

端到端KAN卷积在房颤心音识别中的应用

邓洪, 陈燕, 杨宏波, 赵峰, 蒋永卓, 郭涛, 王威廉

引用本文

邓洪, 陈燕, 杨宏波, 赵峰, 蒋永卓, 郭涛, 王威廉. 端到端KAN卷积在房颤心音识别中的应用[J]. 计算机科学, 2025, 52(9): 62-70.

DENG Hong, CHEN Yan, YANG Hongbo, ZHAO Feng, JIANG Yongzhuo, GUO Tao, WANG Weilian.

[Application of End-to-End Convolutional Kolmogorov-Arnold Networks in Atrial Fibrillation Heart Sound Recognition \[J\]. Computer Science, 2025, 52\(9\): 62-70.](#)

相似文献推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于端到端深度学习的数字语音源录音设备确认取证](#)

Source Recording Device Verification Forensics of Digital Speech Based on End-to-End DeepLearning

计算机科学, 2025, 52(6A): 240800028-7. <https://doi.org/10.11896/jsjcx.240800028>

[一种通过增强图像编码和非对称卷积网络的心音分类算法](#)

Heart Sound Classification Algorithm Using Enhanced Image Coding and Asymmetric Convolutional Networks

计算机科学, 2025, 52(6A): 240700195-8. <https://doi.org/10.11896/jsjcx.240700195>

[面向多样化数据清洗任务的证据集智能选择方法](#)

Intelligent Evidence Set Selection Method for Diverse Data Cleaning Tasks

计算机科学, 2024, 51(8): 124-132. <https://doi.org/10.11896/jsjcx.230900003>

[自编码器端到端通信系统后门攻击方法](#)

Backdoor Attack Method in Autoencoder End-to-End Communication System

计算机科学, 2024, 51(7): 413-421. <https://doi.org/10.11896/jsjcx.230400113>

[基于自编码的改进K-means光伏能源数据清洗方法](#)

Improved K-means Photovoltaic Energy Data Cleaning Method Based on Autoencoder

计算机科学, 2024, 51(6A): 230700070-5. <https://doi.org/10.11896/jsjcx.230700070>

端到端 KAN 卷积在房颤心音识别中的应用

邓洪¹ 陈燕² 杨宏波³ 赵峰² 蒋永卓¹ 郭涛³ 王威廉¹

¹ 云南大学信息学院 昆明 650504

² 中国人民解放军总医院第一医学中心 北京 100853

³ 云南省阜外心血管医院 昆明 650102

(dh2904799190@163.com)

摘要 房颤作为一种严重的心律失常疾病,及早诊断至关重要。房颤的传统检查方法是由心脏科医生借助心电图、超声心动图等设备做出诊断结论。为了缓解传统诊断方法检查成本高、过多依赖临床经验和便捷性不足等问题,创新性地应用 Kolmogorov-Arnold Network(KAN)来构建房颤心音分析模型。文中探索了 KAN 卷积在房颤心音识别中的应用,引入了具有灵活线性激活函数和优异参数效率的 KAN 卷积架构,提出了一种基于 KAN 卷积的端到端房颤心音识别模型。为提高信号的可用性,首先对心音信号进行预处理,包括心音分割、心音信号的质量评估和数据清洗;然后利用 KAN 的卷积层、池化层等自动学习;最后采用 KAN 卷积分类器进行识别研究。在特征提取阶段引入了 KAN 卷积的自注意力机制和焦点调制,以高效提取信号特征;在分类器阶段研究了 KAN 卷积的瓶颈结构和正则化手段,以提升模型的识别能力。该模型在中国人民解放军总医院第一医学中心的心音信号数据集上进行了正常和房颤的识别测试,准确率为 97.86%,灵敏度为 98.18%,特异度为 97.43%, F_{β} 值为 98.06%。实验结果表明,KAN 卷积模型在辅助诊断房颤信号上具有显著的优势。

关键词: 房颤;KAN 卷积;心音识别;端到端;数据清洗

中图分类号 TN912

Application of End-to-End Convolutional Kolmogorov-Arnold Networks in Atrial Fibrillation Heart Sound Recognition

DENG Hong¹, CHEN Yan², YANG Hongbo³, ZHAO Feng², JIANG Yongzhuo¹, GUO Tao³ and WANG Weilian¹

¹ School of Information Science and Engineering, Yunnan University, Kunming 650504, China

² The First Medical Center of Chinese PLA General Hospital, Beijing 100853, China

³ Fuwai Yunnan Hospital, Chinese Academy of Medical Sciences, Kunming 650102, China

Abstract Atrial fibrillation(AF) is a severe cardiac arrhythmia that requires early diagnosis. Traditional diagnostic methods typically involve cardiologists using electrocardiograms(ECG) and echocardiograms to make diagnostic conclusions. To address issues such as high costs, excessive reliance on clinical expertise, and limited accessibility, this study proposes an innovative application of Kolmogorov-Arnold Networks(KAN) in AF heart sound analysis. This study explores the application of convolutional KAN (CKAN) in AF heart sound recognition, proposing an end-to-end AF identification framework based on CKAN architecture, which incorporates flexible linear activation functions and exhibits superior parameter efficiency. To enhance the usability of heart sound signals, the methodology first applies preprocessing, including signal segmentation, quality assessment, and data cleansing. Subsequently, the model autonomously learns discriminative features through KAN-based convolutional and pooling layers. Finally, a CKAN-based classifier is employed for classification. During the feature extraction phase, self-attention mechanisms and focus modulation are incorporated into CKAN to efficiently extract signal features. In the classification phase, CKAN's bottleneck structure and regularization techniques are explored to improve the model's recognition performance. The proposed model is evaluated on a heart sound dataset from The First Medical Center of Chinese PLA General Hospital, achieving an accuracy of 97.86%, sensitivity of 98.18%, specificity of 97.43%, and an F_{β} score of 98.06%. The results indicate that the CKAN model provides significant advantages in aiding the diagnosis of AF from heart sound signals.

Keywords Atrial fibrillation, Convolutional KAN, Heart sound recognition, End-to-End, Data cleansing

到稿日期:2025-01-15 返修日期:2025-04-13

基金项目:国家自然科学基金(82172185,81960067);2018年云南省重大科技专项(2018ZF017)

This work was supported by the National Natural Science Foundation of China(82172185,81960067) and Major Science and Technology Projects of Yunnan Province in 2018(2018ZF017).

通信作者:王威廉(wlwang_47@126.com)

1 引言

心房颤动 (Atrial Fibrillation, AF) 是临床最常见的持续性心律失常 (Persistent Arrhythmia, PA), 简称房颤, 其不仅影响患者生活质量, 还会增加心力衰竭、卒中等并发症风险^[1]。过去 20 年, 我国房颤相关疾病的发病率呈上升趋势。房颤与血栓形成和栓塞事件的风险增加密切相关^[2], 其临床治愈的关键在于早期的诊断和治疗。当前, 房颤的临床诊断通常依赖于心电图 (Electro Cardio Gram, ECG)、超声心动图和动态心电图等^[3-5]。然而, 这些方法存在设备携带不便、难以在大规模筛查中应用等局限性。相比之下, 通过记录心音得到的心音图 (Phono Cardio Gram, PCG) 作为一种实时性高、便捷且非侵入性的辅助诊断工具^[6], 极大方便了患者的检查过程, 能够在大规模筛查中得到有效应用。随着深度学习和机器学习技术的迅猛发展, 基于 PCG 的辅助诊断逐渐成为心音研究的热点^[7], 其能够自动提取心音信号中的关键特征, 并进行精准的识别, 为房颤的辅助诊断提供更简便、高效的解决方案。

应用深度学习对心音进行分析研究, 通常包括预处理、特征提取、模型训练评估 3 个步骤^[8]。

预处理包括对心音信号进行降噪、分割等。文献[9]提出一种结合卷积神经网络与底隐马尔可夫的心音分割方法, 在公开数据集上检测 S1 和 S2 的平均灵敏度为 93.9%, 平均阳性预测值为 94%; 文献[10]通过对 S1 和 S2 进行精准定位, 根据基于心音信号统计特征的 3 个标准在时频域中评估了信号的质量, 以消除噪声并对数据进行清洗。

在特征提取中, 文献[11]使用时频融合特征, 在先天性多分类中取得了 92.21% 的准确率; 文献[12]提出了一种基于 MFCC 的新型改进特征提取, 在心音数据集上使用 CRNN 分类器, 取得了 98% 的准确率。相较于将传统的提取复杂的二维谱图作为输入特征的模型, 端到端的分析处理方法直接输入原始心音信号, 通过深度神经网络自动学习, 降低了心音处理的复杂性以及特征提取中人为参与导致的偏差, 增强了心音分类模型的泛化能力与识别性能^[13]。文献[14]提出了一种基于 CLSTM 的新型端到端的方法, 该方法不仅简化了心

音注释任务, 还在数据集上表现出了卓越的性能; 文献[15]提出了一种基于 MDN-MARNN 的新型端到端的心音分类方法, 并使用 SHAP 算法增强模型的可解释性。

分类器的选择通常取决于所提取的特征, 文献[16]从 PCG 中提取了 MFCC, Delta MFCC, FBANK 等多个特征, 然后使用众多机器学习分类器, 如朴素贝叶斯 (NB)、支持向量机 (SVM)、决策树 (DT)、随机森林 (RF)、K 最近邻 (KNN) 和人工神经网络 (ANN), 最高得到了 99.2% 的准确率。

传统的多层感知器 (Multi-Layer Perceptron, MLP) 虽然有效, 但依赖于固定的激活函数和线性变化。Kolmogorov-Arnold Network (KAN)^[17] 利用 Kolmogorov-Arnold 定理将样条函数融入架构中, 具有出色的非线性拟合能力和灵活性。KAN 结合了 MLP 的通用性结构优势和样条函数的高效拟合特性, 既能学习特征, 又能精准优化学习到的特征^[18]。为此, 本文建立了一种基于 KAN 卷积的端到端房颤心音识别模型, 探索了原始 KAN 的各种变化形式, 结合特征增强与模型优化策略提出了 KAN 卷积的高效设计, 进行了充分的实验以验证 KAN 卷积在心音识别领域的应用。综上, 本文的主要创新点如下:

1) 提出了一种基于 KAN 卷积的端到端房颤心音识别模型, 用于辅助房颤诊断。

2) 对原始心音数据进行了质量评估和数据清洗, 以提高心音分析的有效性和鲁棒性。

3) 优化了 KAN 卷积的瓶颈结构, 通过降维-基函数计算-升维的结构, 显著降低了模型复杂度; 在瓶颈架构基础上, 融合 L1/L2 正则化与 Dropout 的正则化手段, 提升了模型对房颤心音特征的判别能力。

4) 结合自注意力机制与焦点调制, 优化了 KAN 卷积特征交互, 用 KAN 卷积层代替传统 QKV 投影, 增强了模型的特征选择和表达能力。

2 方法和模型

2.1 算法框架

本文提出的端到端 KAN 卷积房颤心音识别算法框架如图 1 所示。具体的算法流程如下。

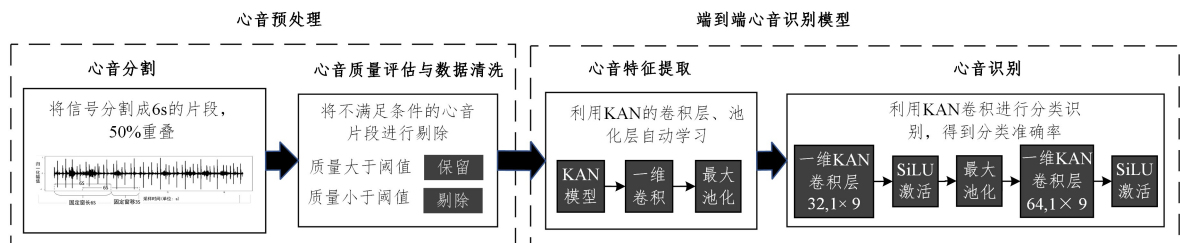


图1 端到端KAN卷积心音识别算法框架

Fig. 1 Framework of end-to-end convolutional KAN heart sound recognition algorithm

1) 心音分割: 为了适应端到端神经网络的自动特征学习, 采用滑动窗口分割策略, 将信号切割成固定长度的心音片段, 以便模型能够学习局部时频域特征。

2) 心音质量评估与数据清洗: 心音信号在采集过程中易受到噪声干扰。噪声与心音信号在时频域上均有重叠, 直接输入低质量信号可能会影响模型的稳定性, 因此选择信号包

络高幅值点数占比和自相关系数的过零数作为信号数据清洗标准。

3) 心音特征提取: 端到端的方法不依赖于人工特征提取, 其利用 KAN 卷积层出色的函数逼近能力, 结合最大池化策略, 以增强特征表示的平滑性。

4) 心音识别: KAN 卷积分类器能够基于所学习到的特

征,在训练时结合交叉损失函数和 AdamW 优化器,在心音识别任务上实现了高准确率。

2.2 心音预处理

2.2.1 心音分割

心音的心动周期包括 S1、收缩期、S2、舒张期,一般是

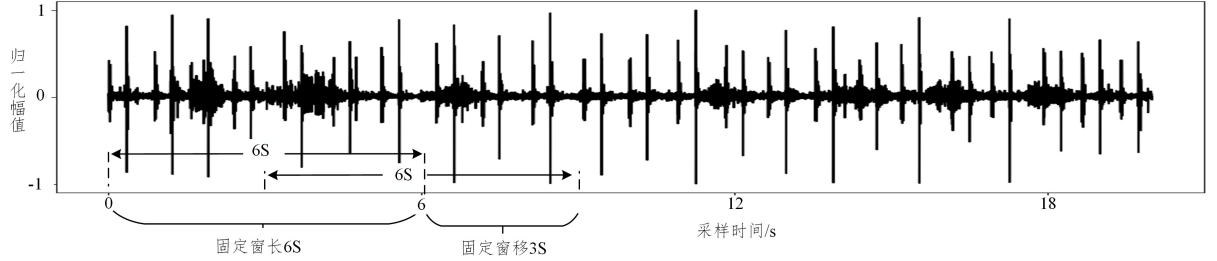


图2 滑窗分割示意图

Fig. 2 Schematic diagram of sliding window splitting

选择 6 s 的信号长度,以涵盖 6~7 个完整的心动周期^[23],有助于端到端网络自适应地提取特征。通过滑动矩形窗口将不同样本的心音信号分割成标签一致的等长片段,滑动窗口长度设置为 6 s,窗移(重叠)设置为 3 s。其中,重叠是为了满足连续性的需求。

2.2.2 心音信号质量评估与数据清洗

受限于采集环境和设备内部杂音,部分心音信号混杂着严重的环境噪声且存在严重的畸变。环境噪声与心音信号在时域与频域上均有重叠,常规方法并不能剔除噪声和复原信号^[24-25],因此,对心音进行数据清洗是不可或缺的一环。通过特定的算法或规则识别并清洗低质量或无关信号,以提高后续分析和模型训练的准确性与鲁棒性。

本文对文献[26]提出的两个心音统计特征进行了改进,并将其用于心音信号数据清洗。这两个统计特征分别是:心音信号波峰插值包络和包络自相关系数。

1)心音信号基于 3 次多项式的波峰插值包络,其定义如式(1)和式(2)所示。通过滑动窗口找出每个窗口内信号的最高峰值,并在峰值之间通过插值的方式构建幅值包络。这种方法无需频域转换,相较于希尔伯特变换求解包络,计算复杂度更小、包络更顺滑。

$$envelope = a * (peak(i))^3 + b * (peak(i+1))^2 + c * peak(i+2) + d \quad (1)$$

$$peak(i) = \max(window(i) * signal) \quad (2)$$

其中,signal 是输入的心音信号;peak(i)是第 i 个窗口的最高峰值;window(i)是第 i 个窗高为 1 的滑动矩形窗;a,b,c,d 是常数;envelop 是输入心音信号的波峰插值包络。

2)心音信号波峰插值包络的自相关系数用于衡量包络 x(t)与其自身滞后包络的相关性,如式(3)所示:

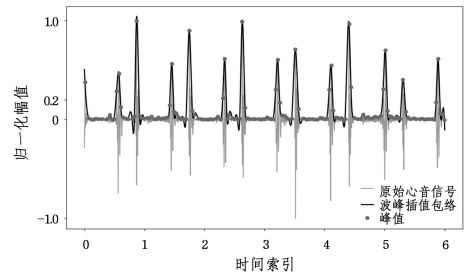
$$\rho(k) = \frac{Cov(x(t), x(t-k))}{\sigma(t)\sigma(t-k)} \quad (3)$$

其中,x(t)是心音信号的波峰插值包络, $\sigma(t)$ 是 x(t)的标准差, $\rho(k)$ 是波峰插值包络的自相关系数。

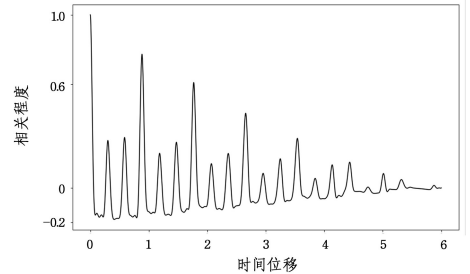
心音信号的波峰插值包络和包络的自相关系数示意图如图 3 所示。原始心音信号指直接从心音采集设备获取的未经过滤波或处理的时序信号,通常包含心音的主要成分(S1,S2)以

0.6~1.0 s。房颤的病理性杂音主要集中在收缩期,信号质量和个体差异使分割算法变得更具挑战性。为了提供一致且可靠的数据进行准确的诊断和评估,在临床应用中需要观察多个心动周期^[19-20]。本文采用简单的滑窗分割方法,无需复杂的分割算法^[21-22],滑窗分割示意图如图 2 所示。

及可能的噪声干扰;峰值是在设定的滑动窗口内,局部幅值最高的信号点;波峰插值包络是采用 3 次多项式插值的方法,对心音信号的局部峰值进行拟合,以构建平滑的幅值包络曲线。



(a)心音信号的波峰插值包络



(b)心音信号峰值包络的自相关系数

图3 心音信号波峰插值包络和包络的自相关系数

Fig. 3 Peak interpolation envelope of heart sound signal and autocorrelation coefficient of envelope

包含严重噪声的信号在包络归一化后,幅值大于 0.2 的样点数占整段信号样点数的比例过高,而比例过低的信号缺失过多的有效信息。因此,清洗条件之一是将该比例限制在 [0.1, 0.3]。此外,高质量心音具备准周期特性,而噪声则缺乏这一特性,致两类信号的自相关系数差异显著。因此,另一清洗条件是相关系数的极大峰值在 0.6 以上。

2.3 KAN 架构及其变体

Kolmogorov-Arnold 定理在神经网络中的应用标志着一项重要的理论整合,增强了神经网络的表达能力和效率。该定理提供了一种将任何多元连续函数表示为单变量函数和加法的组合的方法^[27]。该定理如式(4)所示:

$$f(x_1, \dots, x_n) = \sum_{q=1}^{2q+1} \Phi_q \left(\sum_{p=1}^n \varphi_{q,p}(x_p) \right) \quad (4)$$

其中, $\varphi_{q,p}$ 是映射每个输入变量 x_p 的单变量函数, Φ_q 是连续函数。这使 KAN 能通过组合更简单的单变量函数来模拟更复杂的高维数据。

KAN 的网络架构示意图如图 4 所示。KAN 拥有和 MLP 一样的全连接的结构, MLP 将激活函数放置在节点神经元上, KAN 则将可学习的激活函数放置在边缘权重上。KAN 能够提供更大的灵活性, 并通过减少参数复杂性来有效提升模型的训练效率和泛化能力, 从而解决了 MLP 中参数冗余和计算负担过重的问题。

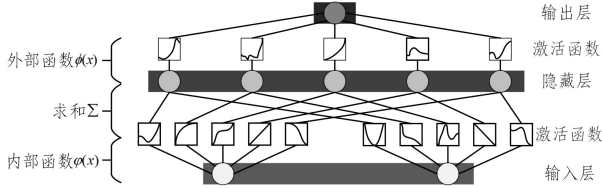


图 4 KAN 的网络架构

Fig. 4 KAN's network architecture

本文应用 Gram 多项式在函数逼近理论中的正交完备性原理, 针对心音信号的非平稳特性和病理特征敏感性需求展开深入探索, 研究用 Gram 多项式替代传统 B-spline 的方法, 提出一种改进型的 KAN 架构——Gram-KAN。Gram 多项式的定义如式(5)所示:

$$\int_a^b G_n^N(x) G_m^N(x) \omega(x) dx = 0, n \neq m \quad (5)$$

其中, $\omega(x)$ 是权函数, 通常是定义在 $[a, b]$ 上的正定函数; $G_n^N(x)$ 是 x 的 n 次多项式, x 的取值范围是 $[0, N-1]$, 可以通过正交性定义。

Gram-KAN 适用于处理一维心音数据集, 为神经网络的数据处理提供了更新颖的选择, 有助于提升模型的特征表达能力和识别性能。

此外, 还探索了使用修正线性单元 (Rectified Linear Unit, ReLU) 和逐点乘法的 KAN 变体 (ReLU-KAN), 用 $R_i(x)$ 多项式替代原始的 B-spline 函数。其中, $R_i(x)$ 的定义如式(6)和式(7)所示:

$$R_i(x) = \frac{16[\text{ReLU}(e_i - x) \times \text{ReLU}(x - s_i)]^2}{(e_i - s_i)} \quad (6)$$

$$\text{ReLU}(x) = \max(0, x) \quad (7)$$

其中, s_i, e_i 是索引 i 决定的常数, 且 $s_i < e_i$; $R_i(x)$ 是定义在区间 $[s_i, e_i]$ 非零的钟形函数。

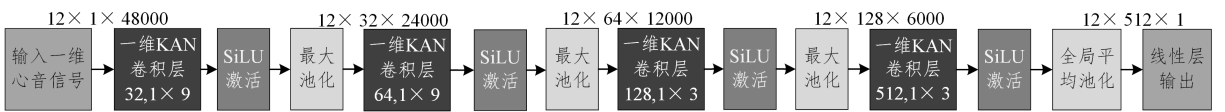


图 5 4 层 KAN 卷积的网络架构

Fig. 5 Network architecture of 4-layer convolutional KAN

输入 6s 的一维心音信号片段 ($12 \times 1 \times 48000$), 首先, 经过 32 个通道数 KAN 卷积层 (1×9) 提取局部特征, 结合 SiLU 激活与最大池化降维至 ($12 \times 32 \times 24000$); 其次, 第二层 64 通道数的 KAN 卷积 (1×9) 进一步提取时序模式, 经 SiLU 激活与最大池化后变为 $12 \times 64 \times 12000$; 再次, 第三层 128 通道数

ReLU-KAN 解决了传统 KAN 中基函数计算的复杂性问题, 能够在提高计算效率的同时保持稳定的训练和卓越的拟合能力; 保留了 KAN 的“灾难性遗忘避免”特性, 其非线性特征的有效建模和稀疏激活特性在处理心音信号时展现了出色的能力。

2.4 KAN 卷积

KAN 的高维非线性函数逼近能力能提升模型的表达能力和泛化性, 卷积层能高效提取局部特征并捕捉多尺度模式, 将 KAN 与卷积层相结合能有效兼顾两者的优势。本文探讨了卷积层在 KAN 中的拓展应用——KAN 卷积, 即 Convolutional Kolmogorov-Arnold Networks (CKAN)。KAN 卷积^[28]可以表示为一组单变量非线性函数。对于内核大小为 k 的一维 KAN 卷积, 可以将其定义为式(8):

$$x_i = \sum_{d=1}^c \sum_{a=0}^{k-1} \varphi_{a,d}(y_{d,i+a}), i \in [1, n-k+1] \quad (8)$$

其中, $\varphi_{a,d}$ 是一个单变量可学习函数, 有自己的学习函数 a 和 d ; c 是通道数; n 是心音的采样点数。

在原始的 KAN 论文中, 作者建议使用式(9)~式(11)所示的函数:

$$\varphi = w_b \cdot b(x) + \tilde{\varphi}(x) \quad (9)$$

$$\tilde{\varphi}(x) = w_s \cdot \text{Spline}(x) \quad (10)$$

$$b(x) = \text{SiLU}(x) = x / (1 + e^{-x}) \quad (11)$$

其中, w_b 是函数 $b(x)$ 的权重, w_s 是样条函数的权重, φ 是激活函数, $\tilde{\varphi}(x)$ 是基函数。

与原始 KAN 相似, 可以选择 B-spline 之外的函数作为基函数 $\tilde{\varphi}(x)$: Gram 多项式和 $R_i(x)$ 多项式, 以构建 KAN 卷积模型。在 Gram 多项式作为样条的情况下, KAN 卷积中的基函数 $\tilde{\varphi}(x)$ 应重新定义, 如式(12)所示, 而在 $R_i(x)$ 多项式中应重新定义, 如式(13)所示:

$$\tilde{\varphi}_g(x) = \sum_{i=0}^{N+1} w_i \cdot G_i^N(x) \quad (12)$$

$$\tilde{\varphi}_r(x) = \sum_{i=0}^{N+1} w_i \cdot R_i(x) \quad (13)$$

重新定义基函数 $\tilde{\varphi}(x)$ 有利于对预训练模型进行参数微调, 有效减少了可训练的参数数量。

KAN 卷积网络架构选择具有 4 个或 8 个的一维 KAN 卷积层来识别心音信号。为了在端到端模型中使网络更有效地学习和提取特征, 提升模型性能和稳定性, 在卷积层之间添加最大池化和 SiLU 激活函数。8 层是在 4 层的基础上增加了 4 个相同的内核为 1×3 的 KAN 卷积层、最大池化和 SiLU 激活函数的组合。4 层的网络架构如图 5 所示。

的 KAN 卷积 (1×3) 加强短时局部特征学习, 结合 SiLU 激活与最大池化降维至 $12 \times 128 \times 12000$; 然后, 第四层 512 通道数的 KAN 卷积 (1×3) 提取全局信息, 并通过 SiLU 激活和全局平均池化压缩至 $12 \times 512 \times 1$; 最后, 在线性层输出识别结果。该网络利用 KAN 的函数逼近能力, 结合多层 KAN 卷

积、SiLU 激活、最大池化与全局平均池化,在降低计算复杂度的同时增强特征表达能力,实现了高效精准的心音识别。

2.5 KAN 卷积的瓶颈结构

KAN 卷积的主要问题在于模型的样条函数,无论选择哪种样条函数,均会给模型引入大量参数,增加对计算资源的需求以及过拟合的可能性。

在 KAN 中引入瓶颈(Bottle)卷积网络可以有效解决上述问题^[29],特别是在处理复杂心音信号时,瓶颈卷积网络可以显著提高计算效率并减少模型的计算负担。此外,KAN 卷积的瓶颈结构可以适应更深层次的网络模型,增强其在心音识别领域的泛化性和实时性。KAN 卷积的瓶颈结构如图 6 所示。

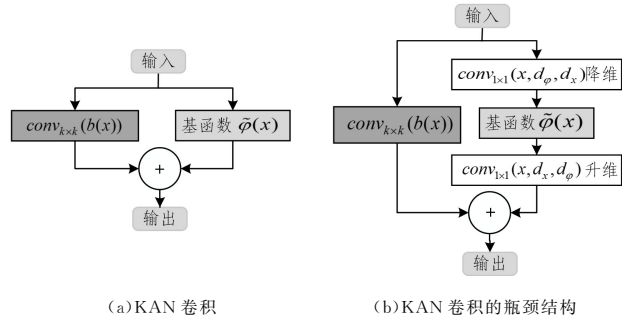


图 6 KAN 卷积和 KAN 卷积的瓶颈结构

Fig. 6 KAN convolution and bottleneck structure of KAN convolution

瓶颈卷积网络在基函数应用到输入之前,使用 1×1 卷积核降维,随后使用 1×1 卷积核升维,通常配合批归一化和激活函数来增加网络的非线性和稳定性。可以将其直观看成一个单层编码器,有助于在输入之前提取到高质量特征,再由单层解码器将其解码。区分 KAN 卷积有无瓶颈网络结构的重点在于右侧的 KAN 包含编码器-解码器卷积层。

为了平衡模型的复杂度和性能,有效控制模型的计算复杂度、特征压缩以及网络的表达能力,定义瓶颈 KAN 中的衰减参数如式(14)所示:

$$attenuation = \frac{d_x}{d_\varphi} \quad (14)$$

其中, d_x 是一维卷积的输入通道数; d_φ 是输出的通道数; $attenuation$ 是 d_x 和 d_φ 的比值,对于控制模型的效率和扩展性至关重要。

表 1 列出了各种 KAN 卷积模型的可训练参数的量。

表 1 不同 KAN 模型的可训练卷积参数量

Table 1 Number of trainable convolutional parameters for different

KAN models		
Model	128 filters	256 filters
KAN	0.442×10^6	1.769×10^6
Gram-KAN	0.245×10^6	0.983×10^6
ReLU-KAN	0.444×10^6	1.773×10^6
Bottle Gram-KAN, $attenuation=2$	0.114×10^6	0.458×10^6
Bottle Gram-KAN, $attenuation=4$	0.069×10^6	0.278×10^6
Bottle Gram-KAN, $attenuation=8$	0.056×10^6	0.225×10^6
Bottle ReLU-KAN, $attenuation=2$	0.164×10^6	0.657×10^6
Bottle ReLU-KAN, $attenuation=4$	0.082×10^6	0.328×10^6
Bottle ReLU-KAN, $attenuation=8$	0.059×10^6	0.238×10^6

可以看出:Gram-KAN 相较于 KAN 减少了约 44.5% 的参数,具有更高的参数效率;瓶颈架构能够将参数量压缩至原始模型的 20% 左右,同时保留计算能力;Gram-KAN 在瓶颈结构中比 ReLU-KAN 更高效,适合参数敏感型任务;选择合适的 $attenuation$ 参数(如 4 或 8)可以在计算效率和模型性能之间取得良好平衡。

$attenuation$ 参数的设置能够通过降低计算开销、减少特征冗余以及促进模型稀疏性等^[30],为实施混合专家(Mixture of Experts, MoE)提供支持。在处理大规模复杂一维心音时,为了在保持计算效率的同时提高模型的判别能力和泛化能力,在 KAN 卷积瓶颈架构的编码器-解码器之间引入稀疏混合专家模型^[31],每个专家都是一组 $\tilde{\varphi}_i(x)$ 的单变量函数,如图 7 所示。

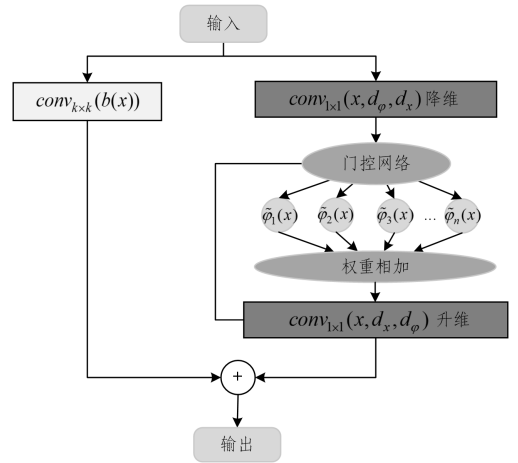


图 7 KAN 卷积瓶颈架构的稀疏混合专家模型

Fig. 7 Sparse hybrid expert model of KAN convolutional bottleneck architecture

首先,KAN 卷积瓶颈架构的稀疏混合专家模型通过 KAN 卷积的瓶颈架构对输入心音信号进行通道压缩与动态基函数激活,生成高判别性的低维表征;然后,由轻量级门控网络根据特征显著性自动选择多个专家子网络(每个专家由可解释的单变量 KAN 函数 $\tilde{\varphi}_i(x)$ 堆叠构成),实现条件计算与参数效率的平衡;最后,通过卷积核进行升维重构,恢复高维表征以匹配原始心音信号的特征空间。这种设计利用 KAN 的局部激活稀疏性抑制高频噪声干扰,通过专家间的竞争性协作增强模型对病理特征的多尺度建模能力。

2.6 KAN 卷积的自注意力机制和焦点调制

传统自注意力机制的构建^[32]:令 x 是输入信号; W_q, W_k, W_v 是可学习的权重矩阵; Q 向量用于获取信息; K 向量用于标记数据重要部分; V 向量包含实际信息。使用 SoftMax 函数计算关系权重之后,与 V 向量相乘,得到单头注意力机制输出 Z 。具体如式(15)、式(16)所示:

$$Q = W_q x, K = W_k x, V = W_v x \quad (15)$$

$$Z = \text{SoftMax}(\varphi(Q) \varphi(K)^T) V \quad (16)$$

其中, $\varphi(\cdot)$ 表示核映射函数,可捕捉非线性特征。

在传统的自注意力机制中, Q (Query), K (Key) 和 V (Value) 投影通常通过线性层来实现,这可能导致内存消耗过大和参数冗余。文献^[33]提出了一种创新的设计方法,该方法

使用卷积层替代线性投影,旨在降低内存占用并提高计算效率。本文在此基础上进行进一步改进,提出用KAN卷积层代替传统的卷积层,以构建自注意力KAN卷积层。尽管自注意力机制能够在模型中捕捉到全局依赖关系,但其庞大的内存需求和大量参数的引用也会影响计算效率和性能。为了解决这一问题,本文在自注意力KAN卷积层中引入了瓶颈卷积网络,将自注意力操作置于两个 1×1 瓶颈卷积层之间。这一设计不仅减少了网络中的参数数量,还有效降低了计算复杂度,从而提升了模型在处理大规模一维心音数据时的效率。

文献[34]提出的焦点调制层作为自注意力层的无缝

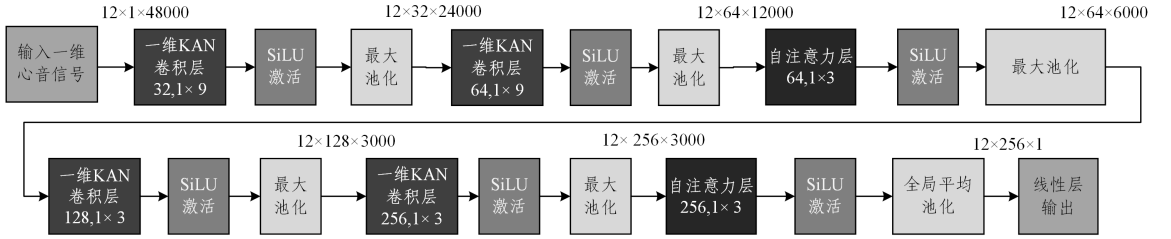


图8 使用注意力模块的KAN卷积

Fig. 8 Convolutional KAN using attention module

输入的一维心音信号首先从两个一维KAN卷积层出发,第一层有32个通道数,第二层有64个通道数,均使用 1×9 的内核;第三层是具有128个通道的注意力层或焦点调制层;下部分由两个KAN卷积层组成,第四层有128个通道数,第五层有256个通道数;第六层是有256个通道数的注意力层或焦点调制层。自第二层之后,均使用 1×3 的内核;最后经过全局平均池化和线性层输出识别结果。

2.7 KAN卷积的正则化

本文引入L1/L2正则化和Dropout正则化技术,以进一步增强KAN模型的泛化能力和稳定性。通过这些正则化手段,有效防止了模型的过拟合,提升了其在不同噪声环境下的表现。

L1正则化在训练过程中通过惩罚权重的稀疏性,自动筛选出最重要的特征,去除冗余特征。L2正则化通过对权重平方和的惩罚,进一步控制了模型的复杂度,从而避免模型过拟合[35]。

此外,Dropout正则化技术的引入也是提升KAN模型鲁棒性的关键。针对卷积层,采用了Inner-Dropout,旨在通过随机丢弃部分卷积核来防止模型对局部特征的过拟合。在全连接层中,引入Out-Dropout能够有效减小神经元之间的过度依赖,增强模型在高层次特征组合上的泛化能力。

3 实验结果

3.1 实验数据说明

本文的实验数据由中国人民解放军总医院第一医学中心临床采集,包括房颤样本129例,正常样本59例,共计188例。每条信号十几秒到十几分钟不等,采样率为8000 Hz,以wav格式保存。库中的标签均经过超声心动图和心血管专家的确认。为了应对数据分布不平衡的问题,采用SVM-SMOTE对正常样本进行数据增强,有效扩充了正常样本的数量,为后续识别模型的训练提供了更均衡的数据集基础。

替代,具有较强的可解释性和特征权重调整功能,能够使模型在特定区域更加敏感。针对焦点调制层的实现,本文提出了焦点调制KAN卷积网络,使用KAN卷积层替代焦点调制的原始结构,进一步提升了特征选择与表达能力。同时,为了减小过多参数对焦点调整精度的影响,本文在焦点调制层中也引入了瓶颈结构,将焦点调制层置于两个 1×1 的瓶颈卷积层之间,进而优化模型的计算效率和性能。

KAN卷积的自注意力机制和焦点调制模块基于4个卷积层,并辅以2个自注意力层或焦点调制层。使用注意力模块的KAN卷积如图8所示。

考虑到数据样本数量有限,本文仅将数据集划分为训练集和测试集,以确保模型能够在可用数据上进行有效训练和评估。实验数据按8:2的比例划分为训练集和测试集。此外,在本实验中,来自同一患者的数据不会被分到不同子集中,确保了实验的有效性和结果的可靠性[36]。

课题研究通过了云南大学人体研究材料伦理委员会和中国人民解放军总医院第一医学中心伦理委员会的审查与批准。

3.2 评价指标

采用4个广义评价指标评估算法的有效性,分别是准确率(Accuracy, Acc)、灵敏度(Sensitivity, Se)、特异度(Specificity, Sp)和 F_β 值,其中 Acc , Se , Sp 的计算式如式(17)~式(19)所示:

$$Se = \frac{TP}{TP + FN} \times 100\% \quad (17)$$

$$Sp = \frac{TN}{TN + FP} \times 100\% \quad (18)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (19)$$

其中, TP 表示真阳性结果的数量, TN 表示真阴性结果的数量, FP 表示假阳性结果的数量, FN 表示假阴性结果的数量。

此外,采用 F_β 来评判并识别模型的综合指标。 F_β 是精确率(Precision, P)和召回率(Recall, R)之间的调和平均数,在临床上比假阳性对假阴性的破坏性更大,通过检查可以排除假阳性样本,而假阴性样本则会延迟治疗时间,根据文献[37]将 β 调整为1.2,以增加敏感性的重要性。定义如式(20)~式(22)所示:

$$R = Se \quad (20)$$

$$P = \frac{TP}{TP + FP} \times 100\% \quad (21)$$

$$F_\beta = (1 + \beta^2) \frac{PR}{\beta^2 P + R} \times 100\% \quad (22)$$

3.3 实验结果对比

3.3.1 多层 KAN 卷积实验

KAN 卷积架构选择具有 4 层或 8 层的一维 KAN 卷积层进行对比实验,所有模型的 KAN 卷积通道数均相同:

1) 4 layers: 32, 64, 128, 512。

2) 8 layers: 16, 32, 64, 128, 256, 256, 512, 512。

所有实验均在 NVIDIA RTX 2080 上进行,具备相同的训练参数。各个 KAN 卷积模型的性能如表 2 所列,包括各个 KAN 卷积模型的对房颤和正常心音信号识别的准确率、灵敏度、特异度、 F_{β} 值和参数量。

表 2 不同层多类型 KAN 卷积在实验数据上的结果

Table 2 Results of different layers of multi-type convolutional KAN on experimental data

Model	Acc/%	Se/%	Sp/%	F_{β} /%	Params
KAN Conv, 4 layers	88.31	90.59	86.85	90.38	8.640×10^6
Gram-KAN Conv, 4 layers	88.65	91.84	81.94	91.64	3.840×10^6
Bottle Gram-KAN, 4 layers	88.57	93.63	77.78	92.03	1.930×10^6
MOE Bottle Gram-KAN, 4 layers	88.32	90.30	85.66	89.52	2.246×10^7
ReLU-KAN, 4 layers	93.21	95.29	89.05	94.94	8.200×10^6
Bottle ReLU-KAN, 4 layers	90.00	94.61	79.62	93.15	2.980×10^6
KAN Conv, 8 layers	94.95	96.85	93.94	95.99	5.426×10^7
Gram-KAN Conv, 8 layers	96.07	96.67	94.90	96.89	2.417×10^7
Bottle Gram-KAN, 8 layers	96.22	97.58	94.49	96.73	8.780×10^6
MOE Bottle Gram-KAN, 8 layers	90.79	89.70	91.91	91.05	5.031×10^7
ReLU-KAN, 8 layers	95.71	95.08	96.99	96.47	5.432×10^7
Bottle ReLU-KAN, 8 layers	96.19	97.15	95.97	96.51	1.132×10^7

在 4 层和 8 层网络中, Gram-KAN 和 ReLU-KAN 的准确率均优于原始 KAN, 说明多项式扩展能够增强 KAN 的特征提取能力; 在使用瓶颈卷积结构后, 参数量降低的同时取得了更高的准确率, 8 层的 Bottle Gram-KAN 仅用 8.78×10^6 的参数量就达到 96.22% 的准确率, 优于参数量更大的 Gram-KAN, 体现了瓶颈结构的高效性; 增加卷积层深度有助于提高模型的识别准确率, Gram-KAN 从 4 层 (88.65%) 增加到 8 层 (96.07%), 准确率提升了 7.42 个百分点, 说明更深的网络能够学习到更丰富的特征, 提高了模型的识别能力。

3.3.2 正则化研究结果

为了缩短计算时间, 提高网络模型正则化过程中对抗过

拟合的能力, 根据表 2 的实验结果, 选择具有 8 个 Bottle Gram-KAN 的卷积层进行研究, 探索了两种不同宽度的 KAN 卷积:

1) Slim: 16, 32, 64, 128, 256, 256, 512, 512。

2) Wide: 32, 64, 128, 256, 512, 512, 1024, 1024。

网络使用 8 个 KAN 卷积层, 通过全局平均池化压缩特征维度, 经线性输出层得到最后的识别结果。实验过程中, 除了正则化的参数外, 其余网络参数均相同, 且未在全局池化使用 Dropout。此外, 在 KAN 卷积的全连接层使用 L1/L2 正则化, 在 KAN 卷积的卷积层、全连接层使用 Dropout 正则化。不同正则化手段在 Slim 层和 Wide 层的表现如表 3 所列。

表 3 不同正则化在实验数据上的实验结果

Table 3 Experimental results for different regularizations on experimental data

Regularization	Slim				Wide			
	Acc	Se	Sp	F_{β}	Acc	Se	Sp	F_{β}
none	96.22	97.58	94.49	96.73	96.35	98.20	92.55	97.41
L1 Activation, 1×10^{-9}	96.55	98.48	94.12	97.16	94.24	96.64	91.27	95.09
L1 Activation, 1×10^{-8}	95.89	97.58	93.75	96.50	96.71	97.58	95.59	97.09
L1 Activation, 1×10^{-7}	94.90	95.45	94.12	95.34	97.37	99.09	95.22	97.88
L1 Activation, 1×10^{-6}	95.56	96.36	94.49	96.01	97.53	97.58	97.43	97.70
L2 Activation, 1×10^{-9}	96.37	97.88	96.59	96.64	97.37	98.18	96.32	97.70
L2 Activation, 1×10^{-8}	96.88	98.18	95.22	97.34	96.88	98.79	94.49	97.46
L2 Activation, 1×10^{-7}	96.38	97.27	95.22	96.79	97.04	98.18	95.59	97.46
L2 Activation, 1×10^{-6}	95.72	96.36	94.85	96.12	96.88	96.97	96.69	97.09
Inner-Dropout, 0.10	97.20	97.58	96.69	97.45	97.53	98.18	96.69	97.82
Inner-Dropout, 0.20	97.37	98.48	95.96	97.76	96.05	96.67	95.22	96.43
Inner-Dropout, 0.30	96.88	97.46	94.25	97.02	97.53	98.52	94.49	98.17
Out-Dropout, 0.10	96.88	96.97	96.69	97.09	97.86	98.18	97.43	98.06
Out-Dropout, 0.20	97.37	98.18	96.32	97.70	96.71	97.27	95.96	97.03
Out-Dropout, 0.30	95.89	98.18	93.01	96.62	97.53	98.18	96.69	97.82

与 L1/L2 正则化相比, Dropout 正则化在对抗模型过拟合方面表现最优秀, 尤其是在 Slim 层中, Inner-Dropout (0.10) 使得模型的 Acc, Se, Sp 分别提高到 97.20%, 97.58%, 96.69%, 明显优于 L1/L2 正则化方法; 在更宽的模型 (Wide) 中, 正则化的影响更为显著, 尤其是在 L1 和 L2 正则化中, 正

则化力度较强时会对准确率产生较大影响。L1 Activation (1×10^{-7}) 在宽模型中仅为 97.37%, 而 Se 和 Sp 的平衡表现较差, 与无正则化的 96.35% 相比, 未能有效提升性能。对比表 2 和表 3 可知, 增加模型的深度 (通过增加层数) 比增加模型的宽度 (通过增加过滤器) 更能带来显著的准确率提升, 从

4层卷积层扩展到8层,准确率提升了7.5个百分点左右,而无论使用哪种正则化手段,准确率均只有不到2%的提升。

3.3.3 自注意力机制和焦点调制实验结果

本文实证研究了Bottle Gram-KAN卷积层的自注意力机制和焦点调制,使用AdamW优化器^[38],并将分类交叉熵作为损失函数,batch size控制在12,得到不同的自注意力机制和焦点调制结果,如表4所列。

表4 不同注意力模块的实验结果

Table 4 Experimental results of different attention modules

Model	Acc/%	Se/%	Sp/%	F_{β} /%	Params
Self-Gram-KAN	86.23	90.33	81.88	82.32	4.77×10^6
Bottle Self-Gram-KAN	87.88	94.86	84.21	86.02	1.36×10^6
Focal-Gram-KAN	87.50	90.17	87.11	88.70	2.67×10^6
Bottle Focal-Gram-KAN	92.45	95.68	90.36	92.16	1.18×10^6

焦点调制模块在性能上显著优于自注意力模块,无论是否引入瓶颈架构,焦点调制的准确率均比自注意力模块高出约6个百分点。此外,注意力机制与瓶颈结构的结合表现优于无瓶颈结构的模型,准确率提升了约4个百分点,同时保持了更低的参数量(参数量约缩减到原始的40%)。自注意力机制与KAN卷积的集成效果不如原始KAN卷积层,各项评价指标均有所下降,其原因可能是:对于心音这种时序信号,自注意力机制的计算开销可能并未转化为性能提升,反而降低了训练效率^[39]。

结束语 本文研究集成了KAN与卷积网络,构建了基于KAN卷积的端到端心音识别模型框架,为房颤诊断提供了新型辅助工具。首先,通过对原始心音数据进行质量评估和数据清洗,提高了数据的可用性和模型的鲁棒性;其次,设计了具有降维-基函数计算-升维流程的KAN卷积的瓶颈结构,在维持特征提取能力的同时降低了模型参数量,显著优化了计算效率,同时,在瓶颈架构的基础上提出融合L1/L2正则化与Dropout正则化的方案,使模型在数据集上的准确率和 F_{β} 值均有提升,有效增强了房颤特征的判别鲁棒性;最后,引入自注意力机制与焦点调制的协同优化策略,采用KAN卷积层重构QKV投影过程,构建了适配心音时序特性的特征交互机制。通过在中国人民解放军总医院第一医学中心的心音数据集上进行多模型、多任务的实验,验证了所提方法的有效性。

然而,KAN卷积在心音识别算法中的应用会面临多个挑战:1)KAN卷积的可拓展性主要聚焦于房颤心音识别,未实证研究非房性心血管疾病的多种类别识别;2)在KAN卷积处理类不平衡问题上,尤其是在模型中引入正则化手段S,仍然是一个亟待解决的问题;3)在神经网络计算开销上,尤其是当网络深度和参数量增加时,模型的训练和推理速度可能成为实际部署中的瓶颈;4)在KAN卷积算法的泛化能力方面,心音数据的时变性和个体差异要求模型具备较强的泛化能力,然而,当前可用的房颤心音数据集相对有限,无法进行更广泛的泛化实验,因此可能会影响KAN卷积在面对不同个体和不同临床情境下的表现,尤其是在复杂的网络结构下,其泛化能力的提升面临一定挑战。

未来将会把工作拓展到KAN卷积在多分类识别任务上

的效果;在论证函数的泛化功能上寻找到更优质的答案;探索一维心音特征提取后生成的多通道矩阵,并将其作为KAN卷积的输入模块;侧重于探索其他正则化技术以及自注意力模块,以进一步提升模型的效率和性能。

参考文献

- [1] GAO M Y, HE L, DU X, et al. 20 years of atrial fibrillation epidemiology in China [J]. Chinese Journal of Cardiovascular Disease, 2024, 52(2): 220-226.
- [2] KORNEJ J, BÖRSCHEL C S, BENJAMIN E J, et al. Epidemiology of atrial fibrillation in the 21st century: novel methods and new insights[J]. Circulation Research, 2020, 127(1): 4-20.
- [3] HAGIWARA Y, FUJITA H, OH S L, et al. Computer-aided diagnosis of atrial fibrillation based on ECG Signals: A review[J]. Information Sciences, 2018, 467: 99-114.
- [4] TAE-SEOK K, HO-JOONG Y. Role of echocardiography in atrial fibrillation[J]. Journal of Cardiovascular Ultrasound, 2011, 19(2): 51-61.
- [5] GIBSON C M, CIAGLO L N, SOUTHARD M C, et al. Diagnostic and prognostic value of ambulatory ECG(Holter) monitoring in patients with coronary heart disease: a review[J]. Journal of Thrombosis and Thrombolysis, 2007, 23: 135-145.
- [6] NABIH-ALI M, EL-DAHSHAN E L S A, YAHIA A S. A review of intelligent systems for heart sound signal analysis[J]. Journal of Medical Engineering & Technology, 2017, 41(7): 553-563.
- [7] YANG L, LI S, ZHANG Z, et al. Classification of phonocardiogram signals based on envelope optimization model and support vector machine[J]. Journal of Mechanics in Medicine and Biology, 2020, 20(1): 1-7.
- [8] ZHAO Q H, GENG S J, WANG B Y, et al. Deep Learning for Heart Sound Analysis: A Literature Review [J/OL]. <https://www.medrxiv.org/content/10.1101/2023.09.16.23295653v1.full-text>.
- [9] RENNA F, OLIVEIRA J H, COIMBRA M T. Deep Convolutional Neural Networks for Heart Sound Segmentation [J]. IEEE Journal of Biomedical and Health Informatics, 2019, 23(6): 2435-2445.
- [10] AKRAM, MUHAMMAD USMAN, et al. Analysis of PCG signals using quality assessment and homomorphic filters for localization and classification of heart sounds[J]. Computer Methods and Programs in Biomedicine, 2018, 164: 143-157.
- [11] WANG Y L, SUN J, YANG H B, et al. Classification model of pulmonary hypertension heart sounds based on time-frequency fusion features[J]. Computer Science, 2024, 51(S1): 387-393.
- [12] DENG M, MENG T, CAO J, et al. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks[J]. Neural Networks, 2020, 130: 22-32.
- [13] WHANG S E, TAE K H, ROH Y, et al. Responsible AI Challenges in End-to-end Machine Learning [J]. IEEE Data Engineering Bulletin, 2021, 44: 79-91.
- [14] CHEN Y, SUN Y, LV J, et al. End-to-end heart sound segmen-

- tation using deep convolutional recurrent network[J]. *Complex & Intelligent Systems*, 2021, 7: 2103-2117.
- [15] LI S, SUN J, YANG H, et al. Interpretable End-to-End heart sound classification[J]. *Measurement*, 2024: 237.
- [16] NAHAR K, AL-HAZAIMEH O M, ABU-EIN A, et al. Phonocardiogram classification based on machine learning with multiple sound features[J]. *Journal of Computer Science*, 2020, 16(11): 1648-1656.
- [17] LIU Z, WANG Y, VAIDYA S, et al. KAN: Kolmogorov-Arnold Networks[J]. arXiv: 2404. 19756, 2024.
- [18] DROKIN I. Kolmogorov-Arnold Convolutions: Design Principles and Empirical Studies[J]. arXiv: 2407. 01092, 2024.
- [19] OUYANG D, HE B, GHORBANI A, et al. Video-based AI for beat-to-beat assessment of cardiac function[J]. *Nature*, 2020, 580(7802): 252-256.
- [20] ZHAN J, WU X, FU X, et al. Non-contact assessment of cardiac physiology using FO-MVSS-based ballistocardiography: a promising approach for heart failure evaluation [J/OL]. <https://www.nature.com/articles/s41598-024-53464-8>. pdf.
- [21] ASMARE M H, WOLDEHANNA F, JANSSENS L, et al. Rheumatic heart disease detection using deep learning from spectro-temporal representation of un-segmented heart sounds [C]//2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020: 168-171.
- [22] SINGH S A, MAJUMDER S. Short unsegmented PCG classification based on ensemble classifier[J]. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2020, 28(2): 875-889.
- [23] LI F, TANG H, SHANG S, et al. Classification of Heart Sounds Using Convolutional Neural Network [J]. *Applied Sciences*, 2020, 10(11): 3956.
- [24] SKALICKY D, KOUCKYU, HADRABA D, et al. Detection of Respiratory Phases in a Breath Sound and Their Subsequent Utilization in a Diagnosis [J]. *Applied Sciences*, 2021, 11(14): 6353.
- [25] SATHESH K, KOWSALYA P, ARAVINDRAJ E, et al. Clearer HeartBeats: Enhancement of cardiac sounds using Adaptive Filtering and Wavelet Decomposition[C]//2024 10th International Conference on Communication and Signal Processing (ICCSPP). IEEE, 2024: 947-952.
- [26] TANG H, WANG M, HU Y, et al. Automated Signal Quality Assessment for Heart Sound Signal by Novel Features and Evaluation in Open Public Datasets[J]. *BioMed Research International*, 2021: 7565398.
- [27] BODNER A D, TEPISICH A S, SPOLSKI J N, et al. Convolutional Kolmogorov-Arnold Networks [J]. arXiv: 2406. 13155, 2024.
- [28] ZOU X, XU S, CHEN X, et al. Breaking the von Neumann bottleneck: architecture-level processing-in-memory technology[J]. *Science China Information Sciences*, 2021, 64(6): 1-10.
- [29] SRINIVAS A, LIN T Y, PARMAR N, et al. Bottleneck transformers for visual recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 16519-16529.
- [30] SINGH S, RUWASE O, AWAN A A, et al. A hybrid tensor-expert-data parallelism approach to optimize mixture-of-experts training[C]//Proceedings of the 37th International Conference on Supercomputing. 2023: 203-214.
- [31] LI B, SHEN Y, YANG J, et al. Sparse Mixture-of-Experts are Domain Generalizable Learners[C]//International Conference on Learning Representations (ICLR). 2023.
- [32] LIU C F, SUN H, DONG H. A study of molecular amplification timing prediction combining Transformer and Kolmogorov Arnold network[J]. *Journal of Graphics*, 2024, 45(6): 1256-1265.
- [33] WU H, XIAO B, CODELLA N, et al. Cvt: Introducing convolutions to vision transformers[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 22-31.
- [34] YANG J, LI C, DAI X, et al. Focal modulation networks [J]. *Advances in Neural Information Processing Systems*, 2022, 35: 4203-4217.
- [35] TIAN M D, PENG D T, ZHANG X. Research on sparse optimization with linear inequality constraints based on Huber loss and Capped L1 regularization [J]. *Theory of Mathematics*, 2022, 12(11): 12.
- [36] WANG Q, YANG H B, PAN J H, et al. Heart sound classification algorithm based on time-frequency combination feature and adaptive fuzzy neural network [J]. *Journal of Biomedical Engineering*, 2023, 40(6): 1152-1159.
- [37] WANG Y L, SUN J, YANG H B, et al. Heart sound classification based on improved mel frequency cepstrum coefficient and integrated decision network method [J] *Journal of Biomedical Engineering*, 2022, 39(6): 1140-1148.
- [38] LOSHCHILOV I, HUTTER F. Decoupled Weight Decay Regularization[J]. arXiv. 1711. 05101, 2017.
- [39] KIM H, KO J G. Fast Monte-Carlo approximation of the attention mechanism[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022: 7185-7193.



DENG Hong, born in 2000, postgraduate. His main research interests include biomedical signal processing and deep learning.



WANG Weilian, born in 1947, professor. His main research interests include biological signal processing, digital-analog hybrid IC and ASIC design.