

结合 CNN-LSTM-SVM 的特征融合在肺音分析中的应用^①



赵 静¹, 杜永飞¹, 韦海成², 张志鹏¹, 许 洋³

¹(宁夏大学 信息工程学院, 银川 750021)

²(北方民族大学 医学技术学院, 银川 750021)

³(北方民族大学 电气信息工程学院, 银川 750021)

通信作者: 赵 静, E-mail: zhaojing_nx@nxu.edu.cn

摘 要: 本研究致力于提升深度学习在肺音分析领域的应用效率和准确性。针对现有深度学习模型在肺音分析中表现出的鲁棒性不足和泛化能力有限的问题, 本研究提出了一种方法, 该方法通过整合卷积神经网络 (convolutional neural network, CNN)、长短时记忆 (long short-term memory, LSTM) 网络和支持向量机 (SVM), 实现了对肺音信号的高效和深入分析。首先对肺音信号进行预处理, 提取出重构信号和其对应的希尔伯特谱图; 其次设计并构建了一个集成 CNN、LSTM 和 SVM 的深度学习网络模型; 最后将处理后的信号数据输入到 CNN-LSTM-SVM 的深度学习网络中, 以提取并融合肺音信号的时域和频域特征。实验结果表明, 该方法在召回率、精确率和 $F1$ -score 这 3 个关键性能指标上分别达到 96.20%、96.56% 和 0.96 的高水平。这些结果证实了所提方法的高效性和可靠性, 为肺部疾病的早期诊断提供了一种技术途径, 并有潜力显著提升临床诊断的速度和准确性。

关键词: 肺音分析; 特征融合; 变分模态分解; 卷积神经网络; 长短时记忆网络; 支持向量机

引用格式: 赵静, 杜永飞, 韦海成, 张志鹏, 许洋. 结合 CNN-LSTM-SVM 的特征融合在肺音分析中的应用. 计算机系统应用, 2026, 35(1): 219-227. <http://www.c-s-a.org.cn/1003-3254/10057.html>

Application of Feature Fusion Combined with CNN-LSTM-SVM in Lung Sound Analysis

ZHAO Jing¹, DU Yong-Fei¹, WEI Hai-Cheng², ZHANG Zhi-Peng¹, XU Yang³

¹(School of Information Engineering, Ningxia University, Yinchuan 750021, China)

²(School of Medical Technology, North Minzu University, Yinchuan 750021, China)

³(School of Electrical and Information Engineering, North Minzu University, Yinchuan 750021, China)

Abstract: This study is dedicated to enhancing the application efficiency and accuracy of deep learning in lung sound analysis. In view of the insufficient robustness and limited generalization capabilities of existing deep learning models in lung sound analysis, it proposes a method that integrates the convolutional neural networks (CNN), long short-term memory network (LSTM), and support vector machine (SVM) to achieve efficient and in-depth analysis of lung sound signals. The method begins with the preprocessing of lung sound signals to extract reconstructed signals and their corresponding Hilbert spectra. Secondly, a deep learning network model that integrates CNN, LSTM, and SVM is designed and built. Finally, the processed signal data are input into the CNN-LSTM-SVM deep learning network to extract and fuse the time-domain and frequency-domain features of lung sound signals. Experimental results show that the method achieves high levels of 96.20% for the recall, 96.56% for accuracy, and 0.96 for $F1$ -score. These results confirm the efficiency and reliability of the proposed method, providing a new technological approach for the early diagnosis of

① 基金项目: 宁夏回族自治区自然科学基金 (2022AAC03006); 宁夏先进智能感知与控制技术创新团队支持计划

收稿时间: 2025-06-30; 修改时间: 2025-07-21, 2025-08-13; 采用时间: 2025-08-19; csa 在线出版时间: 2025-12-01

CNKI 网络首发时间: 2025-12-02

lung diseases, and potentially significantly enhancing the speed and accuracy of clinical diagnosis.

Key words: lung sound analysis; feature fusion; variational mode decomposition; convolutional neural network (CNN); long short-term memory (LSTM) network; support vector machine (SVM)

肺部疾病引起的呼吸系统疾病占全球死亡率的7%,其中慢性阻塞性肺病(COPD)是死亡的第3大原因^[1],COPD给患者造成的经济负担已达世界第5位,中国第1位^[2].肺部疾病的早期发现可有效降低致死率^[3].由此可见,对于肺部疾病的预防以及早期诊断的重要性愈发凸显,在维护呼吸健康中扮演着至关重要的角色.

近年来,伴随现代数字信号处理技术的迅猛发展,肺音信号预处理领域吸引了众多学者的研究兴趣,以期提高信号质量、减小噪声干扰,为后续的肺部疾病分析和诊断奠定良好的基础.其中,Lozano等人^[4]利用集合经验模式分解(EEMD),用瞬时频率和包络特征分析肺音信号,连续不定音分类准确率为94.6%.Serbes等人^[5]采用短时傅里叶变换(STFT)和小波分析,利用肺音信号的时频域特征,达到49.86%的准确率.Aras等人^[6]利用Mel频率倒谱系数(MFCC)作为肺音信号的特征,最终LDA和SVM的分类准确率为93.3%.Abougabal等人^[7]基于离散小波变换的动态时间扭曲技术对儿童喘息的识别率为88.16%.Aziz等人^[8]先使用EMD分解肺音信号,再计算每个固有模态分量(IMF)的MFCC,最后用五折交叉验证,得到良好的分类效果.

随着人工智能技术在各种生物学领域的应用取得了重大进展^[9],利用机器学习和神经网络等人工智能方法对生理信号进行分类已成为研究热点^[10].在人工智能技术对肺音信号进行分类中,Sengupta等人^[11]提取肺音信号MFCCs中的统计特征,再采用ANN对正常肺音、喘息、爆裂音这3种肺音具有良好的分类效果.Kochetov等人^[12]提出了一种新的肺音分析方法,利用噪声掩蔽递归神经网络(NMRNN)对肺音信号进行分类,在ICBHI数据集上得分0.648.Demir等人^[13]用卷积神经网络模型提取特征,再用LDA-RSE区分喘息,其平均准确率为71.15%.Zhao等人^[14]利用BiGRU Attention-XGBoost模型来对呼吸音进行分类,准确度为72.45%.Petmezas等人^[15]将短时傅里叶变换(STFT)频谱图中提取的特征作为长短时记忆(LSTM)网络的输入,最终分类准确率为76.39%.Mondal等人^[16]利用

IMF以及人工神经网络(ANN),对正常肺音、喘息和爆裂音的分类准确率为94.16%.Shi等人^[17]先对小波分解对肺音信号进行降噪,再使用线性判别分析(LDA)降维,最后用BP神经网络进行肺音识别,最终干啰音与湿啰音准确率最优为92.5%.Acharya等人^[18]提出了一个深度CNN-RNN模型,该模型根据Mel频谱图对呼吸声音进行分类,得到58.47%的准确度.Altan等人^[19]先用希尔伯特-黄变换(HHT)提取特征,再用改进的深度信念网络(DBN)对COPD和哮喘分析,得到93.67%的准确率.Messner等人^[20]提出了一种利用光谱、时间和频率信息结合的多通道肺音分析的方法,最后利用递归卷积神经网络处理多通道肺音信号,区别正常肺音信号和肺纤维化得分为92%.Choi等人^[21]提出了一个使用注意力模块和深度学习的肺部疾病分类模型,通过改进VGGish并增加应用高效通道注意力模块(ECA-Net)的轻注意力连接模块,得到92.56%的准确率,但研究中未对肺音信号进行预处理,这可能限制了模型在特征提取方面的全面性,并影响了其在噪声环境中的鲁棒性.Srivastava等人^[22]采用CNN(MFCC)对肺音信号分类,最终F1-score为0.93,尽管该研究提出的慢性阻塞性肺病(COPD)检测系统在创新性方面值得肯定,但在网络模型的多样性和泛化能力方面仍有待提高.

针对上述问题,本文提出了一种结合CNN-LSTM-SVM的特征融合的肺音分析方法.针对预处理阶段数据库肺音信号存在噪声问题,采用变分模态分解(VMD)对肺音信号预处理,滤除低频噪声.针对网络模型单一和泛化能力差问题,首先将时域频域的特征融合后输入到CNN-LSTM模型中,然后利用支持向量机(support vector machine, SVM)替换CNN-LSTM模型的Softmax层,目的是提升模型的泛化性和准确率.

1 模型及方法

我们以COPD患者为例,与健康人群进行肺音信号分析,本文总体框图如图1所示.首先采集健康人群肺音信号数据并下载ICBHI公有数据集,然后利用

VMD 算法对肺音信号进行重构得到重构信号, 再对重构信号进行希尔伯特变换 (Hilbert transform, HT) 得到

希尔伯特谱图, 将重构信号与希尔伯特谱图输入 CNN-LSTM-SVM 网络模型中, 最后分析得到分类结果.

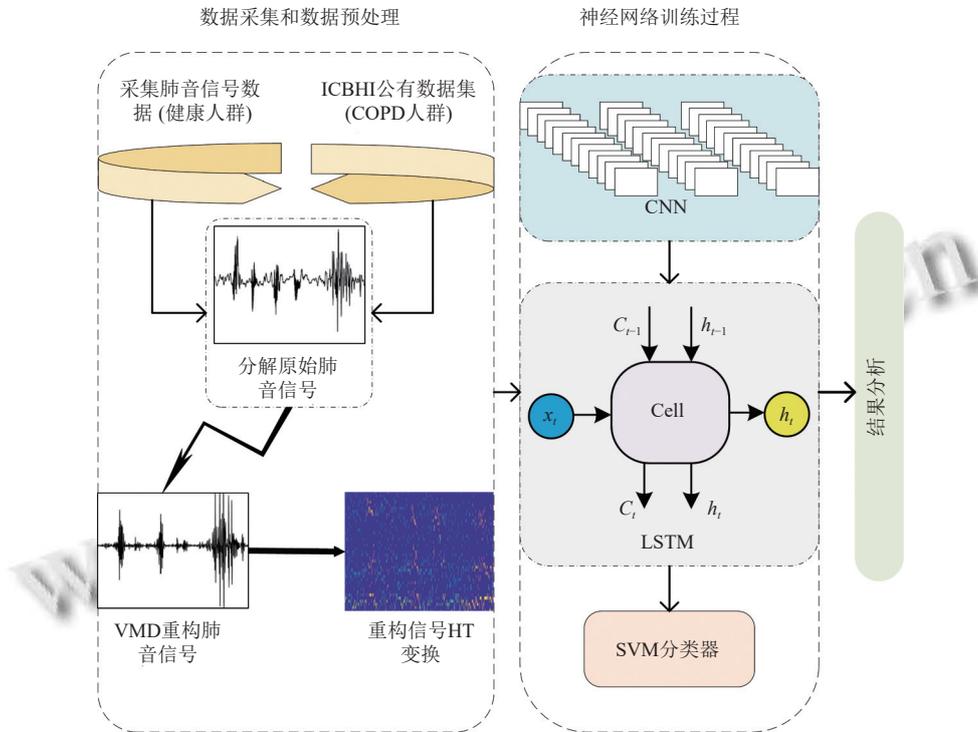


图1 结合 CNN-LSTM-SVM 的特征融合肺音分析框图

1.1 变分模态分解

变分模态分解 (variational mode decomposition, VMD) 是一种先进的信号处理技术, 用于将复杂信号分解为若干个带限 (band-limited) 的模态函数. 这种方法特别适用于处理非平稳、非线性的机械信号和人体生理信号. 该算法将输入的信号通过迭代, 将原始信号分解为设定个数的 IMF. 迭代是为了搜索出变分模型的一个最优解, 通过这个最优解来确定每条分量的带宽和中心频率.

本文利用 VMD 算法对肺音信号进行预处理重构得到重构信号, 再对重构信号进行希尔伯特变换得到希尔伯特谱图, 将重构信号与希尔伯特谱图输入网络模型, 最后优化网络模型.

1.2 长短时记忆网络

长短时记忆 (long short-term memory, LSTM) 神经网络, 是基于 RNN 提出的一种改进的模型^[23], 解决了 RNN 存在的长期依赖问题, 更有效地利用信号中的时间序列信息. 长短时记忆神经网络由 3 个门构成: 输入门、输出门和遗忘门. 遗忘门主要用于决定输入信息

保留和丢弃的比重, 以此保留部分记忆信息. 输入门主要用于将即时状态进行输入, 确定输入信息的比重. 输出门主要用于决定下一时刻隐藏状态的值和实现当前数据的输出, 结构图如图 2 所示.

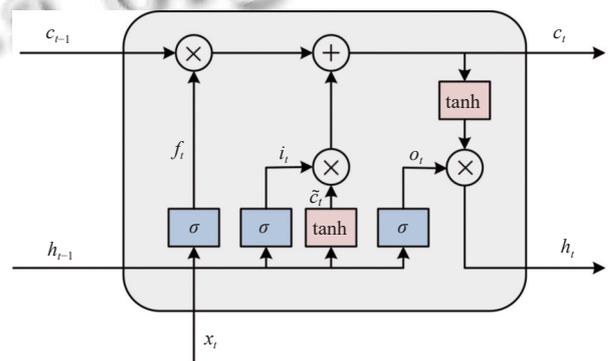


图2 LSTM 结构图

对于第 1 层第 t 个时间步, LSTM 单元的更新公式如式 (1)–式 (6) 所示:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \otimes \tanh(c_t) \quad (6)$$

其中, $x_t \in R^{d_x}$ 表示当前输入向量, h_{t-1} 和 c_{t-1} 分别表示上一时刻的隐状态和细胞状态, σ 为 Sigmoid 函数, \tanh 为非线性激活函数, \otimes 表示 Hadamard 积。

在肺音信号分析这一领域中, 传统信号处理方法通常集中于从频域提取特征, 而未能充分关注时间序列中的动态变化信息。肺音信号作为一类复杂的时间序列数据, 蕴含了大量有关呼吸系统健康状态的关键信息。为了更加全面深入地理解和分析这些信号, 本文采用 LSTM 来捕捉肺音信号中的时间序列特征, 从而能够提取出更为丰富和多层次的特征信息。

1.3 支持向量机

支持向量机 (SVM) 是一种基于监督学习的强大工具, 广泛应用于二元分类问题。作为一种广义线性分类器, SVM 的核心理念是通过寻找一个最佳的超平面来将不同类别的数据进行分隔。与众多分类算法不同, SVM 着重于构建一个能够最大化类别间距的决策边界, 这一过程被称为“最大边距 (maximal margin)”原则, 优化目标如式 (7) 所示:

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (7)$$

式 (7) 的约束条件是 $y_i(w^T x_i + b) \geq 1 - \xi_i$, $\xi_i \geq 0$, 其中, w 为超平面向量, b 为偏置, C 为惩罚参数, ξ_i 为松弛变量。对于非线性数据, 需要通过核函数将特征映射到高维空间, 本研究选择径向基函数 (RBF) 核: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$, 因其对复杂数据分布具有较强的适应性。具体配置如下: 核函数: RBF 核; 惩罚参数 C : 通过五折交叉验证的网格搜索优化, 范围为 [0.1, 1, 10, 100], 最终选定 $C=10$; 核参数 γ : 同样通过网格搜索优化, 范围为 [0.001, 0.01, 0.1, 1], 最终选定 $\gamma=0.01$ 。通过这种方式, SVM 不仅能有效处理线性可分的数据, 还可以通过使用核函数 (kernel function) 处理非线性的数据分布。

参数优化采用网格搜索结合交叉验证, 确保模型在肺音信号分类任务中的泛化性和准确性。本文使用支持向量机代替神经网络的 Softmax 层的原因在于其

最大边距特性可有效减少过拟合风险, 提升模型在小样本数据上的性能。

2 实验与分析

2.1 数据集

本文 COPD 数据采用 ICBHI 2017 呼吸音数据库中的呼吸音样本^[24]。该数据集已被广泛引用, 共 5.5 h, 采集自 126 名受试者的 920 条音频样本, 每条样本由专家注释。健康人群数据来自实验室自采的 489 组肺音信号。

2.2 评估指标

本文采用混淆矩阵进行模型的对比分析。检验网络模型的性能通常通过以下 4 个标准指标来判别模型的优劣: 准确率 (accuracy, Acc)、灵敏度 (sensitivity, Sen)、精确率 (precision, Pre) 和 $F1$ -score ($F1$)。本文通过真阳性 (true positive, TP)、真阴性 (true negative, TN)、假阳性 (false positive, FP) 与假阴性 (false negative, FN) 这 4 个通用的评价指标来更为清晰地对混淆矩阵进行表示。利用以上 4 个指标可以对 Acc 、 Sen 、 Pre 和 $F1$ 进行计算, 计算方法如下。

(1) 准确率: 用于计算正确样本量和样本总量的比值, 如式 (8) 所示:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

(2) 灵敏度: 表示的是所有正例中被分对的比例, 衡量了分类器性能, 如式 (9) 所示:

$$Sen = \frac{TP}{TP + FN} \quad (9)$$

(3) 精确率: 表示的是正例被分对的比例, 衡量了分类器对正例的识别能力, 如式 (10) 所示:

$$Pre = \frac{TP}{TP + FP} \quad (10)$$

(4) $F1$: 是调和平均指标, 反映模型对每种模型识别的综合性能, 如式 (11) 所示:

$$F1 = 2 \times \frac{\frac{TP}{TP + FP} \times \frac{TP}{TP + FN}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}} \quad (11)$$

2.3 实验平台

本文的实验环境配置为: Inter-i7-7700 CPU, GTX1070-8G GPU, 16 GB 内存; 基于 TensorFlow 深度学习框架; 模型结构由连续的卷积 CNN、LSTM 和

SVM 分类器构成; 每层的超参数设置为: Adam 超参数设置为 $\eta=0.001$, $\beta_1=0.9$, $\beta_2=0.999$; 在 CNN 结构中加入 dropout 参数防止模型过拟合.

2.4 数据预处理

原始肺音信号常混有心音和环境噪声, 导致其时域周期性特征不明显, 难以直接进行特征提取. 为消除干扰, 获取更纯净的肺音信号以利于后续分析, 需对原始信号进行预处理.

在诸多研究方法中, VMD 在噪声的处理上具有很好的效果^[25]. 图 3(a) 显示健康人群肺音信号受心音信号及环境噪声干扰, 导致其时域周期性特征被掩盖, 难以直接进行特征提取. 同时, 图 3(b) 则表明 COPD 患者呼吸过程中的咳嗽与喘息会混入肺音, 妨碍对其本质特征的特征. 为减少这些影响, 获取相对纯净的肺音信号, 并为后续分析奠定基础, 需对原始肺音信号进行预处理.

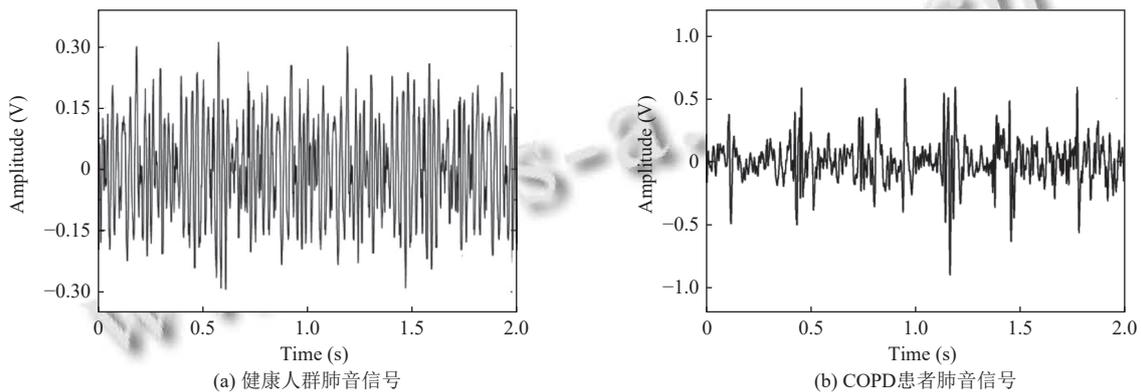


图3 肺音信号原始图

本文选取 IMF3 和 IMF4 进行信号重构, 重构图如图 4(a) 和图 4(b) 所示. 处理后的信号与原始信号相比, 重构后的肺音信号有着明显的周期性, 特征也被较好地保留下来. 对原始肺音信号进行滤波后, 滤

除了大部分环境噪声和心音信号的干扰, 在不同的中心频率上都很好地解决了模态混叠和端点效应的问题, 为下一步对肺音信号进行分类奠定了良好的基础.

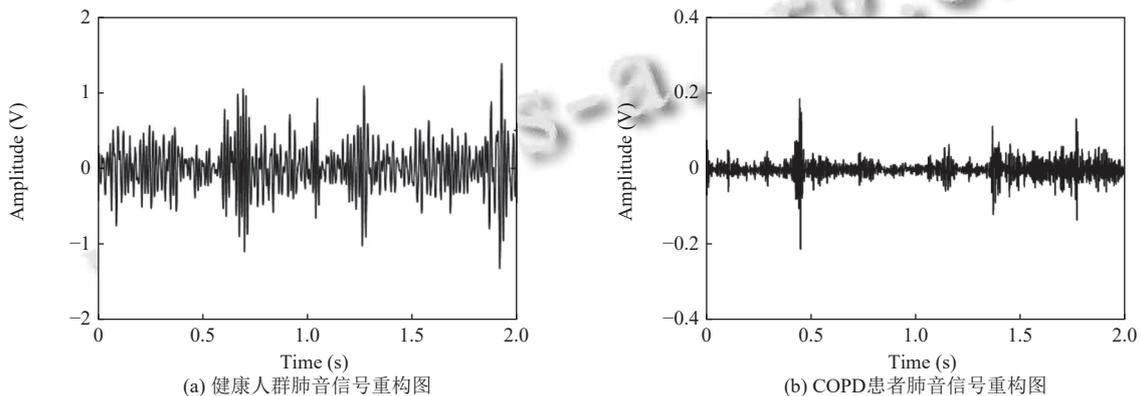


图4 肺音信号重构图

由于预处理后的肺音信号只有时域特征是明显的, 而频域特征不明显, 为了充分利用肺音信号中的信息, 选择从希尔伯特变换得到的希尔伯特谱图中提取频域特征, 继续研究健康人群与 COPD 患者肺音信号在频域特征中的区别. 图 5(a) 和图 5(b) 是健康人群和 COPD

患者的肺音重构后的希尔伯特谱图 (时间-频率-幅值联合图). 为了量化比较两组的时频特征, 计算了以下指标: 频谱中心 (spectral centroid): 表示能量分布的重心, 健康组平均为 28.7 ± 5.2 Hz, COPD 组为 112.4 ± 18.9 Hz ($p < 0.001$); 带宽 (spectral bandwidth): 反映能量分布的

离散程度,健康组为 43.5 ± 8.1 Hz, COPD 组为 178.6 ± 25.3 Hz ($p < 0.001$). 低频能量比 ($E_{<50\text{Hz}}/E_{\text{total}}$): 即谱中 50 Hz 以下能量占比,健康组为 $86.3\% \pm 4.7\%$, COPD 组仅为 $41.2\% \pm 7.5\%$ ($p < 0.001$). 结果表明,健康人群肺音能量重心明显偏低,且绝大部分能量集中在 50 Hz 以下;而 COPD 患者由于气道阻塞与收缩引起的湍流,其肺音能量分布更广、中心频率更高、低频能量比例显著下降. 这些量化差异为临床自动化诊断 COPD 提供了可度量的时-频特征依据.

2.5 结合 CNN-LSTM-SVM 的特征融合的肺音分析

本文提出了一种结合 CNN-LSTM-SVM 的特征融合网络模型:为了提取更多的肺音信号特征,将 LSTM 与 CNN 融合,提取肺音信号的时间序列;为了提升模

型的泛化性,用 SVM 替换网络的 Softmax 层. CNN 结构的最后一层是全连接层,可以整合卷积层、池化层中具有种类区分特点的信息,从而尽可能多地提取出信号的有效特征,实现最终分类. 由于 SVM 作为全连接层只能接收特征向量作为输入,需要卷积层和池化层产生这一组特征. 首先将原始 CNN 的全连接层和 Softmax 分类函数去掉,直接使用最后一层池化层的输出做合并,使之成为一个特征向量作为 SVM 层的输入. 使用 CNN-LSTM-SVM 的方法,就相当于把原来的分类过程分成了两个步骤:第 1 步使用 CNN 训练得到样本的特征向量,第 2 步使用特征向量重新训练 SVM 分类模型. 本文的网络模型图如图 6 所示,结构参数如表 1 所示.

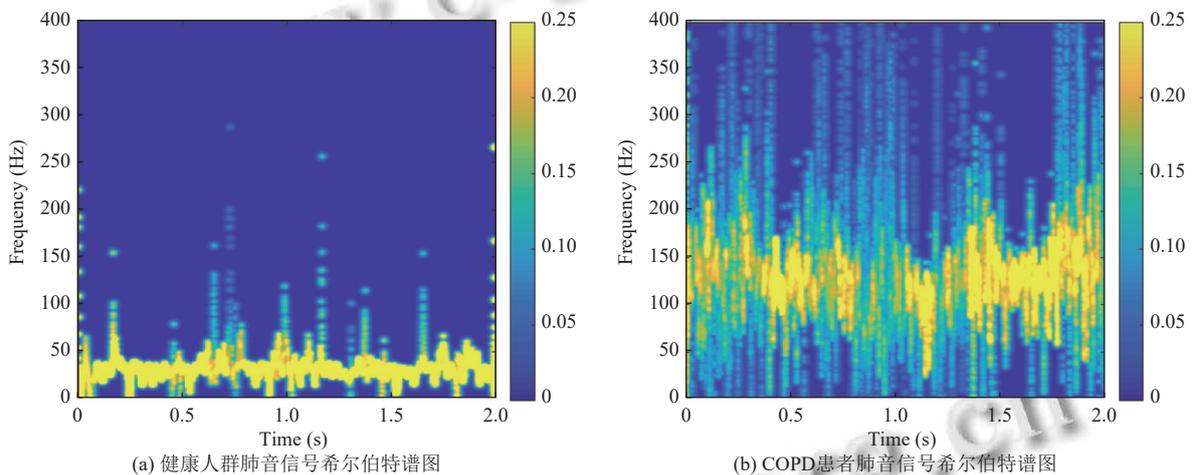


图5 重构信号的希尔伯特谱图

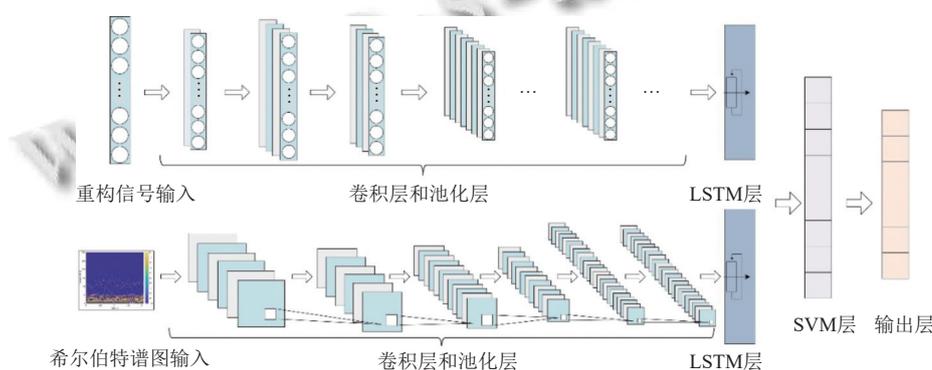


图6 结合 CNN-LSTM-SVM 的特征融合网络模型图

结合 CNN-LSTM-SVM 的特征融合肺音分类模型可以充分提取两种人群的肺音信号特征,更好地完成肺音分类的工作. 该模型通过 CNN 提取二者特征,再

通过 LSTM 提取时序特征,最后通过 SVM 强化细节信息并通过使用此方法作为全连接层,对肺音信号进行分类. 实验结果对比发现结合 CNN-LSTM-SVM 的

特征融合方法提升了模型的泛化能力, 迭代参数的震荡显著减少, 精度也得到了提升. 由此证明结合 CNN-LSTM-SVM 的特征融合模型在精度、泛化性和鲁棒性上面具有优越性.

表1 结合 CNN-LSTM-SVM 的特征融合网络模型结构参数

层数	层类型	输出尺寸	卷积核数量	卷积核大小
1	输入	20000×1	—	—
2	卷积层+池化层	9999×16	16	3×3
3	卷积层+池化层	4998×32	32	3×3
4	卷积层+池化层	2498×64	64	3×3
5	卷积层+池化层	1248×128	128	3×3
6	LSTM层	10	—	—
	(向量、隐藏元维数)	20	—	—
7	SVM层	2	—	—

2.6 结果分析

为了验证本文所提出模型的特征融合方法的有效性, 将肺音信号作为输入与 CNN-LSTM 方法进行对比, 模型的准确率和损失值随迭代次数的变化如图 7

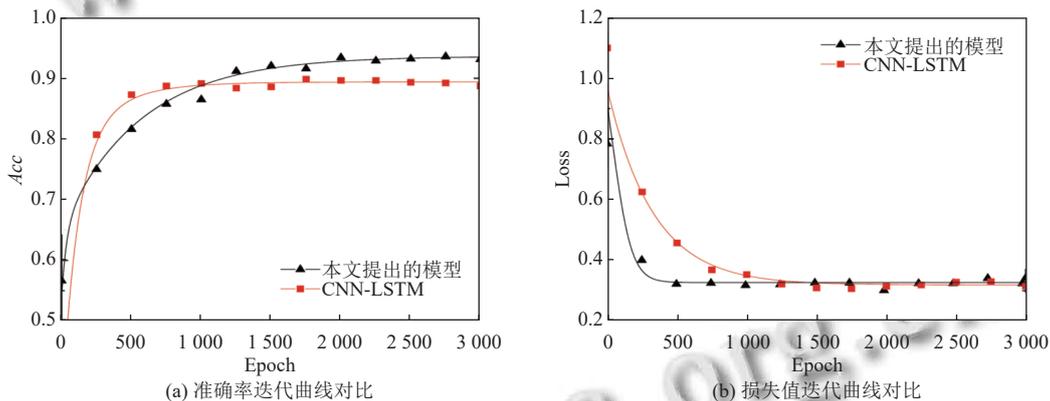


图7 CNN-LSTM 与本文提出的模型迭代曲线对比

为了比较不同输入对模型产生的性能影响, 使用相同的实验数据和硬件条件进行实验. 分别将上述两种人群的重构肺音信号、希尔伯特谱图以及二者的特征融合结果输入至 CNN-LSTM 和结合 CNN-LSTM-SVM 的特征融合模型中, 得到相应的分类模型准确率迭代图和损失迭代图. 然后针对各个模型进行讨论和分析.

图 8 为不同输入信号的分类模型迭代曲线图. 通过图 8(a) 和图 8(b) 可以得出, 分类模型的准确率均随迭代次数的增加而逐渐提高, 最终趋于稳定. 而损失值则随迭代次数的增加而下降. 其中以特征融合作为输入的分类模型准确率最高, 这证明 COPD 患者的肺音信号在时域和频域上均与健康人群有差别, 虽然单一生理信号的输入可以对不同人群的肺音进行分类, 但

所示. 模型优化采用二元交叉熵损失函数, 使用 Adam 优化器进行最小化. 根据两种模型损失值迭代曲线对比可知, CNN-LSTM 网络模型在对肺音信号进行分类时, 损失迭代次数达到 1 000 次时趋于稳定, 而结合 CNN-LSTM-SVM 的特征融合网络模型损失迭代次数达到 260 时就已经趋于稳定. 这说明加入 SVM 后, 网络的特征提取和分类能力有所提升. 其次, 可以发现加入 LSTM 之后, 迭代曲线的震荡程度也有所降低, 这说明加入 LSTM 之后, 网络模型变得更加稳定. 最后, 从准确率来看, 本文所提出的网络模型高于 CNN-LSTM 网络模型.

因此, 在加入 SVM 层之后, 结合了 SVM 泛化能力强的特点之后, 不仅提升了模型的泛化能力, 还得到了具有较高的精确率和较好的鲁棒性的模型. 对模型性能的提升较为明显. 实验结果表明, CNN-LSTM-SVM 模型在性能上要明显优于 CNN-LSTM 模型.

对多种信号进行特征融合的模型在训练时有着更为突出的优势.

本文将所提出的模型与使用 ICBHI 2017 呼吸音数据库对肺音进行分类的相关研究中使用的其他方法进行对比. 对比结果如表 2 所示. 从最终的实验结果来看, 结合 CNN-LSTM-SVM 的特征融合网络模型与文献[22]相比, 灵敏度提升 3.20%, 精确率提升 3.97%, F1-score 提升 3.58%. 可以得出, 我们提出的网络模型对 COPD 患者与健康人群有着较好的区分效果.

总结各项指标提升的原因主要有两个: (1) 预处理过程从肺音中去除了外界噪声和心音信号的干扰; (2) 将时域与频域结合后, 网络可以得到更多的有效特征, 排除干扰特征的影响, 从而进行更精准的分类.

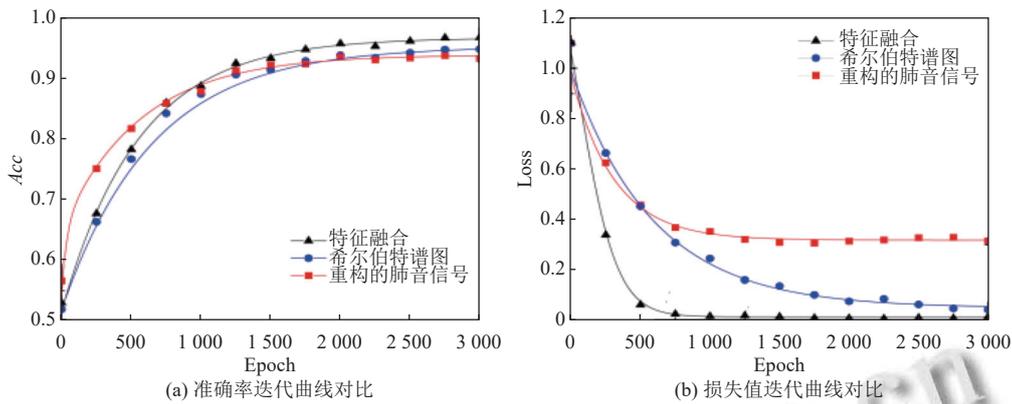


图8 不同输入信号的分类模型迭代曲线图

表2 本文模型与相关研究的对比

文献	方法	Acc (%)	Sen (%)	Pre (%)	F1-score
[13]	LDA-RSE、CNN	71.15	83.00	72.00	0.77
[14]	XGBoost、BiGRU-Attention	72.45	64.24	80.65	0.67
[15]	CNN-LSTM、FL	76.39	52.78	84.26	0.68
[21]	VGGish、Attention	85.61	83.44	83.44	0.84
[26]	Swin Transformer	—	77.3	41.1	0.59
[27]	Bi-ResNet (STFT、Wavelet)	77.81	61.99	90.10	0.71
[22]	CNN (MFCC)	—	93.00	93.00	0.93
本文方法	CNN-LSTM、SVM	96.56	96.20	96.97	0.96

注: 加粗数据为最优结果

3 结论与展望

本文在肺音信号分析的基础上, 提出了一种结合 CNN-LSTM-SVM 的特征融合的肺音分析方法. 以 CNN 为基础模型对肺音信号进行自适应特征提取和分类, 但由于 CNN 模型难以有效捕捉肺音信号作为时间序列所固有的前后时序依赖关系, 进而引入 LSTM, 构建 CNN-LSTM 联合模型, 以同时兼顾特征提取的局部相关性与时序动态特性. 进一步地, 针对 CNN-LSTM 模型在泛化能力方面存在的局限性, 本研究融合了 SVM 分类器, 构建 CNN-LSTM-SVM 特征融合模型, 旨在利用 SVM 优异的泛化性能提升整体模型的鲁棒性. 在此融合模型框架下, 深入探究了输入单一信号与融合信号对分类性能的影响. 实验结果表明, 得益于融合信号蕴含更丰富的有效特征信息, 其分类精度显著优于单一信号输入方案.

在该方法中, 利用 SVM 层将时域特征与频域特征

融合, 关注了信号不同方面的特征, 滤掉了部分无用特征, 使得 COPD 患者被误检的可能性大大降低, 对肺音识别分类具有较好的抗干扰性能, 最终取得了 96.56% 的分类准确率. 本文验证了该方法的可行性. 如何保证预处理后的肺音信号包含更加完整的肺音信息以及实时性便携化是未来研究的重点, 下一步的研究方向将主要考虑如何提升预处理效果, 提取更多特征以及改进网络模型来进一步提升分类准确率, 并把所提模型用于其他疾病的分析.

参考文献

- Gu Z, Sun JJ, Wang LX. mRNA expression insights: Unraveling the relationship between COPD and lung cancer. *The Journal of Gene Medicine*, 2024, 26(9): e3728. [doi: 10.1002/jgm.3728]
- Zhu BF, Wang YF, Ming J, et al. Disease burden of COPD in China: A systematic review. *International Journal of Chronic Obstructive Pulmonary Disease*, 2018, 13: 1353–1364. [doi: 10.2147/COPD.S161555]
- Ajufo E, Rao S, Navar AM, et al. U.S. population at increased risk of severe illness from COVID-19. *American Journal of Preventive Cardiology*, 2021, 6: 100156.
- Lozano M, Fiz JA, Jané R. Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency. *IEEE Journal of Biomedical and Health Informatics*, 2016, 20(2): 486–497. [doi: 10.1109/JBHI.2015.2396636]
- Serbes G, Ulukaya S, Kahya YP. An automated lung sound preprocessing and classification system based on spectral analysis methods. *Proceedings of the 2017 International Conference on Biomedical and Health Informatics*. Thessaloniki: Springer, 2018. 45–49.

- 6 Aras S, Gangal A. Comparison of different features derived from Mel frequency cepstrum coefficients for classification of single channel lung sounds. Proceedings of the 40th International Conference on Telecommunications and Signal Processing (TSP). Barcelona: IEEE, 2017. 346–349.
- 7 Abougabal MM, Diaa Moussa N. A novel technique for validating diagnosed respiratory noises in infants and children. Alexandria Engineering Journal, 2018, 57(4): 3033–3041. [doi: [10.1016/j.aej.2018.05.003](https://doi.org/10.1016/j.aej.2018.05.003)]
- 8 Aziz S, Khan MU, Shakeel M, *et al.* An automated system towards diagnosis of pneumonia using pulmonary auscultations. Proceedings of the 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS). Karachi: IEEE, 2019. 1–7.
- 9 Brown C, Chauhan J, Grammenos A, *et al.* Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2020. 3474–3484.
- 10 许洋. 基于改进 CNN-SVM 的肺音分类方法研究及应用 [硕士学位论文]. 银川: 北方民族大学, 2022.
- 11 Sengupta N, Sahidullah M, Saha G. Lung sound classification using cepstral-based statistical features. Computers in Biology and Medicine, 2016, 75: 118–129. [doi: [10.1016/j.combiomed.2016.05.013](https://doi.org/10.1016/j.combiomed.2016.05.013)]
- 12 Kochetov K, Putin E, Balashov M, *et al.* Noise masking recurrent neural network for respiratory sound classification. Proceedings of the 27th International Conference on Artificial Neural Networks on Artificial Neural Networks and Machine Learning. Rhodes: Springer, 2018. 208–217.
- 13 Demir F, Ismael AM, Sengur A. Classification of lung sounds with CNN model using parallel pooling structure. IEEE Access, 2020, 8: 105376–105383. [doi: [10.1109/ACCESS.2020.3000111](https://doi.org/10.1109/ACCESS.2020.3000111)]
- 14 Zhao XS, Shao YB, Mai JY, *et al.* Respiratory sound classification based on BiGRU-attention network with XGBoost. Proceedings of the 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Seoul: IEEE, 2020. 915–920.
- 15 Petmezas G, Cheimariotis GA, Stefanopoulos L, *et al.* Automated lung sound classification using a hybrid CNN-LSTM network and focal loss function. Sensors, 2022, 22(3): 1232. [doi: [10.3390/s22031232](https://doi.org/10.3390/s22031232)]
- 16 Mondal A, Banerjee P, Tang H. A novel feature extraction technique for pulmonary sound analysis based on EMD. Computer Methods and Programs in Biomedicine, 2018, 159: 199–209. [doi: [10.1016/j.cmpb.2018.03.016](https://doi.org/10.1016/j.cmpb.2018.03.016)]
- 17 Shi Y, Li YQ, Cai ML, *et al.* A lung sound category recognition method based on wavelet decomposition and BP neural network. International Journal of Biological Sciences, 2019, 15(1): 195–207. [doi: [10.7150/ijbs.29863](https://doi.org/10.7150/ijbs.29863)]
- 18 Acharya J, Basu A. Deep neural network for respiratory sound classification in wearable devices enabled by patient specific model tuning. IEEE Transactions on Biomedical Circuits and Systems, 2020, 14(3): 535–544.
- 19 Altan G, Kutlu Y, Allahverdi N. Deep learning on computerized analysis of chronic obstructive pulmonary disease. IEEE Journal of Biomedical and Health Informatics, 2020, 24(5): 1344–1350. [doi: [10.1109/JBHI.2019.2931395](https://doi.org/10.1109/JBHI.2019.2931395)]
- 20 Messner E, Fediuk M, Swatek P, *et al.* Multi-channel lung sound classification with convolutional recurrent neural networks. Computers in Biology and Medicine, 2020, 122: 103831. [doi: [10.1016/j.combiomed.2020.103831](https://doi.org/10.1016/j.combiomed.2020.103831)]
- 21 Choi Y, Lee H. Interpretation of lung disease classification with light attention connected module. Biomedical Signal Processing and Control, 2023, 84: 104695. [doi: [10.1016/j.bspc.2023.104695](https://doi.org/10.1016/j.bspc.2023.104695)]
- 22 Srivastava A, Jain S, Miranda R, *et al.* Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Computer Science, 2021, 7: e369. [doi: [10.7717/peerj-cs.369](https://doi.org/10.7717/peerj-cs.369)]
- 23 Landi F, Baraldi L, Cornia M, *et al.* Working memory connections for LSTM. Neural Networks, 2021, 144: 334–341. [doi: [10.1016/j.neunet.2021.08.030](https://doi.org/10.1016/j.neunet.2021.08.030)]
- 24 Al-Bakaa ZH, Albo-Salih AT. A Review of ICBHI 2017 respiratory sounds analysis using deep learning. AI-Furat Journal of Innovations in Electronics and Computer Engineering, 2024, 3(2): 321–329. [doi: [10.46649/fjiece.v3.2.21a.30.5.2024](https://doi.org/10.46649/fjiece.v3.2.21a.30.5.2024)]
- 25 Wang Y, Bai DY. Application of wavelet threshold method based on optimized VMD to ECG denoising. Proceedings of the 3rd IEEE International Conference on Frontiers Technology of Information and Computer (ICFTIC). Greenville: IEEE, 2021. 741–744.
- 26 Wu CW, Ye N, Jiang JL. Classification and recognition of lung sounds based on improved Bi-ResNet model. IEEE Access, 2024, 12: 73079–73094. [doi: [10.1109/ACCESS.2024.3404657](https://doi.org/10.1109/ACCESS.2024.3404657)]
- 27 Sun WW, Zhang F, Sun P, *et al.* Respiratory sound classification based on swin Transformer. Proceedings of the 8th International Conference on Signal and Image Processing (ICSIP). Wuxi: IEEE, 2023. 511–515.

(校对责编: 张重毅)