

深度学习心音信号分类研究进展

邓漆时超^{1,2}, 吴佳成^{1,2}, 俞键^{1,2}, 张华², 李川涛³, 林冬梅⁴, 陈扶明²

1. 甘肃中医药大学医学信息工程学院, 甘肃 兰州 730000; 2. 中国人民解放军联勤保障部队第九四〇医院医疗保障中心, 甘肃 兰州 730050; 3. 海军军医大学(第二军医大学)海军特色医学中心, 上海 200433; 4. 兰州理工大学电气与信息工程学院, 甘肃 兰州 730050

【摘要】心音是反映心脏健康状况的重要生理信号,心音分类可辅助心脏疾病的早期诊断。相比依赖人工特征提取的传统方法,深度学习能够自动学习心音特征,在分类精度和鲁棒性方面具有明显优势。本文通过系统检索国内外相关文献,围绕深度学习在心音分类领域的研究展开综述。首先对心音信号的特点、常用数据集及预处理方法进行梳理,其次归纳卷积神经网络、循环神经网络、Transformer及多模型融合等深度学习方法在心音分类任务中的应用现状,并进一步从模型结构、分类性能和临床适用性等方面进行比较与分析。最后,讨论深度学习在心音分类任务中的应用前景,以期促进深度学习方法在心音信号分类中的应用。

【关键词】心音信号;深度学习;分类算法;神经网络;综述

【中图分类号】R318;TP183

【文献标志码】A

【文章编号】1005-202X(2026)07-0946-12

Advances in heart sound classification based on deep learning

DENG Qishichao^{1,2}, WU Jiacheng^{1,2}, YU Jian^{1,2}, ZHANG Hua², LI Chuantao³, LIN Dongmei⁴, CHEN Fuming²

1. School of Medical Information Engineering, Gansu University of Chinese Medicine, Lanzhou 730000, China; 2. Medical Support Center, 940th Hospital of Joint Logistics Support Force of Chinese People's Liberation Army, Lanzhou 730050, China; 3. Naval Characteristic Medical Center, Naval Medical University (Second Military Medical University), Shanghai 200433, China; 4. School of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China

Abstract: Heart sounds are vital physiological signals that reflect cardiac status, and heart sound classification plays a crucial role in the early diagnosis of cardiovascular diseases. Compared with traditional methods that rely on handcrafted feature extraction, deep learning techniques can automatically learn representative features from heart sound signals, achieving superior classification accuracy and robustness. This study provides an overview of recent advances in deep learning-based heart sound classification by retrieving and analyzing relevant domestic and international research. This review systematical introduces the characteristics of heart sound signals, commonly-used datasets, and preprocessing methods, summarizes the application status of deep learning approaches, including convolutional neural networks, recurrent neural networks, Transformer-based models, and hybrid architectures, in heart sound classification, and further compares their model architectures, classification performance, and clinical applicability. Finally, the application prospects of deep learning in heart sound classification is discussed, aiming to promote the application of deep learning techniques in heart sound classification.

Keywords: heart sound signal; deep learning; classification algorithm; neural network; review

【收稿日期】2026-03-20

【基金项目】国家自然科学基金(61901515, 62361038);兰州市科技发展指导性计划(2025-5-243);兰州市青年科技人才创新项目(2025-QN-145)

【作者简介】邓漆时超,硕士研究生,研究方向:生物医学信号检测与处理,E-mail: 1095292703@qq.com

【通信作者】陈扶明,正高级工程师,研究方向:生物医学信号检测与处理,E-mail: cfm5762@126.com

前言

近年来,心血管疾病(Cardiovascular Disease, CVD)患者逐年增加,根据世界卫生组织(World Health Organization, WHO)统计,CVD是全球一大死亡原因。在非传染性疾病死亡人数中,CVD占比最大,每年有1 790万人死亡。CVD对全球健康构成严峻挑战,因此,早期诊断和病情监测尤为重要。心音信号反映了心脏运动的规律,其中潜藏的异常特征

往往是心脏疾病的预警,早期识别心音中的异常特征,能够为CVD临床诊断和治疗决策提供关键依据。人耳可听到声音的频率范围为20~20 000 Hz,对1 000~2 000 Hz的声音信号较为敏感,而正常心音信号的频率范围为20~200 Hz,这恰好为人耳敏感度较低的区域,因此,心音信号分类技术在临床诊断中的意义不容小觑,精准的心音分类还能优化临床决策流程,有效提高医疗资源的利用效率。

随着深度学习技术的发展,心音分类的准确度正不断攀升,为临床诊断提供强有力的技术支持。目前国内外已有一些基于深度学习的心音分类研究,但仍存在一些挑战和问题:心音信号具有复杂性和多样性特点,不同个体、不同疾病状态下的心音信号差异较大;心音信号通常带有背景噪声和干扰;大

部分模型仅能实现心音信号的二分类,多分类的准确率有待提高;心音分类模型仍需加强与硬件的结合,最终用在实际中。这些问题会影响模型的分类准确性和泛化能力,因此,如何在复杂噪声环境下有效提取心音信号特征,提高模型的鲁棒性、准确性和进一步对心音信号进行多分类是当前研究的重要方向。

详细行文脉络如图1所示,本文系统性地回顾国内外深度学习技术在心音分类领域应用的相关文献,分析该技术在心音信号分类中的发展现状,并对各类深度学习模型的优势与挑战进行探讨,同时对深度学习技术在心音信号分类中的未来发展趋势进行展望,以期为推动深度学习技术在心音信号分类中的广泛应用提供理论支持和实践指导。

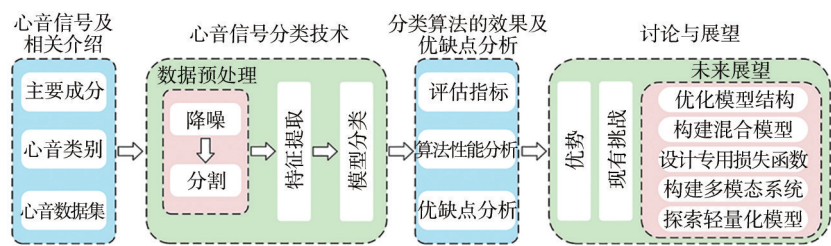


图1 组织结构图

Figure 1 Organizational structure diagram

1 心音信号及相关介绍

1.1 心音信号介绍

心音指由心肌收缩、心脏瓣膜关闭和血液撞击心室壁、大动脉壁进而产生振动所引发的声音,可在胸壁一定部位使用听诊器采集^[1]。一般心音频率为20~200 Hz,当心脏出现病变而异常工作时,心音频率甚至 would 超过1 400 Hz。心音周期包括基础心音以及收缩期、舒张期,基础心音由第一心音(First Heart Sound, S1)和第二心音(Second Heart Sound, S2)组成。S1是由二尖瓣、三尖瓣的关闭产生的,预示着心脏收缩期的开始;S2是由主动脉瓣和肺动脉瓣的关闭产生的,预示着心脏舒张期的开始^[2]。第三心音(Third Heart Sound, S3)和第四心音(Fourth Heart Sound, S4)也是心音信号的主要分段。心音不同分段频率各不相同,S1与S2的频率范围为20~150 Hz,较易被听见。S3和S4频率较弱,为10~50 Hz^[3],不易被听见。正常心音由心脏瓣膜关闭及血流动力学活动产生,主要包括S1和S2,异常心音则多与心脏结构或功能异常相关。

不同的杂音预示着不同的CVD,如图2所示,常

见的CVD主要包括二尖瓣狭窄、二尖瓣反流、主动脉瓣狭窄和主动脉瓣关闭不全等。表1说明正常心音与4种瓣膜病变的听诊特征和成因。此外,心音还可能包含肺音及环境噪声等成分。

1.2 心音数据集介绍

全面的心音数据集有助于训练和优化深度学习模型,提高心脏疾病的诊断准确性和效率。目前已有许多公开可用的数据集,包括PASCAL心音挑战数据集^[4]、Yaseen数据集^[5]、PhysioNet CinC挑战数据集(2016)^[6]、Kaggle竞赛数据集、CirCor DigiScope数据集^[7]、深圳心音数据库(Heart Sounds Shenzhen Corpus, HSS)^[8]以及心音和杂音库等。表2列举了常见心音数据集。

PASCAL数据集包含数据集A和B,每条心音的长度为1~30 s,其中数据集A包含4类心音信号:正常、杂音、额外心音和人工干扰,数据库B包含3类心音信号:正常、杂音和早搏。Yaseen数据集包含正常心音、主动脉狭窄心音、二尖瓣回流心音、二尖瓣狭窄心音和收缩期杂音各200条,每条心音信号时长为2 s,采样频率为8 kHz。PhysioNet CinC挑战数据集(2016)共计3 126条心音记录,持续时间为5~120 s,

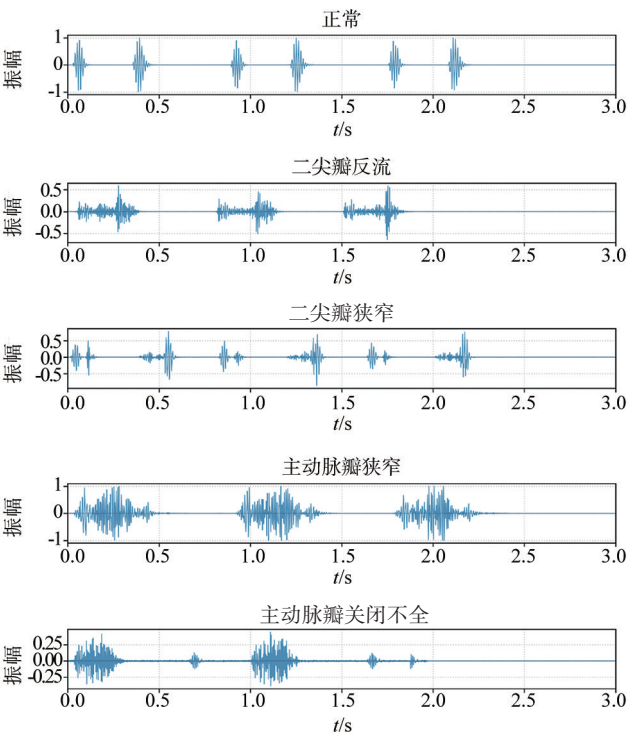


图2 心音波形对比图
Figure 2 Comparison of heart sound waveforms

包含两个标签:正常和异常。Kaggle 竞赛数据集共 620 条心音录音,包含 3 个标签:正常、异常和未分类,持续时间为 5~30 s。CirCor DigiScope 数据集是迄今为止最大的小儿心音集合,共计 5 282 条录音,录音时长为 5~168 s,数据集对心音基础分类为正常、异常和低质量信号,对杂音多维度进行分类,如发生时间、形状和音调等特征进行详细标注。HSS 是由深圳大学总医院录制的心音数据库,共包含 3 个标签:正常、轻度和中度/重度,每条心音信号采样率为 4 kHz,录音时长平均为 30 s,共录得 845 条样本。密歇根大学卫生系统的心音和杂音库包括正常心音、S1、S4 奔马律、S3 奔马律、单侧 S2 等和各种类型的杂音。

2 心音信号分类技术

全球的死亡案例中,CVD 是主要的致死原因之一,心音分类研究已成为医学领域的热点话题,国内外越来越多研究人员开始研究心音的分类技术,如图 3 所示,心音信号分类步骤为数据预处理、特征提取、训练模型和模型分类。

表1 各类心音的声学特征及其病理原因
Table 1 Acoustic characteristics and pathological causes of different types of heart sounds

心音类别	心音特征	产生原因
正常心音	S1 和 S2 清晰,无杂音	二尖瓣与三尖瓣关闭,主动脉瓣与肺动脉瓣关闭
二尖瓣狭窄	S1 增强,舒张期杂音,可能伴有开瓣音	二尖瓣狭窄导致左心房血液流入左心室受阻,产生湍流
二尖瓣反流	S1 减弱,全收缩期杂音	二尖瓣关闭不全导致血液在收缩期反流至左心房,产生湍流
主动脉瓣狭窄	S2 减弱或消失,收缩期喷射性杂音	主动脉瓣狭窄导致左心室血液流出受阻,产生湍流
主动脉瓣关闭不全	S2 减弱,舒张期递减型杂音,可能伴有 Austin Flint 杂音	主动脉瓣关闭不全导致血液在舒张期反流至左心室,产生湍流

表2 常见心音数据集
Table 2 Commonly-used heart sound datasets

数据集	分类数/个	时长/s	音频数/个
PASCAL 心音挑战数据集	数据集 A:4	1~30	数据集 A:176
	数据集 B:3		数据集 B:656
Yaseen 数据集	5	2	1 000
PhysioNet CinC 挑战数据集(2016)	2	5~120	3 126
CirCor DigiScope 数据集	多维度	50~168	5 282
HSS 数据集	3	30	845

2.1 数据预处理过程

2.1.1 心音降噪 心音信号在采集过程中常受到外部环境影 响,如功率频率干扰、随机噪声干扰和呼吸声干扰等,因此,在预处理过程中可以通过一系列方法

减少噪音的影响。作为心音分类的预处理过程,许多去噪方法采用信号滤波器去除心音中的噪声^[9-10],如椭圆滤波器、巴特沃斯滤波器等,常用的心音去噪技术还包括自适应去噪、变分模态分解(Variational

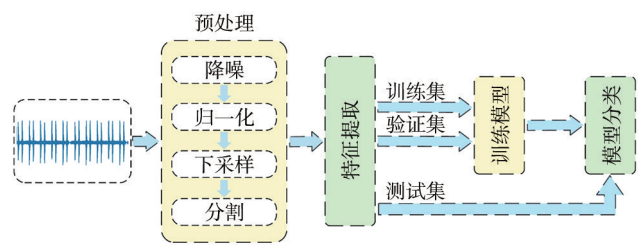


图3 心音信号算法流程

Figure 3 Heart sound signal classification algorithm

Mode Decomposition, VMD)^[11-12]、离散小波变换 (Discrete Wavelet Transform, DWT)^[13-14]、奇异值分解 (Singular Value Decomposition, SVD) 和经验模态分解 (Empirical Mode Decomposition, EMD) 等。

如图4所示,VMD方法因其自适应性强、低耦合性和算法简单等优点用于心音信号去噪研究当中,但也存在参数选择难度大、计算复杂度较高和模态混叠现象等问题,需要合理选择参数。EMD是另一种有效的去噪方法,它由Huang等^[15]于1998年首次提出,该方法是针对非平稳和非线性信号的自适应数据分析方法,无需任何预先已知基函数,但容易出现模态混叠。DWT方法具有计算简便、有效抑制噪声和广泛适用性等优点,其降噪效果如图5所示,在心音信号处理中可有效分离心音成分与高频噪声,但对参数选择和阈值设定要求较高,对非平稳噪声抑制不足。自适应去噪算法通过动态调整滤波器参数,在实时噪声跟踪中表现突出,适应性强,在有效保留原始信号特征的同时实现优异的去噪效果,但算法复杂度较高,在处理非平稳复杂信号时可能导致欠去噪或过去噪。SVD则擅长从矩阵维度去除结构化周期噪声,适用于复杂心音数据集,但对于长信号分解计算复杂度较高,易导致瞬态病理特征丢失。

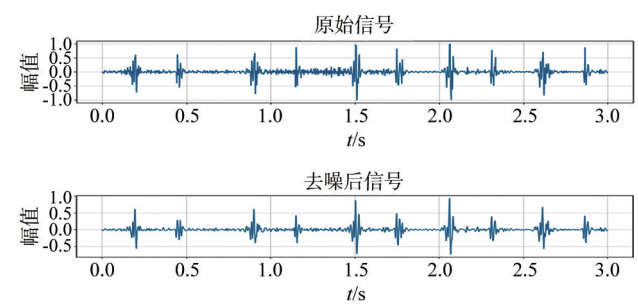


图4 VMD降噪前后心音信号对比

Figure 4 Comparison of heart sound signals before and after VMD denoising

此外,深度学习在心音去噪任务中也展现出良好的效果,并取得一定进展。Ali等^[16]提出结合

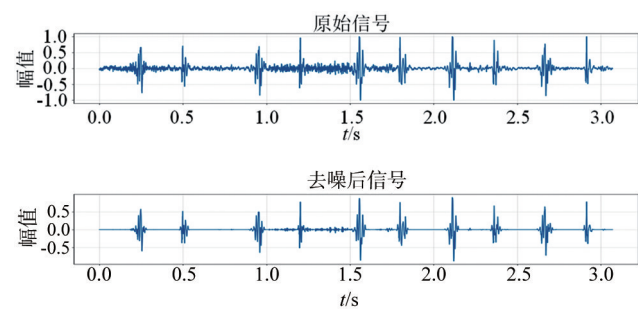


图5 心音信号经DWT降噪前后的对比

Figure 5 Comparison of heart sound signals before and after DWT denoising

LSTM和U-Net端到端深度学习去噪模型,捕捉心音的准周期性时序特征,为低资源嘈杂环境下的CVD自动化诊断提供高效解决方案。Liang等^[17]针对传统心音去噪方法易受主观参数影响、导致信号失真的问题,提出基于生成对抗网络 (Generative adversarial network, GAN) 的心音去噪模型。Al-Zaben等^[18]提出运用修正余弦变换和基于长短期记忆网络 (Long Short-Term Memory, LSTM) 的掩码策略分离心音与噪声。由此可见,采用组合方法获得更好的效果,有助于提高信号质量和检测精度。

2.1.2 心音分割 心音分割作为心音信号处理的常用方法,主要是划分心音周期为4个对应状态,也是心音分类的重要处理步骤^[19]。常用的心音分割方法主要有基于包络的分割方法^[20-21]、基于心电图 (Electrocardiogram, ECG) 信号的分割方法、基于特征的分割方法^[22-23]、基于机器学习的分割方法^[24]、基于隐马尔可夫模型 (Hidden Markov Model, HMM) 的分割方法和基于隐半马尔可夫 (Hidden Semi-Markov Model, HSMM) 的分割方法等^[25-27]。近年来,基于深度学习方法在心音分割中已展现出优势,逐渐成为主流技方向。受卷积神经网络 (Convolutional Neural Networks, CNN) 在图像分割中的启发,心音分割研究也开始探索应用CNN技术,优化心音分割的精度。Enériz等^[28]和Park等^[29]借鉴图像分割技术,采用U-Net架构提升分割精度。同样,基于循环神经网络 (Recurrent Neural Network, RNN) 处理时序信号的独特优势,研究者将其引入心音分割领域^[30]。Chen等^[31]提出利用卷积长短期记忆网络捕捉心音数据的局部特征与长程时序关联,提升复杂心音的边界识别精度。Guan等^[32]提出CNN-LSTM的分割方法,先采用CNN提取心音信号的局部时频特征,再通过LSTM捕捉其长时间依赖信息,表现优于传统的阈值法与HMM方法。表3更直观地呈现不同心音分割方法的特性与不足,从多个方面对不同心音分割技术进行梳理和对比分析。

表 3 心音信号分割研究情况
Table 3 Research on heart sound segmentation

类别	方法	文献	特点	局限性
传统信号处理	心音包络	[20-21]	能清晰反映心音信号的整体轮廓,便于定位心动周期;实现简单、计算速度快,易于实时处理;对边界特征的突出效果较好,可用于快速的初步分割与筛查	易受环境噪声干扰;平滑过程可能掩盖心音信号中的细微或瞬态病理特征;对包络提取算法的参数选择高度敏感;临床解释性不足
特征驱动方法	时频特征	[22-23]	同时反映心音信号的时域和频域变化;对信号的短时动态特征敏感,能准确识别心音周期;可捕捉心音中瞬态及异常特征,有利于病理信号识别;在心音分割中可突出周期性边界信息,提升心音各阶段的识别精度	时频分析过程计算复杂度较高,耗费计算资源,不适用于实时或可穿戴设备;对噪声敏感,信号质量差时特征提取精度显著下降;特征的有效性取决于参数设定;对前处理依赖性强,处理不当将降低分割性能
机器学习	HMM	[25-26]	善于捕捉心音信号的时序依赖特性;能够有效处理信号中的不确定性与随机波动;参数具有较强的可调节性,适用于多种临床应用场景;模型结构相对清晰,具备一定的生理学解释性	依赖于训练数据的质量和数量;模型训练及参数估计过程复杂,计算开销较大;对信号中特殊或异常模式敏感
	K-Means	[24]	以无监督方式实现心音信号的自动分割,方法简单高效;对初始参数和算法实现要求较低,便于快速部署;适用于大规模心音数据的快速聚类分析	易受心音信号中噪声和异常值的干扰,鲁棒性不足;聚类结果缺乏明确的生理学解释,临床可信度不足;特征选择对分割性能影响显著,泛化性受限
深度学习	CNN	[28-29, 32]	能够自动、高效地提取心音信号的时域与频域特征;对不同的心音信号具有良好的适应性与泛化能力;擅长捕捉复杂的局部模式,从而提升心音分割与分类精度;对心音信号中的随机噪声具有一定的鲁棒性	训练过程需要大量高质量的心音数据;计算复杂度较高,对硬件资源和训练时间要求大;对小样本数据易产生过拟合,泛化能力受限;模型决策过程为黑盒,缺乏直观的临床生理解释
	RNN	[30-32]	有效处理心音时序和周期性特征;对于信号中的非线性关系建模能力强;对心音异常节律、噪声干扰表现出较好的鲁棒性	训练时间较长,计算复杂度较高;对训练样本规模和质量要求较高;缺乏明确的生理学解释,可解释性弱

2.2 特征提取

心音信号分类过程指的是通过信号的音频特征来推理信号的类别^[33],即在每个心音片段上提取用于分类的特征来完成分类任务。如表 4 所示,常用的提取心音特征方法有短时傅里叶变换(Short-Time Fourier Transform, STFT)^[34],梅尔频率倒谱系数(Mel-Frequency Cepstral Coefficient, MFCC)^[35-36],频谱图^[37-38]、DWT^[39]和连续小波变换(Continuous Wavelet Transform, CWT)^[40-41]等。Telceken 等^[42]提出一种名为双频倒谱系数的新型特征提取方法,通过结合 Mel 滤波器和 Gammatone 滤波器结构,并利用 DFT 保留振幅和相位信息,实现更全面的时频域分析。Patwa 等^[43]使用小波散射变换在特征提取方面发挥关键作用,为后续神经网络分类提供重要特征基础。Yang 等^[44]提出一种基于 EMD 和 MFCC 相结合的模糊匹配特征提取策略,在低计算开销下能够有效提高心音信号的识别精度,且特征具有良好的可解释性和类间区分度,尤其适用于可穿戴医疗设备等资源受限场景。在心音特征提取中,基于深度学习的方法也逐渐成为主流方法^[45],如 Bahreini 等^[46]提出基于预训练 CNN 模型的心音信号时频特征提取方法,能够更好地识别 CVD。Li 等^[47]采用一维卷积

神经网络与去噪自动编码器相结合的策略,与传统 MFCC 特征相比,该方法提取的深度特征能更好地保留心音信号的关键细节和抗噪能力。Khan 等^[48]通过特征融合技术,拼接 MobileNet 与 DenseNet201 的特征并结合残差学习优化层级特征传递,实现异常心脏节律检测中特征提取能力与模型效率的协同提升。

2.3 基于深度学习的分类算法

传统机器学习方法受限于特征质量,依赖研究者对心音生理机制领域知识的掌握程度,而深度学习通过构建和训练多层神经网络来自动提取和学习特征,降低对人工经验与先验知识的依赖性。目前已证明深度学习技术在分析心音方面的能力高于传统机器学习^[49]。深度学习通常接受原始音频信号或时频表示作为输入从而跳过手工声学特征提取来提升效率^[50]。

心音分类研究中,CNN 凭借其特有的结构优势展现出强大的潜力和广泛的应用前景,CNN 模型显著提升了心音的鉴别精度,其模型结构如图 6 所示^[51-54]。根据文献^[51-55]研究,基于 CNN 的心音信号分类方法在提升分类性能的同时,还显著提高了计算效率。CNN 通过局部感受野高效捕获心音信号

表4 心音信号特征提取研究情况
Table 4 Research status of heart sound feature extraction

方法	文献	特点
STFT	[34]	计算高效,实时性强,但窗长固定,无法同时捕捉瞬态细节和低频特征
MFCC	[35-36]	基于人耳听觉特性,提取梅尔频带的倒谱系数
频谱图	[37-38]	无需人工设计特征,直接输入深度学习模型,但数据维度高
DWT	[39]	多分辨率特性适配心音瞬态特征,如S1/S2边界,但基函数选择依赖经验
CWT	[40-41]	精准定位非稳态成分,但计算复杂度高,难以实时处理
深度学习	[45-48]	能够自动提取复杂特征、处理非线性信号、提高分类精度且可融合多模态数据,但存在数据需求大、模型复杂、解释性差和过拟合风险等缺点

的时频细节,提取心音信号中的局部特征和全局特征,能够应对不同个体、不同环境条件下心音信号的差异,具有较好的泛化能力和鲁棒性。通过训练大量的心音数据学习心音信号的统计规律和特征分布,CNN在新的、未见过的心音样本上也能取得较好的分类效果。Bao等^[56]围绕心音信号分类展开,对比多种时频分布方法与CNN组合的效果填补性能比较的空白,提升心音分类的准确性与效率。此外,结合残差网络(Residual Neural Network, ResNet)模型开展心音分类工作也具有不可忽视的重要性与独特价值^[37]。Qiao等^[57]提出联合学习四元数动态表示(Quaternion Dynamic Representation with Joint Learning, QDRJL)神经网络,通过动态卷积和联合学习优化多通道特征表达,在保证高精度的同时参数减少75%。

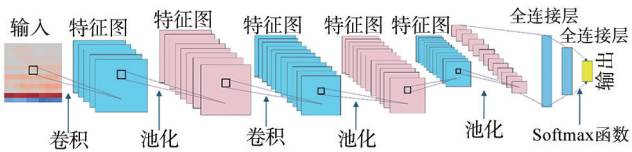


图6 用于心音信号处理的CNN模型架构
Figure 6 CNN model architecture for heart sound processing

RNN则擅长建模心音的长时序依赖关系,捕捉心动周期节律变化,RNN以其独特的循环结构,能够有效捕捉序列数据中的时间依赖性,这对于心音信号这种具有明显时序特征的数据来说尤为关键。众多研究者利用RNN及其变种,如LSTM^[58-59]、门控循环单元(Gate Recurrent Unit, GRU)和最小门控单元等构建心音分类模型进而实现对心音类别的准确分类^[60]。

此外,混合算法模型能够同时捕捉心音信号的空间特征和时序特征,在整合局部与全局特征方面展现出卓越性能,为心音信号分类提供新的解决思

路^[61]。混合模型通常结合多种神经网络,如CNN、RNN、LSTM和GRU^[62-63]。如图7所示,Zhang等^[64]提出一种多特征决策融合网络以解决传统心音诊断方法的不足。Choudhary等^[65]提出一种CNN和GRU的混合模型,在公开数据集的总体准确率达到99.3%。Choudhary等^[66]提出结合CNN和LSTM的心脏诊断系统,为心脏疾病诊断提供新的有效方法。Li等^[67]提出一种基于多尺度密集网络和多头循环神经网络技术的端到端心音分类方法,具有较高的准确率和可解释性。Ma等^[68]提出一种参数高效的密集连接双注意力网络,在性能和参数效率之间取得较好的平衡。Lin等^[69]提出一种多尺度CNN特征融合Bi-LSTM系统,为冠心病无创诊断提供实用方案。

Transformer模型凭借其强大的自注意力机制^[70-72],能够高效地捕捉心音信号中的时序依赖关系,从而实现高精度的心音分类^[73-74]。Liu等^[75]提出一种新的基于双频特征和Vision Transformer模式的心音分类,其模型结构如图8所示,在测试集上准确率为0.91。Cheng等^[71]提出卷积与Transformer编码器神经网络的心音分类方法,在二分类和多分类任务中表现出色。Wang等^[76]提出可解释小波卷积Transformer框架,在公开数据集实现99.98%准确率。

迁移学习采用在大规模数据集上优化且预先训练深度学习模型^[77],在心肺音分类领域也有一定的应用。Barnawi等^[78]通过数据选择和迁移学习实现高效的心音分类。Almanifi等^[79]比较ResNet50、VGG16和VGG19迁移学习模型,发现频谱图-ResNet50组合性能较以往有提升。Ding等^[80]采用迁移学习训练GoogLeNet模型,在检测心脏瓣膜疾病方面表现较好。Harimi等^[81]提出一种基于CyTex的变换和CNN迁移学习方法。Jumphoo等^[82]提出一种基于迁移学习的Conv-DeiT混合架构,通过预训练Transformer模型实现瓣膜性心脏病的有效识别。此外,在迁移学习中微调预训练模型也表现出良好的心音分类性能^[83-84]。

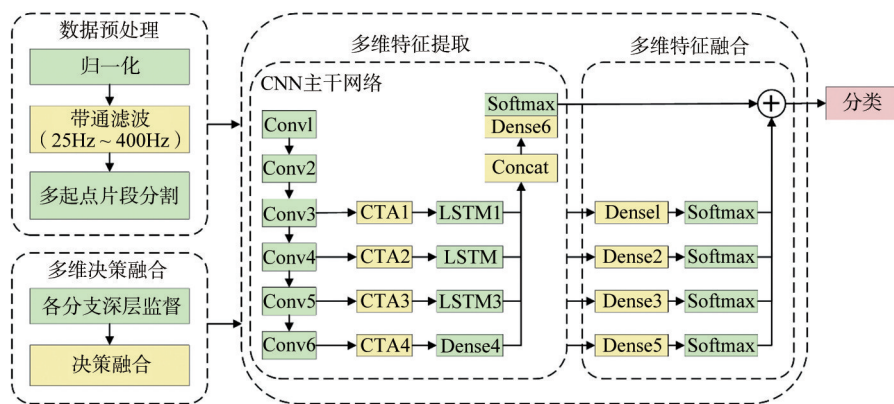


图7 多特征决策融合网络

Figure 7 Multi-feature decision fusion network

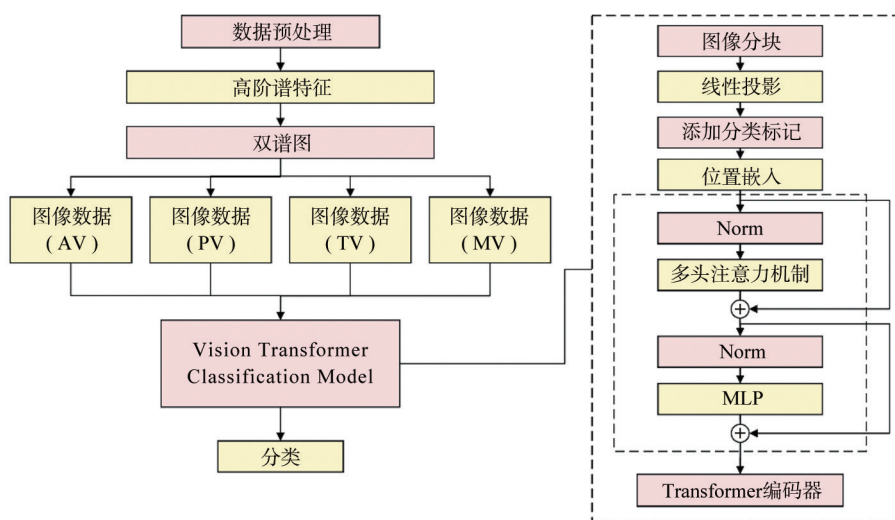


图8 双谱-ViT心音诊断框架流程图

Figure 8 Workflow of the bispectrum-based vision Transformer (ViT) framework for heart sound diagnosis

3 心音分类算法的效果及优缺点分析

3.1 算法性能评估指标

模型评估对于构造可靠的分类模型至关重要,因为它可以对模型性能提供量化衡量^[85]。混淆矩阵是一种表示精度评价的标准格式,在深度学习的分类任务中,用于展示模型预测结果与实际标签之间的差异^[86]。在二分类问题中,混淆矩阵通过4个基础指标系统量化模型的分类性能:真阳性(True Positive, TP),指实际为阳性的样本被正确识别为阳性;真阴性(True Negative, TN),表示实际为阴性的样本被准确判定为阴性;假阳性(False Positive, FP),指阴性样本被错误归类为阳性;假阴性(False Negative, FN),指阳性样本被误判为阴性。这4个指标共同揭示了模型在敏感度与特异性之间的权衡关系,是评估医学分类任务可靠性的核心依据。以心音诊断为例,正确识别异常心音为TP,准确排除健康

心音为TN;FP会导致不必要的复查,FN则会延误病情诊断。

通过混淆矩阵,可以进一步引出多种指标来评估分类器的性能,准确率、灵敏度、特异性和F1值是算法模型评估的核心评价指标,从不同角度反映模型的分类能力。

准确率是指模型预测正确的样本数占总样本数的比例,反映模型对整体样本的预测能力:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

灵敏度是指模型正确识别出的阳性样本数占实际阳性样本数的比例,反映模型对阳性样本的识别能力:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

特异性是指模型正确识别出的阴性样本数占实际阴性样本数的比例,反映模型对阴性样本的识别能力:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

(3)

F1 值是精确率和召回率的调和平均数,用于综合评估模型对阳性样本的识别性能:

$$\text{F1} = 2 \times \frac{\frac{\text{TP}}{\text{TP} + \text{FP}} \times \frac{\text{TP}}{\text{TP} + \text{FN}}}{\frac{\text{TP}}{\text{TP} + \text{FP}} + \frac{\text{TP}}{\text{TP} + \text{FN}}}$$

(4)

3.2 分类算法性能分析

尽管各算法在准确率、泛化性等指标上存在一定差异,但总体而言,以 CNN、ResNet 为代表的卷积神经网络,以及 GRU、LSTM 等循环神经网络,在心音信号处理任务中都表现出良好的性能,展现出对心音特征的有效表征能力。Transformer 架构作为近年来兴起的深度神经网络模型,凭借其全局建模能力和自注意力机制,在心音分类等序列建模任务中

展现出极高的潜力。

不同算法适用于不同类型的心音信号数据集和任务,各分类算法性能总结具体如表 5 所示^[37, 52, 54, 59-61, 63, 65-66, 71, 73-74, 80]。CNN 在处理图像或静态特征时表现出色,通过卷积核的局部感受野特性,提取心音时频谱图中的空间局部特征,而 RNN 则更适合处理时间序列数据。因此,在选择算法时应考虑数据集的特点和任务需求。同时,随着算法复杂度的增加,如引入注意力机制、结合多种网络结构等,模型的性能往往会有所提升,但也会带来更高的计算需求和资源消耗。一些算法在多个数据集上表现出良好的泛化能力,这意味着它们可能更适用于不同类型的心音信号处理任务,但这并不意味着其他算法缺乏泛化能力,需要根据具体情况进行评估。

表 5 算法性能对比(%)

Table 5 Comparison of algorithm performance (%)

算法	文献	发表时间	数据库	准确率	灵敏度	特异性	F1
ResNet	[37]	2023	CirCorDigiScope PCG	-	86.30	-	81.60
CNN	[52]	2024	PhysioNet CinC 挑战数据集(2016)	96.30	90.60	97.90	91.10
	[54]	2024		98.06	97.74	98.39	-
LSTM	[59]	2024	PhysioNet CinC 挑战数据集(2016)	96.93	93.50	99.63	-
MGU	[60]	2024	PhysioNet CinC 挑战数据集(2016)	97.20	96.61	96.67	96.82
CNN-RNN	[61]	2023	多个数据集	91.17	98.80	98.20	-
RNN-LSTN	[63]	2022	PASCAL 心音挑战数据集	99.77	98.60	99.31	98.30
CNN-GRU	[65]	2024	Yaseen 数据集	99.30	99.29	99.81	99.29
CNN-LSTM	[66]	2024	Yaseen 数据集	98.90	99.20	99.80	99.00
Transformer	[71]	2023	PhysioNet CinC 挑战数据集(2016)	96.40	92.87	97.45	-
	[73]	2025	PASCAL 心音挑战数据集	97.52	-	-	98.00
	[74]	2023	Yaseen 数据集	99.36	-	-	99.30
GoogLeNet	[80]	2023	Yaseen 数据集	97.50	96.88	99.22	97.99

3.3 优缺点分析

与传统方法相比,深度学习方法避免了手动特征提取带来的问题,确保更大样本量和更好泛化性,有助于为未来的机器辅助听诊做出实时决策。心音分类的研究可以进一步开发使用具有更合理有效的模块和超参数的 CNN 或 RNN 结构或对更多的音频数据增强来提高分类准确性。此外,随着对心音听诊的进一步理解,可以通过设置更合理的从心音片段判断到心音样本判断的传递规则来提高模型的分类效果。目前,心音分类的研究仍存在局限性,存在部分模型无法区分不同类型心脏病、模型训练时间

长、输入心音段时域未对齐等问题。表 6 总结了对深度学习心音分类算法的综合评价。

4 讨论与展望

听诊是针对听呼吸声的普通而简单的方法,但也有缺点,因为该方法是一项高度主观的工作,主要依赖于观察者的经验和训练,因此呼吸声的采集本身将会成为高度非线性的噪声源^[87]。随着时代的快速发展,未来智能医疗设备如电子听诊器将逐步进入家庭,心音分类技术的发展将十分广阔,如构建多模态数据融合的心音与其他生理信号融合分类模

表6 深度学习心音分类算法综合评价

Table 6 Evaluation of deep learning-based heart sound classification algorithms

算法	优势	局限性
CNN	自动提取局部特征,适合处理心音信号的时频表示,如波形图、频谱图;对基线漂移和局部噪声有良好的鲁棒性;计算效率较高,适合实时心音监测和床旁设备应用	对长时间依赖的心音节律变化建模不足;当心音数据较少时,深层CNN容易过拟合;对于复杂的跨周期依赖或多音重叠情况,表达能力有限
ResNet	具备强大的深层特征提取能力,能处理复杂、非平稳心音信号;有效缓解梯度消失问题,提高分类精度;在噪声较强或心音形态复杂的场景下表现优异	参数量和计算复杂度较高,训练耗时长,对硬件资源依赖大,不利于可穿戴或实时心音监测应用;不具备显式的时间建模能力,难以捕捉心音随时间演变的动态信息
GRU	能有效捕捉心音信号的时间序列特性,尤其适合处理心音周期中的时序依赖;相比LSTM,训练速度更快,参数更少,适合资源有限的嵌入式设备	在长时间依赖建模中可能不如LSTM稳定;在复杂模式识别中表现可能略逊一筹
MGU	能够捕捉时间序列特征,与多模态数据兼容性好,适合轻量化部署,如移动设备、可穿戴监测器等;在小数据集条件下对过拟合有一定缓解作用	相较于GRU和LSTM,对复杂模式的学习能力仍相对较弱,难以捕捉复杂心音变化,对数据量依赖性较大
LSTM	擅长捕捉心音信号中的长时间依赖与节律特征,能够捕捉长时间依赖关系,适合处理心音信号等时间序列数据;在杂音识别、病理性心音分类等任务上具有较强表现	模型大、计算耗时,在大数据集上训练代价高;对超参数敏感,训练调优复杂;难以直接应用于实时、低功耗的可穿戴设备
CNN-LSTM	CNN提取空间/频域特征,LSTM建模时间依赖,能够同时捕捉心音信号的时频特征和心动周期性;在处理复杂任务如多类别心脏病识别时,效果突出	参数量大;模型复杂度较高,需要更多的计算资源和内存;训练和推理时间较长,实时性受限
CNN-RNN	结合了CNN和RNN的优点,适用于处理复杂的心音信号;表现稳定,适用于多分类任务	模型复杂度较高,训练时间较长;超参数多,调优成本较高;心音数据集不平衡,直接训练容易过拟合
Transformer	在处理长序列数据时表现出色,适用于心音信号时间序列数据;并行计算效率高,适合大规模数据训练;在跨模态任务中具有较大潜力。	计算复杂度 $O(n^2)$,内存消耗大;模型复杂度极高,需要大量的计算资源和内存,对小样本学习不友好

型,提升鲁棒性和性能;将心音分类技术嵌入穿戴设备^[88-89],进行实时监测和分类,帮助医生更加准确地评估疾病进展,提高治疗效果。

目前,深度学习技术在心音分类领域已取得显著进展,凭借自动化、大数据适配、高效响应、灵活扩展与强鲁棒性等优势,正在重塑心音分类的技术范式,为CVD的早期筛查、远程监测与精准诊断提供了强有力的技术支撑,推动心音分析从经验依赖的传统模式向数据驱动的智能模式跨越。深度学习模型通过多层次非线性变换自动挖掘从低级到高级的抽象特征,实现端到端的特征提取与分类,不仅降低对专业领域知识的依赖,还能捕捉到人工难以察觉的细微模式,从而减少漏诊误诊风险,提高诊断效率和准确性。由于实际采集的心音信号常受环境噪声、个体差异、设备差异等多种干扰,深度学习通过特殊设计增强抗干扰能力,通过数据增强模拟真实场景中的信号变异;采用注意力机制忽略非关键时段的干扰信息。这些设计使模型在信噪比低、信号质量差的情况下仍能保持稳定性能,尤其适用于基层医疗或家庭监测等非理想采集环境,为广泛推广智能心音诊断技术奠定了基础。其次,深度学习的

模块化特性使其能够根据不同临床任务灵活调整架构:针对二分类任务可采用轻量级CNN模型实现快速部署;面对多类别分类可通过增加网络深度或引入迁移学习提升分类粒度;对于跨场景适配则可通过领域自适应技术减少数据分布差异带来的影响。

虽然深度学习在心肺音分类方面具有许多优点,但仍存在一些挑战:部分算法在特定数据集上表现出色,但仍存在泛化能力不足、对噪声敏感等问题;目前仍需要大量数据来训练网络,但数据集不足,数据收集过程受设备差异、环境噪声和操作者经验影响不同医疗机构的数据格式和标准不统一,难以实现大规模数据共享;处理不同患者、呼吸状态和噪声环境下的信号比较困难;目前对心音信号的分类多为二分类任务(正常和异常),很少有研究将几种心音一起区分;深度学习模型通常具有大量参数,训练依赖于高性能计算资源,且模型可解释性较差,不易直观呈现分类依据,这在临床应用中可能影响医生的信任与接受程度。

目前深圳心音数据库已统一了采样标准,未来可在此基础上增加数据库样本类型,明确听诊位置对分类效果的影响^[90],进一步加强研究,开发更加智

能化的心音分类算法。未来的研究可以从以下几个方面着手改进:(1)优化模型结构,探索更加复杂的网络结构。除ResNet、胶囊网络(Capsule Networks)、DenseNet外,Transformer模型在语音和时序信号分析中表现突出,如Conformer(Convolution-augmented Transformer)和ViT(Vision Transformer)等能够在捕捉心音信号的时序依赖关系与全局上下文方面展现出优越性^[91]。Mamba模型通过引入选择性状态空间机制,能够在保持与Transformer相当建模能力的同时显著提升推理效率和内存利用率,也特别适合心音信号这类长时序、非平稳的生理数据,其改进版本Vision Mamba将该思路扩展至频谱图等二维表示,能够在心音时频图建模中替代传统CNN或ViT,从而更高效地提取全局与局部特征。还可考虑将自注意力机制、多头注意力引入心音分类任务,从而在提升特征提取能力的同时兼顾计算效率和模型解释性。(2)结合新的信号处理技术或机器学习算法,如时频分析与深度学习结合、自监督学习与迁移学习等,充分挖掘心音信号的时-频-病理多维特征关联。通过混合不同的深度学习模型和模块,能够提升整体性能。(3)针对心音分类任务,可设计专用损失函数,如基于类别不平衡的加权交叉熵损失,以提升模型在不平衡数据集上的分类性能。(4)采用先进的优化算法,如Muon、LAMB、QHAdam、Swats和RAdam等,配合动态学习率自适应调度策略,以提高训练效率和模型收敛速度。此外,MambaOut等正则化技术提供了更好的泛化能力,有助于解决心音数据普遍存在的小样本、噪声和跨中心差异问题。(5)与其他医疗数据融合,构建多模态分类系统,进一步提升诊断的全面性,如Calzoni等^[92]提出一种融合ECG与PCG信号的双模态深度学习模型。(6)继续探索轻量化模型,以便能够高效部署在资源有限的硬件设备上,满足实时性和低功耗的需求。除早期的MobileNet、SqueezeNet、MnasNet和EfficientNet等模型外,知识蒸馏模型^[93]、LSNet^[94]、MobileViT^[95]、MobileMamba^[96]、Mobile-Former^[97]、MobileNet和Tiny-Transformer^[98]等模型为在边缘设备上部署心音分类提供了新的可能性。对于智能设备、穿戴式设备和便携式医疗设备上,采用这些轻量化模型有助于实现高效的心音分类,推动远程健康监测和心脏疾病的早期诊断。(7)对于数据集的类别不平衡问题,除常规的过采样方法(如SMOTE),GAN也是一种非常有效的解决方案,Chakraborty等^[99]使用Progressive Wasserstein生成对抗网络合成冠心病心脏声音片段,以解决数据集的类别不平衡问题。

综上所述,基于深度学习的心音分类技术在未来越有着广阔的发展前景,但模型的可解释性和特征选择仍面临挑战,数据的有限性和质量问题也限制了研究进展,未来可从多模态信号融合、标准化验证、自动化决策支持以及实时监测和纵向数据分析等方向展开研究,开展多分类研究、优化模型结构并部署到边缘设备,聚焦于算法创新、临床应用拓展以及技术挑战的应对。通过跨学科合作与技术融合,不断推动心音分类技术的进步,将为心脏疾病的早期发现与精准治疗提供强有力的支持,更好地应用于实际心脏病筛查工作。

【参考文献】

- [1] 孙文慧,陈扶明,张乙鹏,等.心肺音分离方法研究进展[J]. 中国医疗设备, 2024, 39(3): 154-159.
Sun WH, Chen FM, Zhang YP, et al. Research progress of heart-lung sound separation methods[J]. China Medical Devices, 2024, 39(3): 154-159.
- [2] 李金萌. 基于深度学习的心音信号分类方法研究[D]. 济南: 齐鲁工业大学, 2024.
Li JM. Research on heart sound signal classification method based on deep learning[D]. Jinan: Qilu University of Technology, 2024.
- [3] 吴延军,徐泾平,赵艳. 心音的产生与传导机制[J]. 生物医学工程学杂志, 1996(3): 280-288.
Wu YJ, Xu JP, Zhao Y. Mechanism of generation and transmission of heart sounds[J]. Journal of Biomedical Engineering, 1996(3): 280-288.
- [4] Gomes EF, Bentley PJ, Pereira E, et al. Classifying heart sounds-approaches to the Pascal challenge [C]//Proceedings of the International Conference on Health Informatics-Volume OBIOSTEC, Barcelona, Spain, 2013: 337-340
- [5] Yaseen, Son GY, Kwon S. Classification of heart sound signal using multiple features[J]. Appl Sci, 2018, 8(12): 2344.
- [6] Liu CY, Springer D, Li Q, et al. An open access database for the evaluation of heart sound algorithms[J]. Physiol Meas, 2016, 37(12): 2181-2213.
- [7] Oliveira J, Renna F, Costa PD, et al. The CirCor DigiScope dataset: from murmur detection to murmur classification[J]. IEEE J Biomed Health Inform, 2022, 26(6): 2524-2535.
- [8] Dong FQ, Qian K, Ren Z, et al. Machine listening for heart status monitoring: introducing and benchmarking HSS-the heart sounds Shenzhen corpus[J]. IEEE J Biomed Health Inform, 2020, 24(7): 2082-2092.
- [9] Centracchio J, Parlato S, Esposito D, et al. Accurate localization of first and second heart sounds via template matching in forcecardiography signals[J]. Sensors, 2024, 24(5): 1525.
- [10] Shervagar VM. A robust sliding window adaptive filtering technique for phonocardiogram signal denoising[J]. Expert Syst, 2025, 42(1): e13361.
- [11] Liu ZC, Shao TP, Zhang X. BCG signal analysis based on improved VMD algorithm[J]. Measurement, 2024, 231: 114631.
- [12] Chen YL, Wu DP, Wei KY, et al. Joint VMD and PSSVM optimization algorithm for seizure prediction based on ECG signals[J]. Biomed Signal Process Control, 2025, 103: 107478.
- [13] Fang Y, Leng HX, Wang WB, et al. Multi-level feature encoding algorithm based on FBPSI for heart sound classification[J]. Sci Rep, 2024, 14(1): 29132.
- [14] Lin MF, Hong YZ, Hong SC, et al. Discrete wavelet transform based ECG classification using gcForest: a deep ensemble method[J]. Technol Health Care, 2024, 32(S1): 95-105.
- [15] Huang NE, Shen Z, Long SR, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proc R Soc A, 1998, 454(1971): 903-995.
- [16] Ali SN, Shuvo SB, Al-Manzo MIS, et al. An end-to-end deep learning framework for real-time denoising of heart sounds for cardiac disease

- detection in unseen noise[J]. IEEE Access, 2023, 11: 87887-87901.
- [17] Liang MF, Hu JQ, Zhou XB, et al. A GAN based heart sound denoising model [C]//2022 12th International Conference on Information Technology in Medicine and Education (ITME). Piscataway, NJ, USA: IEEE, 2022: 666-670.
 - [18] Al-Zaben A, Al-Fahoum A, Ababneh M, et al. Improved recovery of cardiac auscultation sounds using modified cosine transform and LSTM-based masking[J]. Med Biol Eng Comput, 2024, 62(8): 2485-2497.
 - [19] Dissanayake T, Fernando T, Denman S, et al. A robust interpretable deep learning classifier for heart anomaly detection without segmentation[J]. IEEE J Biomed Health Inform, 2021, 25(6): 2162-2171.
 - [20] Arjoune Y, Nguyen TN, Doroshow RW, et al. A noise-robust heart sound segmentation algorithm based on Shannon energy[J]. IEEE Access, 2024, 12: 7747-7761.
 - [21] Xiao PZ, Wang KP. Segmentation of heart sound signals using improved Hilbert transform and wavelet packet transform[J]. Circuits Syst Signal Process, 2025, 44(7): 4752-4773.
 - [22] Li LY, Huang M, Dao L, et al. Construction and validation of a method for automated time label segmentation of heart sounds[J]. Front Artif Intell, 2024, 6: 1309750.
 - [23] Liu QS, Wu XM, Ma XJ. An automatic segmentation method for heart sounds[J]. Biomed Eng Online, 2018, 17(1): 106.
 - [24] Xu XC, Geng XG, Gao ZX, et al. Optimal heart sound segmentation algorithm based on K-mean clustering and wavelet transform[J]. Appl Sci, 2023, 13(2): 1170.
 - [25] Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition [J]. Proc IEEE, 1989, 77(2): 257-286.
 - [26] Martins ML, Coimbra MT, Renna F. Markov-based neural networks for heart sound segmentation: using domain knowledge in a principled way[J]. IEEE J Biomed Health Inform, 2023, 27(11): 5357-5368.
 - [27] Sadeghi Nia P, Danandeh Hesar H. Abnormal heart sound detection using time-frequency analysis and machine learning techniques[J]. Biomed Signal Process Control, 2024, 90: 105899.
 - [28] Enériz D, Rodriguez-Almeida AJ, Fabelo H, et al. Low-cost FPGA implementation of deep learning-based heart sound segmentation for real-time CVDs screening[J]. IEEE Trans Instrum Meas, 2024, 73: 1-16.
 - [29] Park C, Shin K, Seo J, et al. Enhancement of phonocardiogram segmentation using convolutional neural networks with Fourier transform module[J]. Biomed Eng Lett, 2025, 15(2): 401-413.
 - [30] McDonald A, Gales MJF, Agarwal A. A recurrent neural network and parallel hidden Markov model algorithm to segment and detect heart murmurs in phonocardiograms[J]. PLOS Digit Health, 2024, 3(11): e0000436.
 - [31] Chen Y, Sun YN, Lü JC, et al. End-to-end heart sound segmentation using deep convolutional recurrent network[J]. Complex Intell Syst, 2021, 7(4): 2103-2117.
 - [32] Guan TX, Xu DY, Cai SS, et al. A deep-learning-based cardiac sound segmentation method for smart auscultation applications[C]//2023 4th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). Piscataway, NJ, USA: IEEE, 2023: 171-175.
 - [33] 卢官明, 蔡亚宁, 卢峻禾, 等. 基于Bi-LSTM和时序注意力的异常心音检测[J]. 南京邮电大学学报(自然科学版), 2025, 45(1): 12-20. Lu GM, Lu JH, Cai YN, et al. Abnormal heart sound detection using Bi-LSTM and temporal attention mechanism[J]. Journal of Nanjing University of Posts and Telecommunications(Natural Science Edition), 2025, 45(1): 12-20.
 - [34] Deng EQ, Jia YB, Zhu GB, et al. Heart sound signals classification with image conversion employed[J]. Electronics, 2024, 13(7): 1179.
 - [35] Jiang ZH, Song WH, Yan YH, et al. Automated valvular heart disease detection using heart sound with a deep learning algorithm[J]. Int J Cardiol Heart Vasc, 2024, 51: 101368.
 - [36] Abbas S, Ojo S, Al Hejaili A, et al. Artificial intelligence framework for heart disease classification from audio signals[J]. Sci Rep, 2024, 14(1): 3123.
 - [37] Elola A, Aramendi E, Oliveira J, et al. Beyond heart murmur detection: automatic murmur grading from phonocardiogram[J]. IEEE J Biomed Health Inform, 2023, 27(8): 3856-3866.
 - [38] Ranipa K, Zhu WP, Swamy MNS. A novel feature-level fusion scheme with multimodal attention CNN for heart sound classification[J]. Comput Methods Programs Biomed, 2024, 248: 108122.
 - [39] Kumar SS, Ramachandran P. Nested compressive sensing for ECG signals using novel hybrid two-level approach of DWT and DCT sensing and sparsification[J]. Discover Appl Sci, 2024, 7(1): 18.
 - [40] Chen X, Guo XM, Zheng YN, et al. Heart function grading evaluation based on heart sounds and convolutional neural networks[J]. Phys Eng Sci Med, 2023, 46(1): 279-288.
 - [41] Katal N, Garg H, Sharma B. A comparative analysis of CNNs and LSTMs for ECG-based diagnosis of arrhythmia and congestive heart failure[J]. Comput Methods Biomech Biomed Engin, 2026, 29(7): 1608-1636.
 - [42] Telceken M. A new feature extraction method for AI based classification of heart sounds: dual-frequency cepstral coefficients (DFCCs)[J]. Eur Phys J Spec Top, 2025, 234(15): 4699-4710.
 - [43] Patwa A, Mahboob Ur Rahman M, Al-Naffouri TY. Heart murmur and abnormal PCG detection via wavelet scattering transform and 1D-CNN [J]. IEEE Sens J, 2025, 25(7): 12430-12443.
 - [44] Yang WR, Xu JJ, Xiang JH, et al. Diagnosis of cardiac abnormalities based on phonocardiogram using a novel fuzzy matching feature extraction method[J]. BMC Med Inform Decis Mak, 2022, 22(1): 230.
 - [45] Zheng YN, Guo XM, Yang Y, et al. Phonocardiogram transfer learning-based CatBoost model for diastolic dysfunction identification using multiple domain-specific deep feature fusion[J]. Comput Biol Med, 2023, 156: 106707.
 - [46] Bahreini M, Barati R, Kamali A. Cardiac sound classification using a hybrid approach: MFCC-based feature fusion and CNN deep features [J]. EURASIP J Adv Signal Process, 2025, 2025(1): 2.
 - [47] Li F, Liu M, Zhao YJ, et al. Feature extraction and classification of heart sound using 1D convolutional neural networks[J]. EURASIP J Adv Signal Process, 2019, 2019(1): 59.
 - [48] Khan SUR, Khan Z. Detection of abnormal cardiac rhythms using feature fusion technique with heart sound spectrograms[J]. J Bionic Eng, 2025, 22(4): 2030-2049.
 - [49] Ren Z, Cummins N, Pandit V, et al. Learning image-based representations for heart sound classification[C]//Proceedings of the 2018 International Conference on Digital Health. New York, NY, USA: ACM, 2018: 143-147.
 - [50] Chowdhury TH, Poudel KN, Hu YT. Time-frequency analysis, denoising, compression, segmentation, and classification of PCG signals[J]. IEEE Access, 2020, 8: 160882-160890.
 - [51] Riccio D, Brancati N, Sannino G, et al. CNN-based classification of phonocardiograms using fractal techniques[J]. Biomed Signal Process Control, 2023, 86: 105186.
 - [52] Xiao F, Liu HQ, Lu J. A new approach based on a 1D + 2D convolutional neural network and evolving fuzzy system for the diagnosis of cardiovascular disease from heart sound signals[J]. Appl Acoust, 2024, 216: 109723.
 - [53] Xiong WH, Zhang G, Yan DM, et al. Multichannel feature fusion network-based technique for heart sound signal classification and recognition[J]. Expert Syst Appl, 2025, 273: 126839.
 - [54] Mahmood A, Dhahri H, Alhajlah M, et al. Enhanced classification of phonocardiograms using modified deep learning[J]. IEEE Access, 2024, 12: 178909-178916.
 - [55] Li F, Tang H, Shang S, et al. Classification of heart sounds using convolutional neural network[J]. Appl Sci, 2020, 10(11): 3956.
 - [56] Bao XQ, Xu YJ, Lam HK, et al. Time-frequency distributions of heart sound signals: a comparative study using convolutional neural networks[J]. Biomed Eng Adv, 2023, 5: 100093.
 - [57] Qiao LH, Li ZX, Xiao B, et al. QDRJL: quaternion dynamic representation with joint learning neural network for heart sound signal abnormal detection[J]. Neurocomputing, 2023, 562: 126889.
 - [58] Al-Shannaq MA, Nasrawi A, Bsoul AARK, et al. Abnormal heart sound recognition using SVM and LSTM models in real-time mode [J]. Sci Rep, 2025, 15(1): 9129.
 - [59] Latifi SA, Ghassemian H, Imani M. Multi-branch convolutional network and LSTM-CNN for heart sound classification[J]. Phys Eng Sci Med, 2026, 49(1): 263-287.
 - [60] 王慧. 基于最小门控单元的心音分类研究[J]. 智能计算机与应用, 2024, 14(3): 81-85. Wang H. Classification of heart sounds based on minimal gated unit

- [J]. Intelligent Computer and Applications, 2024, 14(3): 81-85.
- [61] Roy TS, Roy JK, Mandal N. Design of ear-contactless stethoscope and improvement in the performance of deep learning based on CNN to classify the heart sound[J]. Med Biol Eng Comput, 2023, 61(9): 2417-2439.
- [62] Rizal A, Fauzi H, Wulandari DR, et al. Classification of normal and abnormal heart sounds using Hjorth features and LSTM-RNN algorithm[J]. J Phys Conf Ser, 2023, 2673(1): 012003.
- [63] Malik H, Bashir U, Ahmad A. Multi-classification neural network model for detection of abnormal heartbeat audio signals[J]. Biomed Eng Adv, 2022, 4: 100048.
- [64] Zhang HB, Zhang P, Wang ZW, et al. Multi-feature decision fusion network for heart sound abnormality detection and classification[J]. IEEE J Biomed Health Inform, 2024, 28(3): 1386-1397.
- [65] Choudhary RR, Singh MR, Jain PK. Heart sound classification using a hybrid of CNN and GRU deep learning models[J]. Procedia Comput Sci, 2024, 235: 3085-3093.
- [66] Choudhary RR, Rani M, Kaur R, et al. Heart signal analysis using multistage classification denoising model[J]. J Electr Comput Eng, 2024, 2024(1): 1502285.
- [67] Li SZ, Sun J, Yang HB, et al. Interpretable end-to-end heart sound classification[J]. Measurement, 2024, 237: 115113.
- [68] Ma KY, Lu JB, Lu BZ. Parameter-efficient densely connected dual attention network for phonocardiogram classification[J]. IEEE J Biomed Health Inform, 2023, 27(9): 4240-4249.
- [69] Lin JY, Wang SK, Xuan WP, et al. Portable ECG and PCG wireless acquisition system and multiscale CNN feature fusion Bi-LSTM network for coronary artery disease diagnosis[J]. Comput Biol Med, 2025, 191: 110202.
- [70] Lin TY, Wang YX, Liu XY, et al. A survey of transformers[J]. AI Open, 2022, 3: 111-132.
- [71] Cheng JW, Sun KX. Heart sound classification network based on convolution and transformer[J]. Sensors (Basel), 2023, 23(19): 8168.
- [72] Liu ZH, Liu W, Gu Z, et al. Efficient anomaly detection algorithm for heart sound signal[J]. IEEE Access, 2024, 12: 139225-139236.
- [73] Han J, Shaout A. ENACT-heart-ENSEMBLE-based assessment using CNN and transformer on heart sounds[EB/OL]. (2025-02-24). <https://arxiv.org/abs/2502.16914>.
- [74] Wang RS, Duan YF, Li YK, et al. PCTMF-Net: heart sound classification with parallel CNNs-transformer and second-order spectral analysis[J]. Vis Comput, 2023, 39(8): 3811-3822.
- [75] Liu ZY, Jiang H, Zhang FW, et al. Heart sound classification based on bispectrum features and vision transformer mode[J]. Alexandria Eng J, 2023, 85: 49-59.
- [76] Wang SY, Hu JH, Du YW, et al. WCFormer: an interpretable deep learning framework for heart sound signal analysis and automated diagnosis of cardiovascular diseases[J]. Expert Syst Appl, 2025, 276: 127238.
- [77] Ren Z, Chang Y, Nguyen TT, et al. A comprehensive survey on heart sound analysis in the deep learning era[J]. IEEE Comput Intell Mag, 2024, 19(3): 42-57.
- [78] Barnawi A, Boulares M, Somai R. Simple and powerful PCG classification method based on selection and transfer learning for precision medicine application[J]. Bioengineering (Basel), 2023, 10(3): 294.
- [79] Almanifi ORA, Ab Nasir AF, Mohd Razman MA, et al. Heartbeat murmurs detection in phonocardiogram recordings via transfer learning[J]. Alexandria Eng J, 2022, 61(12): 10995-11002.
- [80] Ding SJ, Ding H, Kan MF, et al. A computer-aided heart valve disease diagnosis system based on machine learning[J]. J Healthc Eng, 2023, 2023(1): 7382316.
- [81] Harimi A, Ameri MA, Sarkar S, et al. Heart sounds classification: application of a new CyTex inspired method and deep convolutional neural network with transfer learning[J]. Smart Health, 2023, 29: 100416.
- [82] Jumphoo T, Phapatanaburi K, Pathonsuwan W, et al. Exploiting data-efficient image transformer-based transfer learning for valvular heart diseases detection[J]. IEEE Access, 2024, 12: 15845-15855.
- [83] Maity A, Pathak A, Saha G. Transfer learning based heart valve disease classification from phonocardiogram signal[J]. Biomed Signal Process Control, 2023, 85: 104805.
- [84] Wang ZH, Qian K, Liu HG, et al. Exploring interpretable representations for heart sound abnormality detection[J]. Biomed Signal Process Control, 2023, 82: 104569.
- [85] 王林琳. 基于卷积循环神经网络的肺音识别研究[D]. 武汉: 华中师范大学, 2023.
- Wang LL. Research on lung sound recognition based on convolutional recurrent neural network[D]. Wuhan: Central China Normal University, 2023.
- [86] 张乙鹏. 基于MFCC和深度学习的肺音分类方法研究[D]. 兰州: 甘肃中医药大学, 2024.
- Zhang YP. Research on lung sound classification method based on MFCC and deep learning[D]. Lanzhou: Gansu University of Chinese Medicine, 2024.
- [87] 陈茜茹, 钟清华, 张涵, 等. 基于DSP的心肺音分离系统研究[J]. 现代电子技术, 2019, 42(6): 92-95.
- Chen QR, Zhong QH, Zhang H, et al. Research on heart and lung sound separation system based on DSP[J]. Modern Electronics Technique, 2019, 42(6): 92-95.
- [88] Zang JB, An Q, Li B, et al. A novel wearable device integrating ECG and PCG for cardiac health monitoring[J]. Microsyst Nanoeng, 2025, 11(1): 7.
- [89] De Fazio R, Spongano L, De Vittorio M, et al. Machine learning algorithms for processing and classifying unsegmented phonocardiographic signals: an efficient edge computing solution suitable for wearable devices[J]. Sensors (Basel), 2024, 24(12): 3853.
- [90] 丁思吉, 丁皓, 阙孟菲, 等. 心音信号分类识别算法综述[J]. 软件导刊, 2022, 21(12): 272-278.
- Ding SJ, Ding H, Kan MF, et al. A review of classification methods of heart sound signal[J]. Software Guide, 2022, 21(12): 272-278.
- [91] Gulati A, Qin J, Chiu CC, et al. Conformer: convolution-augmented transformer for speech recognition[C]//Interspeech 2020. Baixas, France: ISCA, 2020: 5036-5040.
- [92] Calzoni A, Savardi M, Silvestri M, et al. Bimodal ECG and PCG cardiovascular disease detection: exploring the potential and modality contribution[J]. J Med Syst, 2025, 49(1): 113.
- [93] Li JQ, Wang C, Chen J, et al. Explainable CNN with fuzzy tree regularization for respiratory sound analysis[J]. IEEE Trans Fuzzy Syst, 2022, 30(6): 1516-1528.
- [94] Wang A, Chen H, Lin ZJ, et al. LSNet: see large, focus small[C]//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2025: 9718-9729.
- [95] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. (2021-06-03). <https://arxiv.org/abs/2010.11929>.
- [96] He HY, Zhang JN, Cai YX, et al. MobileMamba: lightweight multi-receptive visual mamba network[C]//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2025: 4497-4507.
- [97] Chen YP, Dai XY, Chen DD, et al. Mobile-former: bridging MobileNet and transformer[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 5260-5269.
- [98] Wu K, Zhang JN, Peng HW, et al. TinyViT: fast pretraining distillation for small vision transformers[C]//Computer Vision-ECCV 2022. Cham: Springer Nature Switzerland, 2022: 68-85.
- [99] Chakraborty S, Kochhar P, Patil S, et al. Generative adversarial network augmented data for improved heart sound abnormality detection[J]. Comput Biol Med, 2025, 195: 110623.

(编辑:薛泽玲)