

基于一维卷积神经网络与循环神经网络 串联的心音分析方法

肖 斌¹, 陈嘉博¹, 毕秀丽¹, 张俊辉², 李伟生¹, 王国胤¹, 马 旭³

(1. 重庆邮电大学图像认知重庆市重点实验室, 重庆 400065; 2. 重庆医科大学附属第一医院, 重庆 400042;
3. 国家卫生健康委员会科学技术研究所, 北京 100081)

摘 要: 面向心脏疾病计算机辅助诊断, 本文提出一种基于一维卷积神经网络和循环神经网络混合深度学习结构的心音分析方法. 本结构首先利用卷积神经网络学习心脏病在心音信号上的表征, 然后通过循环神经网络处理心音信号中的时序信息进行分类, 在提升心音分类正确率的同时, 大幅度降低了网络参数. 为验证本深度学习结构所学特征的有效性, 除已有的成人心音数据集外, 本文还专门构建了一个面向婴幼儿先天性心脏病的心音数据集, 并通过端到端的类别响应图证明了本方法在室缺诊断时学习到的心音信号特征符合临床医师的心音听诊经验. 实验结果表明, 本文方法能在 3 153 例成人心音数据分类上达到 92.56% 的正确率, 在 528 例婴幼儿心音数据分类上达到 97.48% 正确率, 模型参数仅有 0.05 M.

关键词: 心音听诊; 一维卷积神经网络; 循环神经网络; 类别响应图

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2022)10-2425-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20200713

A Method of Heart Sound Analysis Based on One-Dimensional Convolutional Neural Network and Recurrent Neural Network

XIAO Bin¹, CHEN Jia-bo¹, BI Xiu-li¹, ZHANG Jun-hui², LI Wei-sheng¹, WANG Guo-yin¹, MA Xu³

(1. Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. The First Affiliated Hospital of Chongqing Medical University, Chongqing 400042, China;

3. Institute of Science and Technology, National Health Commission, Beijing 100081, China)

Abstract: For the computer-aided heart disease diagnosis, this paper proposes a method of heart sound analysis based on the mixed structure of one-dimensional convolutional neural network and recurrent neural network. The proposed structure uses the convolutional neural network to learn the representation of heart disease on the heart sound signal, and then processes the time sequence information in the heart sound signal through the recurrent neural network for classification, which greatly reduces the network parameters while improving the accuracy of heart sound classification. The experimental results show that the proposed method can achieve the accuracy of 92.56% on the classification of 3153 cases of normal and abnormal heart sounds for adults, and 97.48% on the classification of 528 cases of normal and abnormal heart sounds for infants and children. The parameter of the proposed method is 0.05M, which is suitable for portable application situations.

Key words: heart sound analysis; one-dimensional convolutional neural network; recurrent neural network; class activation map

1 引言

心血管病 (Cardiovascular Disease, CVD) 是目前世界上较常见的一类疾病. 当前全世界每年有 1 770 万人

死于心血管疾病. 预计到 2030 年每年因心血管疾病死亡的人数将增加至 2 300 万人. 快速准确地对心血管疾病做出诊断有着重大的意义. 心脏的生理状况在不同

收稿日期: 2020-07-14; 修回日期: 2021-06-10; 责任编辑: 孙瑶

基金项目: 国家自然科学基金 (No.61806032, No.61976031); 国家重点研发计划 (No.2016YFC1000307-3); 重庆市基础与前沿项目 (No.cstc2018jcyjAX0117); 重庆市教委科学技术研究计划重点项目 (No.KJZD-K201800601)

模态的心脏相关信号上都有所反应,如超声(Ultrasound)、心音图(PhonoCardioGram, PCG)、心电图(ElectroCardioGraph, ECG)和心脏振动描记图(CardioPlegia)等. 心音听诊作为心脏声音信号分析方法,不仅方便快捷,而且具有非侵入性,是一种可以被广泛应用于心血管疾病诊断的方法. 然而,心音听诊的准确性严重依赖医生的技能和主观经验,而这些技能和主观经验需要医生从长期的听诊经验中获得. 因此,迫切需要一种客观、高效的针对心音的辅助分析工具. 目前,基于生物信号处理和人工智能技术的心音自动分析正成为一个热门的研究点,有望在临床诊断上应用,从而减少昂贵和耗时的人工检查.

基于人工智能技术的计算机辅助心音分析中,心音数据集的构建是关键,目前存在的公开心音数据集数量相对较少,并且大部分数据集中的心音质量和数量都不足以达到训练分类模型(特别是深度学习模型)的要求. 在2016年,PhysioNet/CinC 2016国际会议上^[1]首次公开了一个超过3 000条心音信号的大规模心音数据集,且每条心音信号的长度在5~120 s之间. 但深度学习模型对心音数据质量要求极高,现有的心音数据集还存在着如下的问题:(1)数据集的规模依旧不够,仍然无法满足深度学习模型的数据规模的要求;(2)该数据集中的心音记录的标签仅有正常和异常,还需更为具体的心脏疾病类别标记;(3)由于病理原因,成人和婴幼儿的心脏病种类不同,婴幼儿的心脏疾病以先天性心脏病居多,应分开建立成人和婴幼儿的心音数据集.

现有的计算机辅助心音分析方法通常由心音分割、特征提取和分类器训练3个步骤组成. 心音分割作为心音分析中的第一步,分割效果的好坏一定程度上将影响后续步骤的效果. 心音分割在对心音信号进行降采样、去噪等预处理之后,根据心音的周期性,通过心音分割算法,将心音信号分割成一系列重复的心动周期. 相关研究总结了现有的心音分割方法,将各类方法大体上分为4类^[2-6]. Springer等人^[7]提出的基于隐马尔科夫的心音分割方法,训练出了能够高效分割出包含完整心动周期的心音片段的模型,但该模型在遇到较长的心音周期和异常的窦性心律时,会出现分割错误的情况. 近年来,随着心音数据的积累,一些基于深度学习的方法^[8-10]被陆续提出. Chen等人^[8]引入了梅尔频谱系数对心音信号进行识别,再通过深度神经网络和K-means聚类结合的方法对得到的心音特征进行分类. Messner等人^[9]提取了心音中的幅度谱以及包络特征. Renna等人^[10]提取心音中的包络特征,作为卷积神经网络(Convolutional Neural Network, CNN)的输入,实现了心音的分割. 然而,为了准确地分割心音信号,

大多数的心音分割算法都需要同步心电图作为参考信号,带来了心音信号与心电信号同步获取困难的问题. 另外,由于心音信号的分割标记非常困难,而心音信号分割的效果会一定程度上影响后续辅助诊断的准确率,如何训练有效的心音分割模型以提升准确率,是现有心音分析模型所面临的现实问题.

在心音信号被分割成心音片段之后,利用不同的特征提取方法提取心音信号中的有效特征,再由提取到的特征训练分类器进行心音分类. 现有的心音信号特征提取方法主要包括时域特征、频域特征和时频域特征,如短时傅里叶变换(STFT)^[11,12]和小波变换^[6,13]等. 后级分类器的选择主要取决于所选取的特征类型和心音信号本身的特性,常用的分类器有神经网络(ANN)^[4]、支持向量机(SVM)^[5]、K-最近邻分类器(K-NN)^[6]、隐马尔可夫模型(HMM)^[7]等. 虽然传统的基于特征提取的心音分类方法有使用简单、训练速度快等特点,但存在着需要比较复杂的前处理和后处理过程、实际应用中鲁棒性不足等问题. 近年来,深度学习的方法在许多应用领域都取得了巨大的进展,一些研究者开始探索基于深度学习的心音分析方法. 基于深度学习的心音分析方法是一种端到端的方法,特征提取和后续的分类都在同一个框架中完成,常采用的模型有卷积神经网络(CNN)、循环神经网络(RNN)等^[14]. 而在文献[15~17]中,一维的心音信号首先被转化成二维信号然后送入二维CNN进行特征提取和分类. 其中,Dominguez-Morales等人^[15]使用神经形态听觉传感器(Neuromorphic Auditory Sensor, NAS)等硬件设备将心音信号转化成声波图(sonagram);Rubin等人^[16]将心音信号转化为二维梅尔频谱系数(MFCCs);Nilanon等人^[17]将一维心音信号转化成了功率谱密度(Power Spectral Density, PSD). 另外,也有研究者提出了将深度学习与传统特征提取相结合的方法,如Pote等人^[18]集成了一维CNN和使用时域、频域特征的Adaboost共同决策最终的分类结果. 虽然现有的深度学习模型在心音分析上取得了一定的效果,但还存在模型复杂、参数量大、训练和分类时间长等问题. 如何设计一种参数量低、分类效果好、可以应用于便携式场合的心音深度学习模型具有十分重要且现实的意义.

针对上述问题,本文提出了一种基于一维CNN和RNN混合深度学习结构的心音分析方法,除了已有的成人心音数据集外,还构建了一个面向婴幼儿先天性心脏病的心音数据集来验证本文方法的有效性. 本文方法首先通过滑动窗口将心音信号分割成片段作为深度学习模型的输入;然后,构建了一个基于一维CNN和RNN串联的混合深度学习模型,利用CNN学习心音信号的生理和病理特征,通过RNN提取所学习特征的

上下文关联信息,对心音片段进行分类诊断;最后,使用投票机制将心音片段的的结果转换成心音信号的诊断结果. 本方法具有在极低的参数量的情况下保证较高的准确率的的优势,适合应用于便携式场合. 为验证本文所提出的模型在室间隔缺损疾病的诊断中是否利用了有效的生理和病理特征,从可视化模型的类别响应图(Class Activation Map, CAM)角度出发,探索深度学习模型在室间隔缺损疾病智能辅助诊断领域的可解释性.

2 基于一维卷积神经网络和循环神经网络串联的心音分析方法

本节将具体阐述基于一维 CNN 和 RNN 串联的心音分析方法. 首先介绍心音信号的预处理和一种固定长度的心音分割方法,接着详细描述提出的一维 CNN 和 RNN 串联结构,之后给出将片段级结果转化为记录级诊断结果的决策规则,最后给出基于类别响应图的心音深度学习模型的可解释性方法,具体流程如图 1 所示.

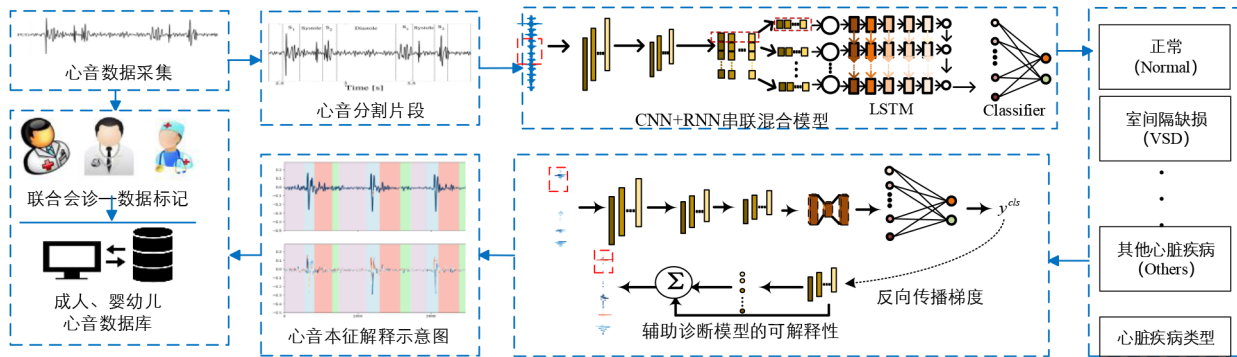


图1 心音分类方法总体流程图

2.1 心音信号预处理与分割

将原始的一维心音信号通过短时傅里叶变换转化为梅尔频率倒谱系数、功率谱密度等二维时频特征表示,是当前心音分类算法中最常采用的方式. 这种方式通常能够很好地提取出声音里的一些特征,但是需要额外的转化操作和许多需要人为优化的超参数,这将带来额外的计算负担和超参数优化负担. 为了克服以上的问题,本方法直接用一维原始心音信号作为网络的输入. 由于心音分割的效果在一定程度上影响后续的分类性能,而现有的心音分割模型很难实现精确、快速的分割,本文的方法中采取一种更为简单、直接的分割方法,利用滑动窗口将不同长度心音信号分割成固定长度的心音片段,分割后的每条心音片段的标签与所属心音信号的标签相同.

分割方法,利用滑动窗口将不同长度心音信号分割成固定长度的心音片段,分割后的每条心音片段的标签与所属心音信号的标签相同.

2.2 一维卷积神经网络与循环神经网络串联混合网络模型

本文方法将分割后的心音片段作为网络模型的输入,然后,构建一个基于一维卷积神经网络和循环神经网络串联的混合模型,结合卷积神经网络提取局部特征与循环神经网络提取上下文关联信息的优势,对心音片段自动提取特征并分类,网络结构如图 2 所示.

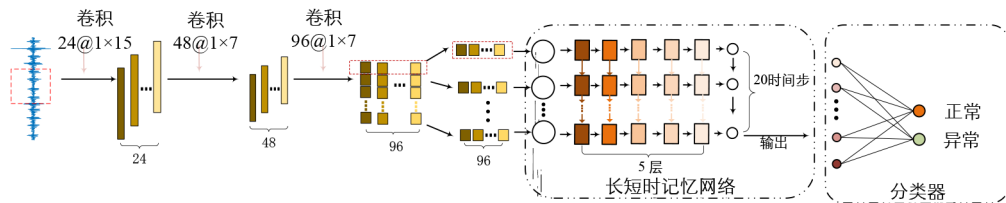


图2 一维卷积神经网络和循环神经网络串联网络结构图

本文方法使用了3个卷积层用以提取心音信号的局部特征,每层特征图的个数分别是24,48和96. 第一层使用大小为 1×15 的卷积核,在前期采用较大的卷积核,有利于提取较长时间片段的特征,使得模型有更好的时序处理能力. 为了兼顾模型轻量化的特性,后2层使用 1×7 大小的卷积核,后期采用较小的卷积核,使得模型能够在有效学习特征的同时降低模型的参数. 紧

跟卷积层的是最大池化层,池化的大小为3,步长为2,在卷积层之后通过最大池化层对心音信号进行数据压缩,有助于后续循环神经网络的处理. 与传统的基于CNN的心音分析方法相比,本文方法中使用到的CNN有网络结构简单、参数量低、有较强的时序处理能力等优势.

心音信号作为一种时间序列,内部存在一些与时

序相关的特征, 仅仅使用卷积神经网络难以提取到相关的时序信息. 因此本文在使用卷积神经网络从原始的心音片段中学习相对局部的特征之后, 使用循环神经网络提取心音片段中的时序关联特征, 例如在心音听诊中在某一个心动周期当中出现了一些异常, 但是医生并不能凭借仅出现一次的异常进行判断, 需要集合前面及后面的心动周期的情况进行诊断, 而在深度学习网络模型中, 需要用到序列前后的关联信息进行判断, 在这方面循环神经网络有很好的表现. 本文使用一种变形的 RNN, 即长短时记忆神经网络(LSTM)提取心音中的时序特征. 首先通过 3 个卷积层来减少输入信号的频率方差, 经过卷积神经网络的输出输入到循环神经网络的向量是一个二维向量, 维度分别是时间长度以及特征图的个数. 原始的心音信号是一个通道数为 1 的一维信号, 其维度就是心音信号的时间长度, 心音信号是根据时间连续采集的, 因此在时间的维度上信号是平滑的. 经过卷积神经网络处理之后, 心音信号的特征图个数发生了变化, 但卷积神经网络的卷积操作是沿着心音信号的时间轴进行的, 因此心音信号在时间轴进行了压缩. 由卷积神经网络输出的向量的维度是时间长度和特征图的个数.

本方法使用 5 层 LSTM 结构, 每一层有 20 个单元. 在训练时 LSTM 展开 20 个时间步, 如图 2 所示, 输入 LSTM 结构的共有 96 个特征图, 设输入的特征图的时间长度为 N , 沿着时间维度将其划分为 20 组, 每组有 96 个时间长度为 $N/20$ 的特征图, 同一组的特征图都是由同一段时间长度的心音信号经过不同的卷积核操作得到, 因此同一组的特征图在时间维度上是一致的, 每组依次输入到 LSTM 当中进行处理, LSTM 根据每组特征对最终结果的影响程度, 选择是否将当前的特征输入到下一个时间步当中. 对于每个时间步要同时考虑传入的特征和上一个时间步传递下来的内容, 从而得到当前时间步的计算结果. 最后将 LSTM 的输出经过批标准化(Batch Normalization)以及非线性激活(ReLu)等操作后输出到全连接层, 通过 softmax 函数得到各个类的概率, 最终得到心音片段的分类结果.

对于心音异常往往会反复出现在多个心动周期当中这一特点, 本文在设计网络模型时着重利用较长时间中包含了相当数量的心动周期的特征, 主要通过前级使用较大的卷积核实现. 本方法中使用的是 1×15 的卷积核, 有利于模型更好地处理时序信息. 后级则使用 1×7 的小卷积核, 在保证有较好的特征提取能力之外, 相比大卷积核有降低模型参数量等优势. 本文模型为处理心音中时序特征引入了 LSTM, 针对心音数据的特点, 这是因为将 LSTM 的时间步设置得很小, 仅有 20 个单元, 这在以往的心音辅助诊断模型中是很少见的, 将

LSTM 的时间步设置较小, 则对于每个时间步内的心音长度较长, 其中会包括多个完整的心动周期. LSTM 共有 5 层结构, 能够充分利用时序信息. 在模型的前级使用的一维卷积神经网络当中使用预先激活的策略, 简而言之, 就是将批标准化和非线性激活等操作放在卷积操作之前, 有助于模型的正则化以及加快拟合的速度.

2.3 最终决策规则

本方法的最终目的是对心音信号进行分类, 因此使用一维 CNN 和 RNN 串联混合网络结构对心音片段的诊断结果需要转换为记录级的结果. 考虑到医生在临床上进行心音听诊不会单凭某个心动周期的异常进行诊断, 而往往是发现某个心动周期的异常之后会继续看接下来的心动周期中是否还存在异常, 如果异常出现次数较多才会做出判断. 本文使用投票机制, 通过心音分割将一段心音信号分割成多个心音片段, 再通过神经网络对每个心音片段进行分类. 如图 3 所示, 统计一段心音信号中所有片段被诊断的类别, 如果异常的类别多于正常的类别则该心音信号诊断为异常, 否则诊断为正常.

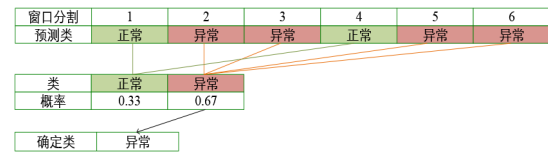


图 3 心音片段投票机制示意图

2.4 最终决策规则基于 CAM 的心音深度学习模型可解释性分析方法

对于常见的 CAM 生成方法来说, 假设串联混合神经网络结构输出的特征图为 $S \in \mathbf{R}^{1 \times H \times K}$, 其中 $1 \times H$ 为空间尺寸, K 为通道数, 将特征图送进一个全局池化层(GAP), 后续的全连接层和 softmax 层进行 C 个类别的心音分类, 第 k 个特征图平均池化后的结果表示为

$$S_k(i) = \frac{\sum_i S_k(i)}{1 \times H}; \quad k=0, 1, \dots, K-1 \quad (1)$$

其中, $S_k(i)$ 为第 k 个通道的特征图 S_k 中的第 i 个元素, i 的取值为 0 到特征图的长度. 同时, 全连接的权值矩阵可以被表示为 $W^{fc} \in \mathbf{R}^{K \times C}$. 因此, 对于目标类别 c , 第 c 个 softmax 节点的输入 y_c^{fc} 可以被定义为

$$y_c^{fc} = \sum_{k=0}^{K-1} S_k \cdot W_{k,c}^{fc} = \frac{\sum_{k=0}^{K-1} \sum_{i=0}^{H-1} S_k(i) \cdot W_{k,c}^{fc}}{1 \times H} \quad (2)$$

其中, $W_{k,c}^{fc} \in \mathbf{R}$ 表示权值矩阵 W^{fc} 中第 k 行、第 c 列的元素. 对类别 c 可视化的类别响应图 A_c^{fc} 可以被定义为

$$A_c^{fc}(i) = \sum_{k=0}^{K-1} S_k(i) \cdot W_{k,c}^{fc} \quad (3)$$

那么 $y_c^e = \sum_i A_c^e(i)$, 其中 A_c^e 直接体现了第 i 个元素在将心音分类为类别 c 的影响程度. 这样的可视化方式虽然可以产生类别响应图, 但是在网络正向传播后需要额外的步骤才能得到, 需要修改网络结构并对模型进行重新训练, 这样的方式应用很不方便. 本文采用一种更为方便的方式, 无需进行结构上的调整, 如图 4 所示, 通过梯度的全局平均来计算每个特征图的权重. 第 k 个特征图对类别 c 的权重可通过式(4)计算:

$$\alpha_k^c = \frac{\sum_i \sum_j \frac{\partial y_c^e}{\partial A_{ij}^k}}{Z} \quad (4)$$

其中, Z 为特征图像素个数; y_c^e 是对应类别 c 的分数; A_{ij}^k 表示第 k 个特征图中 (i, j) 位置的像素值, i 的取值为 0 到特征图的长度, j 的取值为 0 到特征图的宽度. 求得类别对所有权重值的权重后, 求特征图和对对应权重的加权就能得到类别响应图, 即

$$L^c = \text{RELU} \left(\sum_k \alpha_k^c A^k \right) \quad (5)$$

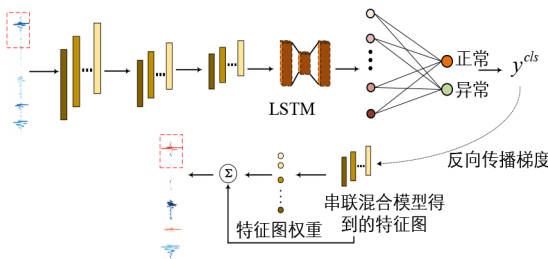


图 4 端到端的类别响应图生成示意图

3 实验结果与分析

3.1 数据集

3.1.1 PhysioNet/Cinc 2016 心音数据集

PhysioNet/CinC 2016 心音数据集首次公布于 2016 年举办的心脏病学挑战赛 (PhysioNet/CinC 2016), 总共包含来自 764 个患者的长度在 5~120 s 之间共 3 153 条心音记录. 心音记录的标签有正常和异常两类. 标签为正常的心音记录采集自健康人; 标签为异常的心音记录采集自被诊断为心血管疾病的患者, 所患心脏病的种类主要为冠状动脉疾病.

3.1.2 婴幼儿心音数据集

深度学习在心音分析领域应用和发展受限的首要原因是训练数据缺乏. 尤其在面向婴幼儿先天性心脏病的心音分析领域, 目前没有公开的婴幼儿心音数据集. 基于这一情况, 本文构建了一个包含心音信号、对应标签和个体年龄的婴幼儿先天性心脏病心音数据集, 使基于深度学习的婴幼儿心音分析成为了可能. 所有的心音数据由合作儿童医院的专业医生从 137 名年

龄为 1 个月至 12 岁之间的婴幼儿个体中采集, 手工裁剪了心音中类似父母对话、小儿啼哭等噪声片段^[14]. 最终数据集保留了 528 条婴幼儿心音数据, 每条心音持续时间为 3~249 s, 并用最后确诊的疾病类型作为每条心音信号的标签. 心音标签分为正常、心房间隔缺损、心室间隔缺损、同时含有心房间隔缺损与心室间隔缺损、法洛四联症以及其他心血管相关疾病. 所有小儿心音数据都是使用采样率为 44.1 kHz、精度为 16 位的 Thinklabs one 数字听诊器在临床环境中采集^[14].

3.1.3 婴幼儿室间隔缺损心音数据集

室间隔缺损是最常见的先天性心脏病. 在婴幼儿心音分析中, 目前没有公开的特定心脏疾病的数据集, 对特定的心脏疾病的研究和诊断产生了巨大的阻碍. 因此, 本文从新构建的婴幼儿心音数据集中挑选出室间隔缺损和正常的样本, 从而构建了针对室间隔缺损疾病的心音数据集, 为婴幼儿室间隔缺损疾病的专业化诊断提供了数据基础. 该数据集共有 348 条心音数据, 分为正常个体和室间隔缺损患者两类, 两类标签的数据量相对平衡.

3.2 评价指标

本文使用了多种评价指标来验证对比方法的性能, 分别是准确率 (Accuracy, Acc)、查准率 (Precision, Pr)、敏感性 (Sensitivity, Se)、特异性 (Specificity, Sp)、总体得分 (Overall score, Score) 和神经网络中可训练参数的数量 (Parameters, Params). 总体得分为 PhysioNet/CinC 2016 中使用的评价指标, 为敏感度和特异性的平均值, 用来衡量方法的综合诊断能力.

3.3 本文算法的实验效果

3.3.1 本文算法实验细节

本文算法在单个 NVIDIA 2080Ti 显卡环境下进行, 所提出的混合神经网络在 TensorFlow 深度学习框架上实现, 并使用十折交叉验证的方法进行实验. 本文算法使用了加权因子为 0.25 的加权交叉熵作为损失函数, 用以解决存在的数据不平衡的问题, 选取带 Nesterov 动量为 0.9 的随机梯度下降 (SGD) 对加权交叉熵进行优化. 训练数据的批大小设置为 64, 早停法训练 80 个 epoch. 训练的初始学习率为 5×10^{-4} , 通过指数衰减策略调整学习率, 模型最终收敛. 表 1 给出了本文提出的混合神经网络的具体配置.

3.3.2 在 PhysioNet/Cinc 2016 心音数据集上的对比实验

为了证明本方法的有效性, 本文主要与两类方法进行对比: 基于 CNN 的心音分类方法 (1-D CNN^[14], MFCC-CNN^[16], PSD-CNN^[17], AdaBoost-CNN^[18], DGRE^[19], 1-D Clique^[20]), 以及基于手工提取特征的心音分类方法 (SS-PLSR^[2] 和 SSTD^[11]). 与这些方法对比的实验结果

表1 神经网络结构具体配置

层类型	输出大小	配置
输入	1×6 000×1	—
卷积层1	1×2 993×24	卷积核大小15,步长2
池化层1	1×1 495×24	最大池化大小3,步长2
卷积层2	1×744×48	卷积核大小7,步长2
池化层2	1×371×48	最大池化大小3,步长2
卷积层3	1×182×96	卷积核大小7,步长2
LSTM层	1×182×20	5层,时间步大小20
分类器	2	批标准化,非线性激活,全连接层

如表2所示,从中可以发现本文方法的心音诊断的特异性是最好的,同时在准确率和总分上也取得了较好的成绩,最重要的是,与其他几种用于心音诊断的基于卷积神经网络的方法相比,本方法使用的参数量是最

少的,仅为0.05 M,对设备资源的需求最低.从表2中可以看出MFCC-CNN也有很好的诊断能力,但是在该方法中首先需要进行信号的维度转换,同时很难保证在这样的信号转换的过程中不会丢失一些重要的生理特征,与本方法相比参数量超过200倍.而对于1-D Clique,该网络结构较为复杂,参数量是本文方法的3倍以上,心音诊断耗时长,且很容易产生过拟合的问题.在传统的手工提取特征的方法中,需要将心音信号精确地分割为完整的心动周期,因此心音分割效果的好坏直接影响这类方法的诊断效果.而本文方法在不需精确分割的情况下仍然能取得很好的诊断效果,这反映了本方法具有较高的特征提取和表示能力,不需要对心音进行精细地分割,就能够实现自动地特征提取并高效分类.

表2 在PhysioNet/Cinc 2016数据集上进行心音诊断的对比实验结果

评价指标	1-D CNN	MFCC	PSD-CNN	AdaBoost	1-D Clique	DGRE	SS-PLSR	SS-TD	Proposed
精确分割	否	是	否	是	否	否	是	是	否
ACC	0.893 3	0.933 1	0.890 5	—	0.932 8	—	—	—	0.929 6
Pr(正常)	0.936 4	0.935 8	0.949 4	—	0.962 7	—	0.95	0.92	0.946 7
Pr(异常)	0.739 6	0.853 6	0.710 3	—	0.829 0	—	0.82	0.88	0.857
Se	0.760 8	0.826 6	0.815 0	0.88	0.862 1	0.83	—	—	0.799 2
Sp	0.928 2	0.951 6	0.910 2	0.82	0.951 6	0.96	—	—	0.963 7
Score	0.844 5	0.889 1	0.862 6	0.85	0.906 9	0.90	—	—	0.881 4
Params	0.19 M	12.41 M	0.24 M	—	0.19 M	—	—	—	0.05 M

3.3.3 在婴幼儿心音数据集上的对比实验

为了进一步验证本方法在婴幼儿心音数据集上的有效性,本文对3.1.2节中所介绍的婴幼儿心音数据集中与其他基于CNN的心音诊断方法进行了对比实验. AdaBoost-CNN, DGRE, SS-PLSR和SS-TD等模型未找到开源代码,因此在婴幼儿心音数据集的对比实验中不加入上述模型.如表3所示所有指标,本方法均取得了最优结果.

表3 在婴幼儿心音数据集上进行心音诊断的对比实验结果

评价指标	1-D CNN	MFCC	PSD-NN	1-D clique	Proposed
精确分割	否	是	否	否	否
ACC	0.911 0	0.923 3	0.856 6	0.916 8	0.974 8
Pr(正常)	0.925 8	0.911 4	0.888 2	0.889 6	0.973 5
Pr(异常)	0.909 2	0.932 2	0.842 2	0.936 6	0.985 4
Se	0.960 9	0.955 8	0.959 7	0.934 0	0.994 8
Sp	0.810 6	0.882 5	0.664 4	0.891 7	0.951 4
Score	0.885 8	0.919 2	0.812 1	0.912 9	0.973 1
Params	0.19 M	12.41 M	0.24 M	0.19 M	0.05 M

3.4 评价指标面向婴幼儿室间隔缺损的心音分析和模型可解释性

由于不同的心脏疾病的病理特征是不同的,为探索本文提出的模型在心音分析上是否学习到了有效的

生理和病理特征,本文选择了在先天性心脏病中占比较大的室间隔缺损疾病来进行可解释性分析.面向婴幼儿室间隔缺损的心音分析模型在3.3.3节所训练的模型基础上进行迁移学习,使模型能够拟合3.1.3节的婴幼儿室间隔缺损心音数据集.同时也与其他的基于卷积神经网络的心音诊断方法进行了对比实验,1-D Clique模型较为复杂,在婴幼儿室间隔缺损疾病的心音数据集上发生了过拟合,因此在对比实验中没有加入1-D Clique.从表4所示的对比实验结果可以看到本方法在除查准率和敏感性以外所有指标上都取得了最好的结果,在查准率和敏感性这两列指标上取得次好

表4 在婴幼儿室间隔缺损疾病心音数据集上进行心音诊断的对比实验结果

评价指标	1-D CNN	MFCC	PSD-CNN	Proposed
精确分割	否	是	否	否
ACC	0.922 6	0.962 4	0.905 1	0.967 6
Pr(正常)	0.891 7	0.957 8	0.876 9	0.978 9
Pr(异常)	0.969 7	0.959 1	0.935 7	0.969 3
Se	0.963 7	0.953 1	0.937 9	0.957 8
Sp	0.890 2	0.961 3	0.882 4	0.978 5
Score	0.926 9	0.957 2	0.910 2	0.968 1
Params	0.19 M	12.41 M	0.24 M	0.05 M

结果.

为进一步验证本文模型在室间隔缺损疾病诊断上的可解释性,利用 2.4 节介绍的基于 CAM 的心音深度学习模型可解释性分析方法对本文模型进行了分析,标记出心音模型提取出的有效的生理和病理特征区域^[21,22].如图 5 所示,类别响应图中红色部分表示响应值高,红色区域对神经网络的最终的诊断有较大的贡献,而蓝色部分表示响应值低,对神经网络的诊断影响较小.如图 5(a)所示,该心音片段的标签为室间隔缺损疾病,而本文方法对其预测的标签也是室间隔缺损疾病,图中红色区域为响应值较高的区域,即对模型将其预测为室间隔缺损疾病的贡献较大.

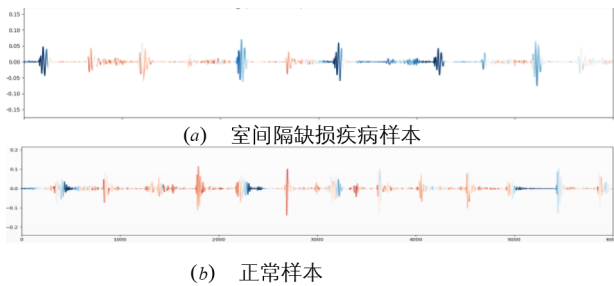


图 5 心音片段类别响应图

图 5 中呈现的热度图的结果和医生在临床上心音听诊的两个规律是一致的.第一个规律是医生在室间隔缺损疾病的临床诊断中主要是依据第一心音、第二心音尤其是舒张期是否存在杂音判断的,图 5(a)中该片段中部分第一心音、第二心音,以及舒张期呈现红色,特别是舒张期呈现红色的区域较多,符合这一规律.第二个规律是在心脏呈现病变时,异常往往不止在心音的其中某一个节拍中才能够听到.而图 5(a)中可以看到在诊断为室间隔缺损疾病的这个心音片段中有多个心动周期都出现了异常的情况.

而对于被模型预测为正常的心音片段的类别响应图,如图 5(b)所示,该片段的标签为正常心音,而模型对其的预测标签也是正常的,可以看到和被诊断为室间隔缺损疾病的心音片段类别响应图不同的是,该片段的类别响应图大部分区域均呈现红色,即对于预测为正常的心音片段整体区域均对最终的预测有较大的贡献.

模型诊断结果呈现出的类别响应图与医生诊断室间隔缺损疾病的心音听诊规律是一致的,也从另一个角度证明了本文模型学习到了有效的生理和病理特征从而完成了心音听诊.

4 总结

本文提出了一种基于一维卷积神经网络与循环神经网络混合的心音分析方法,利用卷积神经网络提取

特征的能力结合循环神经网络处理时序信息的能力,能够在极低参数量的情况下实现较好的诊断效果.该方法无需进行精确的心音分割,精简了心音处理流程.同时本文对室间隔缺损疾病的心音诊断进行了可解释性的探索,利用了心音特定区域的生理和病理特征,和医生的心音听诊的规律一致,证明了本方法的有效性.

参考文献

- [1] 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): 2181-2213.
- [2] LIANG H, LUKKARINEN S, HARTIMO I. Heart sound segmentation algorithm based on heart sound envelopegram [C]//Computers in Cardiology. Lund: IEEE, 1997: 105-108.
- [3] MOUKADEM A, DIETERLEN A, HUEBER N, et al. A robust heart sounds segmentation module based on S-transform[J]. Biomedical Signal Processing and Control, 2013, 8(3): 273-281.
- [4] CHOI S, JIANG Z W. Comparison of envelope extraction algorithms for cardiac sound signal segmentation[J]. Expert Systems With Applications, 2008, 34(2): 1056-1069.
- [5] YAN Z H, JIANG Z W, MIYAMOTO A, et al. The moment segmentation analysis of heart sound pattern[J]. Computer Methods and Programs in Biomedicine, 2010, 98(2): 140-150.
- [6] ARI S, HEMBRAM K, SAHA G. Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier[J]. Expert Systems With Applications, 2010, 37(12): 8019-8026.
- [7] SPRINGER D B, TARASSENKO L, CLIFFORD G D. Logistic regression-HSMM-based heart sound segmentation [J]. IEEE Transactions on Biomedical Engineering, 2016, 63(4): 822-832.
- [8] CHEN T N, YANG S I, HO L T, et al. S1 and S2 heart sound recognition using deep neural networks[J]. IEEE Transactions on Biomedical Engineering, 2017, 64(2): 372-380.
- [9] MESSNER E, ZOHRER M, PERNKOPF F. Heart sound segmentation-an event detection approach using deep recurrent neural networks[J]. IEEE Transactions on Biomedical Engineering, 2018, 65(9): 1964-1974.
- [10] RENNA F, OLIVEIRA J, COIMBRA M T. Convolutional neural networks for heart sound segmentation[C]//2018 26th European Signal Processing Conference(EUSIPCO). Rome: IEEE, 2018: 757-761.

- [11] ZHANG W J, HAN J Q, DENG S W. Heart sound classification based on scaled spectrogram and tensor decomposition[J]. *Expert Systems With Applications*, 2017, 84: 220-231.
- [12] SOETA Y, BITO Y. Detection of features of prosthetic cardiac valve sound by spectrogram analysis[J]. *Applied Acoustics*, 2015, 89: 28-33.
- [13] NIVITHA VARGHEES V, RAMACHANDRAN K I. 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.
- [14] 徐韵秋. 基于深度卷积神经网络的心音智能分析[D]. 重庆: 重庆邮电大学, 2019.
XU Y Q. Heart Sound Intelligent Analysis Based on Deep Convolutional Neural Networks[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2019. (in Chinese)
- [15] DOMINGUEZ-MORALES J P, JIMENEZ-FERNANDEZ A F, DOMINGUEZ-MORALES M J, et al. Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors[J]. *IEEE Transactions on Biomedical Circuits and Systems*, 2018, 12(1): 24-34.
- [16] RUBIN J, ABREU R, GANGULI A, et al. Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients[C]//2016 Computing in Cardiology Conference(CinC). Vancouver: IEEE, 2016: 813-816.
- [17] NILANON T, YAO J, HAO J. Normal/abnormal heart sound recordings classification using convolutional neural network[C]//2016 Computing in Cardiology Conference(CinC). Vancouver: IEEE, 2016: 585-588.
- [18] POTES C, PARVANEH S, RAHMAN A, et al. Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds[C]//2016 Computing in Cardiology Conference(CinC). Vancouver: IEEE, 2016: 621-624.
- [19] THOMAE C, DOMINIK A. Using deep gated RNN with a convolutional front end for end-to-end classification of heart sound[C]//2016 Computing in Cardiology Conference(CinC). Vancouver: IEEE, 2016: 625-628.
- [20] XIAO B, XU Y Q, BI X L, et al. Follow the sound of children's heart: A deep-learning-based computer-aided pediatric CHDs diagnosis system[J]. *IEEE Internet of Things Journal*, 2020, 7(3): 1994-2004.
- [21] 王磊, 徐涛, 宋传东, 等. 基于深度学习的 miRNA 与疾病相关性预测算法[J]. *电子学报*, 2020, 48(5): 870-877.
WANG L, XU T, SONG C D, et al. Prediction algorithm of association between miRNAs and diseases based on deep learning[J]. *Acta Electronica Sinica*, 2020, 48(5): 870-877. (in Chinese)
- [22] 周涛, 霍兵强, 陆惠玲, 等. 残差神经网络及其在医学图像处理中的应用研究[J]. *电子学报*, 2020, 48(7): 1436-1447.
ZHOU T, HUO B Q, LU H L, et al. Research on residual neural network and its application on medical image processing[J]. *Acta Electronica Sinica*, 2020, 48(7): 1436-1447. (in Chinese)

作者简介



肖斌 男, 1982年出生, 重庆人. 博士, 教授, 博士生导师, CCF 会员. 主要研究领域为图像增强与复原.

E-mail: xiaobin@cqupt.edu.cn



陈嘉博 男, 1996年出生, 浙江人. 硕士研究生. 主要研究领域为图像识别与分析.