^{\mathrm{J}},\boldsymbol{y}^{\mathrm{K}},\hat{\boldsymbol{y}}^{\mathrm{K}}\right)=L_{\text{cross-entropy}}\left(\boldsymbol{y}^{\mathrm{J}},\hat{\boldsymbol{y}}^{\mathrm{J}}\right)+\lambda_1 L_{\text{cross-entropy}}\left(\boldsymbol{y}^{\mathrm{K}},\hat{\boldsymbol{y}}^{\mathrm{K}}\right),\tag{8}$$

其中 λ_1 是超参数, $\hat{\boldsymbol{y}}^{\mathrm{J}}$ 和 $\hat{\boldsymbol{y}}^{\mathrm{K}}$ 是塔网络 J 和塔网络 K 的输出结果, $\hat{\boldsymbol{y}}^{\mathrm{J}}$ 和 $\hat{\boldsymbol{y}}^{\mathrm{K}}$ 是对应的真值, 2 个任务使用的损失函数都为交叉熵.

2.5.5 两阶段训练

为了提高模型的性能, 并让模型能够更好地泛化到不同的人群中(包括健康个体和有不同程度睡眠障碍的患者), 提出了一个两阶段训练框架来训练模型, 具体的内容如算法 1 所示.

算法 1. 两阶段训练框架.

输入: 带睡眠阶段标签的公共数据集 D_1, 本文收集的带睡眠阶段和呼吸暂停事件标签的数据集 D_2, 模型早停轮数 θ;

输出: 睡眠阶段预测模型 M.

① /*第 1 阶段: 预训练*/设置学习率为 l_1, 用 D_1 训练模型的 CNN 编码器、序列编码器和对比学习模块, 得到预训练模型 M_1;

② f=$F1$–$score$(M_1);

③ if f 已经连续 θ 个轮次都没有改善

④ 跳转到行⑧;

⑤ else

⑥ 返回行①;

⑦ end if

⑧ /*第 2 阶段: 微调*/在 M_1 的基础上, 以 l_2 的学习率, 用 D_2 中的睡眠阶段标签微调 CNN 编码器、序列编码器和对比学习模块, 同时以 l_1

的学习率, 用 D_2 中的呼吸暂停事件标签训练新加入的多专家学习模块, 得到最终模型 M;

⑨ f=F1−*score*(M);

⑩ if f已经连续θ个轮次都没有改善

⑪ return M;

⑫ else

⑬ 返回行⑧;

⑭ end if

在第 1 阶段, 基于迁移学习[38]的思想, 用一个公共的射频信号数据集[13](该数据集收集了 25 名健康个体在卧室中的整夜睡眠信号数据, 并带有睡眠阶段标签), 以学习率 l_1 对模型中的 2 个编码器和对比学习模块进行预训练, 此阶段仅进行睡眠阶段分类任务; 在第 2 阶段, 使用采集的 UWB 数据集(同时有睡眠阶段和睡眠呼吸暂停事件标签), 以一个较小的学习率 l_2 对模型中在第 1 阶段已经被预训练的部分进行微调, 同时以一个较大的学习率 l_1 来训练模型中的多专家学习模块, 此阶段同时进行睡眠阶段分类和睡眠呼吸暂停事件识别任务. 即, 先利用公共的健康人数据集来训练模型的主体部分, 从而对健康人整夜的睡眠结构和睡眠阶段转换模式进行建模. 然后, 在微调阶段通过多任务学习的方式向预训练的模型中加入睡眠呼吸暂停信息, 从而调整模型的预测结果. 通过这样的方式, 对在患者的睡眠过程中, 因呼吸暂停事件而导致的异常睡眠阶段转换情况进行了建模, 从而使得模型在非健康人数据集上也能取得较好的性能.

因此, 模型在第 1 阶段(预训练)和第 2 阶段(微调)的损失函数分别表示为:

$$L_{\text{pre-training}} = L_{\text{cross-entropy}} + \lambda_2 L_{\text{contrastive}}, \tag{9}$$

$$L_{\text{fine-tuning}} = L_{\text{MTL}} + \lambda_3 L_{\text{contrastive}}, \tag{10}$$

其中, λ_2和λ_3是超参数, $L_{\text{cross-entropy}}$ 是预训练阶段中睡眠阶段分类的损失函数.

3 实验评估

3.1 实验设置

3.1.1 硬件条件

使用商用的 Xethru X4M05 型号的 UWB 芯片来实现 UWB 信号的发送和接收. 该芯片的中心频率为 7.3 GHz, 带宽为 1.4 GHz. 帧率 f_s 设置为 50. 通过杜邦线将 UWB 芯片和树莓派 4 B(数据存储装置)连接在一起, 然后封装成一个小型设备用于数据的采集. 深度学习模型在配有 NVIDIA GeForce RTX 3090 显卡的 DELL 服务器上进行训练与测试.

3.1.2 软件条件

系统实现语言为 Python, 并使用 Pytorch 1.10.1 框架. 模型中具体的超参数赋值如下: 序列样本的长度 T 设置为 55, 步长 R 设置为 10; CNN 编码器的块数 b 设置为 3; 序列编码器的层数 L 设置为 2, 序列嵌入的维度 d 设置为 64, 注意力头的数量 h 设置为 4, dropout 率设置为 0.1; 对比学习模块的温度参数τ设置为 1; 多专家学习模块中每个专家集群中的专家数量 e 都设置为 1; 模型损失函数中的λ_1, λ_2和λ_3分别设置为 0.5、0.1 和 0.1; 批量大小 N 设置为 64, 采用 Adam 优化器作为模型的优化器, 学习率 l_1 和 l_2 分别设置为 1E−3 和 1E−4; 早停轮次θ设置为 15, 即若模型连续 15 个轮次都没有取得更好的性能, 则训练结束.

3.1.3 数据集

本实验收集了来自 110 名个体的 110 晚睡眠数据(共 843 h). 这些人的年龄为 5~69 岁, 并且有着不同程度的 SAHS, 这些个体均在医院通过佩戴 PSG 设备接受睡眠监测, 同时 UWB 设备被摆放在病房的床头柜上, 并面朝被测个体.

医生根据 PSG 的结果对患者整夜的睡眠阶段进行标注, 从而得到本实验的真值. 考虑到患者的隐私问题, 数据的采集过程全程由医生操作(即, 将设备放置于床头柜, 开启/关闭设备). 整个数据采集过程已经得到了医院伦理委员会的批准(包括签署伦理报告与患者知情同意书).

该数据集的相关统计如表 1 所示, 其中, SAHS 的严重程度越高, 睡眠中出现呼吸暂停事件的次数也越多. 该数据集共包含 101 224 个睡眠阶段的数据, 其中有 34 427 个睡眠阶段中出现了呼吸暂停事件.

Table 1 Statistics of Our Dataset

表 1 本文的数据集统计信息

类别	子类	数量
性别	男	95 人
	女	15 人
年龄	最小值	5 岁
	最大值	69 岁
SAHS 严重程度	健康	29 人
	轻度	15 人
	中度	19 人
	重度	47 人
总数		110 人

整个数据集被随机划分为 10 个子集，每个子集都包含 11 个人. 每次选择 1 个子集为测试集，其他的 9 个子集中随机选择 1 个子集作为验证集，剩下的 8 个子集作为训练集. 上述选择过程被重复进行 10 次，这样每个子集都有 1 次作为测试集的机会，模型最终的性能是这 10 次结果的平均结果. 通过这样的方式，模型对数据集中的每个人都进行了评估，并且不存在数据泄露的情况（即同一个人的数据只会出现在一个集合中）.

虽然人体睡眠阶段的 5 分类是医疗界的共识，然而考虑到已有工作在评估模型性能时经常将某些睡眠阶段进行合并，所以本文分别按照以下分类对模型进行性能评估：2 分类（清醒期/睡眠期）、3 分类（清醒期/NREM/REM）、4 分类（清醒期/浅睡眠期/深睡眠期/REM）、5 分类（清醒期/N1/N2/N3/REM）. 其中，5 分类标签为医生标注的结果，其他分类的标签在 5 分类标签的基础上进行合并. 具体地，睡眠期包含 N1、N2、N3 和 REM；NREM 包含 N1、N2 和 N3；浅睡眠期包含 N1 和 N2，深睡眠期包含 N3.

3.1.4 评价指标

采用了准确率、精确率、召回率、*F*1 分数和卡帕（Kappa）系数作为模型性能的评价指标. 由于涉及到多分类，指标中的精确率、召回率和 *F*1 分数都使用多类加权的方式进行计算，权重由各类别标签的数量决定. Kappa 系数则用于度量模型的分类结果和真值这 2 个序列之间的一致性，根据 Kappa 系数的取值，2 个序列之间的一致性关系包括：轻微的一致性（<0.2）、一般的一致性（0.2~0.4）、中等的一致性（0.4~0.6）、高度的一致性（0.6~0.8）和几乎完全一致（0.8~1.0）.

3.1.5 对比方法

本文选择了 6 个非接触式睡眠阶段分类方法作为对比方法. 这些对比方法的描述如下：

1）Sleep hunter[28] 选取 CRF 模型，在由智能手机收集到的声波数据上实现了 3 阶段睡眠分类任务，并且将光照条件和个人的年龄信息作为特征送入模型，从而提升模型的性能.

2）SMARS[16] 使用 SVM 在 Wi-Fi 信号上实现了 3 阶段睡眠分类.

3）Dopplesleep[17] 采用了 4 种传统机器学习分类器（包括朴素贝叶斯、逻辑回归、SVM 和随机森林），在多普勒雷达信号上实现了睡眠阶段 3 分类，并最终采用性能最优的分类器（随机森林）的分类结果作为最终的结果.

4）WiFi-sleep[29] 利用多尺度 CNN 和双向 LSTM 来联合抽取 Wi-Fi 信号中的特征，从而实现睡眠阶段 4 分类. 此外，基于先验医学知识，向模型中引入了阶段转换的限制和阶段时长的限制，从而提高模型的性能.

5) RF-Sleep[13] 提出了一个具有编码器-解码器结构的模型，从而在射频信号上实现了睡眠阶段 4 分类. 采用 CNN+RNN 的结构作为模型的特征提取器，并结合对抗学习的方式来提高模型的泛化性.

6) NSSPS[14] 提出了一个 CNN+RNN 的模型，并使用 CRF 结合先验医学知识来控制睡眠阶段之间的转换关系，最终在生物雷达信号上实现了 4 阶段睡眠分类任务.

3.2 总体性能

让本文的模型分别执行睡眠阶段 2 分类、3 分类、4 分类和 5 分类任务，实验结果如表 2 所示. 可以看出，本文的模型在 4 个分类任务上都取得了不错的性能. 当分类数量增加时，模型的性能逐渐下降. 这是由于 REM 睡眠（N1、N2 和 N3）之间的生理特征较为相似（相较于其他阶段而言），特别是 N1 和 N2 阶段（同属浅睡眠）之间.

Table 2 Model Performance on Four Classification Tasks

表 2 4 个分类任务上的模型性能

目标	评价指标	5 分类	4 分类	3 分类	2 分类
睡眠阶段分类	准确率	0.695	0.811	0.835	0.923
	精确率	0.686	0.805	0.832	0.918
	召回率	0.695	0.811	0.835	0.923
	*F*1 分数	0.686	0.803	0.832	0.919
	Kappa 系数	0.560	0.630	0.658	0.668
睡眠呼吸暂停事件识别	准确率	0.861	0.865	0.864	0.864
	精确率	0.860	0.863	0.863	0.863
	召回率	0.861	0.865	0.864	0.864
	*F*1 分数	0.860	0.864	0.863	0.863
	Kappa 系数	0.687	0.695	0.700	0.695

其次，将本文模型的睡眠分期性能和对比方法的性能进行比较. 为了公平起见，先自行实现了这些对比方法，然后在本文的数据集上对其进行性能评估. 然而这些方法是为特定类别的分类任务（即 3 分类或 4 分类）设计的. 例如，SMARS 专门根据人体在清醒期、REM 和 NREM 睡眠时期的生理特征设计了基于呼吸和体动的手工特征，将其用于其他数量的分类会影响模型的性能. 因此，分别将本文的模型中的类别数设置为与每个对比方法中的类别数一致，然后进行性能评估，实验结果如表 3 所示.

此外，由于 RF-Sleep 和 NSSPS 都在一个公共的

Table 3 Performance Comparison of Different Sleep Staging Methods on Our Dataset

表 3 不同睡眠分期方法在本文数据集上的性能对比

任务	方法	准确率	*F*1 分数	Kappa 系数
3 分类	Sleep hunter	0.637	0.552	0.178
	SMARS	0.567	0.599	0.280
	Dopplesleep	0.680	0.641	0.355
	本文	**0.835**	**0.832**	**0.658**
4 分类	WiFi-sleep	0.761	0.751	0.540
	本文	**0.811**	**0.803**	**0.630**

注：黑体数值表示最优值.

射频信号数据集[13](该数据集收集了 25 名健康个体在不同卧室中的整夜睡眠数据)上进行性能评估，所以也将本文所提出的模型在此数据集上进行性能评估. 具体地，先将本文的模型在本文的数据集上进行预训练，然后在公共数据集上进行性能评估，实验结果如表 4 所示.

Table 4 Performance Comparison of Different Sleep Staging Methods on the Dataset in Ref [13]

表 4 不同睡眠分期方法在文献 [13] 中的数据集上的性能对比

任务	方法	准确率	Kappa 系数
4 分类	RF-Sleep	0.798	0.700
	NSSPS	0.792	0.679
	本文	**0.806**	**0.703**

注：黑体数值表示最优值.

实验结果表明，本文提出的模型在所有评价指标上都优于对比方法. 并且，可以观察到，传统的机器学习方法(即 Sleep hunter、SMARS 和 Dopplesleep)的性能都低于深度学习模型的性能，这是因为传统的机器学习方法使用手工制定的特征，而简单、有限的手工特征缺乏足够的能力来描述来自医院中不同患者的复杂睡眠结构中所蕴含的非线性关系. 此外，如表 4 所示，本文模型在公开数据集上的性能仍然优于对比方法的性能，这表明除医院场景(病房)外，本文模型在家庭场景(卧室)中也能取得出色的性能.

3.3 睡眠时的不同因素对睡眠分期的影响

本节研究将睡眠时的不同因素(包括呼吸暂停事件和血氧饱和度)作为多专家学习模块中的辅助任务时，模型的睡眠分期性能的变化情况. 这里考虑 4 种情况：无辅助任务、仅将呼吸暂停事件识别作为辅助任务、仅将血氧饱和度预测作为辅助任务、将呼吸暂停事件识别和血氧饱和度预测作为 2 个辅助任务，并分别观察不同情况下模型主任务(即睡眠阶段分类)的性能. 4 种情况的实验结果(以 4 分类为例)见表 5.

Table 5 Impact of Auxiliary Task on Performance of Our Model

表 5 辅助任务对本文模型性能的影响

辅助任务	准确率	*F*1 分数	Kappa 系数
无	0.776	0.767	0.574
睡眠呼吸暂停事件识别	**0.811**	**0.803**	**0.630**
血氧饱和度预测	0.794	0.789	0.608
睡眠呼吸暂停事件识别+血氧饱和度预测	0.785	0.779	0.589

注：黑体数值表示最优值.

从表 5 观察到以下现象：1)相比于不设置辅助任务(即不采用多任务学习)，设置辅助任务的模型都获得了性能上的提升，这说明呼吸暂停事件、血氧饱和度和睡眠阶段之间存在一定的联系，通过多任务联合训练的方式对这些联系进行建模，有助于提升模型的性能；2)相比于血氧饱和度预测，将呼吸暂停事件识别作为辅助任务能够得到更好的睡眠分期结果，这说明呼吸暂停事件与睡眠阶段之间的联系更为紧密；3)将呼吸暂停事件识别和血氧饱和度预测同时作为辅助任务进行训练，取得的性能低于单一辅助任务的情况，仅高于无辅助任务的情况. 这说明过多的辅助任务给模型带来了较大的负担，性能反而较单一辅助任务时下降. 因此，本文最终选取睡眠呼吸暂停事件识别作为多任务学习中的辅助任务.

3.4 SAHS 严重程度对睡眠分期的影响

本节研究 SAHS 严重程度对模型性能的影响. 本文采集的数据集包含了 4 种 SAHS 严重程度的对象，即健康、轻度、中度和重度 SAHS 对象. SAHS 的严重程度越高，人体睡眠过程中出现的呼吸暂停事件越多. 表 6 以 4 分类的结果为例，展示了模型对 4 种 SAHS 严重程度的对象的预测性能. 可以观察到，在不使用多任务学习的情况下，被测对象的 SAHS 严重程度越高，模型的性能越差. 这是因为睡眠中频繁出现的呼吸暂停事件扭曲了信号的波形，并打乱了正常的睡眠模式，使得模型在不借助呼吸暂停信息的情况下，很难准确地识别睡眠阶段的变化情况. 同时，从表 6 中可以看出，在采用了本文多任务学习技术将呼吸暂停信息融合进模型后，在 4 种 SAHS 严重程度的对象上都取得了性能的提升，其准确率分别提高了 0.9 个百分点、2.5 个百分点、7.0 个百分点和 8.4 个百分点. 可以看出，模型在 SAHS 严重程度较高的对象上的性能提升更大，经分析认为原因如下：SAHS 严重程度更高的受试者的睡眠中会更加频繁

Table 6 Impact of SAHS Severity on Performance of Our Model

表 6 SAHS 严重程度对本文模型性能的影响

方法	评价指标	健康	轻度	中度	重度
本文	准确率	0.791	0.803	0.832	0.835
	*F*1 分数	0.781	0.792	0.820	0.830
	Kappa 系数	0.617	0.578	0.623	0.664
无多任务学习	准确率	0.782	0.778	0.762	0.751
	*F*1 分数	0.773	0.769	0.751	0.737
	Kappa 系数	0.611	0.556	0.573	0.553

地出现呼吸暂停事件，这些事件为多专家学习模块提供了更多有用的信息，从而有助于对模型做出的睡眠阶段预测进行矫正. 例如，在发生呼吸暂停事件时减少预测 N3 阶段的概率，或在多次呼吸暂停事件发生后增加转换到清醒阶段的概率.

3.5 模型性能与运行时间的对比

本节比较了不同模型在本文数据集上的性能和运行时间（包括训练时间和推理时间）. 从表 7 的实验结果中可以看出，与传统的机器学习方法（即 Sleep hunter、SMARS 和 Dopplesleep）相比，深度学习模型的训练时间明显较长，但能够取得更加优秀的性能. 这是因为深度学习模型拥有更复杂的网络结构，能够更好地从不同对象的复杂睡眠结构中提取隐含的非线性关系，从而用于睡眠分期，同时也需要更多时间进行模型训练. 而与使用深度学习的其他方法（WiFi-sleep 的主要结构是 CNN+双向 LSTM）相比，本文的模型除了性能更优之外，还拥有更少的运行时间. 这是因为本文使用的注意力网络可以同时计算序列中每个位置的注意力权重，实现并行计算，从而加速训练过程. 而 LSTM 等模型由于其顺序处理方式，无法进行并行训练，导致执行速度较慢. 最后，由于睡眠分期并非实时任务，它需要被测对象先进行整夜的睡眠，然后才能基于整夜的睡眠数据进行分期（即医生根据整夜的 PSG 数据进行阶段标注）. 因此，本文模型的推理时间是可接受的.

Table 7 Performance and Runtime Comparison of Different Sleep Staging Methods on Our Dataset

表 7 不同睡眠分期方法在本文数据集上的性能、运行时间对比

任务	方法	准确率	*F*1 分数	训练时间/s	推理时间/s
3 分类	Sleep hunter	0.637	0.552	1.987	0.018
	SMARS	0.567	0.599	24.232	1.385
	Dopplesleep	0.680	0.641	**1.018**	**0.005**
	本文	**0.835**	**0.832**	2 633.476	1.804
4 分类	WiFi-sleep	0.761	0.751	2 669.269	2.558
	本文	**0.811**	**0.803**	**2 637.644**	**1.813**

注：黑体数值表示最优值.

4 结 论

本文使用 UWB 设备以非接触的方式实现了人体的多阶段睡眠监测. 本文设计了一种基于 seq2seq 的深度学习模型，该模型从 UWB 信号中抽取特征，从而对不同睡眠阶段间的时序关系与转换模式进行建模. 值得一提的是，本文提出了一种基于多任务学习的两阶段训练框架，从信号中联合提取睡眠阶段特征和睡眠呼吸暂停特征，利用它们之间的关联关系来改善睡眠分期任务的性能. 模型已在健康个体和不同程度睡眠障碍患者中进行了实验验证，实验结果表明，本文所提出的模型在多阶段睡眠分类任务上表现出色，并且优于所有对比方法. 后续工作将研究其他睡眠障碍或者共病情况对睡眠分期的影响.

作者贡献声明：李思恒提出了算法思路、实验方案，并完成实验和论文撰写；金蓓弘提出了算法思路、指导意见并修改论文；张扶桑提出指导意见并修改部分论文；王志、马俊麒、苏畅提出论文修改意见；任晓勇、刘海琴提出医学方面的指导意见并协助采集数据.

参 考 文 献

[1] Faraut B, Boudjeltia K Z, Vanhamme L, et al. Immune, inflammatory and cardiovascular consequences of sleep restriction and recovery[J]. Sleep Medicine Reviews, 2012, 16(2): 137−149

[2] Chandrasekaran B, Fernandes S, Davis F. Science of sleep and sports performance–a scoping review[J]. Science & Sports, 2020, 35(1): 3−11

[3] Guo Xiaofan, Zheng Liqiang, Wang Jun, et al. Epidemiological evidence for the link between sleep duration and high blood pressure: A systematic review and meta-analysis[J]. Sleep Medicine, 2013, 14(4): 324−332

[4] Khot S P, Morgenstern L B. Sleep and stroke[J]. Stroke, 2019, 50(6): 1612−1617

[5] Tan Sijing, Qi Zheng. How to rescue young people trapped by poor quality sleep[N/OL]. China Youth Daily. 2024-03-22(005). [2024-05-27]. https://zqb.cyol.com/html/2024-03/22/nw.D110000zgqnb_20240322_2-05.htm (in Chinese)
(谭思静，齐征. 如何解救被低质量睡眠“困”住的年轻人[N/OL].

中国青年报. 2024-03-22(005). [2024-05-27]. https: //zqb.cyol.com/html/2024-03/22/nw.D110000zgqnb_20240322_2-05.htm)

[6] News in Health. Good sleep for good health[EB/OL]. [2024-05-27]. https://newsinhealth.nih.gov/2021/04/good-sleep-good-health

[7] Schulz H. Rethinking sleep analysis: Comment on the AASM manual for the scoring of sleep and associated events[J]. Journal of Clinical Sleep Medicine, 2008, 4(2): 99−103

[8] Zhai Bing, Guan Yu, Catt M, et al. Ubi-SleepNet: Advanced multimodal fusion techniques for three-stage sleep classification using ubiquitous sensing[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2021, 5(4): 1−33

[9] Huang Guanjie, Yuan Ye, Cao Guohong, et al. Accsleepnet: An axis-aware hybrid deep fusion model for sleep stage classification using wrist-worn accelerometer data[C]//Proc of 2022 IEEE Int Conf on Bioinformatics and Biomedicine. Piscataway, NJ: IEEE, 2022: 1005−1012

[10] Yin Wenpeng, Kann K, Yu Mo, et al. Comparative study of CNN and RNN for natural language processing[J]. arXiv preprint, arXiv: 1702.01923, 2017

[11] Bengio Y, Frasconi P, Simard P. The problem of learning long-term dependencies in recurrent networks[C]//Proc of IEEE Int Conf on Neural Networks. Piscataway, NJ: IEEE, 1993: 1183−1188

[12] Bianchi M T, Cash S S, Mietus J, et al. Obstructive sleep apnea alters sleep stage transition dynamics[J]. PLoS One, 2010, 5(6): e11356

[13] Zhao Mingmin, Yue Shichao, Katabi D, et al. Learning sleep stages from radio signals: A conditional adversarial architecture[C]//Proc of the 34th Int Conf on Machine Learning. Cambridge, MA: MIT, 2017: 4100−4109

[14] Zhai Qian, Tang Tingyu, Lu Xiaoling, et al. Machine learning-enabled noncontact sleep structure prediction[J]. Advanced Intelligent Systems, 2022, 4(5): 2100227

[15] Korkalainen H, Aakko J, Nikkonen S, et al. Accurate deep learning-based sleep staging in a clinical population with suspected obstructive sleep apnea[J]. IEEE Journal of Biomedical and Health Informatics, 2019, 24(7): 2073−2081

[16] Zhang Feng, Wu Chenshu, Wang Beibei, et al. SMARS: Sleep monitoring via ambient radio signals[J]. IEEE Transactions on Mobile Computing, 2019, 20(1): 217−231

[17] Rahman T, Adams A T, Ravichandran R V, et al. Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar[C]//Proc of the 2015 ACM Int Joint Conf on Pervasive and Ubiquitous Computing. New York: ACM, 2015: 39−50

[18] Jia Ziyu, Lin Youfang, Wang Jing, et al. GraphSleepNet: Adaptive spatial-temporal graph convolutional networks for sleep stage classification[C]//Proc of the 29th Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2021: 1324−1330

[19] Biswal S, Sun H, Goparaju B, et al. Expert-level sleep scoring with deep neural networks[J]. Journal of the American Medical Informatics Association, 2018, 25(12): 1643−1650

[20] Jia Haohui, Yang Ziwei, Gao Pei, et al. Automatic sleep staging via frequency-wise spiking neural networks[C]//Proc of 2022 IEEE Int Conf on Bioinformatics and Biomedicine. Piscataway, NJ: IEEE, 2022: 1028−1033

[21] Chen Jingrui, Li Yujie, Xiao Jing, et al. MSF-sleepnet: Multi-stream fusion network with contrastive learning for sleep stage classification[C]// Proc of 2022 IEEE Int Conf on Bioinformatics and Biomedicine. Piscataway, NJ: IEEE, 2022: 1493−1496

[22] Advanced Brain Monitoring. Sleep profiler: In-home, patient applied objective assessment of sleep architecture & continuity[EB/OL]. [2024-05-27]. https://www.advancedbrainmonitoring.com/products/sleep-profiler

[23] Chang Liqiong, Lu Jiaqi, Wang Ju, et al. SleepGuard: Capturing rich sleep information using smartwatch sensing data[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(3): 1−34

[24] Han Feiyu, Yang Panlong, Feng Yuanhao, et al. EarSleep: In-ear acoustic-based physical and physiological activity recognition for sleep stage detection[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2024, 8(2): 1−31

[25] Zhai Bing, Perez-Pozuelo I, Clifton E A D, et al. Making sense of sleep: Multimodal sleep stage classification in a large, diverse population using movement and cardiac sensing[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(2): 1−33

[26] De Zambotti M, Goldstone A, Claudatos S, et al. A validation study of Fitbit Charge 2™ compared with polysomnography in adults[J]. Chronobiology International, 2018, 35(4): 465−476

[27] Choe E K, Consolvo S, Watson N F, et al. Opportunities for computing technologies to support healthy sleep behaviors[C]//Proc of the SIGCHI Conf on Human Factors in Computing Systems. New York: ACM, 2011: 3053−3062

[28] Gu Weixi, Shangguan Longfei, Yang Zheng, et al. Sleep hunter: Towards fine grained sleep stage tracking with smartphones[J]. IEEE Transactions on Mobile Computing, 2015, 15(6): 1514−1527

[29] Yu Bohan, Wang Yuxiang, Niu Kai, et al. WiFi-sleep: Sleep stage monitoring using commodity Wi-Fi devices[J]. IEEE Internet of Things Journal, 2021, 8(18): 13900−13913

[30] Yim D H, Cho S H. An equidistance multi-human detection algorithm based on noise level using mono-static IR-UWB radar system[C]// Proc of the 2014 Int Conf on Future Communication, Information and Computer Science. Boca Raton, FL: CRC Press, 2014: 131−134

[31] Mushtaq R. Augmented dickey fuller test[J/OL]. 2011 [2024-05-27]. http://dx. doi.org/10.2139/ssrn.1911068

[32] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. Cambridge, MA: MIT, 2017: 5998−6008

[33] Kang Wangcheng, McAuley J. Self-attentive sequential recommendation[C]//Proc of 2018 IEEE Int Conf on Data Mining. Piscataway, NJ: IEEE, 2018: 197−206

[34] Wang Feng, Liu Huaping. Understanding the behaviour of contrastive loss[C]//Proc of the IEEE/CVF Conf on Computer Vision and Pattern

Recognition. Piscataway, NJ: IEEE, 2021: 2495−2504

[35] Wang Tongzhou, Isola P. Understanding contrastive representation learning through alignment and uniformity on the hypersphere[C]// Proc of the 37th Int Conf on Machine Learning. Cambridge, MA: MIT, 2020: 9929−9939

[36] Sateia M J. International classification of sleep disorders[J]. Chest, 2014, 146(5): 1387−1394

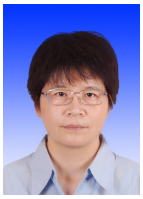
[37] Tang Hongyan, Liu Junning, Zhao Ming, et al. Progressive layered extraction (PLE): A novel multi-task learning (MTL) model for personalized recommendations[C]//Proc of the 14th ACM Conf on Recommender Systems. New York: ACM, 2020: 269−278

[38] Ma Xinyu, Fan Yixing, Guo Jiafeng, et al. An empirical investigation of generalization and transfer in short text matching[J]. Journal of Computer Research and Development, 2022, 59(1): 118−126 (in Chinese)
(马新宇, 范意兴, 郭嘉丰, 等. 关于短文本匹配的泛化性和迁移性的研究分析[J]. 计算机研究与发展, 2022, 59(1): 118−126)



Li Siheng, born in 1997. PhD candidate. Student member of CCF. His main research interests include mobile and pervasive computing, data mining, wireless sensing, and mobile health.

李思恒, 1997 年生. 博士研究生. CCF 学生会员. 主要研究方向为移动与普适计算、数据挖掘、无线感知、移动健康.



Jin Beihong, born in 1967. PhD, professor, PhD supervisor. Distinguished member of CCF. Her main research interests include pervasive computing and machine learning, particularly as applied to healthcare sensing and recommender systems.

金蓓弘, 1967 年生. 博士, 研究员, 博士生导师. CCF 杰出会员. 主要研究方向为普适计算和机器学习, 特别是应用于健康感知和推荐系统方面.



Zhang Fusang, born in 1987. PhD, associate professor. Senior member of CCF. His main research interests include mobile and pervasive computing, smart IoT, wireless sensing, and mobile health.

张扶桑, 1987 年生. 博士, 副研究员. CCF 高级会员. 主要研究方向为移动与普适计算、智能物联网、无线感知、移动健康.



Wang Zhi, born in 1998. PhD candidate. Student member of CCF. His main research interests include mobile and pervasive computing, data mining, wireless sensing, and mobile health.

王　志, 1998 年生. 博士研究生. CCF 学生会员. 主要研究方向为移动与普适计算、数据挖掘、无线感知、移动健康.



Ma Junqi, born in 1999. PhD candidate. Student member of CCF. His main research interests include mobile and pervasive computing, wireless sensing, and indoor localization.

马俊麒, 1999 年生. 博士研究生. CCF 学生会员. 主要研究方向为移动与普适计算、无线感知、室内定位.



Su Chang, born in 1999. PhD candidate. Student member of CCF. His main research interests include mobile and pervasive computing, wireless sensing, and indoor localization.

苏　畅, 1999 年生. 博士研究生. CCF 学生会员. 主要研究方向为移动与普适计算、无线感知、室内定位.



Ren Xiaoyong, born in 1970. PhD, professor, PhD supervisor, chief physician. His main research interests include research on the pathological mechanism and clinical comorbidity of sleep apnea, and clinical and basic research on head, neck tumors and pharyngeal voice diseases.

任晓勇, 1970 年生. 博士, 教授, 博士生导师, 主任医师. 主要研究方向为睡眠呼吸暂停的病理机制及临床共病研究、头颈肿瘤及咽喉嗓音疾病的临床和基础研究.



Liu Haiqin, born in 1975. Bachelor, associate senior technician. Her main research interests include clinical assessment of sleep breathing disorders, sleep monitoring, and therapy of non-invasive positive pressure ventilation.

刘海琴, 1975 年生. 学士, 副主任技师. 主要研究方向为睡眠呼吸障碍临床评估、睡眠监测、无创正压通气治疗.