

人工智能在心音信号检测领域的研究综述

肖裴垚¹, 李轶凯¹, 房子洋¹, 王津梁¹, 张治^{1,2}

(1. 上海理工大学健康科学与工程学院, 上海 200093;
2. 上海交通大学医学院附属第一人民医院心内科, 上海 200080)

摘要: 心音信号分析是心脏智能听诊领域和相关诊断的重要技术之一。近年来, 传统机器学习和深度学习方法在心音信号的预处理、特征提取与分类中发挥了重要作用, 推动了该领域的快速发展。系统总结了现有研究中采用的信号处理方法、机器学习与深度学习技术, 并梳理了常用公开心音数据库的特点及其在研究中的应用价值。最后, 探讨了心音分析技术在数据库建设、模型轻量化及多模态融合分析等方面的未来发展方向, 旨在为相关研究提供参考。

关键词: 心音信号; 深度学习; 机器学习; 心音数据库; 迁移学习

DOI: 10.11907/tjdk.251008

开放科学(资源服务)标识码(OSID):

中图分类号: TP301.6

文献标识码: A

文章编号: 1672-7800(2025)009-0001-08



Review of Research on Artificial Intelligence in the Field of Heart Sound Signal Detection

XIAO Peiyao¹, LI Yikai¹, FANG Ziyang¹, WANG Jinliang¹, ZHANG Zhi^{1,2}

(1. School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. Department of Cardiology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200080, China)

Abstract: Phonocardiogram signal analysis is a crucial technology in intelligent cardiac auscultation and related diagnostic fields. In recent years, traditional machine learning and deep learning methods have played a significant role in the preprocessing, feature extraction, and classification of phonocardiogram signals, driving rapid advancements in this domain. This paper systematically reviews the signal processing methods, machine learning, and deep learning techniques employed in existing research, while summarizing the characteristics of commonly used public phonocardiogram databases and their application value in studies. Finally, it discusses future directions for phonocardiogram signal analysis, including database development, model lightweighting, and multimodal fusion analysis, aiming to provide a reference for related research endeavors.

Key Words: heart sound signals; deep learning; machine learning; heart sound databases; transfer learning

0 引言

心血管疾病(CVD)是全球最主要的死亡原因之一。世界心脏联盟发布的《2023年世界心脏报告》显示, 2021年全球有2 050万人死于心血管疾病, 占全球死亡总数的

近1/3^[1]。其中, 缺血性心脏病和中风占心血管疾病死亡总数的85%。在我国, 心血管疾病的发病率和死亡率也呈持续上升趋势, 已成为严重威胁人民健康的主要公共卫生问题^[2]。

近年来, 电子听诊器在心血管疾病诊断中表现出显著优势。相较于传统听诊器, 电子设备具备信号放大、降噪、

收稿日期: 2025-02-09

扫描二维码阅读全文:

基金项目: 国家自然科学基金项目(82270536)

作者简介: 肖裴垚(2001-), 男, 上海理工大学健康科学与工程学院硕士研究生, 研究方向为心音智能听诊; 李轶凯(2001-), 男, 上海理工大学健康科学与工程学院硕士研究生, 研究方向为人工智能在肺音检测中的应用; 房子洋(2001-), 男, 上海理工大学健康科学与工程学院硕士研究生, 研究方向为基于人工智能的肾上腺分割; 王津梁(1998-), 男, 上海理工大学健康科学与工程学院硕士研究生, 研究方向为心音智能听诊; 张治(1978-), 男, 博士, 上海交通大学医学院附属第一人民医院心内科主任医师、硕士生导师, 研究方向为无线听诊器和人工智能诊断、难治性高血压的器械和药物治疗。



数据存储及远程传输等多项功能,使其不仅在临床环境中得到了广泛应用,还在远程医疗领域发挥了关键作用。例如, Fan等^[3]提出的远程听诊系统能够通过蓝牙技术与智能设备连接,实现心音信号的实时采集与传输。在COVID-19疫情期间,这类设备有效支持了医护人员在佩戴防护装备的情况下进行高效的心肺听诊,显著降低了交叉感染的风险^[4]。这些技术进步为心血管疾病的筛查与诊断提供了创新手段,尤其适用于资源匮乏或高风险的医疗环境。

与此同时,人工智能技术的迅猛发展极大地提升了心音信号分析的精确性与效率。深度学习模型具备自动提取心音核心特征的能力,能够精准划分正常与异常心音,并识别诸如心律失常和心脏杂音等微弱信号。将人工智能与电子听诊器相结合, AI驱动的分析方法不仅克服了传统依赖医生经验的局限性,还通过远程医疗技术有效缓解了偏远地区医疗资源不足的问题。便携式电子听诊器与智能分析模型的结合,既降低了患者的就诊成本,又显著提升了医疗服务的可及性与效率,能够为心血管疾病的早期筛查和及时干预提供有力支持。

综上,本文基于近年来人工智能技术在心音信号分析

领域的快速发展,系统总结了数字化心音技术与深度学习算法结合的研究进展。文章详细介绍了心音信号的预处理技术,包括降噪、信号分割和特征提取方法,分析了传统方法与基于深度学习的改进策略,特别关注了在数据量受限情况下的迁移学习方法。通过梳理相关技术的发展现状和挑战,本文对未来研究的重点方向和应用前景进行了展望,以期为推动心音智能分析的进一步发展提供参考。

1 公开心音数据库及采集设备

1.1 公开心音数据库

心音数据库为心音信号研究奠定了坚实基础,是算法开发、模型训练与技术验证中不可或缺的关键数据来源。近年来,随着人工智能与深度学习技术在心音信号分析领域的日益普及,公开发布的心音数据库在数量与质量上均有所提升。这些数据库蕴含丰富而多元的信号样本,涵盖不同类型的采集设备、多年龄段与多健康状态个体以及多样化的病理特征,为研究人员提供了高价值的资源支持。为便于后续处理与分类工作,表1对近年来广泛使用的典型公开心音数据库进行了概览性汇总,包括其基本信息、数据来源、采样频率等特征指标,以期为未来相关研究与应用提供参考与借鉴。

Table 1 Summary of commonly used publicly available databases

表1 常用公开数据库汇总

数据来源	病例分类	被试者	音频数目	采样频率/Hz	录制时长/s
2016	正常和异常	764	3 240	2 000	5~120
PhysioNet/CinC 心音挑战赛数据库 ^[5]	正常、杂音、额外心音、人为杂音	未知	176	4 000	1~30
PASCAL心音分类挑战赛数据库A ^[6]	正常、杂音、早搏	未知	656	4 000	1~30
PASCAL心音分类挑战赛数据库B	正常、主动脉瓣狭窄、二尖瓣狭窄、二尖瓣返流、二尖瓣脱垂	未知	1 000	8 000	2~5
Yasein等人收集的心音信号 ^[7]	正常、异常、未知	1568	5 282	4 000	5~168
CirCor DigiScope数据库 ^[8]	正常	206	412	11 025	20~70
HSCT-11数据库 ^[9]	正常	26	26	333	1 000~1 400
Fetal数据库 ^[10]	正常、轻度、中/重度	170	845	4 000	25~35
深圳心音数据库 ^[11]	正常	29	29	4 000	2~20
DigiScope2017数据库 ^[12]	正常	24	69	8 000	30
Kazemnejad等人收集的心音信号 ^[13]					

1.2 心音信号采集设备

心音信号采集设备是获取高质量数据的基础,对信号分析与应用具有重要影响。随着电子技术和医疗设备的不断进步,各类电子心音听诊器的性能得到了显著提升,支持高精度采集和多样化应用。表2展示了4款目前较为先进的电子心音听诊器。

2 心音信号预处理

心音信号在采集过程中往往受到环境噪声干扰,且不同设备之间存在采样频率不一致的问题,导致原始数据可能出现偏差与非均匀性。为确保后续分析的准确性与可靠性,信号预处理成为必不可少的步骤。该环节通常包括噪声抑制、信号标准化与有效信息提取等过程,如图1所示。

Table 2 Summary of the features of four electronic heart sound stethoscopes

表2 四款电子心音听诊器功能总结

设备名称	上市时间	功能
Eko CORE 500 数字听诊器	2024	具备全彩显示屏、蓝牙、主动降噪、可视化心音图和3导联心电图捕获功能,音频放大可达40倍,采用Eko软件进行心音录制
Stemscope PRO 数字听诊器	2023	具备蓝牙、主动降噪和可视化心音图功能,音频放大可高达100倍,设备体积小,易于使用,采用DrStemo软件进行心音录制
Popular-3 电子听诊器	2023	具备蓝牙功能,音频放大可达24倍,可使用配套软件进行心音录制
3M Littmann CORE 数字听诊器	2020	具备主动降噪和可视化心音图功能,音频放大可达40倍,采用Eko软件进行心音录制

其主要目标在于提升信号品质与特征的可辨识度,从而为后续任务分类与分析提供稳定且高质量的数据基础。

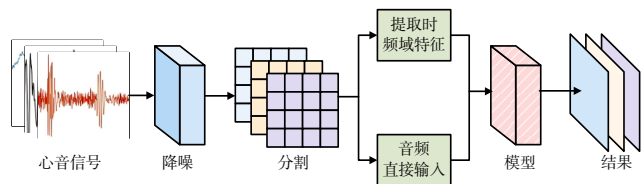


Fig. 1 Heart sound signal processing procedure

图1 心音信号处理流程

2.1 心音信号降噪

2.1.1 传统降噪方法

心音信号常受到外部环境与记录设备的多重噪声干扰,为了提高信号质量并确保后续分析的准确性,有效的降噪策略至关重要。传统的降噪方法涵盖频域滤波、时域滤波以及小波变换等技术手段,通过选择适宜的滤波器与算法,可在有效保留心音信号关键特征的同时显著抑制噪声成分^[14]。常见的去噪方法包括离散小波变换、傅里叶变换和巴特沃斯滤波器等。在此过程中,Alkhodari等^[15]指出心音信号具有明显的非平稳性,其频率成分随时间变化,导致传统的傅里叶变换在提取此类时变频率特征时存在局限性。相比之下,离散小波变换借助其时频局部化特性,既能有效抑制瞬时噪声,又能全面保留信号的主要特征。此外,Singh等^[16]采用了截止频率为25~400 Hz的四阶巴特沃斯低通滤波器,该滤波器因具有平稳的频率响应而广泛应用,可有效避免频率波动与振铃效应,从而减小信号失真。尽管其过渡带衰减相对较为缓慢,但平滑的频率响应有助于最大程度地减少信号干扰。与其他滤波器相比,巴特沃斯滤波器在通带内提供更加均匀的响应,保障了信号的高保真度。同时,其较低的计算复杂度使其更适合实时应用,尤其是在移动设备端的实时信号处理中表现出色。

2.1.2 结合深度学习改进的降噪方法

在深度学习的快速发展趋势下,越来越多的研究者尝试将其引入心音信号降噪领域,以克服传统方法在非平稳、多样化噪声环境中遇到的局限性^[17]。与Alkhodari等^[15]、Singh等^[16]使用小波变换与巴特沃斯滤波器进行信号降噪所取得的有限提升相比,Sharan等^[18]利用DnCNN模型对比传统小波分解与巴特沃斯滤波器时,不仅在信噪比(SNR)、峰值信噪比(PSNR)、均方误差(MSE)和结构相似度(SSIM)等定量指标上取得了显著改善,还展现出更好的特征保留能力。进一步地,Ali等^[19]提出的U-Net结构通过多尺度特征提取与特征融合,有效应对非平稳与非线性噪声挑战,确保在多类噪声环境下仍能稳定去噪,并尽可能减少S1与S2等关键生理特征的损失。同时,Chandel等^[20]利用堆叠式双向长短期记忆网络(Stacked Bi-LSTM)与双信号变换(Dual Signal Transformation)方法,从时域和频域双重视角捕捉心音特征,提升对非平稳、多源噪声的适应性与特征提取能力,使模型在复杂声学条件下仍能获得高保真度的降噪结果。与此相呼应,Liang等^[21]采用生

成对抗网络(Generative Adversarial Network, GAN)将对抗训练机制引入心音降噪中,使模型可不断逼近纯净心音的真实分布,即使在复杂噪声条件下仍能保持较高保真度和稳定性。总体来看,与传统滤波手段相比,这些深度学习模型在不同噪声场景下均展现出更优的鲁棒性与适应性,不仅为特征提取、分类和临床诊断提供了高质量的输入信号,也为心音分析在移动端和实时应用中的实际部署提供了有力支持。

2.2 心音信号分割

2.2.1 传统分割方法

传统的心音信号分割方法依赖于经典的信号处理技术,通过分析心音信号的时间域和频率域特性,准确识别关键的心音成分。这些方法主要包括小波变换^[22]、分形分解^[23](Fractal Decomposition, FD)和香农能量包络^[24]等技术手段。Golpaygani等^[25]提出了一种基于离散小波变换(DWT)的心音分割算法,该算法通过多层分解和小波分量的重构提取心音信号中的关键频率特征,特别是第9层小波细节分量(d9)。结合香农能量包络,该方法增强了信号峰值的识别能力,并通过时间间隔分析和生理验证区分S1和S2心音。在实验中,该算法在93%的病例中准确识别了S1和S2心音,灵敏度为89.58%,精确度达到93.14%。进一步地,Varghees等^[26]提出了一种基于经验小波变换(EWT)和瞬时相位分析的改进方法。该方法采用自适应频率分解策略,有效分离了心音信号与心脏杂音,同时结合香农熵包络进一步提高了信号峰值的放大能力,并利用瞬时相位的零交点自动检测心音与杂音的边界。不同于Golpaygani等^[25]的方法,Varghees等^[26]的研究测试了多个公开的复杂数据集(如PhysioNet和PASCAL Challenge),这些数据集包含了高噪声环境和多种心音模式的信号。在这些更复杂的条件下,该方法在心音信号分割中取得了91.92%的总体准确率和94.38%的灵敏度,展现出在复杂噪声条件下的卓越鲁棒性。相较于传统DWT方法,其泛用性和适应性均有所提升。

此外,Thomas等^[27]通过分形分解方法结合多重分形去趋势波动分析(Multifractal Detrended Fluctuation Analysis, MFDFA),提出了一种创新的心音分割算法。该方法利用心音信号的多重分形特性,通过迭代分解提取S1和S2的关键成分,结合希尔伯特变换包络和瞬时相位分析精确标记S1-S2的起止点。在实验中,该方法在两种数据库中的平均分割准确率达到94.82%,灵敏度为98.2%,并在添加5dB白噪声的环境下保持了较高鲁棒性。与小波变换的分割方法相比,分形分解技术在处理非平稳信号和复杂心音模式时表现出更优的分割性能。此外,Kang等^[28]针对儿童心音的复杂性,提出了一种基于平均香农能量的分割方法,以提高算法鲁棒性。该方法通过计算平均香农能量包络提取信号的关键特征,结合包络信号的时间门限约束,有效解决了儿童心音信号中高心率、节律变化

和杂音干扰带来的分割难题。具体实现上,利用滑动窗口对心音能量信号进行平滑处理,并通过音节长度约束和相邻音节间隔的自适应调整,显著提升了S1和S2的检测精度。实验结果表明,该方法在心音信号分割中达到了96.7%的灵敏度和98.0%的阳性预测值。尽管在心音分类过程中,许多研究已逐渐停止采用传统的分割方法,但這些方法仍能有效地将心音信号分割为不同的片段,从中提取关键信息,以支持进一步的处理步骤,如心音分类和疾病识别。这些分割技术在一定程度上提升了心音分类准确率和疾病诊断精度。

2.2.2 基于深度学习分割方法

近年来,深度学习在心音分割领域逐步展现出其独特的优势,能够更有效地处理复杂信号和噪声。尽管尚处于探索阶段,但这类方法正成为提升分割精度的重要方向。Gaona等^[29]提出了一种基于深度循环神经网络(RNN)的心音分割方法,该方法利用傅里叶同步压缩变换(FSST)提取心音的瞬时频率特征,并采用双向长短期记忆网络(Bi-LSTM)对这些时频特征进行序列建模,从而捕捉S1、S2及

心音周期中其他阶段的时间依赖性。在PhysioNet数据库上的验证中,该模型通过端到端的训练流程实现了高效的自动分割,取得了分割准确率91.3%、灵敏度89.5%的性能,展示了小规模网络在特征选择优化下的潜在优势。此外,Martins等^[30]提出了一种基于马尔可夫神经网络(MNN)的心音分割方法,该方法将隐马尔可夫模型(HMM)的统计先验与一维卷积神经网络(1D-CNN)的判别能力相结合,构建了一个混合框架。通过HMM建模心音信号的潜在状态序列,并利用人工神经网络(ANN)估计后验分布,该方法在多种公开数据集上取得了敏感度约95%、阳性预测值约94%的平均性能。同时,他们还提出了一种无监督微调算法,使预训练的MNN能够适应未见数据的分布偏移。在跨数据集实验中,通过在一个数据集上预训练并迁移到另一个数据集,MNN的阳性预测值提升了3.90%,这项研究展示了结合统计先验与深度学习优势在心音分割任务中的潜力。表3总结了上述研究者在心音分割领域的具体工作细节。

Table 3 Summary of heart sound segmentation methods

表3 心音分割方法汇总

作者	采用数据集	是否分割心音信号	分割效果		
			灵敏度/%	准确率/%	阳性预测值/%
Golpaygani等 ^[25]	离散小波变换结合香农能量包络	自建数据集	89.58	93.14	\
Varghees等 ^[26]	经验小波变换	PhysioNet与PASCAL心音数据库	94.38	91.92	97.25
Thomas等 ^[27]	分形分解	密歇根心音与eGeneralMedicalPCG数据库	98.20	94.82	99.58
Kang等 ^[28]	平均香农能量	自建数据集	96.70	\	98.00
Gaona等 ^[29]	深度循环神经网络	PhysioNet心音数据库	89.50	91.34	89.30
Martins等 ^[30]	马尔可夫模型结合1D-CNN	PhysioNet与CirCor DigiScope心音数据库	95.00	\	94.00

2.3 心音信号特征提取

特征提取和选择是心音信号分析的关键任务,在去噪和分割之后,通常需要人工从原始心音数据中提取特征。一般而言,心音特征的人工提取可以分为时域特征、频域特征和时频域特征。目前研究中常用的方法一般有快速傅里叶变换(FFT)、小波变换(WT)和Mel频率倒谱系数(MFCC)。Rath等^[31]通过从心音图信号中提取MFCC和离散小波变换(DWT)特征,将其作为输入应用于不同的模型,并进行了系统分析。研究结果表明,尽管单独使用MFCC或DWT特征能够捕捉到一定的信号信息,但两者的联合使用显著提升了模型分类性能。这种提升源于两种特征在信号表示方面的互补性:MFCC侧重于提取频率域的全局信息,尤其能反映心音信号的音质特性;而DWT则通过多尺度时频分析捕捉信号的局部特征,特别适用于非平稳信号的处理。通过结合MFCC和DWT特征,模型能够充分利用频率域和时频域的信息,从而在心音分类和心脏病检测任务中实现更高的识别精度和分类性能。

心电图(ECG)记录心脏电活动的变化,而心音图(PCG)反映心音的振动波形特征。通过提取并融合这两种模态的特征,能够从电生理活动和机械振动两方面全面刻画心脏的生理与病理状态。Li等^[32]提出了一种基于

ECG和PCG的双模态心脏功能信号分类算法,采用小波散射变换(WST)提取互补特征,显著提高了心脏疾病诊断的准确性,展示了双模态方法在分类性能上的优越性。Singhal等^[33]则利用高分辨率超级小波变换(ASLT)将ECG和PCG信号转换为二维时频谱图,并结合VGG19深度学习模型,实现了心血管疾病的精确分类。这些研究表明,结合ECG的电活动特征与PCG的机械活动特性,双模态方法为心脏异常的早期检测提供了更全面的信息支持。

3 心音信号分类算法

近年来,基于心音图的心音状态分类及心脏疾病检测模型已被广泛提出。表4对近年来提出的部分深度学习和机器学习方法进行了总结。

3.1 机器学习算法

在机器学习方法中,研究人员设计、改进或应用了多种分类器,以探讨人工提取的特征与心音状态之间的关系。这些分类器涵盖了无监督学习、监督学习、集成学习、人工神经网络以及其他各类分类方法。支持向量机(Support Vector Machine, SVM)凭借其卓越的分类性能和优良的泛化能力,在心音分类任务中得到了广泛应用。SVM通

Table 4 Summary of heart sound classification methods

表 4 心音分类方法总结

作者	数据集	输入特征	模型	分割效果	
				灵敏度/%	准确率/%
Taneja 等 ^[35]	2016 PhysioNet/CinC 数据库	从 Gammatonegram 图像提取纹理特征	SVM	93.61	94.00
Singh 等 ^[36]	2016 PhysioNet/CinC 数据库	小波分解特征、希尔伯特包络特征、短时傅里叶变换、同态包络提取	KNN	93.00	90.00
Tao 等 ^[37]	2016 PhysioNet/CinC 数据库	Mel 频率倒谱系数(MFCC)	2D ViT-1D CRNN	97.35	97.33
Singh 等 ^[39]	2016 PhysioNet/CinC 数据库	短时傅里叶变换	AlexNet+SqueezeNet+VGG19	99.47	99.20
Niizumi 等 ^[40-41]	CirCor DigiScope 数据库	M2D	M2D	89.90	83.20
Ranipa 等 ^[42]	2016 PhysioNet/CinC 和 Yaseen 数据库	MFCC+Mel 频谱+一般频域特征	MACNN	98.41	97.65
Li 等 ^[43]	2016 PhysioNet/CinC 和自收集数据库	MDN-MARNN	MDN-MARNN	94.27	94.19
Choudhary 等 ^[44]	Yaseen 数据库	离散小波变换(DWT)	CNN+GRU	99.67	99.30
Nguyen 等 ^[45]	Yaseen 数据库	Log-Mel 频谱图	CNN	99.67	99.67
Chen 等 ^[46]	2016 PhysioNet/CinC 和 Yaseen 数据库	短时傅里叶变换	CNN	92.50	95.50
Lee 等 ^[47]	2016 PhysioNet/CinC 和 PASCAL 数据库	连续小波变换和小波散射变换	1D-CNN+2D-CNN	92.00	92.00
Ghosh 等 ^[48]	2016 PhysioNet/CinC 和密歇根州数据库	基于高斯窗的改进 Stockwell 变换	SAE-DNN	98.93	97.49

过寻找最大化类别间间隔的最优超平面,并利用径向基函数核等方法处理非线性数据,显著提升了模型适用性和抗干扰能力。这使其在高维特征空间和样本量有限的情况下依然能保持稳定的分类效果,减少了过拟合风险。因此,SVM 已成为心音信号分类中的主流机器学习方法之一。首次将 SVM 系统应用于心音分类领域的研究由 Maglogiannis 等^[34]开展,通过多阶段分类方法,成功识别出特定心脏瓣膜疾病,验证了 SVM 在处理高维、非线性数据时的卓越性能和泛化能力,展示了 SVM 在心音分类任务中的应用潜力。在近期研究中,Taneja 等^[35]利用 Gammatonegram 图像提取纹理特征,并使用 SVM 算法对心音信号进行分类,基于 PhysioNet CinC 2016 数据库,取得了 94.00% 的分类准确率。

除 SVM 外,K 最近邻算法(KNN)也是首选的心音分类方法之一。KNN 通过计算样本间的距离,根据最近的 K 个邻居决定类别,具有实现简单和对多类别问题灵活的优点。一些传统的机器学习方法需要对心音信号进行分割,以提取特定阶段的特征,而 KNN 作为一种基于实例的分类算法,依赖于整个特征空间中的样本相似性,可以直接使用整体或全局提取的特征,无需对信号进行明确分割,极大简化了处理流程。Singh 等^[36]利用提取的时间-频率特征,通过 KNN 分类器实现了未分割心音的自动分类,在 PhysioNet CinC 2016 数据集上取得了 90% 的准确率。研究提出,基于未分段 PCG 信号的分类方法比基于分段 PCG 信号的分类方法快 11 倍,证明了 KNN 算法在无需复杂分割处理的情况下,依然能够实现高效、稳定的分类性能,简化了计算流程,适合实时应用场景。

3.2 深度学习算法

随着深度学习技术的迅猛发展,基于深度神经网络的算法在心音信号分类任务中展现出显著优势。当前,许多最先进的模型(SOTA)均基于卷积神经网络(CNN)与循环神经网络(RNN)的改进与结合。作为深度学习的两大核

心架构,CNN 和 RNN 在心音分类领域得到广泛应用。CNN 通过层级卷积操作,能够自动提取心音信号中的局部特征,其擅长处理具有空间结构的输入数据,如心音的时频图谱,其深层次的特征表示能力使模型在复杂模式识别任务中表现优异。RNN,尤其是其变种长短期记忆网络(LSTM)和门控循环单元(GRU),在处理具有时间依赖性的心音序列数据方面表现出色。RNN 能够捕捉心音信号中的时间动态信息,增强了模型对心音变化趋势的理解能力。当前,许多先进模型通过结合 CNN 的特征提取能力与 RNN 的序列建模能力,进一步提升了心音分类准确性。例如,Tao 等^[37]设计了一种结合一维卷积循环神经网络(1D-CRNN)和二维视觉 Transformer(2D-ViT)的混合模型,将时间域与时间频率域特征进行融合,提出了 2D ViT-1D CRNN 模型。该模型在 PhysioNet/CinC 2016 和 PASCAL A 数据集上的分类任务中取得了显著成果,其中在 PhysioNet/CinC 2016 数据集上达到了 97.33% 的分类准确率,与现有的 SOTA 方法相比性能卓越。在消融实验中,单独使用 1D-CRNN 相比单独使用 CNN 在准确性上显著提升(提高约 2.14%),表明序列建模能力在心音信号分类任务中具有关键作用。同时,结合两者的混合模型进一步展现出卓越的性能优势。

在心音信号分类任务中,标注数据量通常较少,这限制了从头训练深度神经网络模型的可行性。为了应对数据不足的问题,迁移学习被广泛应用。迁移学习通过利用在大规模数据集上预训练的深度学习模型,如 VGG19 和 ResNet50^[38],将其学到的特征和知识迁移到心音分类任务中,如图 2 所示。这不仅提升了模型在小数据集上的性能,还缩短了训练时间,提高了泛化能力。通过微调这些预训练模型,研究人员能够在有限的心音数据上实现高效且准确的分类,充分发挥迁移学习在数据稀缺环境下的优势。Singh 等^[39]通过结合迁移学习和集成学习的方法显著优化了心音信号分类模型的性能,其利用预训练的深度学习

神经网络,包括 AlexNet、SqueezeNet 和 VGG19,提取频谱图中蕴含的时间频率特征,并通过加权集成策略有效增强了模型对异常心音的敏感性。迁移学习在此过程中发挥了重要作用,不仅有效缓解了数据稀缺的限制,还通过减少模型训练所需参数和加速收敛过程,大幅提高了分类准确性和效率。最终,所提出的模型在 PhysioNet-2016 数据集上实现了 99.20% 的分类准确率和 99.47% 的敏感性。

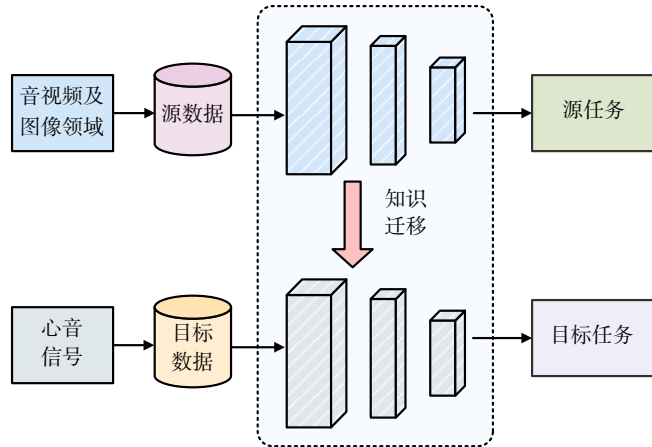


Fig. 2 Schematic diagram of the principle of transfer learning

图2 迁移学习原理示意图

近年来,一些基于 Transformer 架构的预训练模型也被尝试应用于心音分类任务。尽管这些模型尚未在性能上实现显著提升,但它们通过迁移在其他领域积累的丰富知识,展示了迁移学习在解决数据不足问题上的广泛适用性和潜力,表明这一方向具有良好的研究前景。Niizumi 等^[40-41]的研究验证了基于大规模数据集预训练的通用音频表征在心脏杂音检测任务中的迁移学习潜力。他们提出的 M2D 模型是基于 Transformer 架构,并通过自监督学习在大规模音频数据集 AudioSet 上进行预训练,充分利用了遮掩预测任务捕捉全局特征的能力。实验结果表明,M2D 在迁移到 CirCor DigiScope 数据集后,显著提升了心脏杂音检测的加权准确率,达 83.2%,在该数据集上超越了以往最先进方法。这不仅验证了通用音频表征在处理心音信号中的有效性,也展示了在大规模数据集上预先训练的通用音频表征在心脏杂音检测中迁移学习的潜力。

随着智能化心音分析技术在心血管疾病早期诊断领域应用前景的不断拓展,未来研究应聚焦于解决当前面临的关键挑战,以促进该技术在临床实践中的转化与广泛应用。

4 未来研究展望

研究发现,数据库的不足仍然是限制心音智能听诊技术发展主要瓶颈之一。现有的心音数据库样本量较为有限,且缺乏足够多样的病理特征和特殊人群数据。为此,未来研究应优先建设涵盖大规模、多样病理特征的高质量心音数据库,并优化数据采集与标注流程。通过跨地

域、多中心的合作,确保数据库能够全面反映全球各地区心血管疾病的多样性。尤其是对于特殊人群及罕见疾病的数据采集,未来应给予更多关注,以提高模型泛化能力,确保其在不同临床环境中能够表现出较高的准确性与稳定性。

在技术层面,模型轻量化是实现心音智能分析广泛应用的另一项关键任务。当前,许多深度学习模型计算量庞大,且对硬件资源要求较高,这在资源受限的设备(如便携式医疗设备和远程监测平台)上运行时,存在较大困难。为了推动智能心音分析技术在基层医疗和远程医疗中的应用,未来研究应重点关注设计高效、轻量化的模型。例如,通过模型压缩和参数优化等技术手段,可以显著降低模型计算复杂度,使其适应便携设备和移动平台。结合边缘计算技术,可将数据分析任务移至终端设备进行处理,从而减少数据传输延迟,提高响应速度,减少对云端计算资源的依赖,并进一步提升系统可用性和普适性。这些技术的应用不仅可以增强基层医疗机构的服务能力,还将显著拓展技术的应用范围,尤其是在偏远地区及发展中国家的推广应用。

此外,随着多模态信号融合技术的进步,心音、心电图(ECG)和血氧信号等多种生理数据的综合分析将为心血管健康管理提供更加精准的信息。未来研究可集中于开发集成心音、心电图和血氧信号的多模态诊断设备,实现更全面的生理状态监测与异常检测。例如,通过融合心音与心电图信号,可以在临床中实现对心脏疾病的早期精确诊断,同时结合血氧饱和度数据,为心血管疾病的风险评估提供更多维度的支持。这不仅能够提升医院和诊所的诊断效率,还将为社区医疗中心和家庭健康管理提供更加实用的智能化解决方案,进一步满足居民健康管理需求。

总结来看,未来研究方向应集中在数据库建设、模型优化以及多模态信号的融合分析等方面,这些举措将有助于进一步推动心音信号分析技术的临床转化和普及应用,为全世界尤其是发展中国家和地区的心血管健康管理提供更加高效、可负担的解决方案。通过解决现有技术瓶颈并不断创新,心音智能听诊技术将成为推动全球心血管疾病预防和治疗的重要工具。

5 结语

本文系统总结了近年来人工智能技术在心音信号分析领域的研究进展,围绕心音信号的预处理、特征提取与分类方法等关键环节,梳理了现有技术的特点与应用成效。在预处理阶段,总结了降噪与信号分割的常用方法,包括传统技术与基于深度学习的改进策略;在特征提取方面,汇总了时域、频域和时频域特征的主要提取方法及其在不同研究中的表现;在分类方法上,分别总结了机器学习和深度学习技术的应用,特别是在深度学习部分强调了

迁移学习在数据受限条件下的应用,展示了其在提升模型性能和适应性方面的潜力。本文通过总结现有的先进方法与应用经验,全面梳理了心音信号分析技术在疾病诊断中的研究现状,展现了深度学习驱动下该领域的快速发展。

参考文献:

- [1] World Heart Federation. World heart report 2023: confronting the world's number one killer [EB/OL]. <https://world-heart-federation.org/resource/world-heart-report-2023/>.
- [2] ZENG W W. Status of cardiovascular disease in China[J]. *Journal of Geriatric Cardiology*, 2023, 20(6): 397-398.
- [3] FAN M, WANG Q, LIU J, et al. Real-world evaluation of the stethoscope electronic tele-auscultation system [J]. *BioMedical Engineering OnLine*, 2022, 21(1): 63.
- [4] ZHUGE Y, RONG L, YE L, et al. The value of the wireless stethoscope in patients with COVID-19 infection in a makeshift hospital [J]. *Biomedical Engineering Online*, 2023, 22(1): 71.
- [5] LIU C, SPRINGER D, LI Q, et al. An open access database for the evaluation of heart sound algorithms [J]. *Physiological Measurement*, 2016, 37(12): 2181.
- [6] GOMES E F, BENTLEY P J, PEREIRA E, et al. Classifying heart sounds—approaches to the PASCAL challenge [C]// *Proceedings of the International Conference on Health Informatics*, 2013: 337-340.
- [7] YASEEN, SON G Y, KWON S. Classification of heart sound signal using multiple features [J]. *Applied Sciences*, 2018, 8(12): 2344.
- [8] OLIVEIRA J, RENNA F, COSTA P D, et al. The CirCor DigiScope dataset: from murmur detection to murmur classification [J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 26(6): 2524-2535.
- [9] SPADACCINI A, BERITELLI F. Performance evaluation of heart sounds biometric systems on an open dataset [C]// *2013 18th International Conference on Digital Signal Processing (DSP)*, 2013: 1-5.
- [10] CESARELLI M, RUFFO M, ROMANO M, et al. Simulation of foetal phonocardiographic recordings for testing of FHR extraction algorithms [J]. *Computer Methods and Programs in Biomedicine*, 2012, 107(3): 513-523.
- [11] QIAN K, DONG F Q, REN Z, et al. Opportunities and challenges for heart sound recognition: a brief on the heart sounds Shenzhen corpus [J]. *Journal of Fudan University (Natural Science)*, 2020, 59(3): 354-359.
钱昆,董逢泉,任昭,等. 心音识别的机遇与挑战:深圳心音数据库简介 [J]. *复旦学报(自然科学版)*, 2020, 59(3): 354-359.
- [12] OLIVEIRA J, RENNA F, MANTADELIS T, et al. Adaptive sojourn time HSMM for heart sound segmentation [J]. *IEEE Journal of Biomedical and Health Informatics*, 2018, 23(2): 642-649.
- [13] KAZEMNEJAD A, KARIMI S, GORDANY P, et al. An open-access simultaneous electrocardiogram and phonocardiogram database [J]. *Physiological Measurement*, 2024, 45(5): 055005.
- [14] LUBAIB P, BEEGUM D, MUHSINA N, et al. Development of an efficient cardiac auscultation system using proper denoising techniques [C]// *2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, 2023: 1-8.
- [15] ALKHODARI M, FRAIWAN L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings [J]. *Computer Methods and Programs in Biomedicine*, 2021, 200: 105940.
- [16] SINGH S A, MAJUMDER S, MISHRA M. Classification of short unsegmented heart sound based on deep learning [C]// *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2019: 1-6.
- [17] BHANGALE K B, KOTHANDARAMAN M. Survey of deep learning paradigms for speech processing [J]. *Wireless Personal Communications*, 2022, 125(2): 1913-1949.
- [18] SHARAN T S, BHATTACHARJEE R, SHARMA S, et al. Evaluation of deep learning methods (DnCNN and U-Net) for denoising of heart auscultation signals [C]// *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, 2020: 151-155.
- [19] ALI S N, SHUVO S B, AL-MANZO M I S, 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.
- [20] CHANDEL A, KUMAR V, MUDULI P R. Stacked Bi-LSTM network and dual signal transformation for heart sound denoising [C]// *International Conference on VLSI, Communication and Signal Processing*, 2022: 123-133.
- [21] LIANG M, HU J, ZHOU X, et al. A GAN based heart sound denoising model [C]// *2022 12th International Conference on Information Technology in Medicine and Education (ITME)*, 2022: 666-670.
- [23] VERMA N, BISWAS A K, POTDAR R. Phonocardiogram signal segmentation and classification using variance fractal dimension [J]. *International Journal of Advanced Research in Computer Science*, 2013, 4(3): 265-270.
- [24] CHEN Y, WANG J, HU J. Heart sound segmentation based on an improved DHMM with Shannon energy [C]// *2023 4th International Conference on Information Science, Parallel and Distributed Systems (ISPDS)*, 2023: 54-59.
- [25] GOLPAYGANI A T, ABOLPOUR N, HASSANI K, et al. Detection and identification of S1 and S2 heart sounds using wavelet decomposition method [J]. *International Journal of Biomathematics*, 2015, 8(6): 1550078.
- [26] VARGHEES V N, RAMACHANDRAN K. Effective heart sound segmentation and murmur classification using empirical wavelet transform and instantaneous phase for electronic stethoscope [J]. *IEEE Sensors Journal*, 2017, 17(12): 3861-3872.
- [27] THOMAS R, HSI L L, BOON S C, et al. Heart sound segmentation using fractal decomposition [C]// *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016: 6234-6237.
- [28] KANG S, DOROSHOW R, MCCONNAUGHEY J, et al. Heart sound segmentation toward automated heart murmur classification in pediatric patients [C]// *2015 8th International Conference on Signal Processing, Image Processing and Pattern Recognition (SIP)*, 2015: 9-12.
- [29] GAONA A J, ARINI P D. Deep recurrent learning for heart sounds segmentation based on instantaneous frequency features [DB/OL]. <https://arxiv.org/abs/2201.11320>.

- [30] MARTINS M L, COIMBRA M T, RENNA F. Markov-based neural networks for heart sound segmentation: using domain knowledge in a principled way [J]. *IEEE Journal of Biomedical and Health Informatics*, 2023, 27(11): 5357–5368.
- [31] RATH A, MISHRA D, PANDA G, et al. Development and assessment of machine learning based heart disease detection using imbalanced heart sound signal [J]. *Biomedical Signal Processing and Control*, 2022, 76: 103730.
- [32] LI J, KE L, DU Q, et al. Multi-modal cardiac function signals classification algorithm based on improved DS evidence theory [J]. *Biomedical Signal Processing and Control*, 2022, 71: 103078.
- [33] SINGHAL S, KUMAR M. Cardiovascular diseases classification using high-resolution superlet transform on ECG and PCG signals [C]//2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023: 1–5.
- [34] MAGLOGIANNIS I, LOUKIS E, ZAFIROPOULOS E, et al. Support vectors machine-based identification of heart valve diseases using heart sounds [J]. *Computer Methods and Programs in Biomedicine*, 2009, 95(1): 47–61.
- [35] TANEJA K, ARORA V, VERMA K. Heart sound classification method using gammatonegram and SVM [J]. *Multimedia Tools and Applications*, 2024, 83: 1–37.
- [36] SINGH S A, MAJUMDER S. Classification of unsegmented heart sound recording using KNN classifier [J]. *Journal of Mechanics in Medicine and Biology*, 2019, 19(4): 1950025.
- [37] TAO Z, REN Z, YANG X, et al. 2D ViT and 1D CRNN-based heart sound signals detection model [C]//2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2023: 2265–2268.
- [38] MASCARENHAS S, AGARWAL M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification [C]//2021 International Conference on Disruptive Technologies for Multi-disciplinary Research and Applications (CENTCON), 2021: 96–99.
- [39] SINGH S A, DEVI N D, SINGH K N, et al. An ensemble-based transfer learning model for predicting the imbalance heart sound signal using spectrogram images [J]. *Multimedia Tools and Applications*, 2024, 83(13): 39923–39942.
- [40] NIIZUMI D, TAKEUCHI D, OHISHI Y, et al. Exploring pre-trained general-purpose audio representations for heart murmur detection [DB/OL]. <https://arxiv.org/abs/2404.17107>.
- [41] NIIZUMI D, TAKEUCHI D, OHISHI Y, et al. Masked modeling duo: towards a universal audio pre-training framework [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024, 32: 2391–2406.
- [42] RANIPA K, ZHU W P, SWAMY M. A novel feature-level fusion scheme with multimodal attention CNN for heart sound classification [J]. *Computer Methods and Programs in Biomedicine*, 2024, 248: 108122.
- [43] LI S, SUN J, YANG H, et al. Interpretable end-to-end heart sound classification [J]. *Measurement*, 2024, 237: 115113.
- [44] CHOUDHARY R R, SINGH M R, JAIN P K. Heart sound classification using a hybrid of CNN and GRU deep learning models [J]. *Procedia Computer Science*, 2024, 235: 3085–3093.
- [45] NGUYEN M T, LIN W W, HUANG J H. Heart sound classification using deep learning techniques based on log-mel spectrogram [J]. *Circuits, Systems, and Signal Processing*, 2023, 42(1): 344–360.
- [46] CHEN J, GUO Z, XU X, et al. A robust deep learning framework based on spectrograms for heart sound classification [J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023, 21(4): 936–947.
- [47] LEE J A, KWAK K C. Heart sound classification using wavelet analysis approaches and ensemble of deep learning models [J]. *Applied Sciences*, 2023, 13(21): 11942.
- [48] GHOSH S K, PONNALAGU R, TRIPATHY R K, et al. Automated heart sound activity detection from PCG signal using time-frequency-domain deep neural network [J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1–10.

(责任编辑:孙娟)