



特征纹理图驱动的鼾声分类

房 玉, 谢 童, 刘栋博, 王维博

西华大学 电气与电子信息学院, 四川 成都 610039

摘要 鼾声分类是辅助医生进行睡眠呼吸暂停综合征(obstructive sleep apnea-hypopnea syndrome, OSAHS)等相关疾病预诊筛查的重要工具。目前基于音频信号时频域特征对病理性鼾声的分类精确度不高,无法满足临床需求,针对此问题,提出了一种基于高维特征纹理(feature texture, FT)图的鼾声信号分类方法。首先,对数据进行预处理,提出连续呼吸事件质量因子确定呼吸信号连续分析的数量范围;然后,提取 51 维时频域特征并应用 ReliefF 算法对特征进行筛选组合;接着,对不同的特征集进行归一化、灰度转换,构建特征纹理灰度图,并求取其水平、垂直梯度矩阵;最后,通过 RGB 三通道串联构建 FT 图。实验采用卷积神经网络对健康受试者和 OSAHS 患者的整晚呼吸音信号提取的 FT 图进行图像分类。结果表明,FT 图对呼吸声、正常鼾声、OSAHS 鼾声进行分类的准确率达到 97.2%,可以为后续睡眠疾病管理和研究提供参考。

关键词 鼾声分类; OSAHS; 特征纹理图

中图分类号: TP399 **DOI:** 10.16152/j.cnki.xdxbzr.2026-01-011

Classification of snore based on feature texture image

FANG Yu, XIE Tong, LIU Dongbo, WANG Weibo

School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, China

Abstract Snoring classification serves as a crucial tool for physicians in prescreening obstructive sleep apnea-hypopnea syndrome (OSAHS) and related sleep disorders. Current research on pathological snoring classification based on time-frequency domain features of audio signals faces limitations in classification accuracy, failing to meet clinical demands. To address this, we propose a high-dimensional snoring signal classification method utilizing feature texture (FT) images. The approach commences with data preprocessing, introducing a continuous respiratory event quality factor to determine the appropriate range of continuous respiratory signals. Subsequently, 51 time-frequency domain features are extracted and refined through the ReliefF algorithm for optimal feature selection and combination. These feature sets undergo normalization and grayscale conversion to construct feature texture grayscale images. Horizontal and vertical gradient matrices are derived from these images, which are then concatenated across RGB channels to form FT images. Experiments employ a convolutional neural network (CNN) to classify FT images derived from overnight respiratory sound signals of healthy subjects and OSAHS patients. Classification tests encompass breathing sounds, normal snoring, and

收稿日期:2025-09-20 修回日期:2025-11-05

基金项目:国家自然科学基金青年基金(61901393);国家自然科学基金(61571371);教育部春晖计划(202201500)。

第一作者:房玉,女,副教授,从事信号与信息处理研究, Yfang_123@163.com。

通信作者:刘栋博,男,副教授,从事模式识别、智能系统研究, liudb@mail.xhu.edu.cn。

OSAHS pathological snoring, revealing that the FT image approach achieves a classification accuracy of 97.2%, offering a promising foundation for subsequent sleep disorder management and research endeavors.

Keywords snoring classification; OSAHS; feature texture image

打鼾是睡眠障碍中最常见的问题之一^[1-2],不仅令人困扰,而且可能引发严重的健康问题。阻塞性睡眠呼吸暂停综合征 (obstructive sleep apnea-hypopnea syndrome, OSAHS) 是一种严重的疾病,其特征为睡眠过程中上气道部分完全闭塞导致呼吸停止^[3]。目前,多导睡眠检测仪 (polysomnography, PSG) 诊断是检测 OSAHS 的金标准^[4],但其存在检测费用过高且操作专业难度大的缺陷^[5]。因此采用低成本、便捷的打鼾分析方法变得尤为重要。

侯丽敏等基于人工神经网络方法,可将整夜录音中的鼾声、呼吸声与其他噪声精准区分,分类准确率超过 95%^[6]。Qian 等在确定鼾声振动位置研究上,采用支持向量机 (support vector machine, SVM) 对小波、共振峰、梅尔倒谱 (Mel frequency cepstral coefficients, MFCC) 特征分类精度达到 79.5%^[7]。本团队前期在对鼾声分析和 OSAHS 识别研究中,从鼾声的时频域中提出了一种新的打鼾检测倒谱系数 (snore detection cepstral coefficient, SDCC),并与 MFCC 特征经筛选融合,结合 KNN 分类器对单纯鼾声、病理性非窒息鼾声、窒息性鼾声的分类准确率达到 97.8%^[8]。常用的时频域特征结合分类器对鼾声分类准确率仍需提高,通过对特征进行筛选和融合等操作能够提升效果以满足临床需求。

此外,许多研究者从语音的图像类特征出发,探索了对语音以及鼾声分类的可能性。Bakhshi 等对 5 种情绪语音做分类研究时,提出用深度卷积神经网络分类所构造的 Cytex 图,分类精度达到 82.25%^[9]。Amiriparian 等研究鼾声不同振动源时,使用 (convolutional neural network, CNN) 对其声谱图做分类,精度达到 67.0%^[10]。Demir 等在对鼾声与噪声分类研究中,使用 SVM 分类器对声谱图的局部二值模式 (local binary pattern, LBP)、方向梯度直方图 (histogram of oriented gradients, HOG) 纹理特征进行分类,精度达到 72.6%^[11]。邓志平等在对鼾声和非鼾声分类研究中,提出一种基于注意力机制的 NewVGG16 双

向门控循环单元算法对鼾声信号的 Mel 谱图进行分类,分类的准确率达到 96.18%^[12]。

以上研究中的图像都是基于单一原信号构建的,但不同图像所携带的特征价值存在局限性,因此本研究在考虑同类信号连续性基础上,将经典的时频域特征构建成二维图像以解决上述难题。本研究将呼吸信号多特征间的连续性纹理图像化并通过 CNN 实现鼾声分类。通过求取连续音频信号多个特征,经筛选后映射为特征纹理灰度图,求取水平、垂直梯度构建三通道的特征纹理 (feature texture, FT) 图,将有助于更准确地识别和分类鼾声,并及时采取干预和治疗措施,保障人类的身体健康。

1 数据采集与处理

1) 数据采集。本研究为实现呼吸声、正常鼾声以及 OSAHS 鼾声分类的研究目的,使用无线蓝牙耳机 (PTM165) 和录音设备 (Lenovo Tablet TB-8804F) 对 4 名年轻健康男性和 2 名中年 OSAHS 男性患者的整晚睡眠声学信号进行采集和录制,采样率为 22.5 kHz。

2) 预处理。本研究对采集到的数据进行了预处理,包括统一采样率、呼吸事件检测和切割。通过实验分析,将音频数据的采样率降采样至 8 kHz,并通过呼吸事件检测和切割算法^[13]切割出长度均为 5 s 的 wav 格式音频数据。实验数据共包含呼吸音 7 860 例,正常鼾声 3 432 例,OSAHS 鼾声 684 例。

2 基于特征纹理图的分类方法

2.1 数据连续性分析

基于特征纹理图的鼾声分类方法框图如图 1 所示,呼吸音频数据经过预处理后,首先进行数据筛选,之后提取其时频域特征并降维,构建特征纹理图,最后输入 CNN 网络进行呼吸和 2 种鼾声的准确分类。

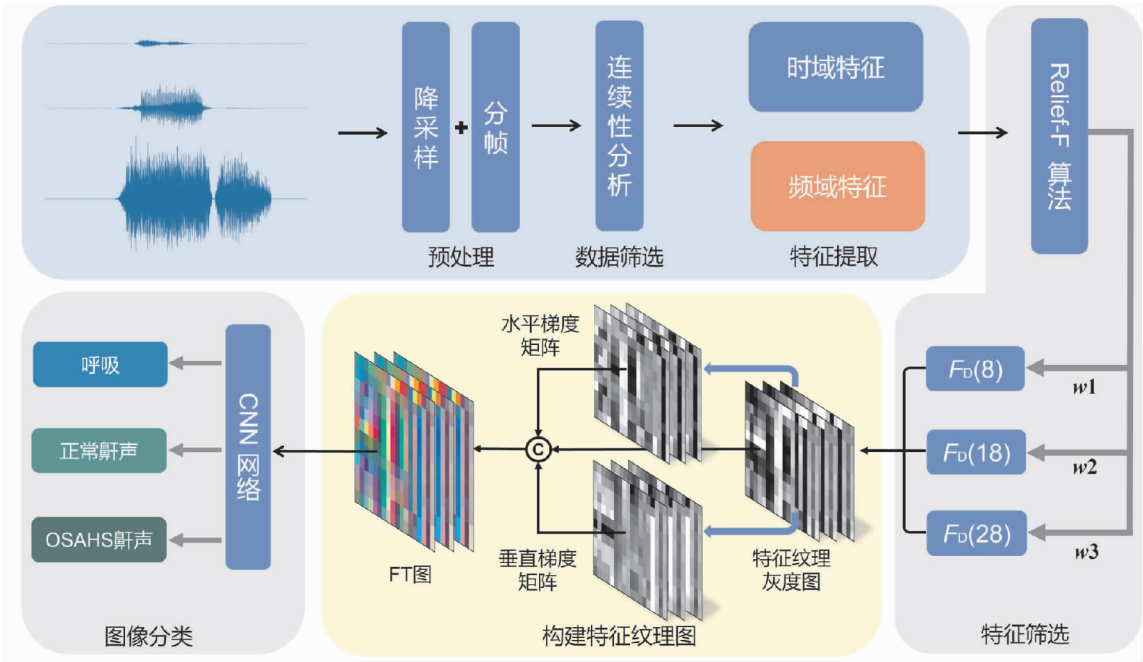


图 1 基于 FT 图的鼾声分类方法整体框图

Fig. 1 Overall block diagram of snore classification method based on FT map

OSAHS 鼾声和呼吸暂停时域波形图如图 2 所示。其中的空白段为 OSAHS 鼾声中出现的呼吸暂停,其打断了打鼾事件的连续性,并且会导致呼吸暂停出现前的打鼾事件和呼吸暂停出现后的打鼾事件的同质性降低,这会使病理性打鼾事件难以被识别,增加与正常人打鼾事件的分类难度。

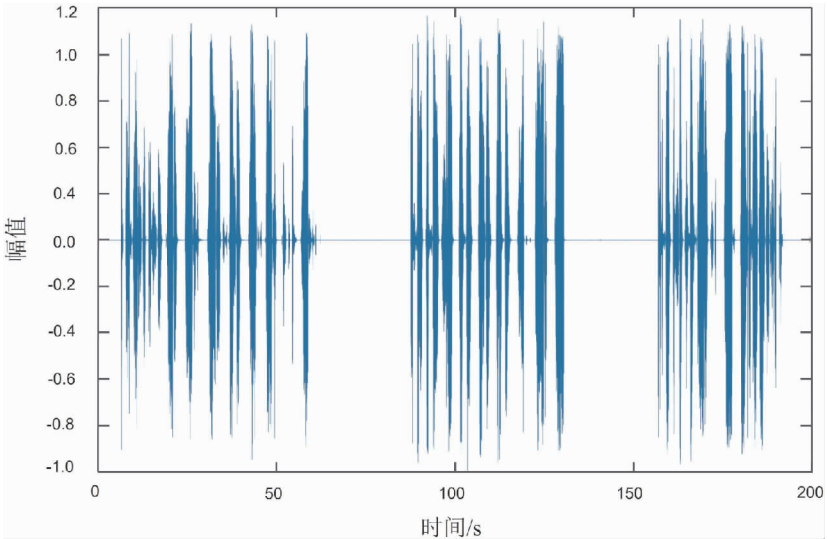


图 2 OSAHS 鼾声和呼吸暂停时域波形图

Fig. 2 Waveform of OSAHS snore and apnea in time-domain

如图 3 所示,呼吸暂停之间的连续打鼾事件持续时间范围为 1 ~ 2 min。功率谱密度质量因子、时间质量因子^[14]分别用于从频域和时间分析信号的连续性。考虑病理性鼾声的时频特性,本文提出了呼吸连续质量因子 $Q_1(k)$ 来确定呼吸信号连续的数量范围。

首先,计算连续的 N 个信号功率谱密度之间

的相似系数:

$$r = \frac{\sum_{i=1}^n P_a(i) P_b(i)}{\sqrt{\sum_{i=1}^n P_a^2(i)} \sqrt{\sum_{i=1}^n P_b^2(i)}} \tag{1}$$

式中: n 为功率谱密度的采样点数目; P_a 、 P_b 分别为计算序列中 2 个呼吸信号的功率谱密度。

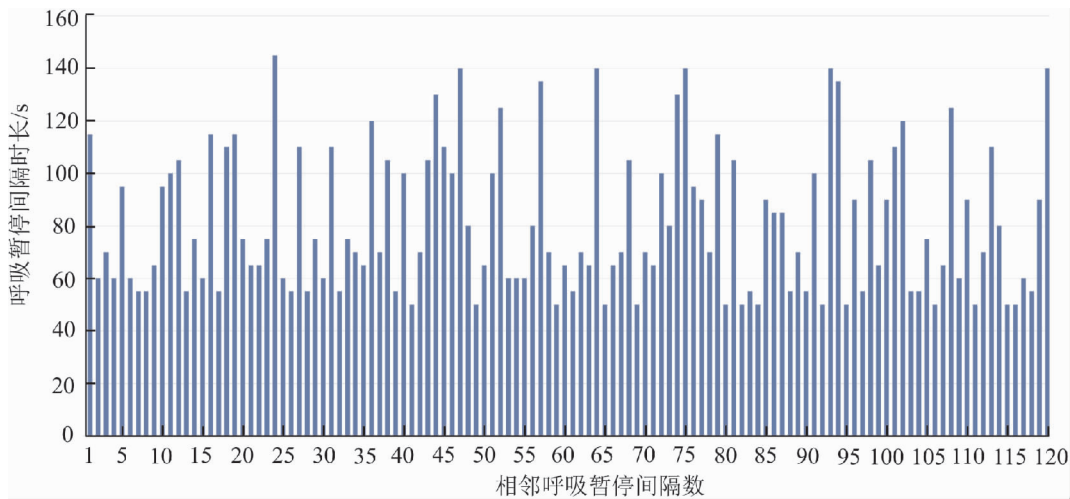


图 3 呼吸暂停之间的连续打鼾事件持续时间统计图

Fig. 3 Statistical chart of continuous snoring event duration between apneas

基于功率谱密度间的相似系数,计算呼吸信号的功率谱密度质量因子:

$$Q_c(k) = \frac{2 \left(\sum_{i=1}^N \left(\sum_{j=1}^N r_{i,j} \right) - \sum_{i=1}^N r_{i,i} \right)}{C_N^{N-2}} \tag{2}$$

式中: $r_{i,j}$ 为第 k 组中信号 i 和 j 的功率谱密度相似系数; $r_{i,i}$ 为第 k 组中信号 i 与其本身的功率谱密度相似系数。

对于每组连续呼吸信号分别计算 N 个事件有效呼吸段在该组的时长占比,利用均分相乘值最大原理,基于时长占比得到时间质量因子:

$$Q_t(k) = N^N \prod_{i=1}^N (R_i(k)) \tag{3}$$

式中, R_i 为第 k 组中信号 i 有效时长与第 k 组总有效时长的比值。

根据功率谱密度质量因子和时间质量因子得到各组呼吸信号的连续质量因子:

$$Q_1(k) = \frac{1}{2} (Q_c(k) + Q_t(k)) \tag{4}$$

连续质量因子结果如图 4 所示, Q_{11} 、 Q_{12} 和 Q_{13} 分别为呼吸音、正常鼾声和 OSAHS 鼾声的连续质量因子,质量因子 Q_1 均随 N 的取值变大而降低, Q_{13} 在改变连续组内事件个数 N 的条件下,始终低于呼吸事件和正常人打鼾事件的连续因子 Q_{11} 和 Q_{12} ,且随 N 的增加差距变大。从医学角度分析,OSAHS 患者的病理性打鼾事件的不稳定性更强,鼾声周期更加凌乱,从而导致连续质量因子 Q_{13} 较呼吸和正常鼾声的更低,连续性更弱。经过实验分析,当 $N = 12$ 时 Q_{12} 、 Q_{13} 的差异明显扩大,但当 N 继续增大后 Q_{12} 、 Q_{13} 的差值没有进一步扩大,且

均在持续减小,故以 $L = 0.3$ 为阈值,将 Q_1 值高于 0.3 的组视作有效连续组,以 $N = 12$ 作为实现鼾声分类研究中连续组内事件数量的基准。

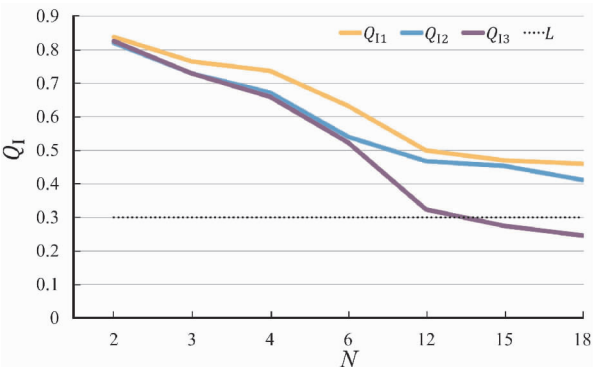


图 4 不同 N 值下连续质量因子 Q_1 结果

Fig. 4 Results of continuous quality factor Q_1 by different N values

2.2 一维时频特征提取

本研究提取 51 维时频域特征表征呼吸声、正常鼾声和 OSAHS 鼾声的特异性,具体特征描述如表 1 所示。

时域特征从原始信号波形、短时能量和短时过零率 3 个角度进行提取。从整体时域信号中提取的 17 维特征描述了音频信号中幅值、峰值、幅度、变化程度、整体强度、尖锐程度、波形特征等性质;8 维短时能量特征可以获取信号的能量分布,刻画音频的强度、变化和发生时刻;7 维短时过零率特征则可以从时域角度反应音频的频率和振动特性。

设置滑动窗口的窗长为 20 ms,滑动步长为 10 ms,计算短时能量 e 和短时过零率 z 公式如下:

$$e = \sum_{n=1}^W (a_n)^2$$

(5)

$$\begin{cases} z = \frac{1}{2} \sum_{n=1}^W | \operatorname{sgn}(a_{n+1}) - \operatorname{sgn}(a_n) | \\ \operatorname{sgn}(a_n) = \begin{cases} 1, a_n \geq 0 \\ -1, a_n < 0 \end{cases} \end{cases}$$

(6)

式中: $W = 400$; a_n 为离散信号的幅值。

短时能量特征除能量差累积(energy differe-

nce accumulation, EDA, 式中简记 A_{EDA})^[15] 外, 与短时过零率特征有 7 维特征都是各自的统计特征^[16]。

$$A_{\text{EDA}} = \sum_{i=1}^K | e_{i+1} - e_i |$$

(7)

式中: 帧数 $K = 481$; e_i 表示第 i 帧的短时能量。EDA 能从帧与帧之间短时能量的变化角度反映不同呼吸状态的特性。

表 1 时频特征描述

Tab. 1 Time-frequency feature description

特征类别	维数	特征名称
时域波形特征	17	完整信号的最大值、最小值、峰值、峰峰值、均值、绝对平均值、方根幅值、方差、标准差、有效值、峭度、偏度、波形因子、峰值因子、脉冲因子、裕度因子、余隙因子
短时能量特征	8	短时能量的均值、中值、方差、标准差、最值、峰度、偏度、EDA ^[15]
短时过零率特征	7	短时过零率的均值、中值、方差、标准差、最值、峰度、偏度
频谱特征	19	频谱熵、频谱质心、频谱拓展、频谱平坦度、频谱衰减点、频谱偏度、频谱斜度、频谱均值、频谱质心变化特征、频谱功率收敛特征

从音频信号中提取的 19 维频域特征中, 有 8 维特征分别为频谱熵、频谱质心、频谱拓展、频谱平坦度、频谱衰减点、频谱偏度、频谱斜度、频谱均值^[17-19]。其他 11 维特征如表 2 所示, $F_{\text{cl}} \sim F_{\text{c4}}$ 描述

的是音频信号的主频率的位置变化程度, $F_{\text{sp1}} \sim F_{\text{sp7}}$ 描述频谱功率的收敛性^[20]。 f_c 是频谱质心; f_i 是频谱幅值 x_i 的中心频率; X_{mean} 为频谱均值。

表 2 主频位置和功率谱收敛性描述特征

Tab. 2 Descriptive features of dominant frequency position and power spectrum convergence

特征	表达式	特征	表达式	特征	表达式	特征	表达式
F_{cl}	$\sqrt{\frac{1}{N-1} \sum_{i=1}^N (f_i - f_c)^2 X_i}$	F_{c2}	$\sqrt{\frac{\sum_{i=1}^N f_i^2 X_i}{\sum_{i=1}^N X_i}}$	F_{c3}	$\sqrt{\frac{\sum_{i=1}^N f_i^4 X_i}{\sum_{i=1}^N f_i^2 X_i}}$	F_{c4}	$\frac{\sum_{i=1}^N f_i^2 X_i}{\sqrt{\sum_{i=1}^N X_i \sum_{i=1}^N f_i^4 X_i}}$
F_{sp1}	$\sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - X_{\text{mean}})^2}$	F_{sp2}	$\frac{\sum_{i=1}^N (X_i - X_{\text{mean}})^3}{(N-1) F_{\text{c2}}^3}$	F_{sp3}	$\frac{\sum_{i=1}^N (X_i - X_{\text{mean}})^4}{(N-1) F_{\text{c2}}^4}$	F_{sp4}	$\frac{F_{\text{cl}}}{f_c}$
F_{sp5}	$\frac{\sum_{i=1}^N (f_i - f_c)^3 X_i}{(N-1) F_{\text{cl}}^3}$	F_{sp6}	$\frac{\sum_{i=1}^N (f_i - f_c)^4 X_i}{(N-1) F_{\text{cl}}^4}$	F_{sp7}	$\frac{\sum_{i=1}^N (f_i - f_c)^{1/2} X_i}{(N-1) F_{\text{cl}}^{1/2}}$		

2.3 一维特征有效性筛选

本研究采用 ReliefF 算法进行特征筛选。该算法通过计算样本之间的邻近关系, 并基于这些邻近关系对特征进行加权, 以评估特征与目标类别之间的相关性, 同时识别出既独立又最相关的特征, 以及那些相关但互不冗余的特征。考虑到数据的连续性, 将最近邻参数 k 设置为 12, 以进一步提高特征筛选的准确性和效率, 从而提升分

类模型的性能。

归一化特征权值排序如图 5 所示, 应用 ReliefF 算法将 51 维特征的特征的权值降序重排, 得到对鼾声分类最佳的多个特征。将权值归一化至 0 ~ 1 后, 分析发现归一化权值 w 分别在 0.5、0.25、0.18 时出现了跳变, 将出现跳变的权值作为阈值, 得到 3 组筛选后的一维特征集, 命名为 $F_D(M)$, $M = 8, 18, 28$ 代表特征集中的特征个数。

每一个音频数据样本的特征数据表示为向量 $S(x,y)$, x 和 y 分别代表样本号和该样本对应特征的序号。为了验证特征筛选步骤的有效性,将未筛选的特征集也作为一组实验的特征数据 ($M=51$), 并作为构建特征向量的基础。

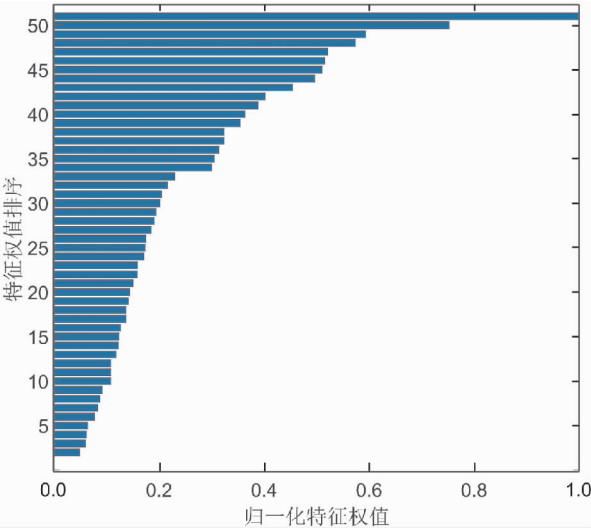


图5 归一化特征权值排序图

Fig.5 Ranking chart of normalized feature weight

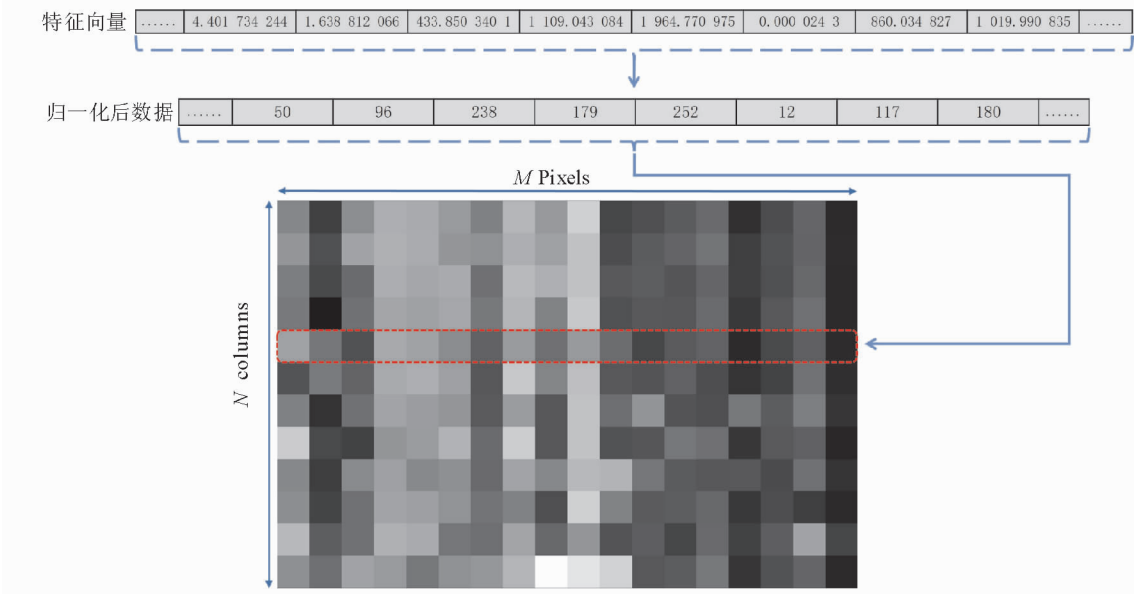


图6 构建特征纹理灰度图过程

Fig.6 Schematic diagram of the process for constructing feature texture grayscale map

将归一化向量中的每个值作为绘制灰度图像的灰度像素值,再调用灰度调色板将其映射为特征灰度像素块,并按位置拼接成灰度条,即将数值域的特征向量映射至图6中红色框内的可视化灰度图像。

2.4.2 求取灰度图水平、垂直梯度

对特征矩阵灰度图像进行分析时,提取每个

2.4 FT图构建

2.4.1 构建特征纹理灰度图

经典的一维时频域特征不能有效区分鼾声信号,特别是病理性鼾声,因此从高维图像的角度切入,进一步提升鼾声分析的准确性。一维特征集 $F_D(M)$ 转换为特征纹理灰度图的过程如图6所示,这是一组连续的12例OSAHS鼾声信号对应的特征纹理灰度图的构建过程。图6中最上方的一行数据代表一个鼾声对应的 $F_D(M)$ 中的一个特征行向量,将数值样本的数据通过如下公式归一化后映射到 $[0 \sim 255]$ 范围,匹配8位灰度图像数据。

$$I(x,y) = \frac{S(x,y) - \min(S_y)}{\max(S_y) - \min(S_y)} \tag{8}$$

式中: x,y 分别表示行数和列数; $S(x,y)$ 、 $I(x,y)$ 分别表示特征矩阵中 x 行 y 列的原数据和归一化数据; S_y 代表 y 列的所有数据。

灰度纹理图像的水平梯度和垂直梯度矩阵。水平梯度 G_x 和垂直梯度 G_y 的求取公式为:

$$G_x = I(x+1,y) - I(x-1,y) \tag{9}$$

$$G_y = I(x,y+1) - I(x,y-1) \tag{10}$$

式中: (x,y) 为灰度图对应灰度矩阵中的像素位置, x 为横坐标, y 为纵坐标。

对于水平梯度,通过计算 (x,y) 处的像素值

与其左侧 $(x-1,y)$ 处的像素值之差 $I(x,y)-I(x-1,y)$,以及 (x,y) 处的像素值与其右侧 $(x+1,y)$ 处的像素值之差 $I(x+1,y)-I(x,y)$,然后将2个差值相加,得到 (x,y) 处的水平梯度 G_x 。同理,也能得到垂直梯度 G_y 。

从横向分析,对每个灰度条而言,水平梯度 G_x 捕获单个数据不同特征之间的动态信息,并将特征之间的数值差异性量化。从纵向分析,则是针对特定的某一特征,垂直梯度 G_y 捕获了同一组内连续信号同一特征间的变化信息,并将这种稳定性量化。

通过求取梯度的方式,将同一数据特征间的差异性和同类别数据间的相似性、稳定性表达为水平、垂直梯度矩阵的形式。

2.4.3 构造 FT 图

按照构造特征纹理灰度图相同的映射原则,将水平、垂直梯度矩阵内的数据映射到 $[0 \sim 255]$ 的灰度像素范围,将2种梯度求得的纹理特性图像化。重新调用调色板,将特征纹理灰度矩阵、水平梯度矩阵、垂直梯度矩阵按照 R、G、B 三通道进行像素块填色。得到3个通道的图像,使用 MATLAB 软件中 cat 函数将三通道进行串联,构建出三通道的 FT 图。

图7为呼吸声、正常鼾声和 OSAHS 鼾声的 FT 图。不难发现,3类 FT 图有明显的差别:呼吸声 FT 图相较于另外两者,像素块的颜色、亮度和纹理非常相似,过渡均匀性和一致性更强。而 OSAHS 鼾声 FT 图相比正常鼾声 FT 图,像素块之间出现了明显的分界,呈现出马赛克般的无序性。



图7 3类呼吸信号 FT 图

Fig. 7 FT maps of three types of respiratory signals

通过将灰度纹理图转化为 FT 图,使不同类型的呼吸数据从二维图像分析的角度上表征性更强。

2.5 CNN 分类网络

本研究使用的 CNN 网络的主要结构由输入层、卷积层、批量归一化层、ReLU 激活函数、最大池化层、分类层等构成。网络架构的设计兼顾了高效性与深度特征提取的能力,具体结构如图8所示。

首先,采用一个16通道的 5×5 卷积层,有效捕获了多样化的局部特征;接着,采用一个步长为2的 2×2 最大池化层,该层通过选取每个 2×2 区域内的最大值作为输出,有效地减小了特征图的尺寸,同时保留了关键的特征信息;在降维之后,堆叠了一个32通道的 3×3 卷积层,这一层通过32个不同的卷积核进一步挖掘和组合特征,生成更为丰富和抽象的特征图;最后,将特征图展平并送入全连接层进行分类决策,通过一个输出维度为3的全连接层得到每个类别的预测得分,并通过 softmax 层将这些得分转换为概率形式。其中,在2个卷积层之后均引入了归一化层和激活层。归一化层用来保持训练过程的稳定性并防止过拟合,激活层则是优化数据分布,加速训练过程,激活层的加入为网络引入了非线性因素,使得网络能够学习更为复杂的映射关系。

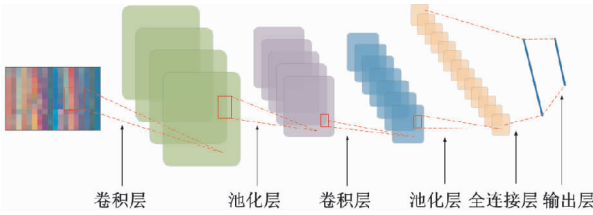


图8 CNN 网络结构图

Fig. 8 CNN network structure diagram

3 实验结果及分析

3.1 实验设置

为了分析参数对病理性鼾声分类的影响程度,本研究通过改变特征集 $F_D(M)$ 的参数 M ,共得到4种不同尺寸的 FT 图,如图9所示。对4种 FT 图设置了3组分类对比实验模型。

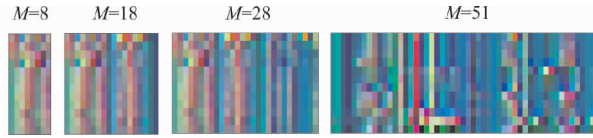


图9 不同 M 取值对应的 FT 图

Fig. 9 FT maps corresponding to different M values

模型1 以单个数据作为样本单位,将特征集 $F_D(M)$ 送入传统机器学习分类器,包括决策树、朴素贝叶斯、支持向量机、K-最近邻分类器进行分类实验。

模型2 以组作为样本单位,采用 CNN 对特征纹理灰度图进行分类实验。

模型3 以组作为样本单位,采用 CNN 对 FT

图进行分类实验。

3.2 评估标准

本研究使用混淆矩阵、受试者工作特性曲线 (receiver operating characteristic curve, ROC) 和二级分类指标准确率 (accuracy, 式中简记 A)、精准率 (precision, 式中间记 P)、召回率 (recall, 式中简记 R) 作为描述分类模型分类效果的评估指标。

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

$$P = \frac{TP}{TP + FP} \tag{12}$$

$$R = \frac{TP}{TP + FN} \tag{13}$$

式中: TP (true positive)、TN (true negative)、FP (false positive) 和 FN (false negative) 分别代表真阳性、真阴性、假阳性和假阴性。

3.3 结果分析与讨论

应用上文所述的 3 种分类模型对本实验的睡眠呼吸声学数据集进行分类实验。

模型 1 的分类结果如图 10 所示,结果表明,在传统分类器中,SVM 分类器分类精度最高。分类效果随参数 M 的减小呈现出先增大后减小的趋势,这与图 11 中模型 2 和模型 3 分类精度随 M 取值而变化的趋势是一致的。这说明从 $M = 51$ 筛选到 $M = 18$ 的过程中消除了冗余特征,而从 $M = 18$ 到 $M = 8$ 的筛选过程则存在一定程度的特征丢失。

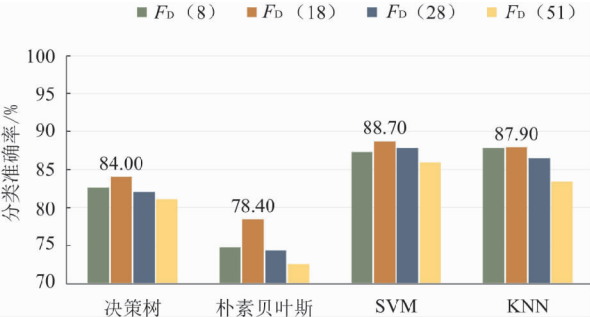


图 10 模型 1 分类实验结果

Fig. 10 Classification results of model 1

模型 2 和模型 3 的分类结果如图 11 所示,可以看出,当参数 M 固定时,模型 3 的分类效果均高于模型 2。这验证了求取灰度纹理图的水平梯度 G_x 和垂直梯度 G_y 构建的 FT 图在鼾声分类中更有效,也表明了同组呼吸事件之间的关联特性能够被量化用于进一步分类研究。

3 种实验模型的鼾声分类结果如表 3 所示,

呼吸声、正常鼾声和 OSAHS 鼾声三分类方案中,模型 3 在 $M = 18$ 时,取得最佳的分类效果,准确率达到 97.2%,精准率和召回率达到了 96.7% 和 93.3%,其混淆矩阵结果和 ROC 曲线结果如图 12 和 13 所示。

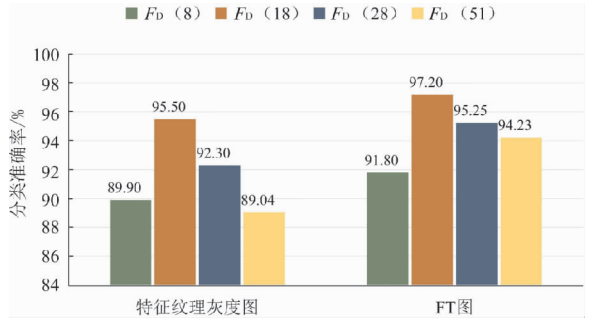


图 11 模型 2、模型 3 分类实验结果

Fig. 11 Classification results of model 2 and 3

表 3 不同分类模型实验结果

Tab. 3 Classification results of different models

分类模型	特征集	准确率 /%	精准率 /%	召回率 /%
模型 1	$F_D(51)$	85.9	73.1	82.9
	$F_D(28)$	87.8	73.7	84.7
	$F_D(18)$	88.7	77.0	87.4
	$F_D(8)$	87.3	73.3	84.3
模型 2	$F_D(51)$	89.0	86.0	80.3
	$F_D(28)$	92.3	89.7	85.3
	$F_D(18)$	95.5	93.8	90.7
	$F_D(8)$	89.9	87.6	82.5
模型 3	$F_D(51)$	94.2	91.6	87.7
	$F_D(28)$	95.1	93.4	89.4
	$F_D(18)$	97.2	96.7	93.3
	$F_D(8)$	91.8	89.3	84.3

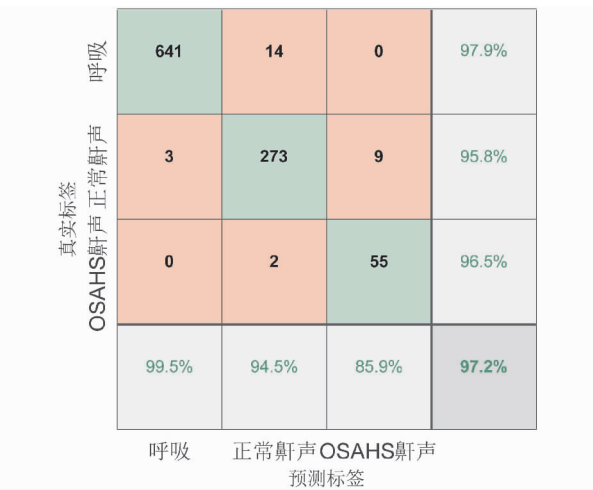


图 12 最佳分类模型混淆矩阵

Fig. 12 Confusion matrix of the optimal classification model

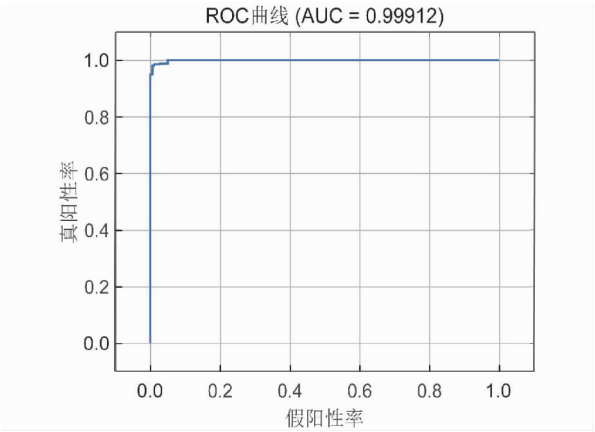


图 13 最佳分类模型 ROC 曲线图

Fig. 13 Confusion matrix of the optimal classification model

表 4 FT 图分类法与其他方法对比

Tab. 4 Comparison of the FT map classification method with other methods in the literature

分类模型	准确率/%	精准率/%	召回率/%
SDCC、MFCC 特征 + KNN ^[8]	95.3	95.5	95.1
Cytex 图 + CNN ^[9]	70.7	70.0	73.6
声谱图 + Alexnet、AGG19 ^[10]	84.9	85.0	85.8
声谱图纹理特征 + SVM ^[11]	73.5	73.9	75.4
Mel 谱图 + NewAGG16 ^[12]	94.1	94.2	94.1
FT 图 + CNN	97.2	96.7	93.3

4 结语

本文针对呼吸音、正常鼾声和 OSAHS 鼾声音频信号,研究其时频域特征和连续呼吸事件间的关联特性,提出一种基于 FT 图的鼾声分类方法,实现了对鼾声精确分类的目的。

从实验结果来看,将一维特征转化为二维特征图的方法切实有效,对鼾声分类准确率可达 97.2%。这是因为特征纹理灰度图的水平梯度 G_x 和垂直梯度 G_y 能刻画特征之间和连续呼吸事件之间的关联特性并用于提升鼾声分类效果。特征的筛选过程对分类结果也有直接的影响,特征权值的阈值设置过大会丢失部分有价值的特征,设置过小会出现特征冗余的现象。这是特征挖掘和优化技术的关键切入点,为鼾声和 OSAHS 病理性鼾声的分类诊断提供了新思路。

在未来的工作中,将通过获取更多受试者的呼吸和鼾声数据,验证和提升本研究方法的鲁棒性、泛化能力,并进一步优化整个分类模型。同时,将继续探索从 FT 图中挖掘更多对于鼾声识

别和分类有价值的信息,并研究是否能从图像纹理特征分析角度对 OSAHS 鼾声进行识别和分类。

将基于 FT 图的鼾声分类模型与其他研究者的分类模型进行比较,在本实验室自测数据集上进行验证,结果如表 4 所示。可以看出,基于前 5 种模型得到的呼吸声、正常鼾声、OSAHS 鼾声三分类的分类准确率分别为 95.3%、70.7%、84.9%、73.5% 和 94.1%,均低于本文提出的基于 FT 图和 CNN 的分类模型的准确率 97.2%。对比经典的一维特征、二维图像及基于图像的特征分类模型,本研究所提出的基于一维有效特征构建的 FT 图和 CNN 结合的分类模型对病理性鼾声分类效果最优。

别和分类有价值的信息,并研究是否能从图像纹理特征分析角度对 OSAHS 鼾声进行识别和分类。

参考文献

[1] MIROSHNICHENKO N A, PIKHTILEVA N A. Snoring and obstructive sleep apnea syndrome as possible causes of auditory and tubal disorders[J]. The Bulletin of Contemporary Clinical Medicine, 2021,14(6): 46-50.

[2] 黄志华,赵智慧,赵青. 心血管疾病患者阻塞性睡眠呼吸暂停评估与管理专家共识(2024 版)[J]. 中国循环杂志,2024,39(5):417-432.

HUANG Z H, ZHAO Z H, ZHAO Q. Expert consensus on the evaluation and management of obstructive sleep apnea in patients with cardiovascular disease (2024 edition) [J]. Chinese Circulation Journal, 2024, 39(5): 417-432.

[3] LEEY H. Implications of obstructive sleep-related breathing disorder in dentistry: Focus on snoring and obstructive sleep apnea[J]. Dental Research and Oral Health, 2022, 5(4): 74-82.

[4] 郑彤,姚侃,钱铁峰,等. 成年打鼾人群患阻塞性睡眠呼吸暂停的危险因素分析[J]. 中国口腔颌面外科杂志,2024,22(2):137-142.

- ZHENG T, YAO K, QIAN Y F, et al. Risk factors for obstructive sleep apnea in snoring adults[J]. Chinese Journal of Oral and Maxillofacial Surgery, 2024, 22(2): 137-142.
- [5] 赵小强,潘毅,侯雨函,等. 老年人阻塞性睡眠呼吸暂停综合征的诊断方法研究进展[J]. 中国实用医药, 2023, 18(23): 165-169.
- ZHAO X Q, PAN Y, HOU Y H, et al. Research progress on diagnostic methods for obstructive sleep apnea syndrome in the elderly[J]. China Practical Medicine, 2023, 18(23): 165-169.
- [6] 侯丽敏,刘焕成,施晓宇,等. 基于人工神经网络的鼾声相关信号的分类[J]. 复旦学报(自然科学版), 2019, 58(4): 530-536.
- HOU L M, LIU H C, SHI X Y, et al. Classification of snore-related signals based on artificial neural networks[J]. Journal of Fudan University (Natural Science), 2019, 58(4): 530-536.
- [7] QIAN K, SCHMITT M, JANOTT C, et al. A bag of wavelet features for snore sound classification[J]. Annals of Biomedical Engineering, 2019, 47(4): 1000-1011.
- [8] FANG Y, LIU D B, ZHAO S X, et al. Improving OS-AHS prevention based on multidimensional feature analysis of snoring[J]. Electronics, 2023, 12(19): 4148.
- [9] BAKHSHI A, HARIMI A, CHALUP S. CyTex: Transforming speech to textured images for speech emotion recognition[J]. Speech Communication, 2022, 139: 62-75.
- [10] AMIRIPARIAN S, GERCZUK M, OTTL S, et al. Snore sound classification using image-based deep spectrum features[C] // Interspeech 2017. Stockholm, Sweden; ISCA, 2017: 3512-3516.
- [11] DEMIR F, SENGUR A, CUMMINS N, et al. Low level texture features for snore sound discrimination[C] // 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI; IEEE, 2018: 413-416.
- [12] 邓志平,王冬霞,马晓冬,等. 基于注意力机制的 NewVGG16-BiGRU 鼾声分类[J]. 计算机应用, 2023, 43(S1): 276-280.
- DENG Z P, WANG D X, MA X D, et al. NewVGG16-BiGRU snoring classification based on attention mechanism[J]. Journal of Computer Applications, 2023, 43(S1): 276-280.
- [13] ZHAO S X, FANG Y, WANG W B, et al. Analysis of sleeping respiratory signal utilizing frequency energy features[C] // 2022 5th International Conference on Information Communication and Signal Processing (ICICSP). Shenzhen, China; IEEE, 2022: 86-90.
- [14] 王佳明,张安琪,郝宏燕,等. 基于功率谱密度相似性的最佳心音信号选择[J]. 中国医学物理学杂志, 2022, 39(11): 1401-1406.
- WANG J M, ZHANG A Q, HAO H Y, et al. Optimal heart sound signal selection based on power spectral density similarity[J]. Chinese Journal of Medical Physics, 2022, 39(11): 1401-1406.
- [15] DENG B Y, XUE B, HONG H, et al. Decision tree based sleep stage estimation from nocturnal audio signals[C] // 2017 22nd International Conference on Digital Signal Processing. London, UK; IEEE, 2017: 1-4.
- [16] XUE B Y, DENG B, HONG H, et al. Non-contact sleep stage detection using canonical correlation analysis of respiratory sound[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(2): 614-625.
- [17] 董武亮,吴宁,彭勇,等. 基于频谱特征的变压器音频监测方法研究[J]. 电声技术, 2025, 49(4): 170-172.
- DOND W L, WU N, PENG Y, et al. Research on transformer audio monitoring method based on spectrum characteristics[J]. Audio Engineering, 2025, 49(4): 170-172.
- [18] 李哲,陈宇,张博,等. 基于深度学习的减压音乐重构研究[J]. 计算机应用与软件, 2022, 39(8): 158-162.
- LI Z, CHEN Y, ZHANG B, et al. Research on decompression music reconstruction based on deep learning[J]. Computer Applications and Software, 2022, 39(8): 158-162.
- [19] 李伟峰,周文进. 基于频谱特征的舰船机械设备振动数据挖掘与故障定位研究[J]. 自动化与仪器仪表, 2019(7): 10-13.
- LI W F, ZHOU W J. Research on vibration data mining and fault location of marine machinery based on spectrum features[J]. Automation & Instrumentation, 2019(7): 10-13.
- [20] LEI Y G, HE Z J, ZI Y Y, et al. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs[J]. Mechanical Systems and Signal Processing, 2007, 21(5): 2280-2294.

(编辑 李 静)