

DOI:10.19651/j.cnki.emt.2518950

融合睡眠结构与个体先验的 OSA 严重程度估计^{*}张航¹ 刘方姿¹ 罗婉婷¹ 李奇^{1,2,3} 王泽¹(1. 长春理工大学计算机科学技术学院 长春 130022; 2. 吉林省脑信息与智能科学国际联合研究中心 长春 130022;
3. 长春理工大学中山研究院 中山 528400)

摘要: 针对阻塞性睡眠呼吸暂停(OSA)严重程度估计在直接量化呼吸暂停-低通气指数(AHI)及融合多源信息方面的不足,本文提出一种融合睡眠结构与个体先验的 OSA 严重程度估计方法。该方法首先整合从整夜鼻气流、胸腹运动和血氧饱和度信号中提取的多维特征,并创新性融入睡眠结构参数与临床先验知识,随后采用梯度提升回归模型对提取的多源特征进行 AHI 预测。在 MESA 数据集上的验证结果显示,模型取得 0.695 的 R^2 , MAE 与 RMSE 分别为 7.46 和 10.57 次/h。在与多种基准模型的对比中,本文方法表现最优,相较于性能次优的随机森林模型,其 R^2 指标相对提升了 12.46%,充分证明了其优越性。特征重要性分析揭示了血氧指标、N1 期睡眠占比及 BMI 等多源信息对 AHI 预测的关键贡献。结果表明,该方法为 OSA 严重程度的直接量化评估提供了有效方案,并为临床辅助诊断提供更准确的连续量化指标。

关键词: 阻塞性睡眠呼吸暂停; AHI 预测; 多源信息融合; 睡眠结构

中图分类号: TN911 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Estimation of OSA severity by integrating sleep structure and individual prior

Zhang Hang¹ Liu Fangzi¹ Luo Wanting¹ Li Qi^{1,2,3} Wang Ze¹(1. School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China;
2. Jilin Provincial International Joint Research Center of Brain Informatics and Intelligence Science, Changchun 130022, China;
3. Zhongshan Institute of Changchun University of Science and Technology, Zhongshan 528400, China)

Abstract: This study proposes an innovative method for estimating obstructive sleep apnoea (OSA) severity by integrating sleep structure and individual priori. This approach aims to overcome limitations in current OSA assessment, particularly the direct quantification of the apnoea-hypopnoea index (AHI) and the integration of multi-source information. The proposed method initially integrates multidimensional features derived from all-night nasal flow, thoracic/abdominal movements, and oxygen saturation signals. It then distinctively incorporates sleep structure parameters with clinical a priori knowledge. Subsequently, a gradient boosted regression model predicts the AHI using these multi-source features. Validation on the MESA dataset demonstrated the model's performance, achieving R^2 of 0.695, MAE of 7.46 events/h, and RMSE of 10.57 events/h. The proposed method outperformed multiple baseline models, and specifically, its R^2 score showed a relative improvement of 12.46% compared to the next-best model, Random Forest, demonstrating its superiority. These results significantly surpassed those of conventional assessment methods. Feature importance analysis highlighted that parameters such as the oxygen desaturation index, N1 sleep stage percentage, and BMI were key contributors to AHI prediction. These findings indicate that the proposed method offers an effective tool for the direct, quantitative assessment of OSA severity. Furthermore, it provides a more accurate, continuous quantitative index to support clinical diagnosis and decision-making.

Keywords: obstructive sleep apnea; apnea-hypopnea index prediction; multi-source information fusion; sleep structure

0 引言

阻塞性睡眠呼吸暂停(obstructive sleep apnea, OSA)

以睡眠期上呼吸道反复塌陷为核心病理,引发间歇性低氧与睡眠结构紊乱^[1]。OSA 不仅损害睡眠质量与日间功能,还可能诱发心脑血管等并发症,严重威胁公众健康。多导

收稿日期:2025-05-27

* 基金项目:吉林省科技发展计划项目(20240101344JC, 20230203098SF)、中山市社会福利与基础研究项目(2023B2015)资助

睡眠图(polysomnography, PSG)是 OSA 诊断与严重程度评估的金标准,通过监测生理信号来准确量化呼吸暂停-低通气指数(apnea-hypopnea index, AHI)^[2]。AHI 指平均每小时睡眠发生的呼吸暂停和低通气总次数,是临床衡量 OSA 严重程度的核心指标,AASM 临床指南将其分为正常($AHI < 5$)、轻度($5 \leq AHI < 15$)、中度($15 \leq AHI < 30$)和重度($AHI \geq 30$)^[3]。然而,PSG 诊断虽精确,但其高成本、依赖人工判读及专业实验室监测等缺陷,限制了其广泛筛查与基层医疗应用。因此,开发经济、便捷、自动化的 OSA 辅助诊断方法以弥补 PSG 局限,具有重要临床价值与公共卫生意义。

近年来,传感器技术与人工智能的进步推动了基于便携式监测设备和少量生理信号的 OSA 辅助诊断研究^[4]。研究者常运用机器学习或深度学习模型分析 Flow、Thor/Abdo 等易获取信号来评估 OSA 严重程度,以提高诊断效率、降低成本并扩大筛查范围^[5]。例如,Bricout 等^[6]利用加速度计信号间接评估气流,通过机器学习识别异常呼吸模式片段并汇总结果来计算 AHI。Choi 等^[7]采用雷达信号捕捉呼吸运动,借助混合 CNN-Transformer 网络自动检测呼吸暂停/低通气事件以量化 OSA 严重程度。Song 等^[8]则通过分析夜间鼾声,利用融合模型识别呼吸紊乱相关鼾声以估计 AHI。然而,这些方法普遍依赖对 30 s 或 60 s 独立片段的分析。这种“分段-判断-汇总”策略,即将整夜生理信号切分为多段,逐段提取特征、判断呼吸事件,再汇总评估 OSA 严重程度,其弊端在于可能忽视睡眠的生理连续性。睡眠是动态连续的生理过程,呼吸事件的发生、持续时间及其生理累积效应常存在跨片段依赖,片段化分析难以捕捉此类长程信息。同时,逐一分析大量片段也增加了计算负担与任务繁琐度。

为克服分段处理的局限,部分研究开始探索直接分析整夜或更长时程的生理信号,以捕捉睡眠的整体动态。例如,Chen 等^[9]利用整夜未分段的 SpO₂ 信号进行 OSA 严重程度分类,Khan 等^[10-11]则基于整夜未分段的脑电信号进行 OSA 受试者分类。尽管这些研究关注了睡眠整体动态,但其 OSA 严重程度评估仍局限于分类任务,未能直接回归预测 AHI 值,而是将 AHI 按临床标准划分为离散类别进行等级判定。将连续 AHI 值粗略映射到有限类别标签,存在一定的局限。首先,这可能会损失精细的病情信息,无法反映患者病情的细微差异和动态变化,例如,AHI 分别为 31 与 80 的重度 OSA 患者,其临床意义和干预策略可能有所不同。其次,离散分类标签难以捕捉治疗效果评估中 AHI 值的具体改善。例如,患者治疗后 AHI 从 40 降至 35,病情虽有所改善,但其 OSA 严重等级标签却未变。这不利于临床医生精准判断疗效和后续治疗方案调整,尤其在长期随访中难以追踪 AHI 的细微变化,使 OSA 诊断和疗效评估缺乏细致的量化标准。

此外,现有研究在特征提取与信息融合层面亦有优化

空间。OSA 的发生不仅与即时的呼吸生理参数相关,也受睡眠结构特征和个体先验知识的影响^[12-13]。睡眠结构(如各睡眠阶段占比、睡眠效率等)能反映个体睡眠质量与稳定性,特定睡眠阶段(如 REM 期)可能与 OSA 事件集中相关,而睡眠结构紊乱本身也可能加剧 OSA 严重程度。同时,年龄、性别及体重指数(body mass index, BMI)等先验知识已被大量临床研究证实为 OSA 的重要影响因素,例如高龄、男性及肥胖人群 OSA 患病率和严重程度通常更高^[14]。将这些易获取的睡眠结构参数及临床信息作为先验知识融入模型,不仅符合临床逻辑,也可能提升模型预测性能与个性化水平。然而,目前鲜有研究系统地将睡眠结构参数、个体先验知识与生理信号特征有效融合,以实现 OSA 严重程度估计。

基于上述分析,为克服当前 OSA 严重程度估计在睡眠整体信息利用不足、依赖分类而非回归以及多源信息融合欠缺等局限,本文提出一种融合睡眠结构与个体先验知识的 OSA 严重程度估计方法,即 SSPK-GBR 模型(sleep structure and prior knowledge-gradient boosted regression, SSPK-GBR)。该方法首先从整夜同步记录的 Flow、Thor/Abdo 和 SpO₂ 信号中提取多维特征。其次,提取反映个体睡眠质量的睡眠结构参数与关键临床先验知识。最终,通过特征拼接融合这些多源异构信息,并采用梯度提升回归(gradient boosted regression, GBR)模型直接预测 AHI 连续值。SSPK-GBR 旨在通过多源信息融合与回归建模策略,实现 OSA 患者严重程度的精细化、个体化量化评估,为临床诊断、疗效评价及个性化治疗提供更可靠的辅助。

1 方法

针对当前 OSA 严重程度估计的局限性,本文提出一种 SSPK-GBR 模型框架。该方法的核心在于系统性地融合从整夜生理信号中提取的多维度生理特征、反映个体睡眠质量的睡眠结构参数以及临床先验知识,以构建能够直接、准确预测 AHI 连续值的 GBR 回归模型,实现更精细化、个体化的 OSA 严重程度评估。SSPK-GBR 的整体框架如图 1 所示,整体方法流程主要包括数据预处理、多维度特征提取、多源特征融合以及基于 GBR 模型的 AHI 预测 4 个核心阶段。

1.1 数据预处理

为确保输入信号的质量并突出与 OSA 病理生理相关的核心信息,本研究对整夜生理信号进行了系统的预处理。所有关键参数均基于信号的物理特性和公认的信号处理原则,其有效性通过后续实验(详见 2.3 节参数敏感性分析)进行了定量验证。

1)Flow 信号处理

采用 0.1~0.8 Hz 三阶巴特沃斯带通滤波器。该频带的设定旨在精确捕捉从慢速呼吸到代偿性快呼吸的完整生

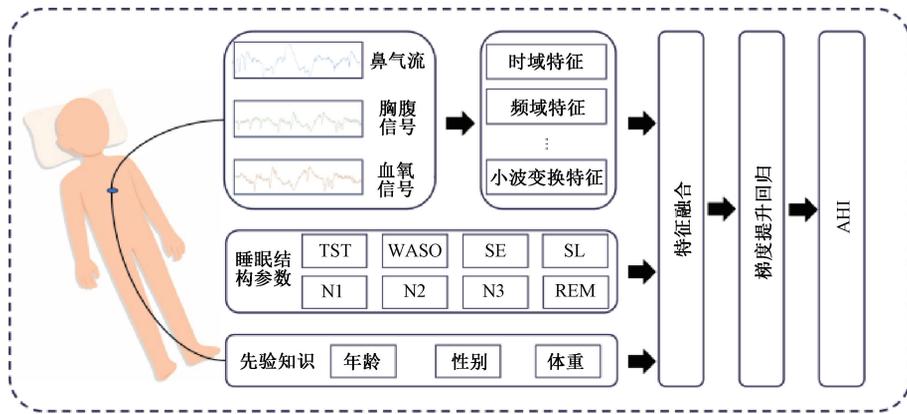


图 1 SSRK-GBR 的整体框架图

Fig. 1 The overall framework of SSRK-GBR

理与病理频率范围,同时有效滤除由身体运动引起的低频基线漂移和心冲击等高频噪声伪影。

2) Thor/Abdo 信号处理

采用截止频率为 2 Hz 的三阶巴特沃斯低通滤波器。此举在完整保留呼吸努力的核心低频动态信息的同时,能有效抑制由肌肉颤动或设备噪声产生的高频干扰,从而清晰地反映呼吸努力的变化。

3) SpO2 信号处理

采用 5 点窗口的中值滤波器进行平滑。该非线性滤波方法特别适用于消除由运动伪影引起的孤立脉冲噪声,且相比线性平滑,它能更好地保留血氧饱和度快速下降事件的边缘特征,避免波形失真。

1.2 多维度特征提取

为全面表征 OSA 严重程度相关的多维信息,本文从预处理后的生理信号、个体睡眠结构以及临床先验知识 3 个维度提取特征。

1) 生理信号特征提取

为全面捕捉 OSA 相关的呼吸模式、呼吸努力程度和机体氧合状态的动态变化,本文针对整夜 Flow、Thor/Abdo 和 SpO2 信号,设计并提取了一系列涵盖时域、频域及小波域等多个角度的量化指标^[15-17]。表 1 详细列举了各项生理信号特征及其具体描述。这些特征均作为基于整夜记录的全局统计量或平均值,旨在从宏观层面揭示个体在整个睡眠期间的整体生理概况。

表 1 提取的生理信号特征

Table 1 Characteristics of physiological signals extracted

信号类型	特征子类	特征名称
Flow	时域统计	气流的均值、标准差、峰度
	频域特征	总功率、生理频段主导频率、低频/高频功率比、0.03~0.5 Hz/10~20 Hz/80~100 Hz 频带平均功率谱密度
	小波能量	基于 db4 小波分解后,不同层级(D1-D5 细节, A5 近似)系数能量
	呼吸模式	平均峰值吸气流速、平均吸气/呼气时长、平均吸呼比、平均吸气上升斜率、平均吸气段曲线下面积
Thor/Abdo	时域统计	信号的均值、标准差、中位数、方差、绝对值均值、绝对值标准差
SpO2	时域统计	信号的最小值、均值、标准差、峰度
	低氧指标	低于特定阈值(如 95%, 90%, 88%)的时间占比、3%氧减指数、缺氧积分

2) 睡眠结构参数提取

基于 PSG 的专业睡眠分期结果,提取了一系列反映个体夜间睡眠宏观结构的参数。这些参数均由业睡眠技师遵循 AASM 判读标准,通过人工判读整夜 PSG 记录获得。所提取的核心睡眠结构参数及其定义如表 2 所示。

这些睡眠宏观结构参数为模型提供了关于个体睡眠质量、连续性和稳定性的重要背景信息。例如,N1 期睡眠占比增高常与 OSA 导致的睡眠结构紊乱相关,且 N1 期上

气道肌肉张力降低的生理特性也可能使其更易发生气道塌陷。因此,将这些参数融入模型,有助于更全面、准确地评估 OSA 严重程度。

3) 个体临床先验知识提取

整合以下易于获取且已被广泛证实为 OSA 重要影响因素的个体临床先验知识:年龄、性别(男性编码为 1,女性为 0)和 BMI。将这些已知的宏观风险因素融入模型,旨在利用这些信息辅助和校准基于生理信号的 AHI 预测,提升模型的临床相关性和个性化评估能力。

表 2 提取的睡眠结构参数

Table 2 Extracted sleep structure parameters

参数名称	英文缩写	计算公式/定义
总睡眠时间	TST	—
睡眠中觉醒时间	WASO	—
睡眠效率	SE	(TST/床上总时间) × 100%
入睡潜伏期	SL	—
N1 期睡眠占比	N1%	(N1 期总时长/TST) × 100%
N2 期睡眠占比	N2%	(N2 期总时长/TST) × 100%
N3 期睡眠占比	N3%	(N3 期总时长/TST) × 100%
REM 期睡眠占比	REM%	(REM 期总时长/TST) × 100%

1.3 多源特征融合

为充分利用多维度特征信息,本文采用特征层面的直接拼接融合策略。具体而言,将生理信号特征集 $\mathbf{F}_{dynamic}$ 、睡眠结构参数集 $\mathbf{F}_{sleep_structure}$ 及个体先验知识集合 \mathbf{F}_{priors} 横向拼接,构成统一的特征向量 \mathbf{F}_{fused} :

$$\mathbf{F}_{fused} = [\mathbf{F}_{dynamic}, \mathbf{F}_{sleep_structure}, \mathbf{F}_{priors}] \quad (1)$$

该融合策略的优势在于保留了所有原始特征的完整信息,避免了早期特征选择或降维可能导致的信息损失。使得后续的 GBR 模型能自主学习不同特征间的复杂相互作用及其对 AHI 的相对贡献。该融合特征向量 \mathbf{F}_{fused} 将作为后续回归模型的输入。

在输入模型前,对 \mathbf{F}_{fused} 中的所有数值型特征进行 Z-score 标准化,以消除不同特征间的取值范围差异,计算公式为:

$$x_{scaled} = (x - \mu) / \sigma \quad (2)$$

其中, x 为原始特征值, μ 和 σ 分别为该特征在数据集上的均值和标准差。

1.4 基于 GBR 的 AHI 预测

采用 GBR 模型学习标准化后的融合特征,以预测 AHI 的连续值。GBR 通过迭代地构建决策树,每棵新树专注于拟合先前集成模型的残差。模型优化采用平方损失函数:

$$L(y, F(x)) = \frac{1}{2} (y - F(x))^2 \quad (3)$$

第 m 轮迭代中,构建决策树 $h_m(\mathbf{X})$ 以拟合损失函数关于上一轮集成模型 $H_{m-1}(\mathbf{X})$ 预测值的负梯度。对于平方损失,此负梯度即为当前轮次的残差 $r_{i,m}$:

$$r_{i,m} = - \left. \frac{\partial L(y_i, F(x_i))}{\partial (F(x_i))} \right|_{F(x)=H_{m-1}(\mathbf{X})} = y_i - H_{m-1}(x_i) \quad (4)$$

其中, y_i 代表第 i 个样本的真实 AHI 值, $H_{m-1}(x_i)$ 是前 $m-1$ 轮集成模型对该样本的预测。

新生成的决策树 $h_m(\mathbf{X})$ 经学习率 ν (亦称收缩率) 加权后,加到上一轮模型 $H_{m-1}(\mathbf{X})$ 上以更新集成模型 $H_m(\mathbf{X})$:

$$H_m(\mathbf{X}) = H_{m-1}(\mathbf{X}) + \nu \cdot h_m(\mathbf{X}) \quad (5)$$

其中, $H_m(\mathbf{X})$ 为第 m 轮的集成模型, $h_m(\mathbf{X})$ 是第 m

轮训练的、拟合残差的决策树, ν ($0 < \nu \leq 1$) 为学习率,用以控制单棵树的贡献并防止过拟合。

2 实验及结果分析

本文方法是用 Python 3.7.0 和 torch 1.7.1 + cu110 实现的。所有实验都在一台配有 Intel Silver - 4210R CPU、32 G 内存和 Nvidia GeForce RTX 3080 GPU 的计算机上进行。

2.1 数据集与实验设置

本文数据源自美国国立心肺血液研究所资助的动脉粥样硬化多种族研究 (multi-ethnic study of atherosclerosis, MESA) 的睡眠研究子项目^[18]。为构建一个均衡的数据集以避免模型偏向特定严重等级,本文从原始数据中筛选并纳入了 1 200 名受试者,确保了正常级 (AHI < 5)、轻度级 ($5 \leq \text{AHI} < 15$)、中度级 ($15 \leq \text{AHI} < 30$) 和重度级 ($\text{AHI} \geq 30$) 四个临床严重等级的样本量完全均衡,各包含 300 人 (占总数的 25.0%)。纳入的生理信号包括 Flow、Thor/Abdo 和 SpO₂。同时,基于 PSG 睡眠分期结果提取了睡眠结构参数,并收集了受试者的年龄、性别和 BMI 等个体先验信息。所有受试者均具有明确的 AHI 作为模型预测的目标。

在该均衡数据集中,用于建模的目标变量 AHI 范围为 1.0~104.2 次/h,呈现显著的偏态分布 (均值 20.8 ± 19.1 次/h,中位数 15.0 次/h)。AHI 均值与中位数的显著差异,以及较大的标准差,共同揭示了数据中存在极端高值,这进一步凸显了本研究构建均衡数据集以避免模型产生偏差的必要性。为确保模型评估的客观性,本研究采用 5 折交叉验证策略。将 1 200 名受试者的数据按个体随机划分为 5 个互不相交的子集。在每一轮验证中,使用 4 个子集 (960 人,占 80%) 作为训练集,剩余的 1 个子集 (240 人,占 20%) 作为验证集。该过程重复 5 次,以保证每个样本都有且仅有一次被用于验证。最终报告的模型性能是这 5 次验证结果的平均值,这种方法能更充分地利用数据并提供更稳健的性能评估。

2.2 评价指标

为全面评估 SSPK-GBR 模型的 AHI 预测性能,本文采用 5 折交叉验证策略,并根据以下指标评估模型:

$$\text{平均绝对误差 (mean absolute error, MAE):} \quad \text{MAE} = \frac{1}{N} \sum_{i=1}^N | \text{AHI}_i^{\text{pred}} - \text{AHI}_i^{\text{true}} | \quad (6)$$

$$\text{均方根误差 (root mean square error, RMSE):} \quad \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{AHI}_i^{\text{pred}} - \text{AHI}_i^{\text{true}})^2} \quad (7)$$

$$\text{决定系数 (R}^2\text{):} \quad R^2 = 1 - \frac{\sum_{i=1}^N (\text{AHI}_i^{\text{pred}} - \text{AHI}_i^{\text{true}})^2}{\sum_{i=1}^N (\text{AHI}_i^{\text{true}} - \text{AHI}^{\text{mean}})^2} \quad (8)$$

其中, N 为样本总数, AHI_i^{pred} 为预测 AHI, AHI_i^{true} 为真实 AHI, AHI^{mean} 为真实值 AHI 均值。 R^2 越接近 1 且 MAE 和 RMSE 越小, 预测效果越好。

除核心回归指标外, 还进行了以下补充评估: 首先, 通过 Bland-Altman 分析评估模型预测 AHI 值与真实 AHI 值之间的一致性, 包括平均偏差和 95% 一致性界限。此外, 将预测 AHI 映射至临床分级(正常/轻/中/重度), 计算准确率与 Cohen's Kappa, 以评估其在临床分级场景的应用潜力, 并传统的直接分类方法进行间接比较。

2.3 结果与分析

1) SSPK-GBR 模型性能

图 2 展示了 SSPK-GBR 模型在 MESA 数据集的上 AHI 预测散点图。结果显示, 数据点整体沿理想对角线($y=x$)分布, 表明模型预测值与真实值间存在显著的正相关性。模型 R^2 为 0.695, 表明其能解释约 69.5% 的 AHI 变异性。同时, MAE 和 RMSE 分别为 7.46 和 10.57 次/h, 量化了平均预测误差。然而, 观察图 2 可以发现, 尽管整体趋势良好, 但在高 AHI 值区域(例如 $AHI > 60$ 次/h), 预测点的离散程度有所增加, 且可能存在低估现象。这种现象在基于生理信号的 AHI 估计中较为常见, 可能与极端事件特征复杂性、样本不足或模型平滑效应有关。

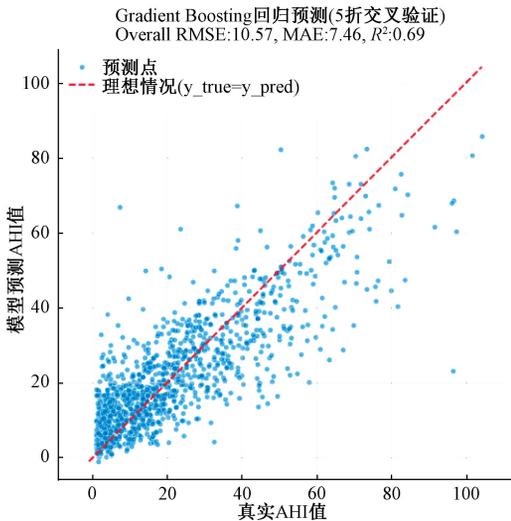


图 2 AHI 预测散点图

Fig. 2 Scatterplot of AHI prediction

为进一步评估预测的准确性与一致性, 图 3 呈现了 Bland-Altman 分析结果。结果显示, 平均偏差仅为 0.19 次/h, 接近于 0, 表明模型整体上无显著的系统性高估或低估偏差。但 95% 一致性界限($-20.53 \sim +20.90$ 次/h)范围较宽, 表示个体预测的随机误差较大, 意味着部分个体预测 AHI 与真实 AHI 差异可能跨越临床分级阈值。尽管模型的平均性能良好, 但个体化精准预测方面仍有提升空间, 尤其在需精确划分临床等级的边缘病例中。此外, 由图 3 可见, 一致性界限的宽度似乎在高平均 AHI 值处略有

增大的趋势, 提示模型对高 AHI 个体预测的不确定性可能更大。

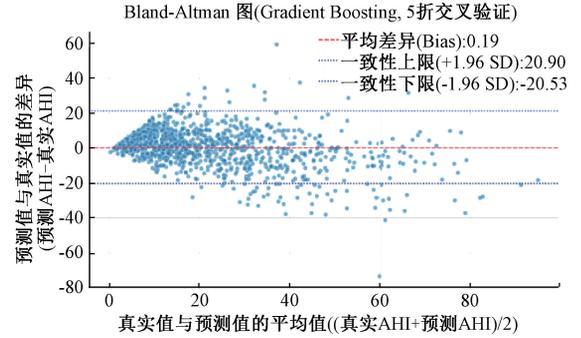


图 3 AHI 预测 Bland-Altman 分析图

Fig. 3 Bland-Altman analysis of AHI predictions

2) 消融实验

为探究睡眠结构参数和个体临床先验知识对 AHI 预测性能的贡献, 本文进行了一系列消融实验。所有模型均采用相同的 5 折交叉验证策略进行评估, 其性能比较如图 4 所示。

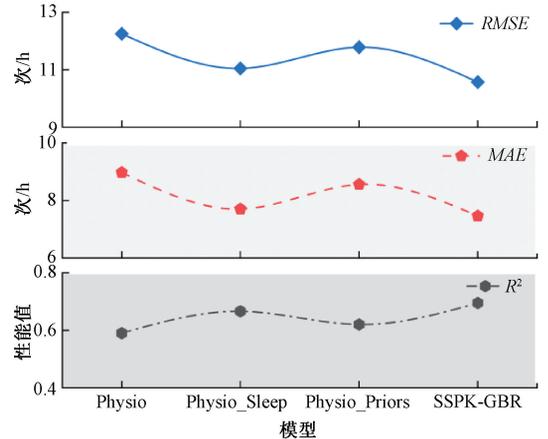


图 4 消融研究中各模型性能比较

Fig. 4 Comparison of the performance of the models in the ablation study

观察图 4, 以仅使用生理信号特征的 Physio 模型为基准, 在分别融入睡眠结构参数形成 Physio_Sleep 模型, 或融入个体先验知识形成 Physio_Priors 模型后, 预测性能均得到改善, 表现为 RMSE 和 MAE 的降低以及 R^2 的提升。值得注意的是, 当整合全部 3 种特征集, 即完整的 SSPK-GBR 在所有评估指标上均展现出最优性能($R^2 = 0.695$, $MAE = 7.46$, $RMSE = 10.57$)。结果表明, 睡眠结构参数与临床先验知识均为 AHI 预测提供了关键的补充信息, 有效融合这些多源信息能显著提升模型预测准确性与解释能力, 从而验证了本文多源特征融合策略的优越性。

3) 特征重要性分析

为探究各输入特征对 SSPK-GBR 模型 AHI 预测的贡

献程度,本文分析了特征重要性,其排名前 15 的特征如图 5 所示。结果显示,血氧饱和度相关的指标占据主导地位,其中血氧低于 90%的时间占比 spo2_ct_lt90_ratio 的重要性远超其他特征,其平均重要性分数约为 0.347,直接反映了间歇性低氧作为 OSA 核心病理的关键性。紧随其后的是睡眠结构参数,特别是 N1 期睡眠占比 stage_N1_percent_tst 位列第二,揭示了睡眠结构紊乱与 OSA 严重程度的密切关联。此外,个体临床先验信息如体重指数 clinical_bmi 和男性性别 clinical_sex_male 亦具有显著贡献,这与它们作为 OSA 已知风险因素的临床认知高度一致,支持了将临床先验纳入模型的合理性。部分呼吸气流特征及腹动标准差 abdo_std_abs_mean 也位列重要特征。

总体而言,此特征重要性分布印证了所选多源特征与 OSA 病理生理的高度相关性,也证明了本文融合策略在捕捉 OSA 复杂影响因素方面的有效性,为精确估计 AHI 提供了多维度信息支撑。

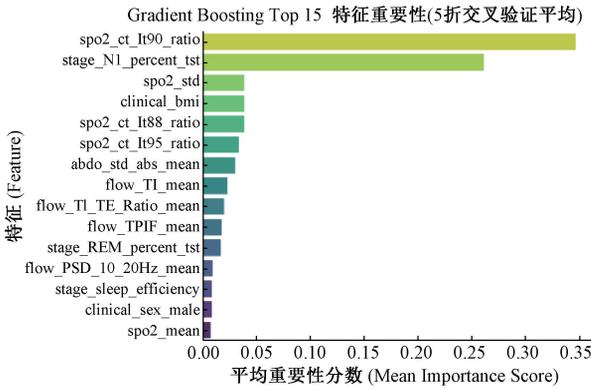


图 5 SSPK-GBR 模型 top 15 特征重要性排序

Fig. 5 SSPK-GBR model top 15 feature importance ranking

4) 与不同模型的比较

为进一步验证 SSPK-GBR 模型在 AHI 预测任务中的有效性,本文将其与多种常用机器学习回归模型进行了性能比较。所有模型均采用相同的融合特征集,并在统一的 5 折交叉验证框架下评估。表 3 展示了各模型在 RMSE、MAE 和 R^2 指标上的对比结果。具体而言,本文提出的 SSPK-GBR 在所有评估指标上均表现最优,取得了最低的 RMSE(10.57 次/h)、MAE(7.46 次/h)以及最高的 R^2 (0.695)。与表现次优的随机森林模型($R^2=0.618$)相比,SSPK-GBR 的 R^2 提升了约 12.46%,且误差显著更低。相比之下,传统线性模型、K 近邻及支持向量回归在本任务中表现逊色,其 R^2 值均低于 0.6,误差亦显著高于 SSPK-GBR,表明这些模型可能难以充分捕捉 AHI 与多源异构特征间的复杂非线性关系。XGBoost 模型也表现出较好的性能,但仍略逊于本文模型。对比结果有力地证明了 GBR 算法结合本文提出的多源特征融合策略在 AHI 回归预测中的优越性。

5) 临床严重程度分类评估

尽管本文聚焦于 AHI 的连续回归预测以精确量化病情,但仍评估了 SSPK-GBR 预测结果在临床分级中的应用潜力。将 5 折交叉验证预测的 AHI 值依据临床标准映射至 OSA 四个严重等级,混淆矩阵如图 6 所示。结果显示,模型在等级分类上总体准确率为 71.25%,Cohen's Kappa 系数为 0.617,表明预测与真实分级间具有实质性的一致性。图 6 显示,模型在识别正常及重度 OSA 受试者方面表现相对稳健,主要误分类发生在相邻等级(如轻度与中度)间,极少将正常个体误判为重度,反之亦然,显示其在宏观严重性判断上具有一定的可靠性。

表 3 不同机器学习模型在 AHI 预测任务上的性能比较

Table 3 Performance comparison of different machine learning models on AHI prediction task

模型	RMSE	MAE	R^2
SVR_RBF	15.79	10.89	0.319
KNN	14.99	11.22	0.386
Linear Regression	12.63	9.00	0.564
Ridge	12.51	8.94	0.572
Lasso	12.45	8.91	0.576
XGBoost	12.01	8.42	0.606
Random Forest	11.82	8.38	0.618
SSPK-GBR	10.57	7.46	0.695

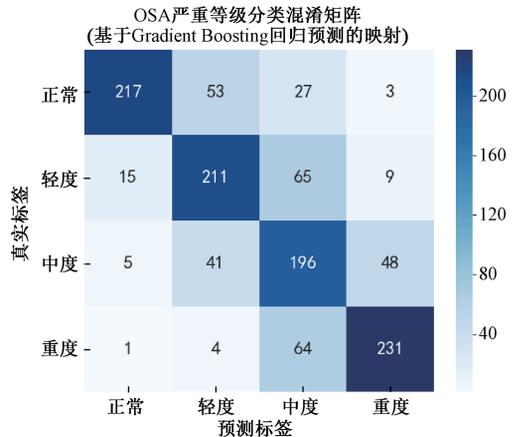


图 6 OSA 严重等级分类混淆矩阵

Fig. 6 OSA severity level classification confusion matrix

SSPK-GBR 的核心优势在于提供连续 AHI 值,这对监测疾病细微进展和精确评估治疗效果至关重要,这是直接分类模型难以实现的。为供参考,表 4 对比了本文回归映射分类结果与 Chen 等^[9]的直接分类结果。从表 4 可见,虽然本文回归映射分类的准确率略低于 Chen 等^[9]的方法,但本文能够输出临床意义更丰富的连续 AHI 值,且基于多源信息融合,为评估 OSA 严重程度提供了更全面的视角。因此,SSPK-GBR 在需要精细化病情管理和疗效追踪的场景下具有独特优势。

表 4 OSA 四级分类性能对比

Table 4 Comparison of OSA four-level classification performance

模型	准确率/%	Cohen's Kapp
Chen	73.75	0.651
SSPK-GBR	71.25	0.617

6) 参数敏感性分析

数据预处理中的参数选择对模型性能有直接影响。为定量验证本文所选参数的合理性,进行了一项关键预处理参数的敏感性分析,系统评估了不同参数设置对模型预测性能的影响,结果如表 5 所示。具体而言,本文所采用的 0.1~0.8 Hz 带通设置在 Flow 信号处理中表现最佳,相较于其他频带,它在有效保留呼吸信息与滤除噪声伪影之间取得了最优平衡。同样地,对于 Thor/Abdo 信号,2 Hz 低通滤波设置也展现出最佳性能。尽管 3 Hz 的性能具有竞争力,但 2 Hz 作为更保守的设定能更可靠地滤除高频干扰而未损失关键信息。在 SpO₂ 信号处理方面,5 点中值滤波的优势尤为明显,其性能显著优于其他窗口大小,这表明较小的窗口可能平滑不足,而较大的窗口则可能过度平滑并导致生理波形失真。综上所述,该敏感性分析为本研究的预处理流程提供了客观且定量的数据支持,证实了所选参数的鲁棒性和有效性,从而确保了后续特征提取和模型训练的可靠性。

表 5 参数敏感性分析结果

Table 5 Parameter sensitivity analysis results

信号类型	测试参数	R ²	MAE	RMSE
Flow	0.05~1.0 Hz 带通	0.681	7.65	10.81
	0.1~0.8 Hz 带通 (our)	0.695	7.46	10.57
	0.15~0.6 Hz 带通	0.675	7.78	10.95
Thor/Abdo	1 Hz 低通	0.688	7.59	10.72
	2 Hz 低通 (our)	0.695	7.46	10.57
	3 Hz 低通	0.692	7.51	10.64
SpO ₂	3 点中值滤波	0.690	7.55	10.68
	5 点中值滤波 (our)	0.695	7.46	10.57
	7 点中值滤波	0.687	7.61	10.75

3 结 论

本文构建的 SSPK-GBR 模型,通过创新性地系统融合睡眠宏观结构、个体临床先验及整夜生理信号特征,并采用 GBR 模型直接预测关键连续指标 AHI,在 OSA 严重程度量化评估领域取得了显著进展。与现有方法相比,SSPK-GBR 凭借多源信息的协同效应,在预测准确性($R^2=0.695$)上超越了多种基准模型,其关键优势更在于克服了分类标签难以捕捉病情细微变化与治疗效果的局限,从而实现了更精细的病情刻画。虽然模型在高 AHI 值预测

和对专业睡眠分期的依赖上仍需完善,但 SSPK-GBR 已能为 OSA 的个体化风险分层和精准疗效评估提供有力的量化支持,为未来更智能、全面的 OSA 管理及融合异构数据深化疾病理解提供有效方案。

参考文献

- [1] LI CH J, SHI ZH H, LI N, et al. OSAMamba: An adaptive bidirectional selective state-space model for OSA detection [J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 1-16.
- [2] HOU L X, ZHUANG Y, ZHANG H, et al. Time-hybrid OSAformer(THO): A hybrid temporal sequence transformer for accurate detection of obstructive sleep apnea via single-lead ECG signals[J]. Computer Methods and Programs in Biomedicine, 2025, 260: 108558.
- [3] 付卓志, 吴亚岑, 李媚希, 等. 阻塞性睡眠呼吸暂停严重度评估参数及诊断技术[J]. 生物化学与生物物理进展, 2025, 52(1): 147-61.
- [4] FU ZH ZH, WU Y C, LI M X, et al. Severity assessment parameters and diagnostic technologies of obstructive sleep apnea[J]. Progress in Biochemistry and Biophysics, 2025, 52(1): 147-161.
- [5] GANGULY B, DASGUPTA R, DEY D. A residual deep learning framework for sleep apnea diagnosis from single lead electrocardiogram signals: An explainable artificial intelligence approach [J]. Engineering Applications of Artificial Intelligence, 2025, 148: 110481.
- [6] 韩亮, 罗统军, 蒲秀娟, 等. 使用异质集成学习和心电信号异构特征融合的睡眠呼吸暂停分类方法[J]. 仪器仪表学报, 2024, 45(6):320-327.
- [7] HAN L, LUO T J, PU X J, et al. Sleep apnea classification method utilizing heterogeneous ensemble learning and electrocardiogram heterogeneous feature fusion[J]. Chinese Journal of Scientific Instrument, 2024, 45(6):320-327.
- [8] BRICOUT A, FONTECAVE-JALLON J, PÉPIN J, et al. Accelerometry-derived respiratory index estimating apnea-hypopnea index for sleep apnea screening[J]. Computer Methods and Programs in Biomedicine, 2021, 207: 106209.
- [9] CHOI J W, KOO D L, KIM D H, et al. A novel deep learning model for obstructive sleep apnea diagnosis: Hybrid CNN-Transformer approach for radar-based detection of apnea-hypopnea events[J]. Sleep, 2024, 47(12):184.
- [10] SONG Y J, SUN X R, DING L, et al. AHI estimation of OSAHS patients based on snoring classification and fusion model[J]. American Journal

- of Otolaryngology, 2023, 44(5): 103964.
- [9] CHEN J W, LIU CH M, WANG CH Y, et al. A deep neural network-based model for OSA severity classification using unsegmented peripheral oxygen saturation signals [J]. Engineering Applications of Artificial Intelligence, 2023, 122: 106161.
- [10] KHAN A, BISWAS S K, CHUNKA C, et al. Intelligent sleep apnea detection by advanced ML using single-lead EEG signal data [J]. IEEE Sensors Journal, 2025, 25(2): 3859-3866.
- [11] KHAN A, BISWAS S K, CHUNKA C, et al. A machine learning model for obstructive sleep apnea detection using ensemble learning and single-lead EEG signal data[J]. IEEE Sensors Journal, 2024, 24(12): 20266-20273.
- [12] 张硕楠, 王爽, 张思宇. 不同年龄阻塞性睡眠呼吸暂停患者睡眠结构参数差异及其与日间脑功能状态的关系[J]. 中国耳鼻咽喉颅底外科杂志, 2024, 31(8): 519-525.
- ZHANG SH N, WANG SH, ZHANG S Y. Differences of sleep structure parameters in patients with obstructive sleep apnea at different ages and the relationship with daytime brain functional status[J]. Chinese Journal of Otorhinolaryngology Skull Base Surgery, 2024, 31(8): 519-525.
- [13] WAN CH, NNAMDI M C, SHI W, et al. Advancing sleep disorder diagnostics: A transformer-based EEG model for sleep stage classification and OSA prediction[J]. IEEE Journal of Biomedical and Health Informatics, 2025, 29(2): 878-886.
- [14] MENCAR C, GALLO C, MANTERO M, et al. Application of machine learning to predict obstructive sleep apnea syndrome severity[J]. Health Informatics Journal, 2020, 26(1): 298-317.
- [15] HOANG N H, LIANG Z. Detection and severity classification of sleep apnea using continuous wearable SpO2 signals: A multi-scale feature approach [J]. Sensors, 2025, 25(6): 1698.
- [16] LIN Y Y, WU H T, HSU C A, et al. Sleep apnea detection based on thoracic and abdominal movement signals of wearable piezoelectric bands [J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(6): 1533-1545.
- [17] LI X L, LING S H, SU S. A hybrid feature selection and extraction methods for sleep apnea detection using bio-signals[J]. Sensors, 2020, 20(15):4323.
- [18] CHEN X L, WANG R W, ZEE P, et al. Racial/ethnic differences in sleep disturbances: the Multi-Ethnic Study of Atherosclerosis (MESA) [J]. Sleep, 2015, 38(6): 877-888.

作者简介

张航, 博士研究生, 主要研究方向为脑机接口技术、睡眠障碍辅助诊断。

E-mail: 2422755413@qq.com

刘方姿, 硕士研究生, 主要研究方向为睡眠障碍辅助诊断。

E-mail: 3499221517@qq.com

罗婉婷, 硕士研究生, 主要研究方向为睡眠剥夺。

E-mail: 2090567487@qq.com

李奇(通信作者), 博士, 教授, 主要研究方向为脑机接口技术、神经康复工程、类脑计算。

E-mail: liqi@cust.edu.cn

王泽, 硕士研究生, 主要研究方向为情绪识别。

E-mail: 1036144868@qq.com