

DOI: 10.16300/j.cnki.1000-3630.23080401 CSTR: 32055.14.sxjs.1000-3630.23080401

引用格式: 叶自强, 彭健新. 基于鼾声特征的呼吸暂停低通气指数预测[J]. 声学技术, 2024, 43(6): 862-866. [YE Ziqiang, PENG Jianxin. Prediction of apnea hypopnea index based on snoring sound characteristics[J]. Technical Acoustics, 2024, 43(6): 862-866.]

基于鼾声特征的呼吸暂停低通气指数预测

叶自强, 彭健新

(华南理工大学物理与光电学院, 广东广州 510640)

摘要: 打鼾是阻塞性睡眠呼吸暂停低通气综合征 (obstructive sleep apnea hypopnea syndrome, OSAHS) 患者最典型的症状, 患者鼾声中含有重要诊断信息。文章从 OSAHS 患者整晚鼾声中提取了梅尔倒谱系数、线性预测系数、谱熵等声学特征参数, 通过相关分析研究患者鼾声的特征参数与呼吸暂停低通气指数 (apnea hypopnea index, AHI) 的相关性, 探讨基于梯度提升回归的 AHI 预测, 并与其他模型进行对比。结果表明: 谱熵和梅尔倒谱系数的某些维度与 AHI 具有较强的相关性, 其相关系数大于 0.6。与线性模型相比, 梯度提升回归模型表现出更好的预测效果, 其预测的 AHI 与参考测量值之间的相关性高, 相关系数为 0.813。结果表明, 基于鼾声的声学特征预测 OSAHS 患者的 AHI 是可行的, 对 OSAHS 的诊断具有较好的参考价值。

关键词: 阻塞性睡眠呼吸暂停低通气综合征; 鼾声; 呼吸暂停低通气指数; 声学特征; 梯度提升回归

中图分类号: O429

文献标志码: A

文章编号: 1000-3630(2024)-06-0862-05

Prediction of apnea hypopnea index based on snoring sound characteristics

YE Ziqiang, PENG Jianxin

(School of Physics and Optoelectronics, South China University of Technology, Guangzhou 510640, Guangdong, China)

Abstract: Snoring is the most common symptom in the patients with obstructive sleep apnea hypopnea syndrome (OSAHS). It provides important diagnostic information for patients. In this study, acoustic features such as Mel-frequency cepstral coefficients (MFCC), linear predictive coefficients (LPC), and spectral entropy (SE) are extracted from the snoring sounds of patients throughout the night. The correlation between snoring characteristics and the apnea hypopnea index (AHI) is evaluated through correlation analysis. An AHI prediction model based on gradient boosting regression is investigated, and its performance is compared with that of other models. The results indicate a strong correlation between some characteristics of SE and MFCC with AHI, which have a correlation coefficient greater than 0.6. Compared to the linear model, the gradient boosting regression model demonstrates better prediction performance with a correlation coefficient of 0.813 between predicted AHI and measured values. These findings manifest that it is feasible to predict AHI based on the acoustic characteristics of snoring, providing valuable reference for diagnosing OSAHS patients.

Key words: obstructive sleep apnea hypopnea syndrome; snore; apnea hypopnea index; acoustic feature; gradient boosting regression

0 引言

阻塞性睡眠呼吸暂停低通气综合征 (obstructive sleep apnea hypopnea syndrome, OSAHS) 是一种日常生活中常见的睡眠呼吸障碍疾病, 患者的症状为睡眠过程中上呼吸道部分或完全塌陷导致低通气或呼吸暂停, 并伴有频发性血氧饱和度降低及睡眠结

构混乱等现象^[1]。多导睡眠图 (polysomnography, PSG) 是临床诊断 OSAHS 的金标准^[2], 但其昂贵的检测费用、复杂的检测过程和较低的普及率等导致大量潜在的 OSAHS 患者无法得到及时诊断和治疗^[3-4]。因此研究一种低成本、易操作的 OSAHS 诊断方法是当前的一个研究热点。

打鼾是 OSAHS 患者最典型的临床症状之一, 研究表明鼾声与 OSAHS 存在着紧密的联系^[5]。呼吸暂停低通气指数 (apnea hypopnea index, AHI) 是衡量 OSAHS 严重程度的评估指标, 定义为患者睡眠期间平均每小时呼吸暂停事件与低通气事件的次数之和。已有研究利用患者鼾声预测 AHI^[6-9], 以

收稿日期: 2023-08-04; 修回日期: 2023-08-16

基金项目: 国家自然科学基金 (11974121)

作者简介: 叶自强 (1999—), 男, 湖北随州人, 硕士, 研究方向为鼾声信号检测与处理。

通信作者: 彭健新, E-mail: phjxpeng@scut.edu.cn

辅助诊断 OSAHS。Ding 等^[6]提出了一种基于预训练的 VGG19 和长短期记忆网络融合模型来区分正常鼾声和呼吸紊乱相关的鼾声，并检测出呼吸紊乱事件以计算 AHI。侯丽敏等^[7]采用高斯混合模型和基频特征对正常鼾声和 OSAHS 鼾声分别建模，然后判断呼吸紊乱事件并计算 AHI。上述研究基于呼吸紊乱事件前后鼾声特性对 AHI 进行预测，依赖于对呼吸紊乱事件的准确判别。Kim 等^[8]对受试者整晚睡眠录音信号进行分析，使用随机森林回归和多种时频域特征构建 AHI 预测模型。Xie 等^[9]提出了鼾声变异率来描述整晚鼾声的能量变化，采用 XGBoost 回归预测 AHI。此类方法基于受试者整晚鼾声声学特征预测 AHI，无需识别呼吸紊乱事件及其相关鼾声，简化了前期对鼾声的预处理。但已有研究中，重度 OSAHS 患者数量偏少，使其预测结果的应用受限。本文通过对 102 名受试者整晚睡眠鼾声进行分析，提取了多种鼾声声学特征并评估其参数与 AHI 的相关性，分别利用梯度提升回归 (gradient boosting regression, GBR)、支持向量回归 (support vector regression, SVR) 和线性回归 (linear regression, LR) 方法预测 OSAHS 患者的 AHI 值，为临床提供诊断 OSAHS 提供参考。

1 基于鼾声分析的 AHI 预测模型

1.1 鼾声声学特征提取

鼾声是由上呼吸道组织振动产生的，携带了上呼吸道结构的重要信息^[10]。研究表明鼾声的频域特征比时域特征包含更多有效的信息^[11]。本文首先将受试者整晚录音中的所有鼾声片段分割成帧，设置帧长为 20 ms，帧移为 50%，然后对每一帧计算其特征参数，并按时序排列得到特征矩阵，行数代表特征数量，列数代表时间帧数。最后按帧计算特征值的最大值、最小值、均值和方差，拼接后得到代表受试者整晚睡眠鼾声信息的特征向量。文中共提取了 10 种描述鼾声的声学特征，其特征集维度为 244，计算特征参数与 AHI 的皮尔森 (Pearson) 相关系数 (Pearson correlation coefficient, PCC) 以评估其相关性。表 1 为所提取的声学特征以及相应维度，包括梅尔倒谱系数 (Mel-frequency cepstrum coefficients, MFCC)^[12]、线性预测系数 (linear prediction coefficients, LPC)^[12]、谱熵 (spectral entropy, SE)^[11]、800 Hz 功率比 (power ratio 800 Hz, PR800)^[13]、基频 (pitch)^[14]、共振峰 (formant)^[14]、最大功率比 (maximum power ratio, MPR)^[11]、频谱质心 (spectral centroid, SC)^[11]、频谱平坦度 (spectral

表 1 鼾声的声学特征集
Table 1 Acoustic feature set of snoring

声学特征	描述	特征维度
梅尔倒谱系数	包含 13 阶 MFCC 及其一阶、二阶差分	156
线性预测系数	包含 12 阶 LPC	48
谱熵	描述频谱的无序性度量	4
800 Hz 功率比	以 800 Hz 为界的频带能量之比	4
基频	描述声带振动的频率	4
共振峰	提取第一、第二和第三共振峰	12
最大功率比	描述声音的抖动程度	4
频谱质心	一定范围内频谱能量的加权和	4
频谱平坦度	频谱的几何平均值与算术平均值之比	4
频谱滚降点	频谱能量的二阶矩	4

flatness, SF)^[15]、频谱滚降点 (spectral roll-off, SR)^[11]。

1.2 回归模型

为了比较不同模型对 AHI 的预测能力，分别利用 LR、SVR 和 GBR 三种模型构建 AHI 回归预测模型。LR 模型是通过寻找最佳拟合直线，使得预测值与观测值之间的误差平方和最小，一般使用最小二乘法求取最优解。该模型简单高效，能根据权重系数直接判断每个特征对目标值的影响程度，可解释性较强。

SVR 模型的拟合目标是找到一个能够最大化边界区域内样本点与回归函数之间间隔的曲线，同时限制误差不超过预设的阈值。SVR 模型与 LR 模型相比，其决策边界面是一个区间，样本只要落入决策区间范围内就认为没有误差，不计算损失函数，使得模型具备一定的泛化性和鲁棒性。

GBR 模型是一种基于决策树的集成学习模型，通过迭代训练一个序列化的决策树模型来逐步减小预测误差。GBR 模型中每一棵树都是弱预测模型，其预测结果只能解释部分数据的变化。在每轮迭代中，模型基于前面所有树的残差来构建当前树，然后将新模型的预测值与原始目标变量之间的误差进行最小化，从而改进预测能力。

三种模型参数设置如下：LR 模型参数设置为默认；SVR 模型设置核函数为 linear，惩罚因子为 10.0，其余参数为默认；GBR 模型设置决策树数量为 100，最大深度为 3，学习率为 0.1，其余参数为默认。实验中使用的硬件设备为 CPU AMD Ryzen 9 4900H，软件环境为 Matlab 2021a、Python 3.8 和 Sklearn 0.24.1。

1.3 回归模型评估指标

Pearson 相关系数反映了自变量和变量之间线性相关性的强弱程度。假设测试集样本的真实标签为 y ，预测标签为 $\hat{y}$ ，测试集真实值的均值为 μ ，测

试集预测值的均值为 $\hat{\mu}$ 。计算得到测试集真实标签与预测标签的皮尔森相关系数为

$$PCC = -\frac{\sum_{i=1}^n (y_i - \mu)(\hat{y}_i - \hat{\mu})}{\sqrt{\sum_{i=1}^n (y_i - \mu)^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \hat{\mu})^2}} \quad (1)$$

平均绝对误差 (mean absolute error, MAE) 表示预测值与真实值之间的绝对误差的均值, 所有个体差异在平均值上的权重大小都相等。均方根误差 (root mean square error, RMSE) 表示预测值和观测值之间差异的样本标准差, 对高残差给予更多的惩罚, 表示样本的离散程度。

2 数据采集与实验设置

本文选取来自广州医科大学附属第一医院和开源数据集 PSG-Audio 的受试者共 102 例^[2]。在受试者进行多导睡眠图 (PSG)(伟康 Alice-5) 监测时, 将电容式测量传声器 (Behringer ECM8000) 放置于受试者口鼻上方约 1 m 处, 同步录制鼾声。采样频率和采样精度分别为 48 kHz 和 24 bit, 时长为 6~8 h。后续从采集的受试者整晚睡眠录音中提取鼾声片段, 并在耳鼻喉科专家的指导下对所有的鼾声片段进行人工核对以保证数据的准确性。在本研究中采集到的鼾声总数为 101 283 个。受试者的统计信息如表 2 所示。临床诊断根据 AHI 值将 OSAHS 患者划分为 4 类: 正常打鼾者 (AHI<5)、轻度 OSAHS 患者 (5≤AHI<15)、中度 OSAHS 患者 (15≤AHI<30) 和重度 OSAHS 患者 (AHI≥30)。

表 2 受试者的统计信息
Table 2 Statistical information of subjects

AHI区间	人数	年龄/岁	AHI/(次·h ⁻¹)
AHI<5	10	40	3.3
5≤AHI<15	17	42	10.6
15≤AHI<30	20	44	22.0
AHI≥30	55	47	62.9

为了充分利用实验数据, 采用五折交叉验证训练回归模型并验证。将 102 名受试者的共 101 283 个鼾声片段均分为 5 组不相交的子集, 其中 4 组分别为 20 人, 剩余一组为 22 人。随机挑选一组受试者的整晚鼾声数据作为验证集, 其余作为训练集构建回归模型, 该迭代过程重复 5 次, 保证每组受试者的鼾声数据有且仅有一次机会作为验证集。最后将这 5 次预测结果进行合并以遍历数据集, 获得全体受试者的 AHI 预测值, 以评估模型的预测性能。

3 特征评估与 AHI 预测结果

3.1 特征相关性评估结果

图 1 为 PCC 排序前 10 的特征类别, 分别是 SE、MFCC、LPC、PR800、MPR、SC 以及 SF 的统计量, PCC 均大于 0.4 且显著性水平 $P<0.001$, 这表明这些特征参数与 AHI 之间具有较好的相关性。图 1 中纵坐标物理量下标 mean 表示均值, max 表示最大值, min 表示最小值。其中 PCC 较大的两个特征 (SE 均值、MFCC14 的最大值) 与 AHI 的相关性更强, 其 PCC 大于 0.6。Sun 等^[6]的研究中提到 OSAHS 患者的 SE 均值大于正常打鼾者, 表明 AHI 越高的受试者, 其 SE 均值也越高。这与图 1 中 SE 均值与 AHI 呈现较强的正相关这一结果吻合。

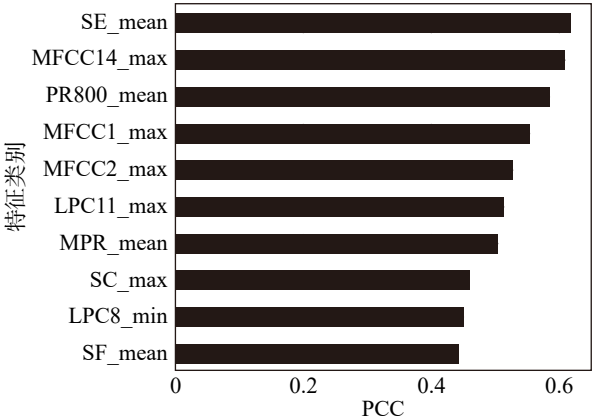


图 1 Pearson 相关系数排序前 10 的特征
Fig.1 Top-10 features ranking by Pearson correlation coefficients

3.2 AHI 预测结果

表 3 给出了三种回归模型下 AHI 的预测结果, LR 和 SVR 模型下 AHI 的预测值与 PSG 测量值 AHI_{PSG} 之间的 PCC 分别为 0.712 和 0.727, MAE 分别为 15.67 和 15.47, 表明 LR 模型和线性核 SVR 模型可以较好地拟合鼾声特征与 AHI 之间的线性关系。SVR 模型对容忍偏差内的样本点不计算损失函数, 相较于 LR 模型更能降低异常值对模型的影响。虽然两者的预测效果差距不大, 但 SVR 模型更具有稳定性和鲁棒性。GBR 模型 AHI 预测值与测量值 AHI_{PSG} 之间的 PCC 达 0.813, MAE

表 3 三种回归模型的 AHI 预测结果
Table 3 AHI prediction results with three regression models

回归模型	PCC	MAE/(次·h ⁻¹)	RMSE/(次·h ⁻¹)
LR	0.712	15.67	19.91
SVR	0.727	15.47	19.55
GBR	0.813	14.15	18.77

为 14.15。对比 LR 和 SVR 模型, GBR 模型预测的 AHI 相关性最高, 与目标值之间的偏差也最小。这是由于 GBR 模型通过迭代训练多个弱学习器, 并组合成一个强大的集成模型, 使得 GBR 模型的预测能力和泛化能力均优于 LR 和 SVR 模型。

表 4 给出了三种模型下 4 个 AHI 区间内患者的 AHI 预测结果。LR 模型和 SVR 模型对正常和轻度患者的 AHI 预测均值偏离参考区间, 对中度和重度患者的预测均值在参考区间内。GBR 模型

对正常、轻度、中度和重度 OSAHS 患者的 AHI 预测均值分别为 7.19、14.26、27.27 和 43.97, 对 OSAHS 患者能相对准确地预测其 AHI。与各个区间内 AHI_{PSG} 均值的绝对误差分别为 4.03、3.95、6.25 和 18.23, 正常打鼾者、轻度和中度 OSAHS 患者的 AHI 预测均值误差为 4~7, 这表明采用 GBR 模型能较准确地预测非重度患者的 AHI 值, 极个别重度患者的 AHI 预测值不在 PSG 结果对应区间内, 整体预测结果略偏低。

表 4 OSAHS 患者在各个区间内的 AHI 预测结果
Table 4 The prediction AHI results of OSAHS patients in each interval

AHI 区间	人数	AHI_{PSG} 均值/标准差	AHI_{GBR} 均值/标准差	AHI_{LR} 均值/标准差	AHI_{SVR} 均值/标准差
0< AHI <5	10	3.16/1.19	7.19/3.27	19.6/13.3	15.59/6.67
5< AHI <15	17	10.31/3.08	14.26/6.29	22.5/10.99	18.86/4.27
15< AHI <30	20	21.02/4.21	27.27/12.82	29.4/12.66	26.27/12.00
AHI >30	55	62.20/18.45	43.97/12.45	50.3/13.97	42.62/12.51

图 2 给出了采用 GBR 模型 AHI 预测值与 AHI_{PSG} 测量值之间的对应关系散点图。当 AHI 小于 50 时, GBR 模型预测得到的受试者的 AHI 与 PSG 测量得到 AHI 之间的差异较小。但随着受试者的 AHI 增加, GBR 模型预测的 AHI 偏小, 与 PSG 测量得到 AHI 之间的偏离趋势逐渐变大, 这与 Kim 等^[8]和 Xie 等^[9]等得到的实验结果一致。对比 PSG 记录和鼾声录音数据发现, 存在呼吸紊乱事件没有鼾声的情形和呼吸紊乱事件中鼾声过于微弱导致没有被检测出鼾声片段, 加上重度患者发生呼吸紊乱事件的频率较高, 这些情形会导致低估重度 OSAHS 患者的 AHI。

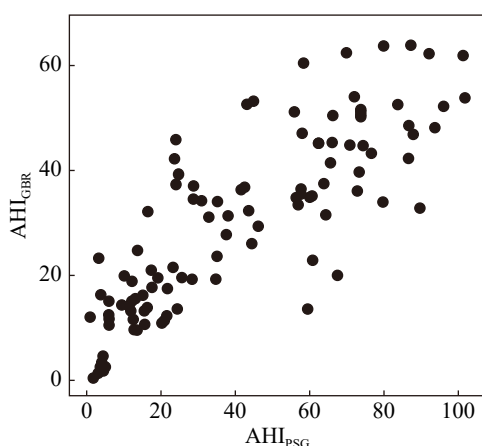


图 2 102 名受试者的 AHI_{GBR} 与 AHI_{PSG} 散点图
Fig.2 The scatter plot of AHI_{GBR} and AHI_{PSG} for the 102 subjects

4 讨论

上述研究结果表明, 根据受试者整晚睡眠鼾声学特征参数预测 OSAHS 患者的 AHI 与采用

PSG 方法得到的 AHI 具有较好的相关性, 验证了方法的可行性。Ding 等^[6]和侯丽敏等^[7]的实验对普通鼾声与呼吸紊乱事件相关的鼾声进行分类, 利用相关鼾声建模来判断呼吸紊乱事件并计算 AHI。这种方法依赖于相关鼾声的识别, 所有的鼾声片段都要经过分类模型筛选出病态鼾声。当鼾声数量较大时, 将导致计算效率低、开销大的问题。本文直接分析受试者整晚鼾声的声学特征信息与 AHI 的关系, 并构建基于回归算法的 AHI 预测模型, 省去繁琐的呼吸紊乱事件的识别。Kim 等^[8]的实验中, 分析了受试者整晚睡眠呼吸声特征与 AHI 的关系, 使用随机森林回归训练 AHI 预测模型并测试, 相关系数达到 0.83。本文直接从受试者整晚睡眠录音中提取鼾声片段, 剔除了环境噪声等其他非鼾声的干扰, 保证了预测模型可靠性。Xie 等^[9]提出了鼾声变异率来描述整晚鼾声能量的动态特征, 并引入了身体质量指数、年龄等患者统计信息作为 AHI 预测的特征变量, 使用 XGBoost 回归预测 AHI 的相关系数 PCC 达到了 0.786, 平均绝对误差为 7.26。但其数据集中 AHI <30 的受试者占比为 81%, AHI <15 的受试者占比为 58%。而临床诊断时重度患者的占比往往更多, 本文使用的数据分布更接近临床实际情况。由于缺乏大数据量的公开数据集, 基于整晚鼾声特征的 AHI 预测研究受到样本数量少及分布不均衡的限制, 其预测精度还有较大的提升潜力。

5 结论

本文探讨了基于受试者整晚睡眠鼾声特征的

AHI 回归预测模型, 从 OSAHS 患者整晚睡眠鼾声中提取了 MFCC、LPC、谱熵等声学特征参数, 采用 Pearson 相关评估声学特征, 并对比了 GBR、LR 和 SVR 三种模型的预测性能。结果显示 SE 和 MFCC 的统计量与 AHI 之间具有较好的相关性。相较于单一模型, 集成模型 GBR 预测 OSAHS 患者的 AHI 效果更好, 其 PCC 为 0.813, MAE 为 14.15。这表明通过 MFCC、SE 等声学特征参数结合 GBR 模型预测 OSAHS 患者的 AHI 是可行的, 对临床借助鼾声辅助诊断 OSAHS 具有重要参考意义。下一步的研究将引入更多的受试者和更多相关的特征训练模型, 进一步提升模型的预测性能。

参 考 文 献

- [1] MALHOTRA A, FORCE F S T, AYAPPA I, et al. Metrics of sleep apnea severity: beyond the apnea-hypopnea index[J]. *Sleep*, 2021, **44**(7): 1-16.
- [2] KOROMPILI G, AMFILOCHIOU A, KOKKALAS L, et al. PSG-Audio, a scored polysomnography dataset with simultaneous audio recordings for sleep apnea studies[J]. *Scientific Data*, 2021, **8**: 197.
- [3] KHAN T. A deep learning model for snoring detection and vibration notification using a smart wearable gadget[J]. *Electronics*, 2019, **8**(9): 987.
- [4] GOTTLIEB D J, PUNJABI N M. Diagnosis and management of obstructive sleep apnea: a review[J]. *JAMA*, 2020, **323**(14): 1389-1400.
- [5] QIAN K, JANOTT C, PANDIT V, et al. Classification of the excitation location of snore sounds in the upper airway by acoustic multifeature analysis[J]. *IEEE Transactions on Biomedical Engineering*, 2017, **64**(8): 1731-1741.
- [6] DING L, PENG J X, SONG L J, et al. Automatically detecting apnea-hypopnea snoring signal based on VGG19 + LSTM[J]. *Biomedical Signal Processing and Control*, 2023, **80**: 104351.
- [7] 侯丽敏, 施晓宇, 童超, 等. 鼾声的基频分布与 SAHS 的关联性[J]. *声学技术*, 2019, **38**(2): 176-181.
- [8] HOU Limin, SHI Xiaoyu, TONG Chao, et al. Correlation between snore pitch distribution and sleep apnea hypopnea syndrome[J]. *Technical Acoustics*, 2019, **38**(2): 176-181.
- [9] KIM J W, KIM T, SHIN J, et al. Prediction of apnea-hypopnea index using sound data collected by a noncontact device[J]. *Otolaryngology-Head and Neck Surgery*, 2020, **162**(3): 392-399.
- [10] XIE J L, FONSECA P, VAN DIJK J, et al. Assessment of obstructive sleep apnea severity using audio-based snoring features[J]. *Biomedical Signal Processing and Control*, 2023, **86**: 104942.
- [11] KIM T, KIM J W, LEE K. Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques[J]. *BioMedical Engineering OnLine*, 2018, **17**(1): 1-19.
- [12] JIANG Y M, PENG J X, SONG L J. An OSAHS evaluation method based on multi-features acoustic analysis of snoring sounds[J]. *Sleep Medicine*, 2021, **84**: 317-323.
- [13] 曾歆, 张雄伟, 孙蒙, 等. 基于 GMM 模型和 LPC-MFCC 联合特征的声道谱转换研究[J]. *声学技术*, 2020, **39**(4): 451-455.
- [14] ZENG Xin, ZHANG Xiongwei, SUN Meng, et al. Research on vocal tract spectrum conversion based on GMM model and LPC-MFCC[J]. *Technical Acoustics*, 2020, **39**(4): 451-455.
- [15] JIANG Y M, PENG J X, ZHANG X W. Automatic snoring sounds detection from sleep sounds based on deep learning[J]. *Physical and Engineering Sciences in Medicine*, 2020, **43**(2): 679-689.
- [16] 褚钰, 李田港, 叶硕, 等. 语音情感识别中的特征选择方法[J]. *应用声学*, 2020, **39**(2): 216-222.
- [17] CHU Yu, LI Tiangang, YE Shuo, et al. Research on feature selection method in speech emotion recognition[J]. *Journal of Applied Acoustics*, 2020, **39**(2): 216-222.
- [18] 郭乐乐, 曹辉, 李涛. 有效特征参数分类正常与病理语音[J]. *声学技术*, 2019, **38**(5): 554-559.
- [19] GUO Lele, CAO Hui, LI Tao. Classification of normal and pathological speech by effective feature parameters[J]. *Technical Acoustics*, 2019, **38**(5): 554-559.
- [20] SUN X R, DING L, SONG Y J, et al. Automatic identifying OSAHS patients and simple snorers based on Gaussian mixture models[J]. *Physiological Measurement*, 2023, **44**(4): 10.1088/1361-10.1088/6579/accd43.