

一种基于集成学习的OSA鼾声监测与分级方法

王者风¹, 周琦斓¹, 徐 曦^{1,2*}

¹湖南工业大学计算机与人工智能学院, 湖南 株洲

²智能信息感知及处理技术湖南省重点实验室, 湖南 株洲

收稿日期: 2026年6月6日; 录用日期: 2026年7月7日; 发布日期: 2026年7月14日

摘 要

针对阻塞性睡眠呼吸暂停(Obstructive Sleep Apnea, OSA)居家筛查需求, 设计分区气囊枕鼾声监测系统, 提出基于Stacking集成学习的鼾声分类方法。系统集成嵌入式鼾声检测模块, 实现鼾声信号采集。针对采集的鼾声信号提取梅尔频率倒谱系数(Mel Frequency Cepstral Coefficients, MFCC)特征, 构建融合卷积神经网络(Convolutional Neural Network, CNN)、双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU)及极端梯度提升(Extreme Gradient Boosting, XGBoost)的异构Stacking模型, 并采用逻辑回归进行二层融合分类。实验结果表明, 所提方法在OSA鼾声四分类任务中取得95.91%的准确率和95.87%的加权平均F1值, 整体性能优于单一模型, 可为OSA家庭化智能筛查提供技术支持。

关键词

阻塞性睡眠呼吸暂停, 鼾声监测, MFCC, Stacking集成学习, 嵌入式系统

A Method for OSA Snore Monitoring and Grading Based on Ensemble Learning

Zhefeng Wang¹, Qilan Zhou¹, Xi Xu^{1,2*}

¹School of Computer Science and Artificial Intelligence, Hunan University of Technology, Zhuzhou Hunan

²Hunan Provincial Key Laboratory of Intelligent Information Perception and Processing Technology, Zhuzhou Hunan

Received: June 6, 2026; accepted: July 7, 2026; published: July 14, 2026

Abstract

To address the need for home-based screening of Obstructive Sleep Apnea (OSA), this paper designs a snore monitoring system based on a partitioned airbag pillow and proposes a snore classification

*通讯作者。

文章引用: 王者风, 周琦斓, 徐曦. 一种基于集成学习的 OSA 鼾声监测与分级方法[J]. 计算机科学与应用, 2026, 16(7): 22-29. DOI: 10.12677/csa.2026.167237

method using Stacking ensemble learning. The system integrates an embedded snore detection module for snore signal acquisition. Mel Frequency Cepstral Coefficients (MFCC) are extracted from the collected snore signals, and a heterogeneous Stacking model is constructed by fusing Convolutional Neural Network (CNN), Bidirectional Gated Recurrent Unit (BiGRU), and Extreme Gradient Boosting (XGBoost), with Logistic Regression employed for second-level fusion classification. Experimental results demonstrate that the proposed method achieves 95.91% accuracy and 95.87% weighted average F1-score on the four-class OSA snore classification task, outperforming individual models overall, and thus can provide technical support for intelligent home-based OSA screening.

Keywords

Obstructive Sleep Apnea, Snore Monitoring, MFCC, Stacking Ensemble Learning, Embedded System

Copyright © 2026 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. 引言

阻塞性睡眠呼吸暂停(Obstructive Sleep Apnea, OSA)是一种睡眠呼吸紊乱疾病, 长期患病可诱发心血管疾病[1]、认知障碍[2]、高血压[3]、糖尿病[4]甚至猝死[5]等, 早期识别和筛查对预防并发症具有重要意义。多导睡眠图(Polysomnography, PSG)是 OSA 诊断的金标准, 但其存在设备复杂、费用较高等问题。目前, OSA 早期筛查主要依赖医学指标问卷及家庭化监测设备。

问卷筛查方面, Laharnar 等[6]构建基于 BMI、颈围等六项客观指标的形态学评分系统, 实现对重度 OSA 的快速判别; Howarth 等[7]提出适用于基层医疗的全科医学睡眠量表, 可由患者自主填写。在家庭化监测方面, 接触式方案如腹部佩戴式传感器[8]、柔性 PVDF 薄膜胸带[9]等信号质量较好, 但长时间佩戴影响舒适性; 非接触式方案如红外摄像头[10]、雷达与热成像[11]等提升了舒适度, 但仍面临光照遮挡、隐私泄露及运动伪影等问题。

鼾声作为 OSA 最直观的早期症状, 蕴含上气道狭窄与振动的生理信息[12], 具有非接触、易采集及环境干扰小等优势。Sebastian 等[13]在距床 1.5 m 的天花板安装麦克风录制整夜鼾声, 利用嵌套交叉验证特征选择策略将鼾声事件分类为 OSA 相关鼾声、单纯鼾声及环境噪声, 识别准确率达 72%。Tan 等[14]通过纳入 16 项利用呼吸音(包括鼾声)诊断 OSA 的相关研究进行贝叶斯荟萃分析, 汇总灵敏度达 90.3%, 性能与家用睡眠监测设备相当。

基于上述研究现状, 本研究设计了一种基于分区气囊枕的鼾声监测系统, 并提出一种基于 Stacking 集成学习的 OSA 鼾声四分类方法。本研究主要贡献如下:

① 设计分区气囊枕, 集成气压传感器、电磁阀及气泵实现动态气压调控, 并嵌入麦克风与鼾声检测模块, 实现鼾声实时采集, 为 OSA 分级诊断及头部姿态干预提供硬件支持。

② 提出融合 CNN、BiGRU 与 XGBoost 的 Stacking 集成分类模型, 通过逻辑回归对局部时频特征、长程时序依赖及全局统计信息进行二层融合, 实现 OSA 鼾声四分类。

2. 鼾声监测系统硬件设计

鼾声监测系统硬件设计结构图如图 1 所示, 由分区气囊枕、鼾声检测模块、USB 麦克风和核心控制

模块组成，核心控制模块集成在智能控制盒内。

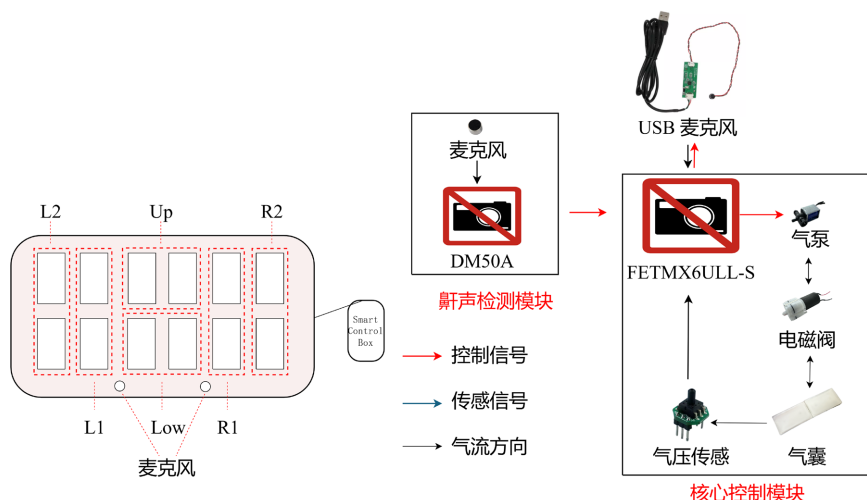


Figure 1. Hardware design structure diagram of snore monitoring system
图 1. 鼾声监测系统硬件设计结构图

分区气囊枕由高密度海绵与六个独立气囊分区构成,麦克风嵌入枕体底端两侧,用于采集睡眠声音。系统可根据鼾声情况调节不同气囊分区气压,从而调整头部姿态,缓解气道阻塞。

鼾声检测模块基于 DM50A 芯片离线工作,采用深度神经网络实现鼾声实时检测,识别准确率达 95% 以上,最快响应 0.1 秒。通过数字 IO 引脚输出结果:持续鼾声时输出一 300 ms 高电平脉冲,待机保持低电平,适用于低功耗、高精度鼾声识别场景。

USB 单麦克风遵循 UCA 标准协议，通过 USB 接口即插即用，兼容 Windows、Linux 及 Android 等系统。内置模数转换，采样率 48 kHz、量化精度 16 位，可高质量采集声音信号，实现数字音频直接输入。

如图 2 所示, 核心控制模块以 FETMX6ULL-S 核心板为中心, 集成蓝牙、气压传感器及相关控制接口。FETMX6ULL-S 核心板采用 NXP 处理器 MCIMX6Y2 开发设计, 集成高性能、低功耗的 ARM Cortex-A7 内核, 运行速度高达 800 MHz, 支持多种外设接口, 并在 Linux 系统完成数据存储与控制管理。当检测到鼾声信号时, 系统自动启动录音并保存 .wav 格式音频, 用于后续分析与处理。

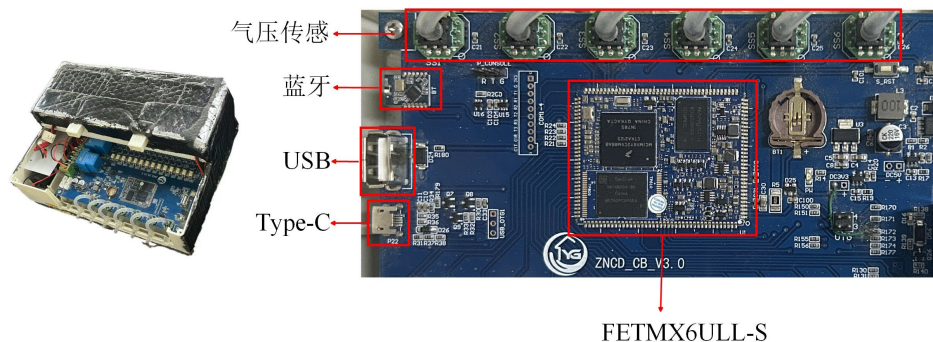


Figure 2. Schematic diagram of core control module
图 2. 核心控制模块示意图

3. 鼾声分类方法

3.1. 数据处理

数据处理部分负责将鼾声处理为纯净、标准化且能有效反应其特性的输入数据，数据处理部分包括了对谱减法去噪、长度标准化及 MFCC 特征提取。

首先，采用谱减法(Spectral Subtraction)对鼾声信号进行降噪。

其次，对降噪后的鼾声信号进行长度标准化。当长度不足 3 s 时采用零填充补齐；当长度超过 3 s 时，采用窗长为 3 s、帧移为 2 s 的滑动窗口进行切分，得到等长鼾声片段集合。

最后，对每个片段提取 MFCC 特征。音频统一重采样至 16 kHz，采用窗长 25 ms、帧移 10 ms 的汉明窗进行分帧。对第 m 帧信号计算功率谱 $P_{i,j,m}(k)$ ，经 40 个梅尔滤波器组取对数能量：

$$E_{i,j,m}(l) = \log \sum_k P_{i,j,m}(k) H_l(k), l = 1, 2, \dots, 40,$$

再通过离散余弦变换得到 40 维 MFCC 向量：

$$c_{i,j,m}(q) = \sum_{l=1}^{40} E_{i,j,m}(l) \cos \left[\frac{\pi q(l-0.5)}{40} \right], q = 1, \dots, 40.$$

将片段内所有帧的 MFCC 按时间排列，得到矩阵 $C_{i,j} \in \mathbb{R}^{T \times 40}$ 。为构造固定尺寸的输入，设定最大帧数 $T_{\max} = 300$ ，最终形成尺寸为 300×40 的特征矩阵。全部片段的特征堆叠为输入张量 $\mathcal{F} \in \mathbb{R}^{N \times 300 \times 40}$ ，其中 N 为总样本数。

3.2. 基于 Stacking 集成学习的鼾声分类方法

为充分挖掘鼾声 MFCC 特征中的多层次信息，本研究提出一种基于 Stacking 集成学习的分类方法，整体框架如图 3 所示。该方法融合 CNN、BiGRU 与 XGBoost 三个异构基分类器，分别从局部时频特征、长程时序依赖及全局统计分布三个角度对鼾声进行建模，并通过逻辑回归完成二层融合分类。

CNN 局部特征提取模型。MFCC 特征矩阵的维度为 300×40 。为沿时间维度捕获局部频谱变化模式，模型采用两层一维卷积结构，卷积核大小均为 3，通道数分别为 64 和 128，并结合 ReLU 激活函数与最大池化操作进行特征提取，最终通过全连接层输出 4 维类别预测向量 $P_{\text{CNN}} \in \mathbb{R}^4$ 。

BiGRU 时序特征建模模型。鼾声信号具有显著的时间动态特性，模型采用单层双向 GRU 结构，隐藏层维度为 128，通过前向与后向时序信息融合获得全局时序特征，并输出 4 维类别概率 $P_{\text{BiGRU}} \in \mathbb{R}^4$ 。

XGBoost 统计特征分类模型。除深度时序特征外，MFCC 序列的统计分布信息同样蕴含判别价值。对每个样本的 MFCC 矩阵沿时间维度分别计算均值 μ 、标准差 σ 、最大值 $f_{\max}$ 与最小值 $f_{\min}$ ，拼接为统计特征向量：

$$\mathbf{f}_{\text{stat}} = [\mu, \sigma, f_{\max}, f_{\min}] \in \mathbb{R}^{160}$$

将该向量输入 XGBoost 分类器，通过梯度提升策略逐步优化弱分类器，输出类别概率 $P_{\text{XGB}} \in \mathbb{R}^4$ 。实验中弱学习器数量设为 200，最大树深设为 6。

为融合不同模型的互补优势，本研究采用 Stacking 策略进行二层学习。首先通过五折交叉验证获得 CNN、BiGRU 与 XGBoost 的预测概率，并将其拼接形成元特征矩阵：

$$\mathbf{P} = [P_{\text{CNN}}, P_{\text{BiGRU}}, P_{\text{XGB}}] \in \mathbb{R}^{N \times 12}$$

其中 $P_{\text{CNN}}, P_{\text{BiGRU}}, P_{\text{XGB}} \in \mathbb{R}^{N \times 4}$ ， N 为训练样本数。随后以逻辑回归作为元分类器进行融合分类，输出最终

的 OSA 严重程度四分类结果。相比单一模型及传统投票方法，该方法能够更有效地融合局部时频特征、时序依赖及统计分布信息，从而提升分类性能。

4. 实验结果与分析

4.1. 实验数据

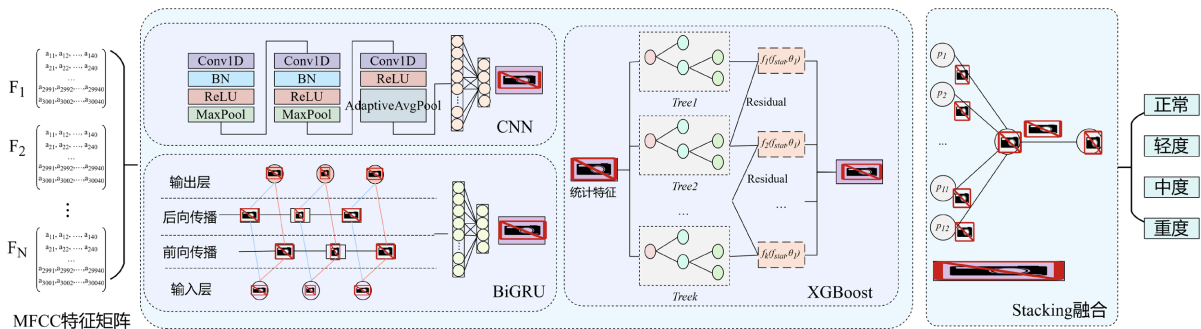


Figure 3. OSA severity classification method based on Stacking ensemble learning

图 3. 基于 Stacking 集成学习的 OSA 严重程度分类方法

实验采用 PSG-Audio 数据集[15]的“MIC”通道信号作为 OSA 鼾声及部分正常鼾声的数据来源，该数据集包含阻塞性呼吸暂停、低通气等多种睡眠事件。同时，利用所设计的气囊枕鼾声监测系统采集 22 名正常人的整夜睡眠录音，作为正常鼾声补充数据。

研究共纳入 74 名受试者，包括 52 名 OSA 患者和 22 名正常人。OSA 鼾声样本由 PSG-Audio 数据集[15]中鼾声与 OSA 事件重叠的片段截取获得；正常鼾声样本则来源于无其他睡眠事件的鼾声片段及正常人整夜录音。所有音频统一处理为 3 s 片段，最终获得 29,003 个样本，具体分布见表 1。

Table 1. Distribution of OSA snore samples by severity

表 1. 不同严重程度 OSA 鼾声样本分布

类别	数量
正常鼾声	12,382
轻度 OSA 鼾声	109
中度 OSA 鼾声	1906
重度 OSA 鼾声	14,606
总计	29,003

4.2. 模型测试与训练

实验部分使用 python3.9 编程语言和 pytorch2.5.1 深度学习框架实现，所有实验在配备 NVIDIA RTX A5000 (40,452 MB) GPU 的设备上进行。数据集按 8:2 划分为训练集与测试集。训练阶段采用五折分层交叉验证生成 Stacking 元特征，并使用逻辑回归完成二层融合分类。为缓解类别不平衡，CNN 与 BiGRU 在交叉熵损失中引入平衡类别权重，共训练 20 轮，采用 Adam 优化器，设定学习率 1×10^{-3} ，设定批大小为 16。最终，Stacking 模型在测试集上取得 95.91% 的准确率和 95.87% 的加权平均 F1 值，其混淆矩阵详见表 2，各类别详细性能见表 3。

Table 2. Confusion matrix of Stacking model on test set**表 2.** Stacking 模型在测试集上的混淆矩阵

真实\预测	正常	轻度	中度	重度
正常	2345	0	26	106
轻度	0	22	0	0
中度	75	0	305	1
重度	27	0	0	2894

Table 3. Classification performance of Stacking model for each category**表 3.** Stacking 模型各类别分类性能

类别	精确率	召回率	F1 值
正常	0.9583	0.9467	0.9525
轻度	1.0000	1.0000	1.0000
中度	0.9215	0.8005	0.8567
重度	0.9643	0.9908	0.9772

其中,轻度类别样本量虽少,但各项指标均为 100%,表明其特征高度可分。重度类别召回率达 99.08%,仅 27 例误判为正常,显示出对严重 OSA 的优异敏感性。正常类别的精确率与召回率分别为 95.83%和 94.67%,分类性能优秀。中度类别召回率为 80.05%,但有 75 例误判为正常、1 例误判为重度,同时正常样本中有 106 例误判为重度,反映出相邻等级间存在一定特征重叠;但中度类别的精确率仍达 92.15%,且 305 个的中度样本被正确识别。综合来看,Stacking 模型在测试集上取得 95.91%的准确率和 95.87%的加权平均 F1 值,各类别性能总体均衡可靠,尤其对重度和轻度判别优势明显。

4.3. 消融实验

为验证 Stacking 融合策略的有效性,本研究在相同数据划分下对 CNN、BiGRU、XGBoost 进行独立评估。各模型在测试集上的表现如表 4 所示,归一化混淆矩阵对比如下:

Table 4. Comparison of accuracy and weighted F1 of each model**表 4.** 各模型准确率与加权 F1 对比

模型	准确率	F1 值
BiGRU	0.8780	0.8798
CNN	0.9267	0.9276
XGBoost	0.9569	0.9563
Stacking	0.9591	0.9587

由表 4 可知,BiGRU 仅依赖时序建模,分类性能相对较低;CNN 通过局部卷积特征提取提升了分类效果;XGBoost 利用统计特征在单模型中表现最佳。相比之下,Stacking 融合模型取得最高准确率和加权 F1 值,整体性能优于各基学习器。

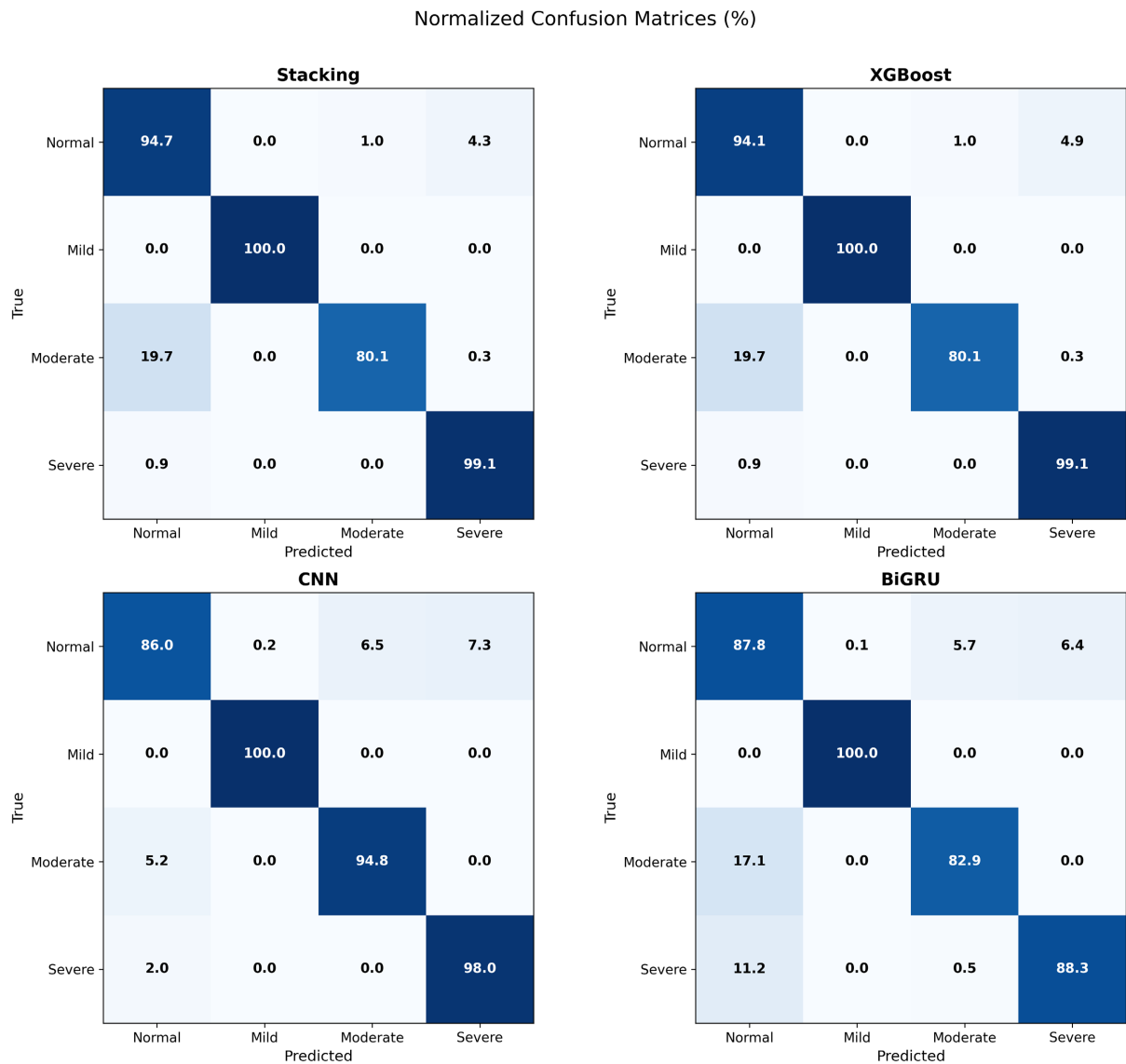


Figure 4. Comparison of normalized confusion matrices of each model

图 4. 各模型归一化混淆矩阵对比

从图 4 的归一化混淆矩阵可以看出，Stacking 模型在保持重度类别高识别率的同时，减少了正常类误判，说明异构模型融合能够有效提升鼾声四分类的准确性与均衡性。

5. 结论

本研究设计了一种基于分区气囊枕的鼾声监测系统，并提出一种融合 CNN、BiGRU 与 XGBoost 的 Stacking 集成分类方法。该方法能够综合学习鼾声 MFCC 特征中的局部时频信息、时序依赖及全局统计特征，并通过逻辑回归实现二层融合分类。实验结果表明，所提方法在 OSA 鼾声四分类任务中取得 95.91% 的准确率和 95.87% 的加权平均 F1 值，整体性能优于单一模型，为 OSA 无创居家筛查提供了新的技术参考。

受限于样本规模与计算开销，模型泛化能力及轻量化部署仍需进一步研究。未来将通过扩充多中心

数据集和探索轻量化集成策略，进一步提升模型的实际应用能力。

基金项目

大学生创新创业训练计划项目(S202411535127)。

参考文献

- [1] Niiranen, T.J., Kronholm, E., Rissanen, H., Partinen, M. and Jula, A.M. (2016) Self-Reported Obstructive Sleep Apnea, Simple Snoring, and Various Markers of Sleep-Disordered Breathing as Predictors of Cardiovascular Risk. *Sleep and Breathing*, **20**, 589-596. <https://doi.org/10.1007/s11325-015-1253-4>
- [2] Grigg-Damberger, M. and Ralls, F. (2012) Cognitive Dysfunction and Obstructive Sleep Apnea: From Cradle to Tomb. *Current Opinion in Pulmonary Medicine*, **18**, 580-587. <https://doi.org/10.1097/mcp.0b013e328358be18>
- [3] Witkowski, A., Prejbisz, A., Florczak, E., Kądziała, J., Śliwiński, P., Bieleń, P., *et al.* (2011) Effects of Renal Sympathetic Denervation on Blood Pressure, Sleep Apnea Course, and Glycemic Control in Patients with Resistant Hypertension and Sleep Apnea. *Hypertension*, **58**, 559-565. <https://doi.org/10.1161/hypertensionaha.111.173799>
- [4] Reutrakul, S. and Mokhlesi, B. (2017) Obstructive Sleep Apnea and Diabetes: A State of the Art Review. *Chest*, **152**, 1070-1086. <https://doi.org/10.1016/j.chest.2017.05.009>
- [5] Cicek, D., Lakadamyali, H., Yağbasan, B., Sapmaz, İ. and Müderrisoğlu, H. (2011) Obstructive Sleep Apnoea and Its Association with Left Ventricular Function and Aortic Root Parameters in Newly Diagnosed, Untreated Patients: A Prospective Study. *Journal of International Medical Research*, **39**, 2228-2238. <https://doi.org/10.1177/147323001103900619>
- [6] Laharnar, N., Herberger, S., Prochnow, L., Chen, N., Cistulli, P.A., Pack, A.I., *et al.* (2021) Simple and Unbiased OSA Prescreening: Introduction of a New Morphologic OSA Prediction Score. *Nature and Science of Sleep*, **13**, 2039-2049. <https://doi.org/10.2147/nss.s333471>
- [7] Howarth, T., Hedger, J., Chen, W., Garg, H. and Heraganahally, S.S. (2025) General Practice Sleep Scale—The “GPSS”—A Proposed New Tool for Use in General Practice for Risk Assessment of Obstructive Sleep Apnoea. *Sleep Medicine*, **125**, 168-176. <https://doi.org/10.1016/j.sleep.2024.11.022>
- [8] Dang, T.H., Kim, S., Choi, M., Hwan, S., Min, H. and Bien, F. (2025) An Automated Algorithm for Obstructive Sleep Apnea Detection Using a Wireless Abdomen-Worn Sensor. *Sensors*, **25**, Article No. 2412. <https://doi.org/10.3390/s25082412>
- [9] Liu, H., Wang, Q., Lan, X., Song, X., Wang, Y., Yuan, H., *et al.* (2025) A Multi-Cascade Deep Learning Method for Physiological Micro-Vibration Signal Enhancement: A Case for OSA Monitoring. *Measurement*, **253**, Article ID: 117796. <https://doi.org/10.1016/j.measurement.2025.117796>
- [10] Choi, Y.R., Eo, G., Youn, W., *et al.* (2023) SIAction: Non-Intrusive, Lightweight Obstructive Sleep Apnea Detection Using Infrared Video. <https://arxiv.org/abs/2309.02713>
- [11] Del Regno, K., Vilesov, A., Armouti, A., *et al.* (2024) Thermal Imaging and Radar for Remote Sleep Monitoring of Breathing and Apnea. <https://arxiv.org/abs/2407.11936>
- [12] Patil, S.P., Schneider, H., Schwartz, A.R. and Smith, P.L. (2007) Adult Obstructive Sleep Apnea: Pathophysiology and Diagnosis. *Chest*, **132**, 325-337. <https://doi.org/10.1378/chest.07-0040>
- [13] Sebastian, A., Cistulli, P.A., Cohen, G., *et al.* (2021) Automatic Classification of OSA Related Snoring Signals from Nocturnal Audio Recordings. <https://arxiv.org/abs/2102.12829>
- [14] Tan, B.K.J., Gao, E.Y., Tan, N.K.W., *et al.* (2025) Machine Listening for OSA Diagnosis: A Bayesian Meta-Analysis. *Chest*, **168**, 520-530.
- [15] Korompili, G., Amfilochiou, A., Kokkalas, L., Mitilneos, S.A., Tatlas, N.A., Kouvaras, M., *et al.* (2021) PSG-Audio, a Scored Polysomnography Dataset with Simultaneous Audio Recordings for Sleep Apnea Studies. *Scientific Data*, **8**, Article No. 197. <https://doi.org/10.1038/s41597-021-00977-w>