

基于多模态特征融合与集成学习的心音信号分类研究

杨震¹, 刘黎², 彭扬¹, 陈孝杰¹, 赵阳¹

¹重庆科技大学数理科学学院, 重庆

²重庆市渝中区大数据局, 重庆

收稿日期: 2025年12月5日; 录用日期: 2026年1月5日; 发布日期: 2026年1月12日

摘要

针对心音分类中存在的样本不平衡、特征提取不足与模型泛化能力有限等问题, 本文提出了一种基于多模态特征融合与高级集成学习的心音异常检测方法。首先, 采用Challenge2016数据集, 通过带通滤波与归一化预处理, 并提出异常类别导向的数据增强策略以解决样本不平衡问题。其次, 综合提取时域统计、频域谱质心及小波域能量熵等多维特征, 构建全方位特征空间。在模型构建上, 设计了CNN-BiLSTM深度神经网络以捕捉局部与长时依赖特征, 并结合支持向量机(SVM)的稳定性优势。最后, 提出包含动态加权、置信度加权及元学习等六种集成策略。实验结果表明, 元学习集成策略性能最佳, 准确率达98.89%, F1-score为0.9872。该方法在保证高精度的同时也具备良好的鲁棒性, 为心脏瓣膜疾病的智能诊断提供了有效方案。

关键词

心音信号, 心脏瓣膜疾病, 特征融合, SVM, CNN-BiLSTM, 集成学习

Heart Sound Signal Classification Based on Multi-Modal Feature Fusion and Ensemble Learning

Zhen Yang¹, Li Liu², Yang Peng¹, Xiaojie Chen¹, Yang Zhao¹

¹School of Mathematical Sciences, Chongqing University of Science and Technology, Chongqing

²Chongqing Yuzhen District Big Data Bureau, Chongqing

Received: December 5, 2025; accepted: January 5, 2026; published: January 12, 2026

文章引用: 杨震, 刘黎, 彭扬, 陈孝杰, 赵阳. 基于多模态特征融合与集成学习的心音信号分类研究[J]. 计算机科学与应用, 2026, 16(1): 56-71. DOI: 10.12677/csa.2026.161006

Abstract

To address the challenges of data imbalance, insufficient feature extraction, and limited model generalization in heart sound classification, this paper proposes a heart sound anomaly detection method based on multi-modal feature fusion and advanced ensemble learning. Firstly, using the Challenge2016 dataset, data preprocessing including band-pass filtering and normalization is applied, and an anomaly-oriented data augmentation strategy is proposed to mitigate sample imbalance. Secondly, multi-dimensional features, including time-domain statistics, frequency-domain spectral centroid, and wavelet-domain energy entropy, are extracted to construct a comprehensive feature space. In terms of model construction, a CNN-BiLSTM deep neural network is designed to capture local and long-term dependency features, combined with the stability of Support Vector Machine (SVM). Finally, six ensemble strategies, including dynamic weighting, confidence weighting, and meta-learning, are proposed. Experimental results show that the meta-learning ensemble strategy performs the best, with an accuracy of 98.89% and an F1-score of 0.9872. This method ensures high accuracy while maintaining good robustness, providing an effective solution for intelligent diagnosis of valvular heart diseases.

Keywords

Heart Sound Signal, Valvular Heart Disease, Feature Fusion, SVM, CNN-BiLSTM, Ensemble Learning

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

心脏瓣膜疾病是全球高致死率的心血管疾病之一，早期诊断对改善预后至关重要。心音信号包含了瓣膜开合与血液流动的病理特征，是临床筛查的核心依据。然而，传统听诊依赖医生经验，存在主观性强、难以识别微弱杂音等局限[1]-[3]。近年来，机器学习与深度学习在心音分类中展现出巨大潜力，能自动提取复杂特征并处理大规模数据[4]。

尽管现有研究取得了一定进展，但仍面临诸多挑战。在特征提取方面，Liang [5]、Springer [6]及 Yaseen [7]等人分别利用小波包分解、隐马尔可夫模型及 MFCC 进行尝试，但多局限于单一特征域，难以全面捕捉非平稳信号特征。在模型方面，Raza [8]和 Khan [9]等人虽然引入了深度学习，但现有方法在处理数据不平衡、捕捉时序长依赖关系以及抗噪鲁棒性方面仍有不足。

针对上述问题，本文提出一种基于多模态特征融合与集成学习的心音异常检测方法，主要贡献如下：

- (1) 数据增强：提出异常类别导向的增强策略，通过针对性扰动异常样本解决数据不平衡问题，提升模型敏感性；
- (2) 特征融合：综合提取时域统计、频域谱质心及小波域能量熵等多层次特征，构建兼顾全局与局部的特征空间；
- (3) 模型架构：设计 CNN-BiLSTM 深度神经网络，利用卷积层提取局部特征，双向 LSTM 捕捉长时依赖，相比单一模型显著提升了分类性能。

2. 特征提取

在心音信号的分析与分类中，特征提取是关键的一步。常见的特征提取方法通常分为时域特征、频

域特征以及时频域特征等[10]-[12]，见表 1。基于这些常见方法，我们提出了一种结合时域、频域和小波域的多特征提取方案，以期全面描述心音信号的多层次特征。

Table 1. Cardiac sound signal feature extraction indicators

表 1. 心音信号特征提取指标

特征类型	具体指标	作用与意义
时域特征	均值、标准差、最大值、最小值、均方根、偏度、峰度	描述信号的整体幅值水平、波动程度及分布形态，反映心音强弱、稳定性以及异常波动情况。
	过零率	衡量信号变化的频繁程度，间接反映心音的复杂度和高频成分。
	包络均值与标准差	反映心音能量包络随时间的变化，敏感于杂音和心音强度不均等病理特征。
频域特征	谱质心	表征频谱能量的“重心”，反映整体频率分布位置，异常心音往往向高频偏移。
	谱带宽	描述能量在频谱上的扩展程度，提示杂音的频率分布范围。
	频带能量	定量分析特定频带的能量分布，用于识别不同心音成分及病理性杂音。
	梅尔频率倒谱系数 (MFCC, 13 维)	模拟人耳听觉感知，对心音“音质”变化敏感，常用于语音和心音分类任务。
小波特征	小波分解系数的均值、标准差、能量、熵	同时提供时域与频域信息，能捕捉心音在不同尺度下的细节与趋势，适合非平稳信号分析。

与以往依赖单一域分析的研究不同，本研究的创新之处在于提出了一种多域特征融合策略。该方法将时域、频域与小波域特征有机结合，实现了从全局统计特性到局部动态细节的多层次表征。通过融合不同域的信息，模型不仅能够强化对心音全局结构与局部异常的联合感知能力，还显著提升了对病理性杂音中微弱异常模式的敏感性。实验结果表明，这种多域特征融合方法在异常心音识别中表现出更高的鲁棒性与分类准确率。

3. 模型构建

本研究旨在针对心音信号的自动分类问题，提出一种基于支持向量机(Support Vector Machine, SVM)[13]与双向长短期记忆网络(Bidirectional Long Short-Term Memory, Bi LSTM)的混合分类方案。该方法充分利用了传统机器学习模型在小样本特征空间上的优势，同时结合深度学习模型在时序建模与特征自动提取上的能力，从而在准确率、泛化能力与鲁棒性之间取得平衡。

3.1. 支持向量机(SVM)模型设计

SVM 在小样本高维特征的分类任务中表现出色，尤其适用于心音信号中经过特征工程后的低维度数据。在医学数据中，样本数量通常有限，传统的深度神经网络容易出现过拟合现象，而 SVM 利用核函数技巧，在有限数据条件下仍能实现较高的分类精度。本研究在 SVM 部分采用了径向基函数核(Radial Basis Function Kernel, RBF)，为了兼顾精度与训练效率，本研究针对超参数搜索与模型拟合进行了优化。SVM 心音分类模型流程见图 1。

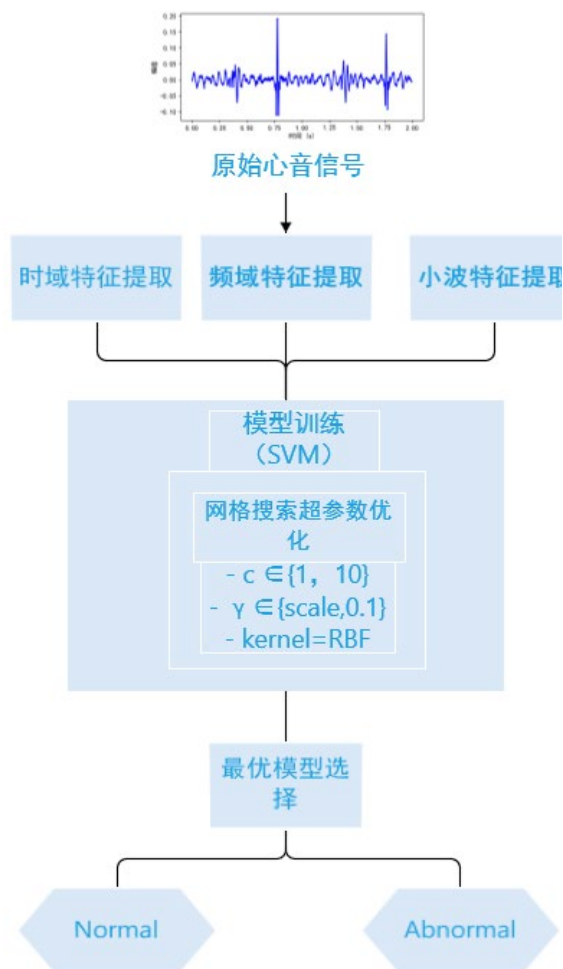


Figure 1. SVM heart sound classification model flowchart
图 1. SVM 心音分类模型流程

3.2. 双向卷积长短期记忆网络(BiLSTM-CNN)模型设计

尽管 SVM 在小样本条件下表现优异,但其依赖人工设计特征,难以充分利用心音信号的原始时序结构。因此,本研究进一步引入了卷积神经网络(CNN)与双向长短期记忆网络(BiLSTM)结合的深度学习架构,能够在端到端的框架下自动提取特征并建模长时间依赖关系。BiLSTM 模型能够同时考虑时间序列的前向与后向信息,在捕捉周期性、非平稳的心音信号方面优于单向 LSTM [14]。同时, CNN [15]模块则用于提取局部频谱特征,从而提升整体性能。

本研究所设计的模型采用 BiLSTM-CNN 的混合架构。首先,在卷积特征提取阶段,使用一维卷积层(Conv1D)对心音信号进行局部特征提取,以削弱原始输入中的噪声干扰。该卷积层设置了 64 个卷积核,卷积核大小为 3,激活函数为 ReLU,并采用 *same* 填充以保证序列长度不变;同时结合 Batch Normalization 与 MaxPooling1D,以提升特征分布的稳定性并有效降低计算复杂度。随后,模型引入两层双向长短期记忆网络(BiLSTM),其中第一层包含 128 个单元,用于捕获全局时序依赖关系;第二层包含 64 个单元,用于进一步抽象高阶特征。在每一层 BiLSTM 之后均加入 Dropout (比例为 0.3),以缓解过拟合问题。最后,网络通过全连接层(Dense, 64 单元, ReLU 激活, 结合 L2 正则化)对特征进行整合,并在输出层采用 Softmax 激活函数实现二分类概率输出,从而完成心音信号的分类任务。BiLSTM-CNN 心音分类模型流程见图 2。

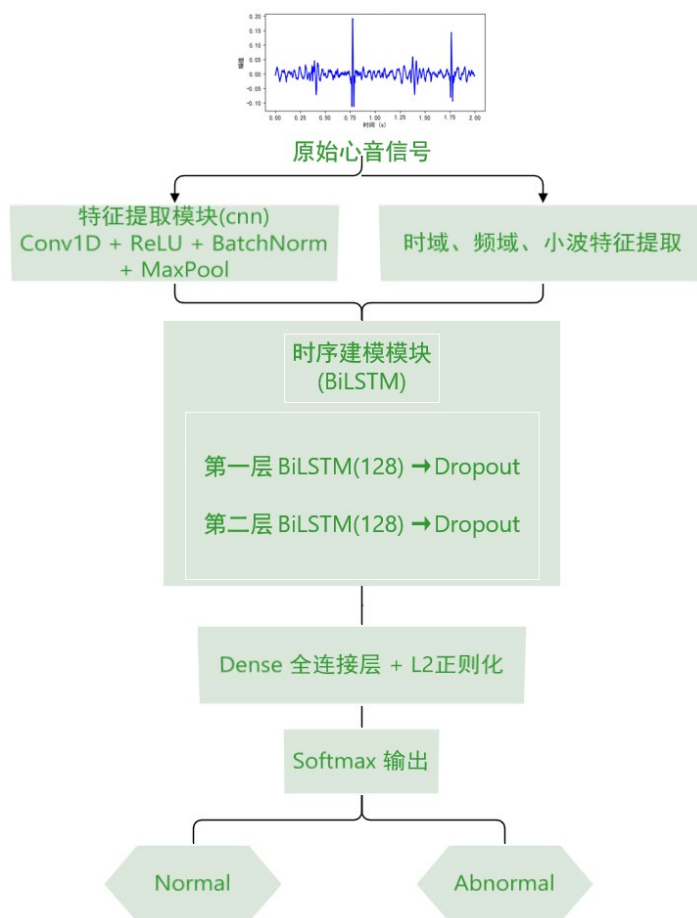


Figure 2. Flowchart of the BiLSTM-CNN heart sound classification model
图 2. BiLSTM-CNN 心音分类模型流程

3.3. 集成模型设计

在本研究中,除了单独构建支持向量机(SVM)和长短时记忆网络(BiLSTM-CNN)两类模型外,我们进一步设计并实现了一个集成模型[16][17] (Ensemble Model),以期在保持较高分类精度的同时兼顾模型的稳定性和泛化能力。集成方法选择不同的策略,通过这些策略,可以使得集成模型在不同情况下做出更为合理的预测,从而实现性能提升。图 3 集成模型设计具体流程。

为兼顾分类精度与泛化能力,本研究设计了六种策略融合 SVM 与 CNN-BiLSTM 的预测结果(记为 P_{SVM} , $P_{CNN-BiLSTM}$)。

动态加权策略[18] (Dynamic Weighted): 基于模型在验证集上的 F1 分数动态分配权重:

$$P_{ensemble} = \frac{F1_{SVM}^{\alpha} \cdot P_{SVM} + F1_{CNN-BiLSTM}^{\alpha} \cdot P_{CNN-BiLSTM}}{F1_{SVM}^{\alpha} + F1_{CNN-BiLSTM}^{\alpha}}$$

其中,指数参数 $\alpha = 1.5$ 是基于验证集网格搜索确定的最优值,旨在适度放大高能模型的贡献,同时避免单一模型主导。

置信度加权策略[19] (Confidence Weighted): 根据模型预测的最大概率值(即置信度 $C = \max(P)$)调整权重,使高置信度模型在决策中占据更大比重。

层次决策策略[20] (Hierarchical Decision): 优先采信高置信度模型。当任一模型置信度高于阈值 θ 时

直接输出其结果；否则采用加权平均。

$$P_{ensemble} = \begin{cases} P_i, & \text{if } C_i > 0.8 \\ 0.75P_{SVM} + 0.25P_{CNN-BiLSTM}, & \text{otherwise} \end{cases}$$

阈值 $\theta = 0.8$ 是通过分析训练集预测概率分布后设定的经验值，能有效区分高确信度样本与模糊样本。

其他策略：包括直接选择置信度最高的最大置信度策略[21]；基于逻辑回归学习模型间交互特征(如预测差值、乘积)的元学习器策略[22]；以及基于硬分类结果多数决定的投票机制[23]。

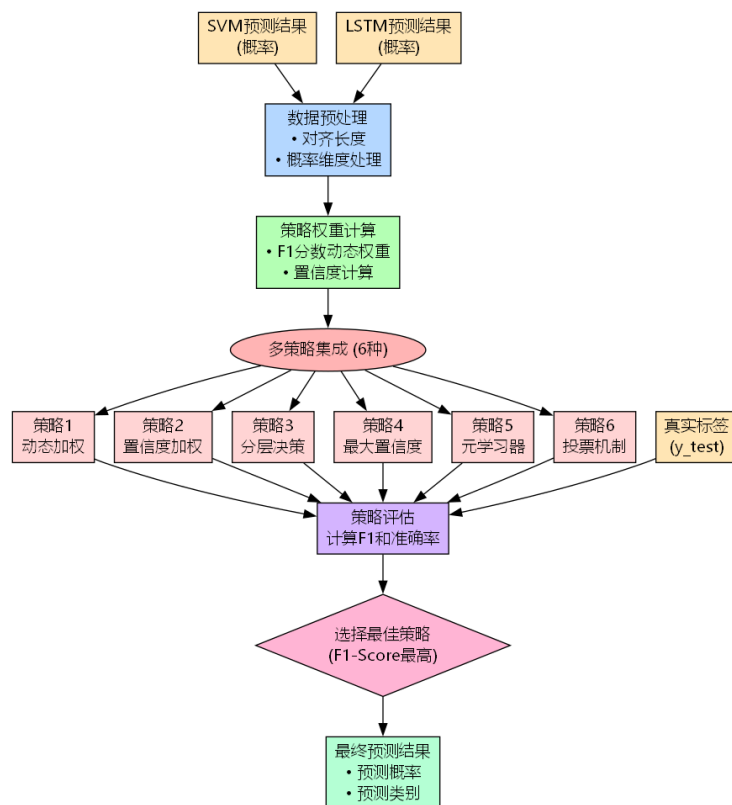


Figure 3. Integrated model flowchart

图 3. 集成模型流程图

4. 评估指标

在心音分类任务中，尤其是针对二分类问题(正常与异常心音的分类)，混淆矩阵是一种有效的工具，用于总结分类模型的预测结果。混淆矩阵以表格的形式展示了模型预测结果与真实标签之间的关系，能够为模型性能提供详细的分析。在该矩阵中，行表示样本的真实标签，列表示模型的预测标签。以二分类问题为例，其中正常心音和异常心音分别作为正类与负类进行分类。表 2 展示了二分类混淆矩阵[24]的结构：

Table 2. Confusion matrix

表 2. 混淆矩阵

	实际为正常心音	实际为异常心音
预测为正类	真正类(TP)	假正类(FP)
预测为负类	假负类(FN)	真负类(TN)

在表 3 中, TP (True Positive)表示模型正确地将实际正常心音样本预测为正常心音的数量; FP (False Positive)表示模型错误地将实际异常心音样本预测为正常心音的数量; FN (False Negative)表示模型错误地将实际正常心音样本预测为异常心音的数量; TN (True Negative)表示模型正确地将实际异常心音样本预测为异常心音的数量。针对该二分类任务,可以基于混淆矩阵的内容计算以下评价指标,以全面衡量分类模型的性能[25]:

5. 数据分析和预处理

5.1. 数据集介绍

本研究采用 Challenge2016 数据集[26], 包含 3240 例采集自临床及家庭环境的多源心音数据(时长 5~180 秒)。该数据集涵盖心脏瓣膜缺陷、冠状动脉疾病等多种病理类型,分为正常与异常两类。其跨设备采集的特性显著增强了数据的临床泛化能力。实验前,所有数据均重采样至 2000 Hz 以统一标准。

5.2. 数据预处理

为确保特征提取的有效性,预处理流程(如图 4 所示)主要包含以下步骤:

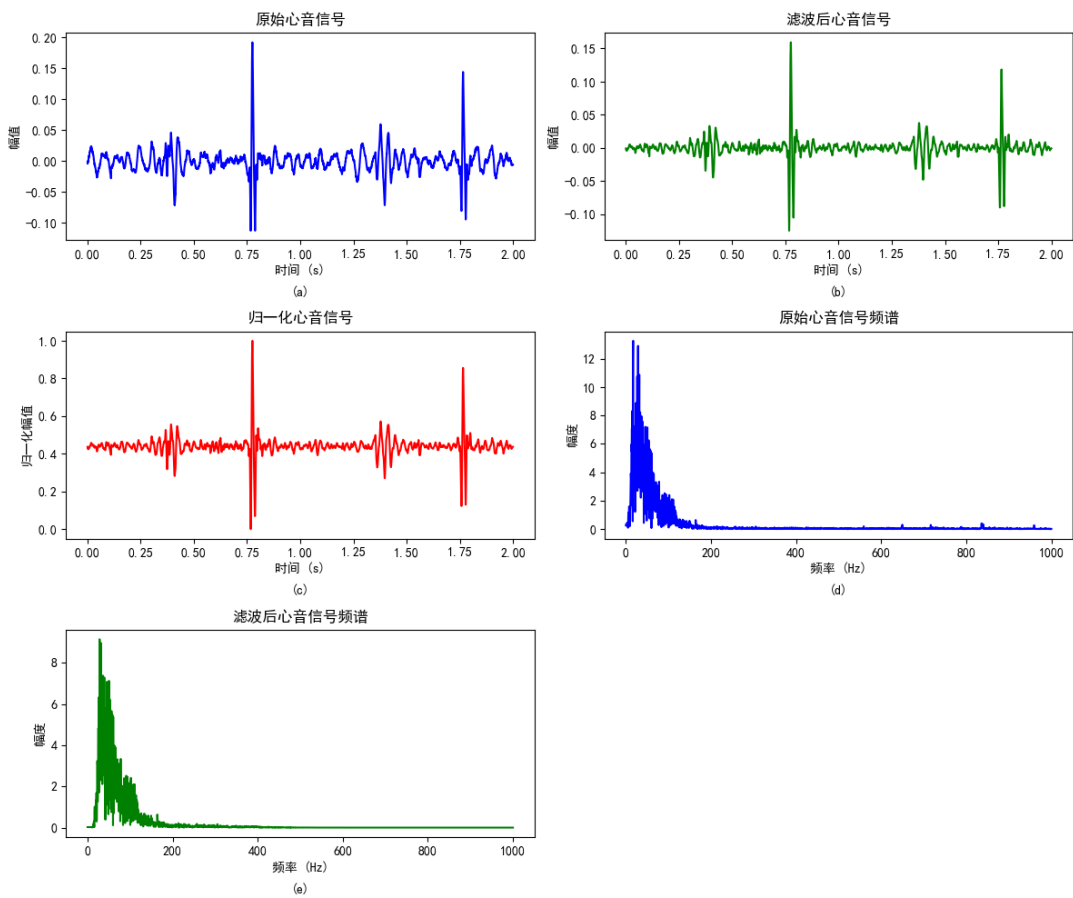


Figure 4. Preprocessing of the heart sound signal: (a) Original heart sound signal; (b) Filtered heart sound signal; (c) Heart sound signal after filtering and normalization; (d) Spectrum of the original heart sound signal; (e) Spectrum of the filtered heart sound signal

图 4. 心音信号的预处理, (a) 原始心音信号; (b) 滤波后的心音信号; (c) 经过滤波和归一化处理的心音信号; (d) 原始心音信号的频谱图; (e) 心音信号经过滤波后的频谱图

滤波降噪：采用三阶巴特沃斯带通滤波器，截止频率设为 25 Hz~400 Hz [27]。如图 4(b)和图 4(e)所示，该处理有效去除了低频基线漂移与高频环境噪声，保留了心音的主要能量区间。归一化：对滤波后的信号进行线性归一化，将幅值映射至[0,1]区间(图 4(c))，以消除不同采集设备导致的幅值差异。定长分割与增强：归一化后的数据将被分割为固定长度片段，并对异常样本实施增强。

在对每个心音信号进行滤波和归一化处理，需要将信号重新采样到 2000 Hz，然后再进行分割。图 5 是心音信号分割前后对比。

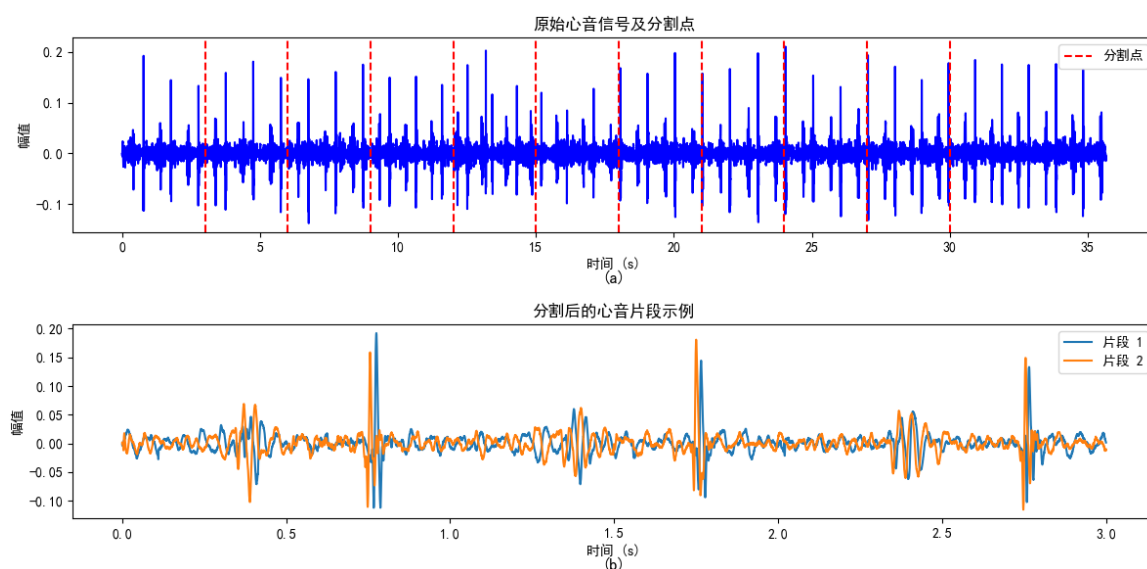


Figure 5. Segmentation and display of heart sound signals

图 5. 心音信号的分割及片段展示

本研究采用固定时长切片策略：首先将采样率统一为 2000 Hz，随后以 3 秒为窗口对信号进行非重叠分割。对于不足 3 秒的样本采用零填充(Zero-padding)补齐，超过 3 秒的长信号则拆分为多个独立片段。该策略不仅保证了模型输入维度的一致性，更通过长信号切分有效扩充了样本数量，从而增强了模型的泛化能力。

5.3. 数据增强

针对数据集严重的类别不平衡问题(正常样本远多于异常样本)，本研究提出一种异常类别导向的混合增强策略。首先，仅对训练集中的异常样本(Label=1)引入受控的高斯噪声进行扰动，为每个原始样本生成 3 个增强副本，在扩充数据的同时提升模型对噪声的鲁棒性。随后，采用 SMOTE 技术对少数类样本进行线性插值。这种“噪声扰动 + 合成过采样”的组合策略有效平衡了类分布，在保留特征空间结构信息的同时避免了模型向多数类过拟合。

6. 实验结果

6.1. 单模型性能表现

从图 6 LSTM 的混淆矩阵可以看出，该模型在识别正常与异常心音时表现较为均衡，其中正常类别的预测正确数为 6732，异常类别的预测正确数为 5197，整体准确率达到 98.93%。同时，训练过程曲线表明模型在收敛过程中训练集与验证集的准确率逐渐接近，最终训练准确率为 99.42%，验证准确率为

98.93%，且无明显过拟合迹象。

相比之下，图 7 为 SVM 模型的混淆矩阵，显示其对正常类(6809 个正确预测)与异常类(5208 个正确预测)的区分效果更为稳定，整体表现略优于 LSTM。其 AUC 值达到 0.9985，与 LSTM 相当甚至略高。图 8 展示了 LSTM 训练过程。

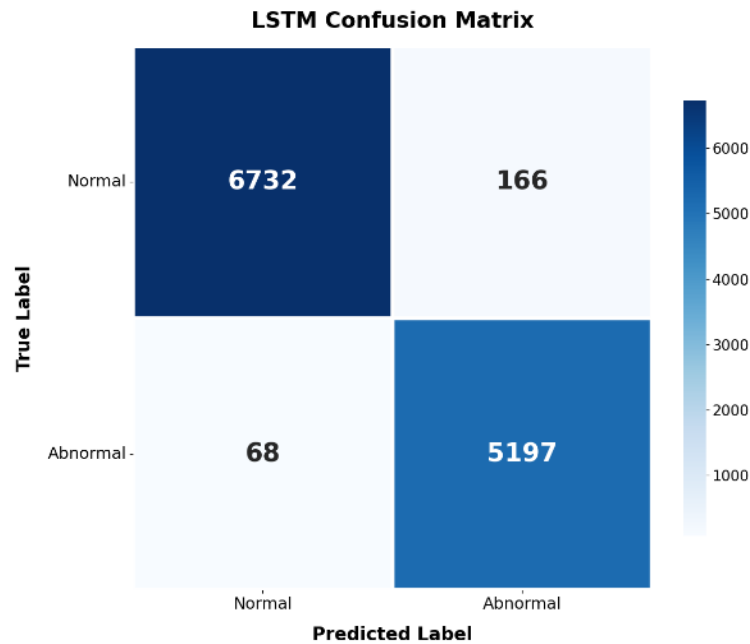


Figure 6. Confusion matrix of the LSTM model
图 6. LSTM 的混淆矩阵

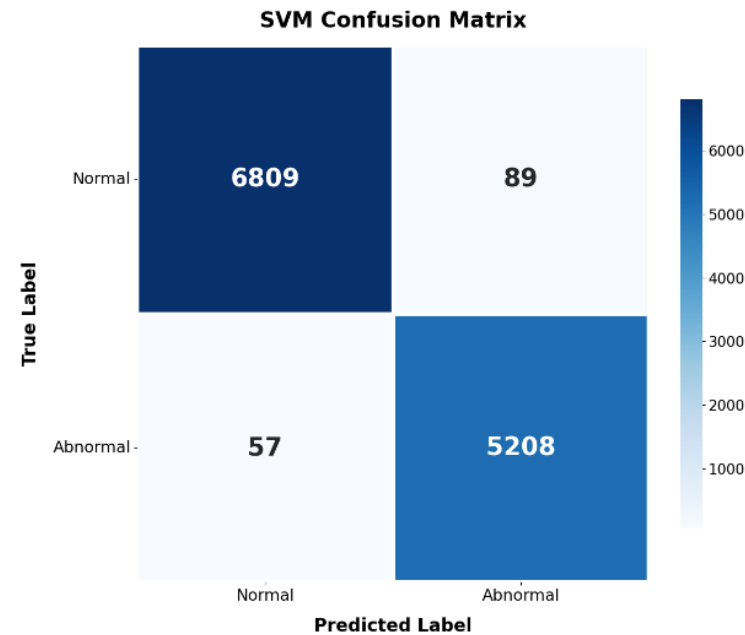


Figure 7. Confusion matrix of the SVM model
图 7. SVM 模型的混淆矩阵

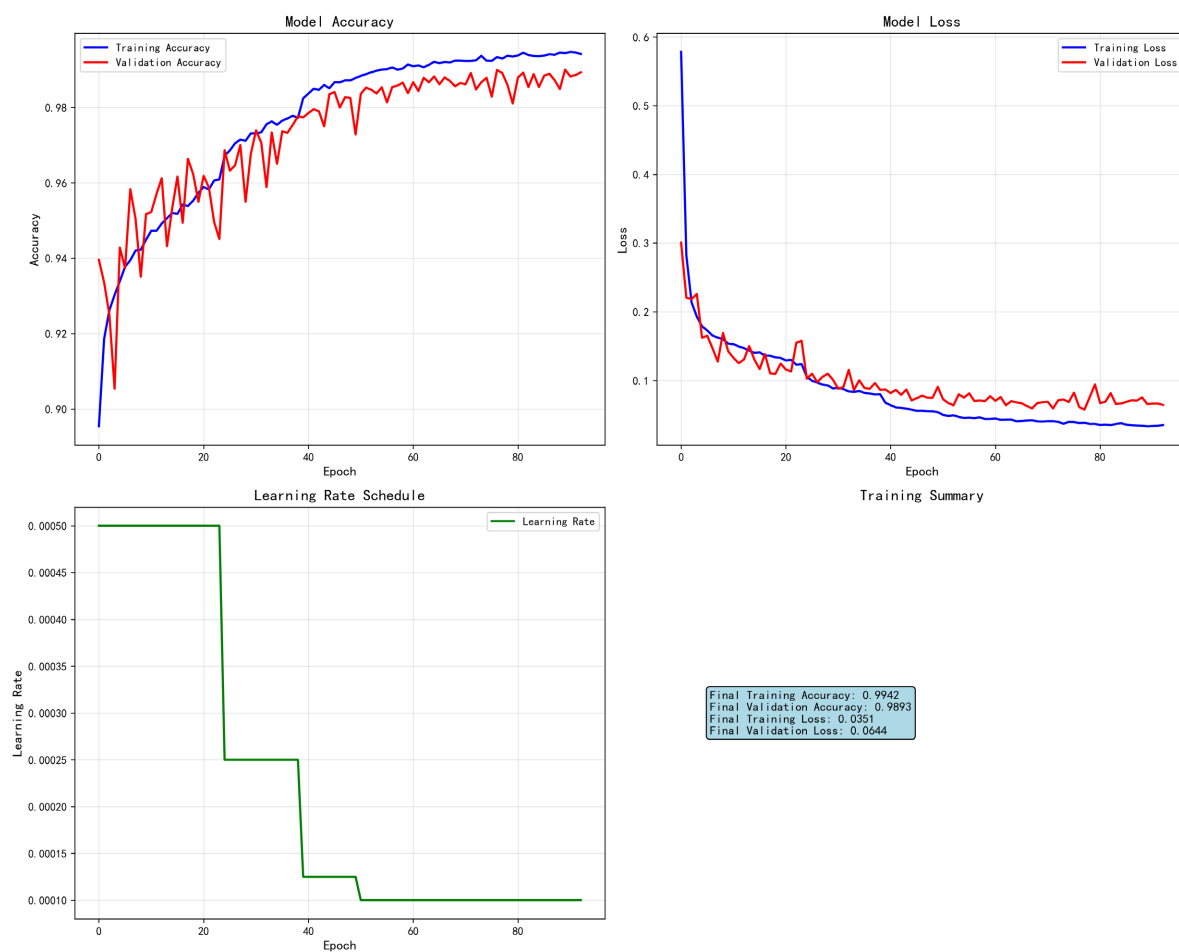


Figure 8. LSTM training process curve
图 8. LSTM 训练过程曲线

6.2. 集成模型性能

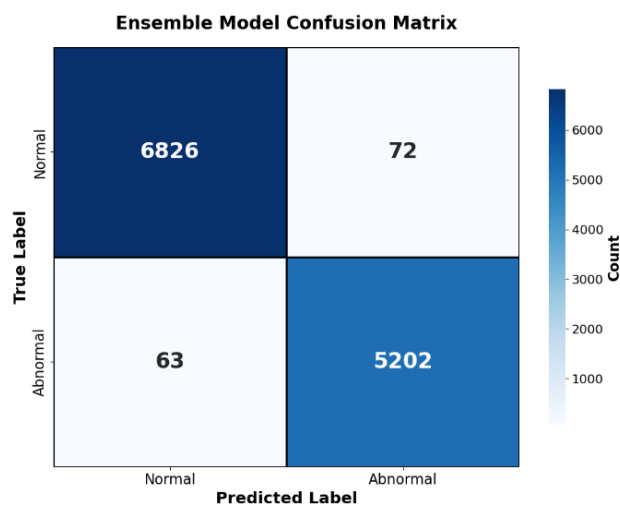


Figure 9. Confusion matrix of the integrated model
图 9. 集成模型混淆矩阵

为了进一步提升模型的鲁棒性与泛化能力，本研究设计了多种集成策略，包括动态加权、置信度加权、层次融合、最大置信度选择以及元学习。

从集成模型的混淆矩阵(图 9)可以看出，该方法在正常类别上预测正确数为 6826，异常类别为 5202，相比单一模型误判数量进一步减少。其整体分类性能在准确率与召回率上均达到 98.8%以上。

Table 3. Performance of each strategy
表 3. 各策略的表现

Ensemble Strategy	F1-Score	Accuracy
voting	0.9795	0.9820
max_confidence	0.9856	0.9875
confidence_weighted	0.9856	0.9875
dynamic_weighted	0.9859	0.9877
hierarchical	0.9863	0.9881
meta_learner	0.9872	0.9889

从表中可以看出，元学习(meta_learner)策略在所有策略中表现最佳，取得了最高的 F1-Score (0.9872) 和准确率(0.9889)。相较之下，投票策略(voting)尽管在实际应用中常被使用，但其性能相对较差，F1-Score 为 0.9795，准确率为 0.9820。此外，层次化(hierarchical)策略表现也较为突出，F1-Score 为 0.9863，准确率为 0.9881，紧随其后。其他策略如动态加权(dynamic_weighted)和置信度加权(confidence_weighted)也表现较好，二者的 F1-Score 分别为 0.9859 和 0.9856，准确率分别为 0.9877 和 0.9875。

基于实验结果，元学习策略在各项指标上都表现优异，因此选择采用元学习策略作为本文的最终集成方法。

6.3. 模型间对比

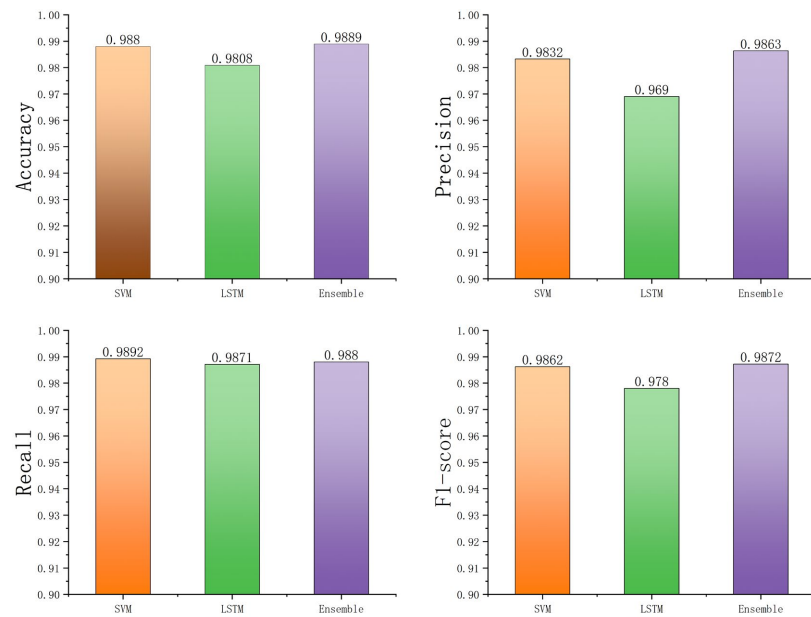


Figure 10. Comparison of SVM, LSTM, and ensemble methods in terms of Accuracy, Precision, Recall, and f1 score
图 10. SVM、LSTM、Ensemble 在 Accuracy、Precision、Recall、F1 的对比

将 SVM、LSTM 与集成模型进行综合比较(图 10 SVM、LSTM、Ensemble 在 Accuracy、Precision、Recall、F1 的对比), 结果表明三种模型在准确率、精确率、召回率与 F1-score 上均表现优异, 且差距极小。其中: 三者准确率均约为 98.8~98.9%; 精确率: LSTM 为 0.9870, SVM 与集成模型分别为 0.9862 与 0.9863; 召回率: LSTM 在异常心音识别上略低(0.9831), 而 SVM 与集成模型均为 0.9880; F1-score: 集成模型略优(0.9872), 体现了其在平衡精确率与召回率方面的优势。

进一步通过图 11 可以直观对比三种模型的整体性能, 发现集成模型在四个指标上均达到最优或次优, 呈现更均衡的表现。

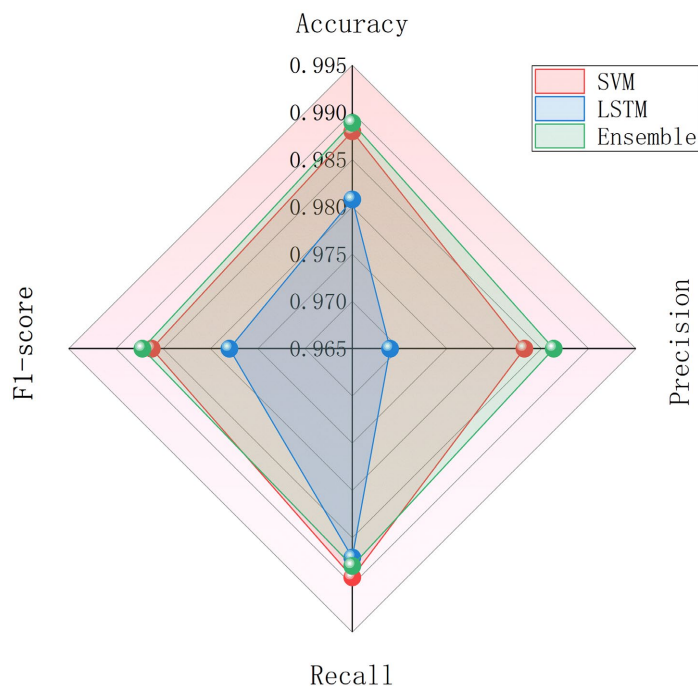


Figure 11. Model performance radar chart
图 11. 性能雷达图

6.4. 综合性能分析

通过 Precision-Recall 曲线(图 12)与 ROC 曲线(图 13)比较进一步验证了模型的稳定性。三种模型的 PR 曲线与 ROC 曲线均接近理想状态, 平均精度(AP)均大于 0.997, AUC 值均大于 0.998, 表明在异常心音检测任务中, 模型几乎无性能瓶颈。

综合而言, 单模型(LSTM 与 SVM)均能提供接近 99%的高准确率, 但在鲁棒性与泛化性上, 集成模型整体表现更优。特别是在减少误判与保持精确率 - 召回率平衡方面, 集成策略展现出明显优势, 这也说明在医疗类异常检测任务中, 集成学习方法具有重要的实际应用价值。

图 14 展示了三篇已有研究与本文所提出的三种模型在相同数据集上的分类准确率对比情况。已有文献中, Al-Shannaq 等的方法准确率为 98.1%, Cheng 和 Sun [28]的模型达到 96.40%, 而 Deng M 等[29]的卷积神经网络模型仅为 98.34%。相比之下, 本文提出的三种模型在相同数据集上均取得了更高的准确率, 其中 SVM 模型为 98.80%, LSTM 模型为 98.08%, 融合模型(Ensemble Model)最高, 达到 98.89%。整体来看, 本文的方法在心音信号分类任务上显著优于其他几篇已有的研究成果, 说明所设计的模型能够更有效地捕捉心音信号的特征, 提高分类性能, 为临床辅助诊断提供更可靠的支持。

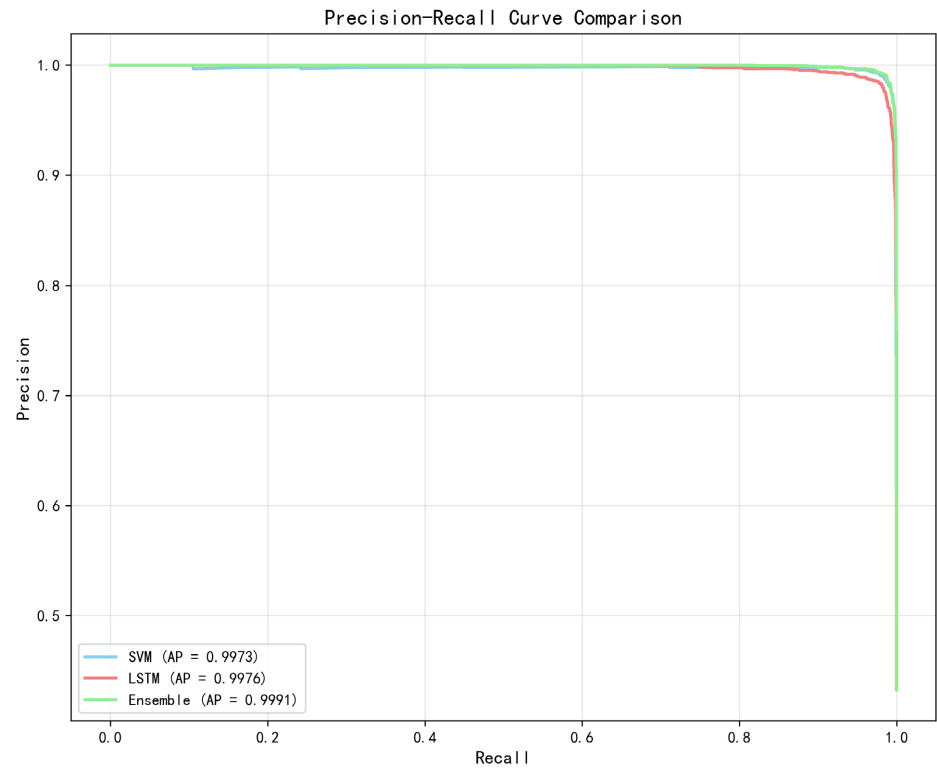


Figure 12. Precision-recall curve comparison
图 12. PR 曲线比较

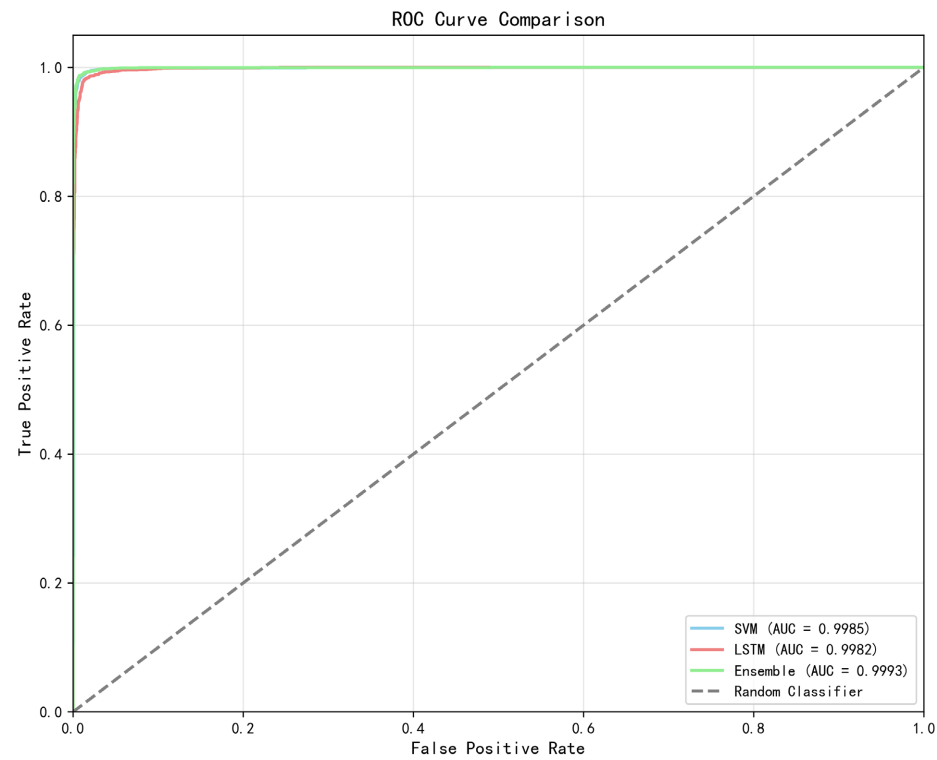


Figure 13. Comparison of ROC curves
图 13. ROC 曲线比较

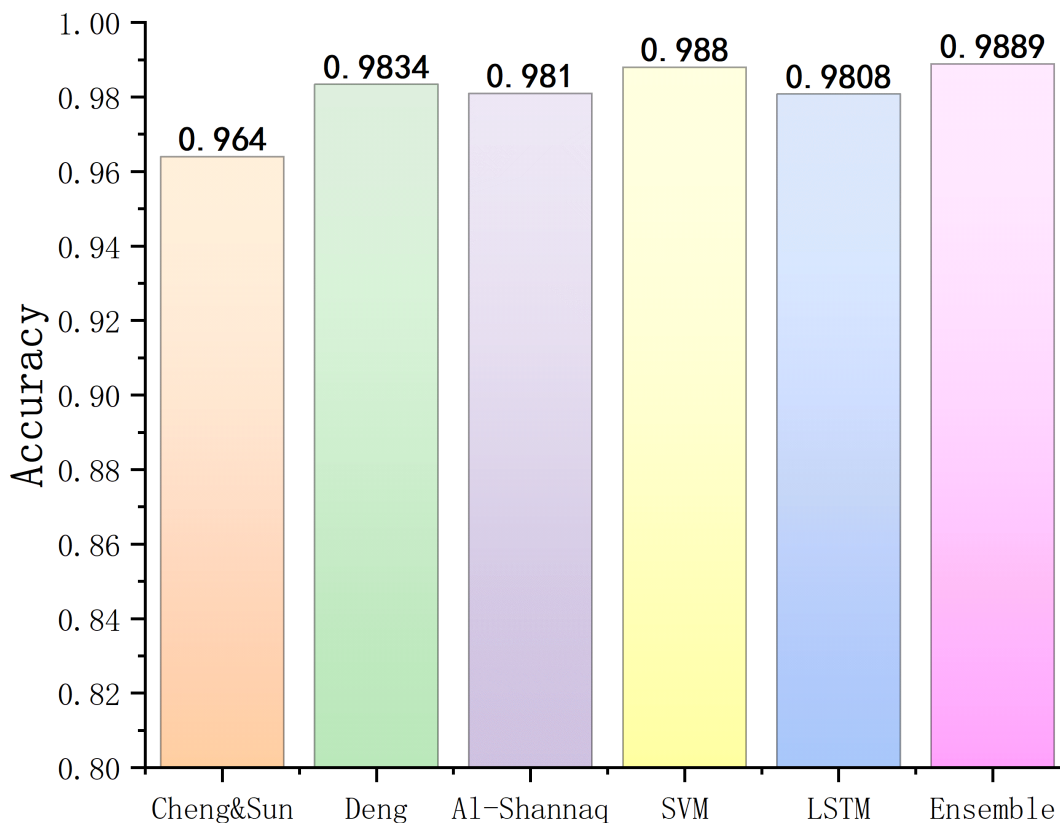


Figure 14. Performance comparison of different methods in normal/abnormal classification

图 14. 不同方法在正常/异常分类中的性能比较

7. 讨论

性能提升与复杂度的权衡：本研究的集成框架虽然因多模型并行导致计算开销高于单一模型，但在医疗诊断中，降低“漏诊率”(False Negative)的价值远高于毫秒级的推理速度。本方法将召回率提升至0.9880，显著降低了漏诊风险。这种以适度增加计算成本为代价换取更高诊断安全性的策略，在临床应用中具有高度的必要性与合理性。

临床挑战与未来方向：尽管模型在跨设备数据集上表现良好，但真实临床部署仍面临设备异构性与非平稳噪声的挑战。此外，深度学习的“黑盒”特性限制了医生的信任度。未来的工作将聚焦于两个方向：一是结合可解释性 AI 与注意力机制，可视化病灶时频区域以辅助医生决策；二是研究模型蒸馏技术，在保持集成性能的同时实现模型轻量化，以适应资源受限的可穿戴设备。

8. 结论

本研究构建了一个基于多模态特征融合与自适应集成学习的心音异常检测系统。通过系统性的实验验证，得出以下主要结论。

多模态特征的互补性：融合时域统计、频域谱质心及小波能谱熵的特征向量，能够从不同维度刻画心音信号的非平稳特性。实验表明，该融合策略比单一特征域提供了更具判别力的特征表示。

异构模型的优势互补：改进的 CNN-BiLSTM 架构利用卷积层提取局部特征并由 BiLSTM 捕捉长时依赖，结合 SVM 在小样本下的分类稳定性，有效克服了单一模型的泛化瓶颈。

集成策略的显著提升：本研究提出的多策略集成框架(特别是元学习器策略)解决了传统加权平均的

局限性。该框架在保持高准确率(98.89%)的同时,显著提升了召回率与 F1 分数,证明了其在降低临床漏诊率方面的实用价值。

综上所述,本文提出的方法通过“特征融合 + 异构模型 + 智能集成”的技术路线,为心脏瓣膜疾病的早期筛查提供了一种高精度、强鲁棒性的自动化解决方案。

参考文献

- [1] 张文波. 心音图在慢性收缩性心力衰竭患者中诊断价值的研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2013.
- [2] Marcus, G., Vessey, J., Jordan, M.V., Huddleston, M., McKeown, B., Gerber, I.L., *et al.* (2006) Relationship between Accurate Auscultation of a Clinically Useful Third Heart Sound and Level of Experience. *Archives of Internal Medicine*, **166**, 617-622. <https://doi.org/10.1001/archinte.166.6.617>
- [3] Lehmann, S.J. (1972) Auscultation of Heart Sounds. *AJN, American Journal of Nursing*, **72**, 1242-1246. <https://doi.org/10.1097/00000446-197207000-00030>
- [4] Yupapin, P., Wardkein, Yupapin, P., Phanphaisarn, Koseeyaporn, Roeksabutr, *et al.* (2011) Heart Detection and Diagnosis Based on ECG and EPCG Relationships. *Medical Devices: Evidence and Research*, **2011**, 133-144. <https://doi.org/10.2147/mder.s23324>
- [5] Liang, H. and Nartimo, I. (1998) A Feature Extraction Algorithm Based on Wavelet Packet Decomposition for Heart Sound Signals. *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis (Cat. No.98TH8380)*, Pittsburgh, 9 October 1998, 93-96. <https://doi.org/10.1109/tfsa.1998.721369>
- [6] Springer, D., Tarassenko, L. and Clifford, G. (2015) Logistic Regression-HSMM-Based Heart Sound Segmentation. *IEEE Transactions on Biomedical Engineering*, **63**, 822-832. <https://doi.org/10.1109/tbme.2015.2475278>
- [7] Yaseen, Son, G. and Kwon, S. (2018) Classification of Heart Sound Signal Using Multiple Features. *Applied Sciences*, **8**, 2344. <https://doi.org/10.3390/app8122344>
- [8] Raza, A., Mehmood, A., Ullah, S., Ahmad, M., Choi, G.S. and On, B. (2019) Heartbeat Sound Signal Classification Using Deep Learning. *Sensors*, **19**, Article 4819. <https://doi.org/10.3390/s19214819>
- [9] Khan, K.N., Khan, F.A., Abid, A., Olmez, T., Dokur, Z., Khandakar, A., *et al.* (2021) Deep Learning Based Classification of Unsegmented Phonocardiogram Spectrograms Leveraging Transfer Learning. *Physiological Measurement*, **42**, Article ID: 095003. <https://doi.org/10.1088/1361-6579/ac1d59>
- [10] Chen, H. and Gu, W.Y. (2025) A Dual Branch Feature Extraction Network for Heart Sound Signal Analysis. *Scientific Reports*, **15**, Article No. 27557. <https://doi.org/10.1038/s41598-025-12303-0>
- [11] Sun, S., Wang, H., Jiang, Z., Fang, Y. and Tao, T. (2014) Segmentation-Based Heart Sound Feature Extraction Combined with Classifier Models for a VSD Diagnosis System. *Expert Systems with Applications*, **41**, 1769-1780. <https://doi.org/10.1016/j.eswa.2013.08.076>
- [12] Li, F., Liu, M., Zhao, Y., Kong, L., Dong, L., Liu, X., *et al.* (2019) Feature Extraction and Classification of Heart Sound Using 1D Convolutional Neural Networks. *EURASIP Journal on Advances in Signal Processing*, **2019**, Article No. 59. <https://doi.org/10.1186/s13634-019-0651-3>
- [13] Guven, M. and Uysal, F. (2023) A New Method for Heart Disease Detection: Long Short-Term Feature Extraction from Heart Sound Data. *Sensors*, **23**, Article 5835. <https://doi.org/10.3390/s23135835>
- [14] Chauhan, V.K., Dahiya, K. and Sharma, A. (2018) Problem Formulations and Solvers in Linear SVM: A Review. *Artificial Intelligence Review*, **52**, 803-855. <https://doi.org/10.1007/s10462-018-9614-6>
- [15] Zha, W., Liu, Y., Wan, Y., Luo, R., Li, D., Yang, S., *et al.* (2022) Forecasting Monthly Gas Field Production Based on the CNN-LSTM Model. *Energy*, **260**, Article ID: 124889. <https://doi.org/10.1016/j.energy.2022.124889>
- [16] He, K., Gkioxari, G., Dollar, P. and Girshick, R. (2017) Mask R-CNN. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/iccv.2017.322>
- [17] Ganaie, M.A., Hu, M., Malik, A.K., Tanveer, M. and Suganthan, P.N. (2022) Ensemble Deep Learning: A Review. *Engineering Applications of Artificial Intelligence*, **115**, Article ID: 105151. <https://doi.org/10.1016/j.engappai.2022.105151>
- [18] Xiao, Y., Wu, J., Lin, Z. and Zhao, X. (2018) A Deep Learning-Based Multi-Model Ensemble Method for Cancer Prediction. *Computer Methods and Programs in Biomedicine*, **153**, 1-9. <https://doi.org/10.1016/j.cmpb.2017.09.005>
- [19] Fernando, K.R.M. and Tsokos, C.P. (2022) Dynamically Weighted Balanced Loss: Class Imbalanced Learning and Confidence Calibration of Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, **33**, 2940-2951. <https://doi.org/10.1109/tnnls.2020.3047335>

-
- [20] Dredze, M., Crammer, K. and Pereira, F. (2008) Confidence-Weighted Linear Classification. *Proceedings of the 25th International Conference on Machine Learning—ICML'08*, Helsinki, 5-9 July 2008, 264-271. <https://doi.org/10.1145/1390156.1390190>
- [21] Douglas, J.M. (1985) A Hierarchical Decision Procedure for Process Synthesis. *AIChE Journal*, **31**, 353-362. <https://doi.org/10.1002/aic.690310302>
- [22] Vilalta, R. and Drissi, Y. (2002) A Perspective View and Survey of Meta-Learning. *Artificial Intelligence Review*, **18**, 77-95. <https://doi.org/10.1023/a:1019956318069>
- [23] Jafar, U., Aziz, M.J.A. and Shukur, Z. (2021) Blockchain for Electronic Voting System—Review and Open Research Challenges. *Sensors*, **21**, Article 5874. <https://doi.org/10.3390/s21175874>
- [24] Heydarian, M., Doyle, T.E. and Samavi, R. (2022) MLCM: Multi-Label Confusion Matrix. *IEEE Access*, **10**, 19083-19095. <https://doi.org/10.1109/access.2022.3151048>
- [25] Ahsan, M., Mahmud, M., Saha, P., Gupta, K. and Siddique, Z. (2021) Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance. *Technologies*, **9**, Article 52. <https://doi.org/10.3390/technologies9030052>
- [26] Clifford, G.D., Liu, C., Moody, B., *et al.* (2016) Classification of Normal/Abnormal Heart Sound Recordings: The PhysioNet/Computing in Cardiology Challenge 2016. 2016 *Computing in Cardiology Conference (CinC)*, IEEE, Vol. 43, 609-612. <https://doi.org/10.22489/CinC.2016.179-154>
- [27] Messer, S.R., Agzarian, J. and Abbott, D. (2001) Optimal Wavelet Denoising for Phonocardiograms. *Microelectronics Journal*, **32**, 931-941. [https://doi.org/10.1016/s0026-2692\(01\)00095-7](https://doi.org/10.1016/s0026-2692(01)00095-7)
- [28] Cheng, J. and Sun, K. (2023) Heart Sound Classification Network Based on Convolution and Transformer. *Sensors*, **23**, Article 8168. <https://doi.org/10.3390/s23198168>
- [29] Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J. and Fan, H. (2020) Heart Sound Classification Based on Improved MFCC Features and Convolutional Recurrent Neural Networks. *Neural Networks*, **130**, 22-32. <https://doi.org/10.1016/j.neunet.2020.06.015>