基于MLFCC特征和P-CRNN的心音分类研究

宋晨翔,张孙杰,刘奥磊,王 哲

上海理工大学光电信息与计算机工程学院,上海

收稿日期: 2025年1月24日; 录用日期: 2025年2月17日; 发布日期: 2025年2月26日

摘要

心音分类是医学领域中一个重要的诊断任务。但由于数据量不足、宽频带信号表征能力欠缺、心音序列 上下文学习存在挑战等问题,该领域依旧存在许多的不足之处。针对以上问题,提出了一种心音分类的 新方法:对每段心音信号进行重叠切片,截取成2s的信号作为样本;随后采用改进的梅尔频率倒谱系数 (MFCC)与线性频率倒谱系数(LFCC)分别提取心音信号相应频率系数,并分别计算其一阶差分作为融合特 征。分类网络使用提出的并行卷积递归神经网络(P-CRNN)方法进行训练。实验表明,相比于其他传统识 别方法,所提方法对心音信号分类有明显提高。

关键词

心音信号, 混合特征, 分类网络

Research on Heart Sound Classification Based on Mixed Features of MFCC and LFCC and Parallel Convolutional Recurrent Neural Network

Chenxiang Song, Sunjie Zhang, Aolei Liu, Zhe Wang

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Jan. 24th, 2025; accepted: Feb. 17th, 2025; published: Feb. 26th, 2025

Abstract

Heart sound classification is a crucial diagnostic task in the medical field. However, due to the insufficient amount of data, the lack of wideband signal representation ability, and the challenges of contextual learning of heart sound sequences, there are still many shortcomings in this field. To address these issues, a novel method for heart sound classification was proposed. In this method, each signal was segmented into 2-second overlapping slices, and the corresponding frequency coefficients were extracted using improved Mel Frequency Cepstral Coefficients (MFCC) and Linear Frequency Cepstral Coefficients (LFCC). The first-order difference was calculated as fusion features, and the proposed Parallel Convolutional Recurrent Neural Network (P-CRNN) method was trained for classification in comparison to traditional recognition methods. The experiment shows that compared to other traditional recognition methods, the proposed method has a significant improvement in heart sound signal classification.

Keywords

Heart Sound Signal, Hybrid Feature, Classification Network

Copyright © 2025 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/

1. 引言

每年全球因心血管疾病而丧生的人数高达数百万,被认为是世界上最常见的致命疾病之一[1]。心血 管疾病严重影响健康,以至于导致心脏病发作、中风以及心力衰竭。它不仅严重影响个人健康,同时也 给医疗保健系统带来巨大的财政压力。因此,对于控制心血管疾病而言早期的发现和诊断便至关重要。 心音听诊是心血管疾病临床检查的重要组成部分,其作为一种不可或缺的非侵入式诊断工具,具有成本 低、效率高、无痛等优点。该方法有助于检测一系列病理性心脏异常,如二尖瓣反流、主动脉瓣狭窄等 瓣膜性心脏病,以及心力衰竭和心律失常等[2]。然而,传统的心音听诊方法仍然严重依赖于医生的经验 水平,心脏病专家的听诊诊断准确率约为 80%,而初级心脏病医生的听诊诊断率在 20%~40% [3]。因此, 为了解决这个问题,研究一种自动心音分类的识别方法迫在眉睫。

近年来,深度学习技术在心音分类领域表现出优异的性能[4]-[6]。特别是自 PhysioNet/CinC Challenge 2016 [7]发布以来,基于深度学习的心音分类方法在数量和质量上都有了质的飞跃。Posts 等人[8]采用卷 积神经网络(Convolutional Neural Network, CNN)与自适应增强算法(Adaptive Boosting, AdaBoost)相结合的 策略,从心音图(Phonocardiogram, PCG)中提取 124 个时频特征。他们在随机抽样的测试集上实现了 86.0% 的未加权平均召回率(Unweighted Average Recall, UAR),在 PhysioNet/CinC Challenge 2016 中获得了第一 名。Yaseen 等人[9]利用从心音信号中提取的梅尔频率倒谱系数(Mel Frequency Cepstral Coefficients, MFCC) 和离散小波变换(Discrete Wavelet Transform, DWT)特征,使用支持向量机(Support Vector Machine, SVM) 和 DWT 进行训练和分类,并进一步将这种方法与基于质心位移的 K 近邻算法(K-Nearest Neighbors, KNN) 和基于深度神经网络(Deep Neural Networks, DNN)的分类算法相结合,对心音进行分类,从而实现了 97% 的分类准确率。Akram MU 等人[10]通过同态包络图的包络提取和辨识突出的峰值,对心音信号中的心跳进行定位,利用径向基函数核的非线性支持向量机根据提取的特征对心音进行分类,最终获得 91%的准确率。LiF 等人[11]将传统的特征工程方法与深度学习算法相结合,从 8 个域中提取 497 个特征,将卷积神经网络中的全局平均池化层取代分类层之前使用的全连接层,最后采用分层五重交叉验证来评价所提出方法的性能,在 PhysioNet/CinC Challenge 2016 数据集上获得的准确率、灵敏度、特异性分别为 86.8%、87%和 86.6%。Zhang 等人[12]提出了一种基于比例谱图与偏最小二乘回归(Partial Least Squares Regression,

PLSR)的 PCG 信号分类方法。采用香农能量包络的短时平均幅度差,计算得到的心循环谱图,随后进行特征提取;并在处理对数谱图时采用双线性插值,得到固定尺寸的缩放谱图;之后采用 PLSR 对标度谱 图进行降维;最后通过 SVM 得到分类结果。虽然已经有大量关于心音分类的方法与工作,但由于心音信 号采集环境的复杂多变,导致大多数方法的性能在实际使用中都有所下降,主要原因在于所选的人工特 征提取方法仅能表征心音信号的局部特征,以及采取的网络模型过于简单使得模型的泛化能力较弱。因 此本文提出的 MFCC 与线性频率倒谱系数(Linear Frequency Cepstral Coefficients, LFCC)相融合的特征和 并行卷积递归神经网络的心音分类模型,通过综合 MFCC 在处理语音信号上的成熟优势,以及引入 LFCC 在提供更高频率分辨率和适应非语音信号处理上的显著优点,并融合两者的一阶差分反映信号的动态特 征,从而实现了更为全面和精确的心音信号特征表征;随后利用提出的并行卷积递归神经网络(Parallel Convolutional Recurrent Neural Network, P-CRNN)对融合特征进行分类学习。本文使用的方法不需要进行 心音分割,降低了算法复杂度的同时提升了模型鲁棒性和泛化能力。

2. 算法原理

2.1. 整体框架

对心音数据集中所采集的 PCG 信号进行分类识别一般分为三个步骤:预处理、特征提取和分类识别, 具体任务步骤示意图如图1所示。





预处理部分一般便是对心音信号进行去噪、分割等操作,从而改善信号质量,本文将采用五阶巴特 沃斯滤波器对初始信号进行滤波处理。特征提取阶段保留更多的心音信号并且简化处理流程,有助于识 别不同心音类别的信息,因此本文将在不进行心音分割的情况下对心音信号进行处理,采用改进后的 MFCC与LFCC提取心音信号相对应的频率特征,并对其计算一阶差分进行特征信号的融合。分类识别 步骤则是使用提取的特征来区分不同类别的心音,本文将采取卷积循环神经网络(Convolutional Recurrent Neural Network, CRNN)与双向长短期记忆网络(Bidirectional Long Short-Term Memory Network, BiLSTM) 并行处理的方式对特征信息进行分类处理。具体的心音信号分类整体框架图如图 2 所示。

2.2. 预处理

首先,利用五阶巴特沃斯滤波器对原始信号进行滤波,相比于心音信号的局部心音片段,整体信息 记录更易受到噪音与杂波的干扰。因此,我们引入重叠窗口的固定长度切片,从而获取更多的数据信息, 以解决数据量不足、数据质量残次不齐等问题,在后续处理过程中,可获得更好的鲁棒性和泛化能力。

由于一个心跳周期的持续时长通常在 0.6 s~1 s,因此我们将滤波后的心音信号切片成 2 s 的切片级 信号,并且重叠窗口每次移动的步长为 1 s,确保每段至少有一个完整的心音周期信号被采集到,具体处 理形式如图 3 所示。

2.3. 特征提取

为了克服单一特征参数在信号表征上的局限性,本研究采用了传统 MFCC 与性能更为灵活优异的 LFCC 结合的改进方法进行心音信号的特征提取,能够更全面地捕捉心音信号的本质特征,为后续的心音 分类和识别提供了更为坚实的基础。





Figure 3. Overall framework diagram of heart sound signal classification 图 3. 心音信号重叠切片方法示意图

在传统方法中,在特征提取过程中常采用汉明窗来进行加窗,可能会导致频谱泄露等缺点。因此本研究采用汉宁自卷机窗(Hanning self-convolution window, HSCW)代替汉明窗对分帧后的信号进行加窗,从心音信号中提取改进后的 MFCC 和 LFCC 特征,通过此类方法,能够在频域分析时,更有效地减少旁

瓣干扰,从而减少频谱泄露,提高频谱估计的精度。

2.3.1. 改进 MFCC 与 LFCC

MFCC 根据人耳听觉感知机理,反应了人耳与声音频率之间的非线性关系,从而提出梅尔刻度频率域,梅尔频率*f*与实际频率之间的关系见下式(1)所示。

$$Mel = 2595 \lg \left(1 + \frac{f}{700}\right) \tag{1}$$

式中,Mel(f)为梅尔刻度频率,单位为Mel,f为心音信号的实际频率,单位为Hz。

LFCC 则采用线性频率域,滤波器在频率空间内以等间距均匀覆盖,以捕捉频谱的连续特性。

进行改进 MFCC 与 LFCC 特征提取具体过程如下:

(1) 预加重

由于在采集过程中高频分量损失,所以采用预加重弥补采集后信号的高频分量损失,以使频谱平滑, 同时,通过一阶预加重滤波器 *H*(*z*),对随机噪音进行有效抑制,滤波器 *H*(*z*)公式如式(2)。

$$H(z) = 1 - \mu z^{-1} \tag{2}$$

式中, µ为预加重系数,取值通常在(0.9~1)之间。

(2)分帧、加窗

虽然心音信号属于非平稳信号,但在 24 ms~256 ms 之间达到稳定状态[13]。因此,本研究首先将心音信号进行分帧,并在各短帧内进行特征提取。为了实现帧与帧之间的平滑过渡,在连续帧之间采取 50% 的重叠,并且采取汉宁自卷积窗代替传统汉明窗,从而达到更好减少旁瓣干扰、抑制频谱泄露等效果,进一步提升特征参数的准确性。经过实验验证,采取 2 阶汉宁自卷积窗作为窗函数对信号进行加窗处理。 2 阶汉宁自卷积窗表达式如式(3)所示。

$$W_{H2}(n) = \left\{\frac{1}{2}W_R(n) + \frac{1}{4}\left[W_R\left(n - \frac{2\pi}{N}\right) + W_R\left(n + \frac{2\pi}{N}\right)\right]\right\}^2$$
(3)

式中, W_R(n)为矩形框的频率响应, 表达式见式(4)所示。

$$W_{R}(n) = e^{-\frac{j(N-1)n}{2}} \left(\frac{\sin\left(\frac{nN}{2}\right)}{\sin\left(\frac{n}{2}\right)} \right)$$
(4)

(3) 快速傅里叶变换(Fast Fourier Transform, FFT)

为了将信号从时域转换到频域,本文将采取快速傅里叶变换,将时域信号 *x*(*n*)转换为频域信号 *X*(*k*), 方程如式(5)。

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi nk/N}, 0 \le n, k \le N-1$$
(5)

其中 x(n) 为经过降噪、加窗、分帧预处理以后的心音信号, n 为每帧的采样个数。

(4) 功率谱图计算

将信号频谱 X(k) 的模平方,得到功率谱 P(k):

$$P(k) = \frac{1}{N} \left| X(k) \right|^2 \tag{6}$$

DOI: 10.12677/mos.2025.142144

(5) 梅尔频率三角滤波器组与线性滤波器组

将功率谱 P(k) 通过一组梅尔三角滤波器组进行滤波与梅尔刻度转换,在每个频率刻度下,计算 P(k) 与滤波器 H_m(k) 的乘积。定义一个有 M 个滤波器的三角滤波器组,三角滤波器 H_m(k)的频率响应如下:

$$H_{m}(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{k - f(m-1)}{f(m) - f(m-1)}, & f(m-1) \le k \le f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)}, & f(m) \le k \le f(m+1) \\ 0, & k > f(m+1) \end{cases}$$
(7)

式中, f(m)为第m个滤波器的中心频率。

同时将功率谱 *P*(*k*) 通过一组线性滤波器组进行滤波,频率响应如上式相同,但与梅尔滤波器组不同的是,线性滤波器组不需要进行梅尔刻度转换,其中心频率 *f*(*m*) 在频率轴上均匀分布。

(6) 对数能谱计算

随后通过对数运算得到每帧信号的对数能谱S(m):

$$S(m) = \ln\left[\sum_{k=0}^{N-1} P(k) H_m(k)\right], \ 0 \le m \le M$$
(8)

式中P(k)为功率谱, $H_m(k)$ 为滤波器组,M为滤波器个数。

(7) 离散余弦变换(Discrete Cosine Transform, DCT)

对上述对数谱进行离散余弦变换,从而得到 MFCC 与 LFCC 系数 C(n):

$$C(n) = \sum_{m=0}^{N-1} S(m) \cos(\pi n(m-0.5)/M), \ n = 1, 2, \cdots, L$$
(9)

式中, L为 MFCC 与 LFCC 系数的阶数, M为滤波器组的个数。

(8) 动态 MFCC 与 LFCC 特征提取

通过上面的特征提取的 MFCC 特征和 LFCC 特征仅反映了心音信号的静态特征,由于人耳对于声音 信号的动态特性更为敏感,心音频谱的动态信息中也包含了丰富的信息,可以进一步提高分类准确率。 为了更好反映心音信号的动态信息,将 MFCC 与 LFCC 分别提取一阶差分系数 ΔMFCC 和 ΔLFCC。计算 公式如下所示:

$$D(n) = \frac{1}{\sqrt{\sum_{i=-k}^{i=k} i^2}} \sum_{i=-k}^{i=k} i \cdot C(n+i)$$
(10)

式中C(n)为心音信号的MFCC系数和LFCC系数,k为差分系数,取值设置为1。

2.3.2. 混合特征融合方法

由于改进 MFCC 与改进 LFCC 采用不同滤波器尺度进行心音信号特征提取,表征心音信号不同的特征信息,为了便于后续处理,因此对两组特征参数混合,构成新的混合特征矩阵,其式如下:

$$MLFCC = (MFCC, \Delta MFCC, LFCC, \Delta LFCC)$$
(11)

2.4. P-CRNN 分类网络

在之前的研究中证明,卷积神经网络(Convolutional Neural Network, CNN)在大规模和高维数据学习中的性能表现优异。递归神经网络(Recurrent Neural Network, RNN)在时间序列学习和长期依赖关系捕获

方面具有较好的效果。因此在本文中,采取将 CNN 与 RNN 结合起来的方式,利用从 CNN 中提取的局部编码特征和 RNN 所捕获的长期依赖关系,对心音信号处理后得到的特征参数进行特征学习,本文提出 了将 CRNN 与 RNN 并行作用的网络模型,结果表明优于其他模型。

传统的 CRNN 模型使用 RNN 学习基于 CNN 输出的时间特征,虽然展现出了强大的特征提取和序列 处理能力,但存在无法保留原始心音信号的时间特征等问题,因此,为了进一步捕获原始信号的时空特 征,本研究提出了一种新的网络框架: P-CRNN。本文提出的 P-CRNN 网络框架示意图如图 4 所示。



Figure 4. P-CRNN Network framework diagram 图 4. P-CRNN 网络框架示意图

如图所示,将此前所提取改进后的 MFCC 与 LFCC 及其一阶差分特征作为网络的输入,并行传送进 CRNN 和 RNN 中分别进行特征学习。随后,将两个并行模块融合成一个统一的特征向量进行后续分类, 采用全连接层和 Softmax 层对不同类间的概率分布进行推导。所提出的 P-CRNN 模型融合了从 CRNN 中 提取的时空特征和从 RNN 提取的时间特征。

在 CRNN 模块中使用了三个卷积层与一个 GRU 层,其中第一个卷积层采用 16 个卷积核,第二个卷 积层采用 32 个卷积核,第三个卷积层采用 64 个卷积核,在每一个卷积核之后使用 PreLU (Parametric Rectified Linear Unit)激活函数、最大池化层和批量归一化层,其中最大池化层分别采用 2×2、2×2和4×4 的窗口。之后采用展平层将数据铺平以便传入 GRU 中进行心音信号的时间特征学习,之后再接入一 个展平层,便于与后续 RNN 中传出的信号进行融合。

在 RNN 模块中,先采取 2X2 的最大池化层对输入信号进行降维,并且使用 LSTM 层进行心音信号时间特征学习。最后将 CRNN 模块与 RNN 模块的输出连接成一个特征向量,并使用 32 个神经元的全连接层学习全局特征。随后采用 dropout 层减少神经元数量,从而防止过拟合,此后再接入由 2 个神经元构成的全连接层。最后采用 Softmax 层推导出正常与异常心音的类别概率。

3. 实验与结果分析

3.1. 实验数据

实验中所使用的心音信号来自 PhysioNet/CinC challenge 2016 数据库,可在 Physio 网站上公开获取。 该数据库来自于七个研究小组贡献的九个独立数据库,包括从临床和非临床环境下健康受试者和各种心 脏病患者中收集的共 3240 份原始心音信号记录。这些心音数据被各种电子听诊器在 4 个主要听诊部位所 采集:主动脉瓣听诊区、肺动脉瓣听诊区、三尖瓣听诊区、二尖瓣听诊区,长度从 5 s 到 120 s 不等,采 样频率为 2000 Hz。其中异常的心音信号来自于心脏瓣膜缺损(包括二尖瓣脱垂、二尖瓣关闭不全、主动 脉瓣反流和主动脉瓣狭窄等)和冠状动脉疾病等各类心脏疾病患者。该数据库共发布心音信号 3240 段, 其中共有 2575 段心音信号被标记为正常,665 段心音信号被标记为异常。实验中将其中 20%作为测试集, 将此外 80%的正常和异常心音信号按比例选取 80%作为训练集,20%作为验证集。具体数据集划分见表 1。

数据集	正常	异常		
训练集	1648	426		
验证集	412	106		
测试集	515	133		

Table 1. Data partitioning 表 1. 数据划分

3.2. 实验环境

实验中所使用的编程软件为 Pycharm 2021,编程语言为 Python 3.9,所使用并行卷积递归神经网络均在 TensorFlow 2.8 框架中实现。硬件环境:中央处理器(Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30 GHz, RAM 为 62 GB),独立显卡 NVIDIA Corporation GP104 [GeForce GTX 1080]。

本文并行卷积递归神经网络模型配置如下:采用 Adam 优化器对目标函数进行优化;学习率初始值为 0.001,衰减率为 1 × 10⁻⁶;损失函数为稀疏分类交叉熵损失函数;Dropout 层的保持概率为 0.5。

3.3. 评价指标

为了评估所提出的方法,每个实验使用了如下分类评估指标:准确率(Accuracy),召回率(Recall),精密度(Precision),F1,灵敏度(Specificity),特异度(Sensitivity)。这些测量的定义如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
(12)

$$\operatorname{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$
(13)

$$Precision = \frac{TP}{TP + FP}$$
(14)

$$F1 = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$
(15)

Specificity =
$$\frac{TN}{TN + FP}$$
 (16)

Sensitivity =
$$\frac{TP}{TP + FN}$$
 (17)

式中, TP 为异常心音样本被正确分类的数量; TN 为正常心音被正确分类的数量; FP 为正常心音被错误 分类的数量; FN 为异常心音被错误分类的数量。

3.4. 结果分析

3.4.1. 与不同网络框架性能对比

为了进一步说明 P-CRNN 的分类性能,本研究与不同网络框架进行了性能对比,如: CNN、RNN 和 CRNN。其具体网络框架如表 2 所示。

表 2. 其他传统网络框架				
模型	网络框架			
CNN	Conv(16,3,3,1)-MaxPool(2,2)-BN-Conv(32,3,3,1)-MaxPool(2,2)-BN -Conv(64,3,3,1)-MaxPool(2,2)-BN-FC(32)-Drop(0.5)-FC(2)			
RNN	LSTM(64)-FC(64)-Drop(0.5)-FC(2)			
CRNN-(LSTM)	Conv(16,3,3,1)-MaxPool(2,2)-BN-Conv(32,3,3,1)-MaxPool(4,4)-BN -Conv(64,3,3,1)-MaxPool(2,2)-BN-LSTM(64)-FC(32)-Drop(0.5)-FC(2)			
CRNN-(GRU)	Conv(16,3,3,1)-MaxPool(2,2)-BN-Conv(32,3,3,1)-MaxPool(4,4)-BN -Conv(64,3,3,1)-MaxPool(2,2)-BN-GRU(64)-FC(32)-Drop(0.5)-FC(2)			

Table 2 Other traditional network frameworks

表中 Conv (x, y, z, n)表示带有 x 个卷积核, 其中 y 和 z 代表二维卷积核的宽度和高度, n 为卷积过程 中的步长; MaxPool (x, y)代表最大池化层, x和 y为池化窗口的宽度和高度; BN表示批归一层; FC(x) 代表有 x 个节点的全连接层; LSTM(x)和 GRU(x)分别表示 LSTM 层和 GRU 层,其中 x 是输出空间的维 数; Drop(x)层表示放弃层,系数为 x。除最后一层采用 Softmax 激活外,其余激活函数都采用 PReLU。 采用稀疏分类交叉熵作为损失函数,优化器采用 Adam 优化器。

表 3 为上述其他传统网络框架与 P-CRNN 网络框架基于改进后 MLFCC 特征下的实验结果。从实验 结果中可以看出, P-CRNN 模型无论是在准确率、灵敏度、特异度还是 F1 分数都取得了最好的结果:准 确率达到 97.25%, 灵敏度为 97.84%, 特异度为 95.29%, F1 分数为 98.21%。

网络框架	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 (%)
CNN	92.34	85.14	94.36	95.06
RNN	91.18	80.38	94.49	94.26
CRNN(LSTM)	94.70	85.94	97.58	96.53
CRNN(GRU)	94.89	90.34	96.20	96.70
P-CRNN	97.25	95.29	97.84	98.21

Table 3. Comparison of classification results of different network frameworks 表 3. 不同网络框架分类结果比较

3.4.2. 不同特征提取方式性能对比

心音分类所采取的特征提取方式同样可以决定分类结果的准确性。为了更好地对比不同特征提取方 式对分类结果的性能影响,使用不同特征提取方式对本研究所提出的改进 MLFCC 进行定性对比,所采 取的特征提取方式为: MFCC、LFCC、GFCC (Gaussian Filtered Cepstral Coefficient,高斯滤波倒谱系数)。 将以上特征提取方式分别作为 P-CRNN 网络的输入特征,从而得到表 4 所示的分类结果。

Table 4. Comparison of classification results of different feature extraction methods 表 4. 不同特征提取方法分类结果比较

特征提取方法	Sensitivity (%)	Specificity (%)	Accuracy (%)
MFCC	89.06	96.21	94.59
LFCC	93.99	96.95	96.30
GFCC	91.44	93.33	92.96
MLFCC	95.29	97.84	97.25

由表 4 可以看出,使用 MLFCC 混合特征参数的分类效果明显优于其他单一特征。使用混合特征参数 MLFCC 的分类准确率为 97.25%,而使用单一特征 MFCC、LFCC、GFCC 的分类准确率则分别为 94.59%、 96.30%和 92.96%,相比之下,分别提高了 2.66%、0.95%和 4.29%。而在特异度方面,相比于其他三种单一特征,使用 MLFCC 混合特征参数分别提升了 6.03%、1.30%、3.85%,灵敏度分别提升了 1.63%、0.89%、 4.51%。

可见,采用混合特征参数 MLFCC 的特征提取方法综合了 MFCC 在处理语音信号上的成熟优势和 LFCC 在高频率分辨率和非语音处理上的优势,从而能够更好地表征心音信号。同时采用汉宁自卷积窗进 行加窗处理,使得在频域分析时更有效地减少旁瓣干扰,提高了心音分类的准确率。

4. 结束语

本文提出了一种基于 MLFCC 特征和并行卷积递归神经网络的自动心音分类算法。首先对心音信号 进行预加重、分帧、重叠切片、加窗,获得预处理后的心音信号;随后使用改进后的特征提取方式分别 提取信号的 MFCC 和 LFCC 特征及其各自的一阶差分系数,并且进行混合特征融合;最后利用所提出的 P-CRNN 网络结构对正常心音和异常心音进行分类识别。通过试验结果表明,相比于其他预处理方法, 采用固定窗口的重叠切片,可获取更多完整的数据信息,解决数据量不足、数据质量参差不齐等问题, 为后续分类处理获得了更好的鲁棒性和泛化能力;相比于其他特征提取方法,利用 MFCC 与 LFCC 相结 合的特征提取方法,综合了 MFCC 在处理语音信号上的成熟特征和 LFCC 在高频率分辨率和非语音处理 上的优势,从而能够更好地表征心音信号;相比于其他网络结构,P-CRNN 模型利用从 CRNN 中提取的 局部编码特征和 RNN 所捕获的长期依赖关系,在测试集中获得最优的分类结果。此分类算法综合性较其 他算法有明显的改进,为心音信号的自动分类研究提供了可行的思路。

基金项目

国家自然科学基金(No.61603255),上海市晨光计划(NO.18CG52)。

参考文献

- [1] WHO (2021) Cardiovascular Diseases (CVDs). https://www.who.int/mediacentre/factsheets/fs317/en/
- [2] Liu, C. and Murray, A. (2017) Applications of Complexity Analysis in Clinical Heart Failure. In: Barbieri, R., Scilingo, E. and Valenza, G., Eds., *Complexity and Nonlinearity in Cardiovascular Signals*, Springer, 301-325. <u>https://doi.org/10.1007/978-3-319-58709-7_11</u>
- [3] Lam, M.Z.C., Lee, T.J., Boey, P.Y., et al. (2005) Factors Influencing Cardiac Auscultation Proficiency in Physician Trainees. Singapore Medical Journal, 46, 11-14.
- [4] Zeinali, Y. and Niaki, S.T.A. (2022) Heart Sound Classification Using Signal Processing and Machine Learning Algorithms. *Machine Learning with Applications*, 7, Article ID: 100206. <u>https://doi.org/10.1016/j.mlwa.2021.100206</u>
- [5] Krishnan, P.T., Balasubramanian, P. and Umapathy, S. (2020) Automated Heart Sound Classification System from Unsegmented Phonocardiogram (PCG) Using Deep Neural Network. *Physical and Engineering Sciences in Medicine*, 43, 505-515. <u>https://doi.org/10.1007/s13246-020-00851-w</u>
- [6] Nguyen, M.T., Lin, W.W. and Huang, J.H. (2022) Heart Sound Classification Using Deep Learning Techniques Based on Log-Mel Spectrogram. *Circuits, Systems, and Signal Processing*, 42, 344-360. https://doi.org/10.1007/s00034-022-02124-1
- [7] Liu, C., Springer, D., Li, Q., Moody, B., Juan, R.A., Chorro, F.J., et al. (2016) An Open Access Database for the Evaluation of Heart Sound Algorithms. *Physiological Measurement*, **37**, 2181-2213. <u>https://doi.org/10.1088/0967-3334/37/12/2181</u>
- [8] Potes, C., Parvaneh, S., Rahman, A. and Conroy, B. (2016). Ensemble of Feature: Based and Deep Learning: Based Classifiers for Detection of Abnormal Heart Sounds. 2016 Computing in Cardiology Conference (CinC) Vancouver, 11-14 September 2016, 621-624. <u>https://doi.org/10.22489/cinc.2016.182-399</u>
- [9] Yaseen, Son, G. and Kwon, S. (2018) Classification of Heart Sound Signal Using Multiple Features. Applied Sciences,

8, Article 2344. https://doi.org/10.3390/app8122344

- [10] Mubarak, Q., Akram, M.U., Shaukat, A., Hussain, F., Khawaja, S.G. and Butt, W.H. (2018) Analysis of PCG Signals Using Quality Assessment and Homomorphic Filters for Localization and Classification of Heart Sounds. *Computer Methods and Programs in Biomedicine*, 164, 143-157. <u>https://doi.org/10.1016/j.cmpb.2018.07.006</u>
- [11] Li, F., Tang, H., Shang, S., *et al.* (2020) Classification of Heart Sounds Using Convolutional Neural Network. *Applied Sciences*, **10**, Article 3956.
- [12] Zhang, W., Han, J. and Deng, S. (2017) Heart Sound Classification Based on Scaled Spectrogram and Partial Least Squares Regression. *Biomedical Signal Processing and Control*, **32**, 20-28. <u>https://doi.org/10.1016/j.bspc.2016.10.004</u>
- [13] 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