

DOI:10.3969/j.issn.1005-202X.2021.10.012

医学信号处理与医学仪器

基于交叉对比神经网络的心音分类

任灵, 黄玉丹, 陈颖

南京大学电子科学与工程学院, 江苏 南京 210023

【摘要】目的:通过交叉对比神经网络(CCNN)实现心音信号的自动分类,从而对心血管疾病进行早期诊断。**方法:**实验基于PhysioNet/Cinc 2016心音数据库。训练集和测试集数据来自互斥的健康受试者/病理患者,并以4:1的比例进行划分,输入CCNN。CCNN利用深度卷积神经网络进行特征提取,结合基于信息的相似度量理论(IFS),对特征向量间的相似性进行度量并分类。**结果:**实验结果得出灵敏度为0.834 6,特异性为0.962 3,最终大赛综合得分为0.898 5。**结论:**CCNN使用交叉对比的输入模式扩充数据量,引入信号间的对比信息,同时在神经网络的训练过程中应用统计学思想,使网络具备良好的泛化性,更加适应医学数据量较少的场景,在心音分类中取得较好的结果。

【关键词】心音分类;交叉对比神经网络;基于信息的相似度量理论;深度学习

【中图分类号】R318.5

【文献标志码】A

【文章编号】1005-202X(2021)10-1251-07

Heart sound classification based on cross-contrast neural network

REN Ling, HUANG Yudan, CHEN Ying

School of Electrical Science and Engineering, Nanjing University, Nanjing 210023, China

Abstract: Objective To automatically classify heart sound signals by cross-contrast neural network (CCNN), thereby realizing the early diagnosis of cardiovascular diseases. **Methods** The experiment was carried out based on PhysioNet/Cinc 2016 heart sound database. The training set and test set data which came from mutually exclusive healthy subjects/pathological patients were divided at a ratio of 4:1 and then input into CCNN. Finally, CCNN used deep convolutional neural network for feature extraction and was combined with information-based similarity measurement theory to measure and classify the similarity between feature vectors. **Results** The sensitivity and specificity of CCNN for heart sound classification in the experiment were 0.834 6 and 0.962 3, respectively, and the overall score reached 0.898 5. **Conclusion** By expanding the amount of data using a cross-contrast input mode, introducing contrast information between signals and applying statistical ideas in the training process of neural networks, CCNN has good generalization and is more suitable for small medical data, having a good performance in heart sound classification.

Keywords: heart sound classification; cross-contrast neural network; information-based similarity measurement theory; deep learning

前言

心血管疾病目前已经成为人类的头号致死病种^[1],医生能够根据心音信号在心血管疾病的早期发

现症状并提供初步诊断^[2]。心音分割^[3-4]和心音分类^[5-6]是心音分析主要方法。心音分类的典型方法有基于人工神经网络的分类^[7]、基于支持向量机的分类^[8]、基于隐马尔可夫模型^[9-10]和基于聚类的分类^[11]。Liang等^[12]采用波包分解特征与人工神经网络的方法,成功区分了儿童心音信号中的天然杂音和病理性杂音,并且取得了80%的敏感性与90%的特异性。Ari等^[13]采用最小二乘支持向量机的方法判别心音信号是否正常,在测试集上的准确率达到86.72%。Bentley等^[14]使用无监督K近邻(K-Nearest Neighbors, KNN)来区分正常心音和异常心音,在一个大小未知的数据库中报告了100%和87%

【收稿日期】2021-04-13

【基金项目】国家自然科学基金(81671777);江苏省重点研发计划(BE2017679, BE2016733)

【作者简介】任灵,硕士研究生,研究方向:生物医学工程, E-mail: MF1823042@smail.nju.edu.cn

【通信作者】陈颖,副教授,研究生导师,研究方向:生物医学工程, E-mail: yingchen@nju.edu.cn

的准确率。

由于先前的研究缺乏标准化、高质量、严格验证的心音数据库,不同研究之间的直接对比分析受到阻碍。同时许多研究并未单独划分测试集,使得模型在训练集上过拟合。本研究的心音信号来自2016年的PhysioNet挑战赛的心音数据库^[15],选择了Huang等^[16]提出的新型交叉对比神经网络(Cross-Contrast Neural Network, CCNN)区分正常与异常的心音信号。CCNN是一种结合了时间序列分析方法与机器学习的神经网络,机器学习的方法弥补了传统特征提取方法中泛化能力不足、抗干扰性较差的问题,时间序列分析又从统计学的角度为深度学习的可解释性提供理论支持,使得CCNN适用于数据量较小的医学问题。

1 基于CCNN的心音分类原理

1.1 基于信息的相似度量理论(IFS)

Yang等^[17]提出了IFS,观测在两种不同的信号中,同一个序列出现的频率特征。如图1所示,每一个点代表某种序列,其所在位置代表其在信号中出现频次的排名,其中X轴代表时间信号1,Y轴代表时间信号2。结合IFS,图1中点的散乱程度表示两个信号的相似程度。图中的点越集中在对角线附近,则表示两个信号越相似,图中的点越散乱,越远离对角线,两个信号的相似程度越低。

目前IFS的应用领域有:SARS冠状病毒的基因组分类^[18]、医学图像配准^[19]、文学作品作者的鉴别等^[20]。

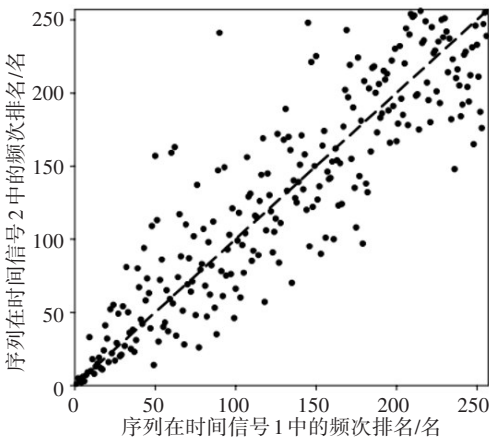


图1 两个时间信号上不同序列特征出现频次秩的散点图
Fig.1 Scatter plot of the rank-frequency distribution of different sequences on two time signals

1.2 CCNN

CCNN由Huang等^[16]在2019年发表文章中首次提出,如图2所示该网络利用VGG的前4层得到特征概率,而后利用改进的IFS方法计算两个输入信号的相似程度,根据众多滤波器判断的结果进行投票,区分正常和异常的心音信号。

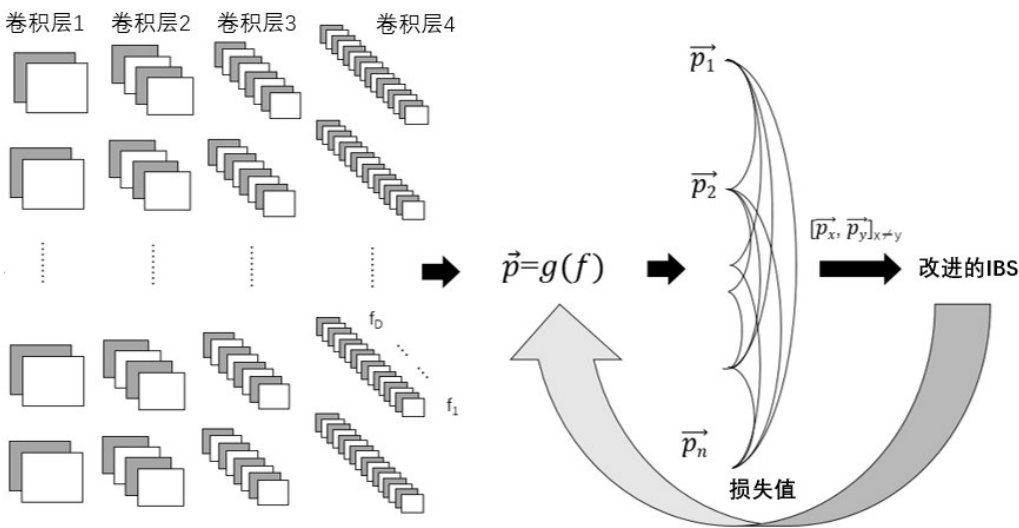


图2 CCNN网络结构
Fig.2 Cross-contrast neural network (CCNN) structure

具体的训练过程如下:(1)将 N 个信号两两配对,组合成 C_N^2 个数据输入对,打上标签(类别相同label=0,不同为label=1);(2)将每批数据输入CCNN的前半

部分,在第四个卷积层获得大小为 $w \times h \times d$ 的特征图($w \times h \times d = 28 \times 28 \times 512$);(3)计算特征图($w \times h$)中大于某个阈值的元素个数,得到 d 维特征向量

$\vec{p} = [p_1, p_2, \dots, p_d]$; (4) 将一个批次 C_N^2 个关系对得到的特征向量 \vec{p}_x, \vec{p}_y 输入改进后的 IBS 模型, 计算损失, 将梯度反向传播给网络的参数, 更新网络的权重; (5) 反复执行过程 (1)~(4), 直到模型在验证集上的损失值不再下降或预测的准确率不再上升为止。

本研究使用 `softsign` 函数代替逻辑判断的阈值计数操作, 同时对特征向量 \vec{p}_n 进行归一化操作, 归一化的方法如下:

$$p_i = g(f_i) = \frac{\sum_j \text{softsign}(\max(f_{ij}, 0))}{\sum_i \sum_j \text{softsign}(\max(f_{ij}, 0))}, i \in [1, d], j \in [1, M]$$

(1)

其中, f_{ij} 代表第 i 号特征图中的第 j 个像素值, d 和 M 分别是最后一层卷积层的维度和输出特征图的像素个数。CCNN 在原有 IBS 方法的基础上提出一种改进算法—Modified IBS 算法 (简称 ModIBS), 将 IBS 的值归一化到 0~1。设原始 IBS 值为 Normal IBS, 简称 NorIBS, 公式如下:

$$\text{NorIBS} = \frac{1}{d} \sum_i |R_{x,i} - R_{y,i}| \frac{F_i}{Z}$$

(2)

$$F_i = E_x(p_i) + E_y(p_i)$$

(3)

$$E(t) = -t \log(t)$$

(4)

$$Z = \sum F_i$$

(5)

其中, $R_{x,i}$ 是第 x 个数据输出概率 \vec{p}_x 里第 i 个概率值的排名, F_i 是第 i 个滤波器在输入 x 和 y 中的概率熵之和, Z 是归一化系数。

随机打乱概率向量 \vec{p}_x, \vec{p}_y 的其中之一, 得到混乱状态下的 IBS 值, 命名为 RevIBS。

$$\text{RevIBS} = \text{NorIBS}(\text{random}(\vec{p}_y))$$

(6)

$$\text{ModIBS} = \exp - \left(\frac{\text{RevIBS} - \text{NorIBS}}{\text{NorIBS}} \right)^\alpha$$

(7)

α 为指数放大因子, 最终的损失函数采用交叉熵计算:

$$\text{loss} = (1 - \text{label}) \times \log(1 - \text{ModIBS}) + \text{label} \times \log(\text{ModIBS})$$

(8)

2 实验方法

2.1 预处理

本研究在 2016 年 PhysioNet/Cinc 挑战赛提供的开源心声数据集上展开^[15], 原始数据为 2 000 Hz 的 wav 格式心音信号。数据预处理流程如下: 首先, 读入 2 000 Hz 采样的一维时间序列, 对原序列进行 1 000 Hz 的降采样。然后, 根据心音信号的频率范围, 使用 25~400 Hz 的带通滤波器进行滤波。最后, 对序列做平滑处理。心音信号的预处理过程如图 3 所示。

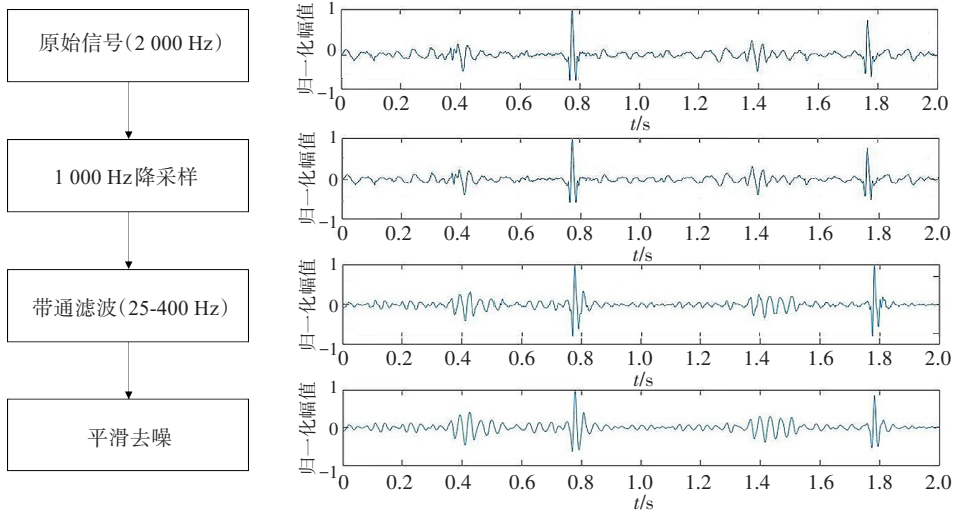


图3 心音信号预处理流程图
Fig.3 Preprocessing of heart sound signals

本研究选取 VGG 网络进行特征提取, 其输入信号的维度为 $224 \times 224 \times 3$ 。采用长度为 512 的窗口截取原始信号并叠加汉明窗, 通过短时傅里叶变换得到频谱图, 将频谱图的前 224 个幅度特征存储为 1×224 大小的数组。以步长 256 滑动窗口, 重复上述步骤, 直到取满 224 个窗口为止。对每一个 1×224 的特征

进行一阶差分和二阶差分并和原数组叠加, 最终输出 $224 \times 224 \times 3$ 的频谱图。在变换后的数据上划分训练集与测试集 (比例为 4:1), 使用训练集进行 CCNN 网络迭代训练。

2.2 训练与测试

本研究的训练与测试流程如图 4 所示。

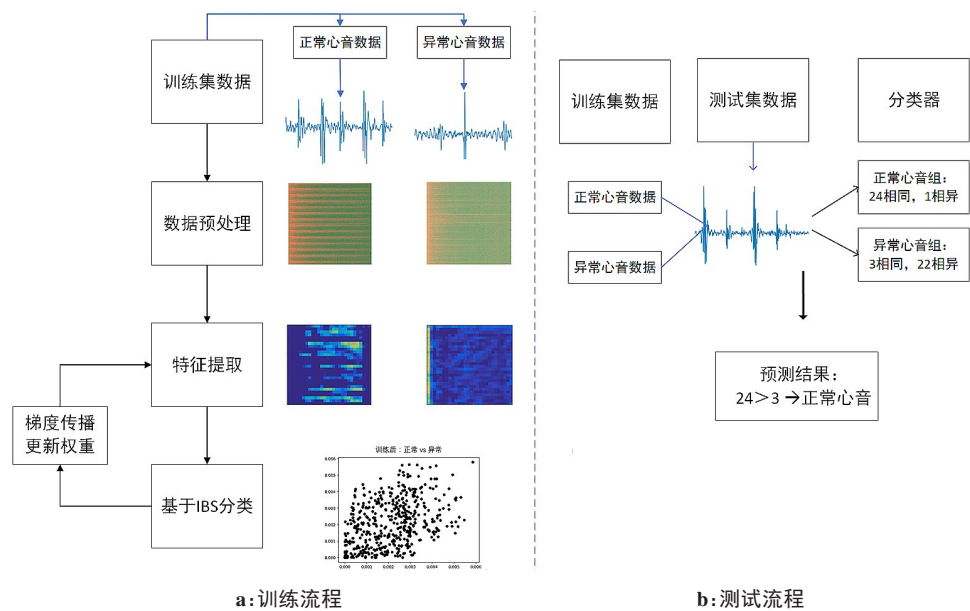


图4 基于CCNN的心音信号分类的训练和测试流程图

Fig.4 Flowchart of training and testing of heart sound signal classification based on CCNN

2.2.1 模型训练 假设一个批次的数据大小为 N 张图像,两两配对得到 C_N^2 个数据输入对。由于计算平台 GPU 内存有限,本文采取批量梯度下降的方法对网络进行训练。为了保证输入数据的 0、1 平衡性,具体数据输入方法如下:(1)每次从训练集中顺序读取 1 个锚点文件,假设其为正常心音;(2)从训练集剩余的文件中分别挑选 $x-1$ 张和 y 张正常、异常心音频谱图,使得以下公式成立:

$$\begin{cases} \frac{C_x^2 + C_y^2}{C_N^2} = \frac{1}{2}, \text{解出 } x = \frac{2N \pm \sqrt{3N^2 + N}}{4} \\ x + y = N \end{cases} \quad (9)$$

(3)选取正常心音和异常心音的个数分别为 x 和 $N-x$ 的近似正整数,组成数据大小为 N 的批次,输入网络;(4)计算损失,将梯度反向传播给网络的参数,更新网络的权重;(5)重复步骤(1)~(4),直到遍历完训练集中的所有锚点文件。

为了解决网络的过拟合与陷入局部最优的问题,本研究在 CCNN 的基础上,加入了可动态调整的学习率和 L2 正则化。其中,学习率决定梯度对 CCNN 网络权重调整幅度的大小;L2 正则化则可以提高模型的泛化能力并且防止过拟合。图 5 和图 6 展示了训练过程。

训练过程中随着训练步数的推进,训练集的准确率逐渐上升,损失值逐渐下降,同时真阴率(TN)、真阳率(TP)也稳步上升。本研究划分出 20% 数据作为验证集,为了避免过拟合问题,当模型在验证集上的准确率开始下降时停止训练。这样的训练方式在维持数据输入均衡性的同时保证训练的稳定性,通

过将信号两两组成信号对的方法,有效填补数据,缓解了深度神经网络难以在小数据集上收敛的问题。

2.2.2 模型测试 将第 2.2.1 小节训练得到的模型用于心音信号的测试。采用与训练集数据完全相同的预处理方式将未知正常或异常的心音信号转化为频谱图,并且将未知心音信号与训练集中的正常和异常心音信号进行两两组合构成信号对,如图 4b 所示,以预测结果为 0 (两幅图像来源于同一分类)最多的类别作为最终分类结果。

3 模型评估

3.1 评价准则

心音记录的分类分为 3 个层次:正常、异常、不确定,总分根据分类的各项指标综合计算^[6]。

总体得分 Score、灵敏度 Se、特异性 Sp 定义如下:

$$\text{Score} = \frac{\text{Se} + \text{Sp}}{2} \quad (10)$$

$$\text{Se} = \frac{wa_1 \cdot Aa_1}{Aa_1 + Aq_1 + An_1} + \frac{wa_2 \cdot (Aa_2 + Aq_2)}{Aa_2 + Aq_2 + An_2} \quad (11)$$

$$\text{Sp} = \frac{wn_1 \cdot Nn_1}{Na_1 + Nq_1 + Nn_1} + \frac{wn_2 \cdot (Nn_2 + Nq_2)}{Na_2 + Nq_2 + Nn_2} \quad (12)$$

针对本研究中的数据, $wa_1=0.8602$, $wa_2=0.1398$, $wn_1=0.9252$, $wn_2=0.0748$ 。其中参数的定义如表 1 所示^[6]。

3.2 测试结果

本研究在测试集上进行了 10 次 5 折交叉验证,结果如表 2 所示。本研究选取了常用的机器学习方法与 CCNN 进行对比,结果如表 3 所示。可以看到,

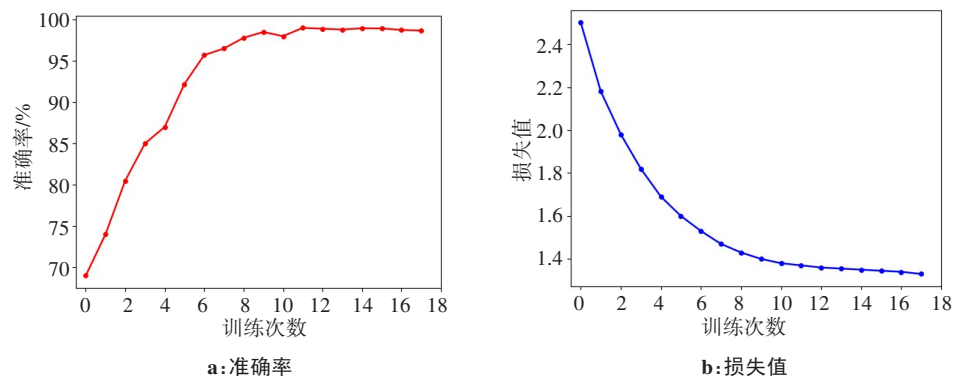


图5 准确率和损失值随着训练步骤进行的变化趋势
Fig.5 Variation trends of accuracy and loss value with the training steps

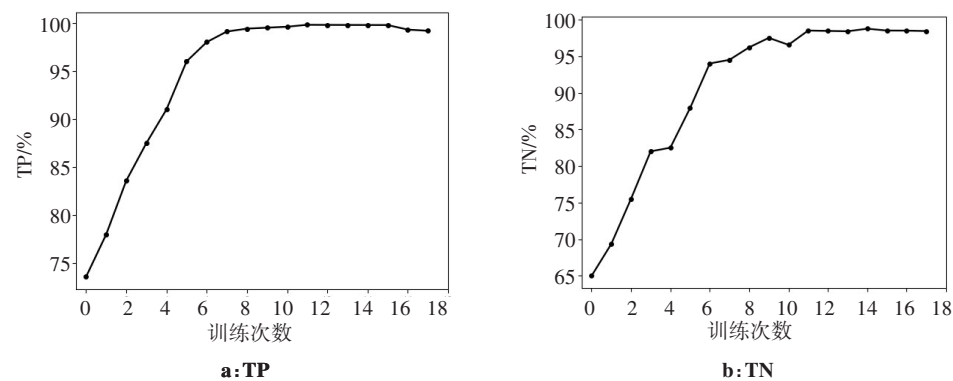


图6 TP和TN随着训练步骤进行的变化趋势
Fig.6 Variation trends of TP and TN with the training steps

表1 分类结果具体定义
Tab.1 Definitions of the classification results

参考标签	信号质量	权重	预测类别		
			异常(1)	不确定(0)	正常(-1)
异常	好	wa_1	Aa_1	Aq_1	An_1
异常	差	wa_2	Aa_2	Aq_2	An_2
正常	好	wn_1	Na_1	Nq_1	Nn_1
正常	差	wn_2	Na_2	Nq_2	Nn_2

上述几种方法中,CCNN对心音分类的效果最好。CCNN的预期是经过训练后,同一类别的两幅图像中的点应当聚集在对角线附近,而对于不同类别的两种信号,其特征分布则呈现出散乱的状态。为了验证实验预期,本文选择3段心音信号,其中2段来自正常心音信号,1段来自异常心音信号。

观察这两组心音信号的特征分布在训练前后的变化。图7展示的为第四层卷积在训练前和训练后的特征分布。可以看到,对于相同类别的信号,从图7a到图7b呈现出往对角线聚集的趋势,而对于不同类别的信号,从图7c到图7d呈现远离对角线的散乱趋势。图8展示了3段不同心音信号的原始频谱图(图像大小224×224)与通过几种滤波器后输出的频

表2 CCNN测试结果
Tab.2 CCNN test results

编号	灵敏度	特异性	得分
1	0.797 0	0.974 8	0.885 9
2	0.819 5	0.963 1	0.891 3
3	0.969 9	0.994 2	0.982 1
4	0.834 6	0.953 4	0.894 0
5	0.932 3	0.937 9	0.935 1
6	0.872 2	0.934 0	0.903 1
7	0.751 9	0.970 9	0.861 4
8	0.661 7	0.982 5	0.822 1
9	0.872 2	0.943 7	0.908 0
10	0.834 6	0.968 9	0.901 8
均值	0.834 6	0.962 3	0.898 5

谱图(图像大小28×28)。
从心音信号的维度来看,同一个滤波器对不同的心音信号有着不同的响应。可以明显地看出,394号滤波器对正常心音信号响应更强,27号滤波器在异常信号上的响应更强,而72号滤波器不管对正常信号还是异常信号的响应都没有明显的强弱关系,代表了某种区分度较小的模式。从这3个典型的滤

表3 常用的机器学习方法与CCNN进行对比
Tab.3 Comparison between commonly-used machine learning methods and CCNN

模型	灵敏度	特异性	得分
随机森林	0.540	0.950	0.745
KNN	0.340	0.930	0.635
朴素贝叶斯	0.560	0.820	0.690
Adaboost	0.590	0.910	0.750
GBDT	0.640	0.930	0.785
XGBoost	0.550	0.920	0.740
CCNN	0.835	0.962	0.899

波可以看出,不同滤波器表现出不同的区分模式,而这些不同的滤波器,在一次次迭代的过程中更新它们的权重,最终可以得到一组最能区分正常与异常心音信号的权重数据,达到了最初对CCNN结合时间序列分析方法的设想。

这种新的网络结构与训练方法,同时利用深度卷积神经网络与IBS方法的优点,针对医学数据做出了改进,一定程度上克服了诸如数据量少、数据集差异不明显的弊端。而与传统机器学习中人工设计特征的方法相比,深度神经卷积网络的引入可以由网络提取出对分类有用的特征,大大简化特征提取流

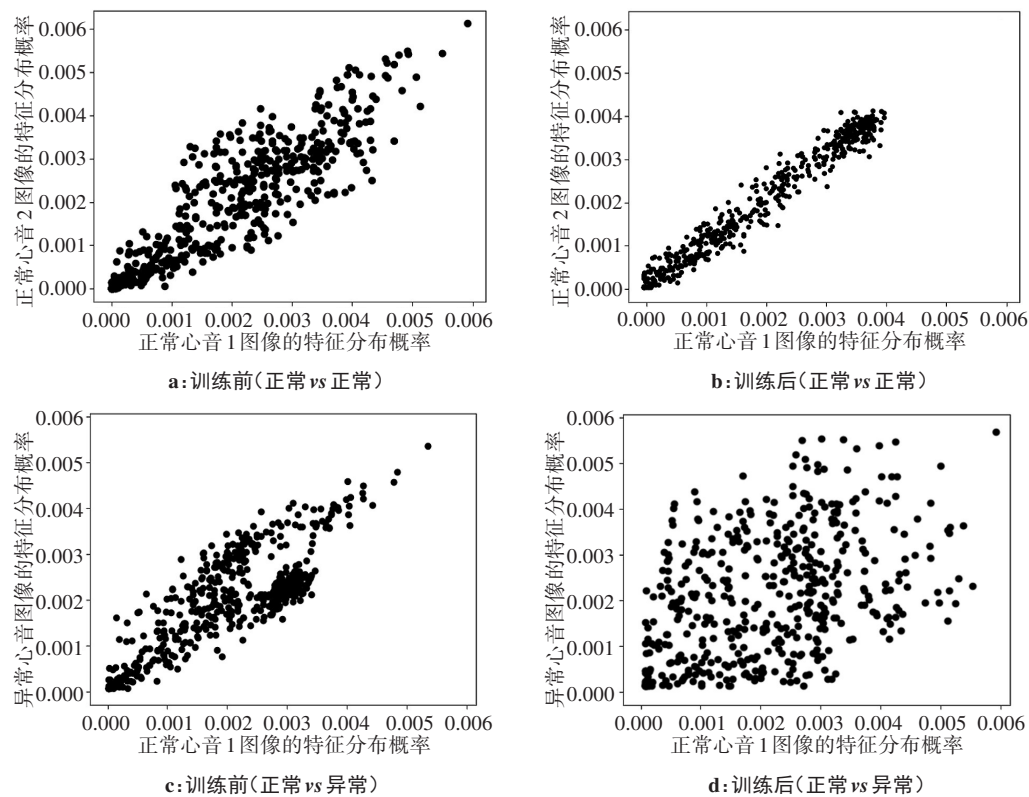


图7 模型训练过程中分布的变化
Fig.7 Distribution changes during model training

程。其中两两组合信号的方式给网络提供更多的信息,放大不同类别心音信号间的差异性,并且一定程度上扩充了心音数据的数据量。这种方法尤其适用于数据难以获取,并且数据数量较少的医学场景。

4 结 语

本文受到IBS与CNN的启发,将新型深度神经网络CCNN用于心音信号的分类。在此基础上,采用了IBS中序列相似性的度量方法,借助CNN的强大特征表现能力,对动态生成的信号进行分类。本

文采用的训练方法主要有以下3个主要特点:(1)结合统计学的思想,将先验知识加入神经网络的特征提取过程,使网络具备良好的泛化性;(2)使用交叉对比的输入模式扩展数据集,引入了除内容信息之外的对比信息;(3)CCNN前半部分的特征抽取网络架构灵活易调整。

综上所述,本文利用CCNN在心音数据集上取得了较好的结果(灵敏度为0.834 6,特异性为0.962 3,综合得分为0.898 5),证明CCNN在医学数据集上的可行性,为小数据深度学习提供一种新思路。

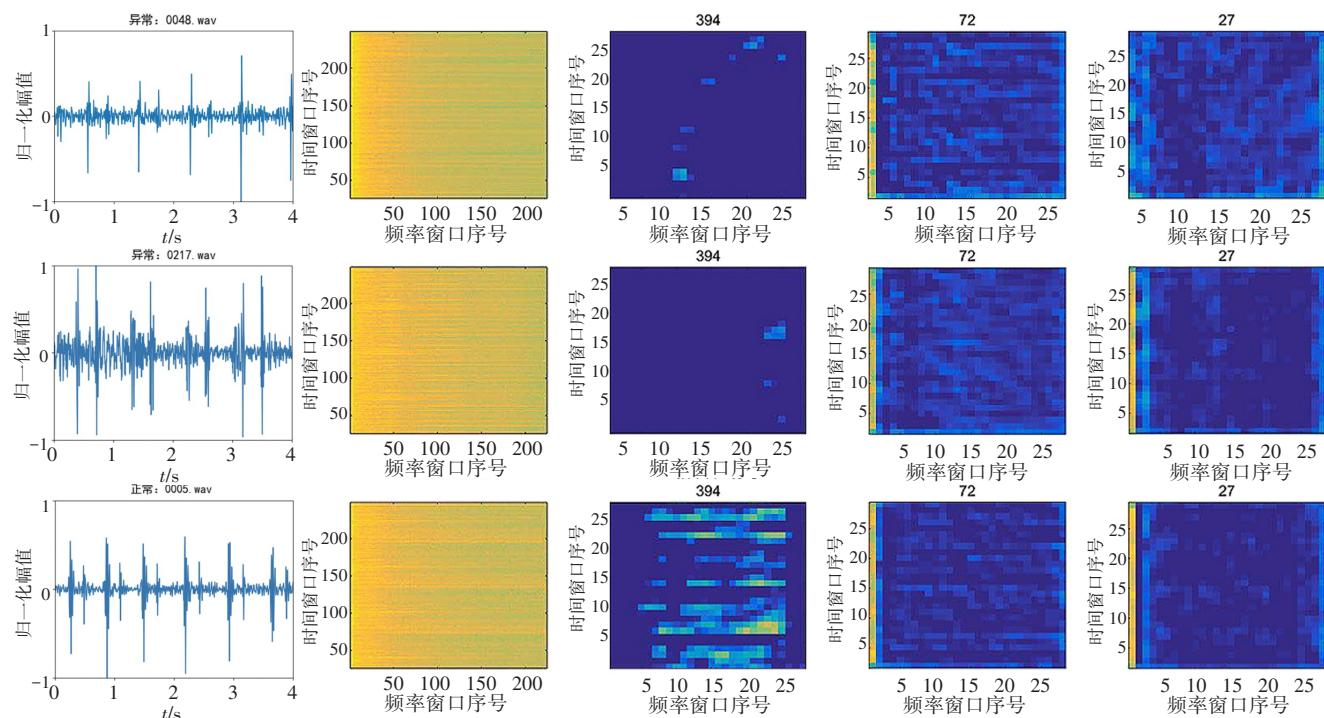


图 8 不同心音信号的原始频谱图与通过几种滤波器后的频谱图

Fig.8 Original spectrograms of different heart sound signals and spectrograms filtered by different filters

【参考文献】

[1] 中国心血管病预防指南(2017)写作组, 中华心血管病杂志编辑委员会. 中国心血管病预防指南(2017)[J]. 中华心血管病杂志, 2018, 46(1): 10-25.
Chinese Cardiovascular Disease Prevention Guidelines (2017) Writing Group, Editorial Board of Chinese Journal of Cardiovascular Diseases. Chinese cardiovascular disease prevention guidelines (2017) [J]. Chinese Journal of Cardiovascular Diseases, 2018, 46(1): 10-25.

[2] RAGHU A, PRAVEEN D, PEIRIS D, et al. Engineering a mobile health tool for resource-poor settings to assess and manage cardiovascular disease risk: SMARThealth study [J]. BMC Med Inform Decis Mak, 2015, 15(1): 36.

[3] PAPADANIIL C D, HADJILEONTIADIS L J. Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features [J]. IEEE J Biomed Health Inform, 2014, 18(4): 1138-1152.

[4] MESSNER E, ZHRER M, PERNKOPF F. Heart sound segmentation-an event detection approach using deep recurrent neural networks[J]. IEEE Trans Biomed Eng, 2018, 65(9): 1964-1974.

[5] YADAV A, DUTTA M K, TRAVIESO C M, et al. Automatic classification of normal and abnormal PCG recording heart sound recording using fourier transform[C]//2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI). IEEE, 2018: 1-9.

[6] CLIFFORD G D, LIU C, MOODY B, et al. Classification of normal/ abnormal heart sound recordings: the physionet/computing in cardiology challenge 2016[C]//Computing in Cardiology. IEEE, 2017: 593-596.

[7] JADHAV S M, NALBALWAR S L, GHATOL A. Artificial neural network based cardiac arrhythmia classification using ECG signal data [C]//International Conference on Electronics & Information Engineering. IEEE, 2010: 228-231.

[8] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[J]. ACM Trans Intel Syst Tech, 2011, 2(3): 1-27.

[9] PING W, CHU S L, CHAUHAN S, et al. Phonocardiographic signal analysis method using a modified hidden markov model [J]. Ann Biomed Eng, 2007, 35(3): 367-374.

[10] SARACOGLU R. Hidden markov model-based classification of heart valve disease with PCA for dimension reduction[J]. Eng Appl Artif Intel, 2012, 25(7): 1523-1528.

[11] AVENDAO-VALENCIA L D, GODINO-LLORENTE J I, BLANCO-VELASCO M, et al. Feature extraction from parametric time-frequency representations for heart murmur detection[J]. Ann Biomed Eng, 2010, 38(8): 2716-2732.

[12] LIANG H, SAKARI L, IIRO H. A heart sound segmentation algorithm using wavelet decomposition and reconstruction [C]//International Conference of the IEEE Engineering in Medicine & Biology Society. IEEE, 1997: 1630-1633.

[13] ARI S, HEMBRAM K, SAHA G. Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier[J]. Expert Syst Appl, 2010, 37(12): 8019-8026.

[14] BENTLEY P M, GRANT P M, MCDONNELL J T, et al. Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds[J]. IEEE Trans Biomed Eng, 1998, 45(1): 125-128.

[15] LIU C, 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.

[16] HUANG Y, CHEN Y, ZHU H, et al. A liver fibrosis staging method using cross-contrast network[J]. Expert Syst Appl, 2019, 130: 124-131.

[17] YANG C C, HSEU S S, YIEN H W, et al. Linguistic analysis of the human heartbeat using frequency and rank order statistics[J]. Phys Rev Lett, 2003, 90(10): 108103.

[18] YANG C C, GOLDBERGER A L, PENG C K. Genomic classification using an information-based similarity index: application to the SARS coronavirus[J]. J Comput Biol, 2005, 12(8): 1103-1116.

[19] REEL P S, DOOLEY L S, WONG P. A new mutual information based similarity measure for medical image registration [C]//Image Processing. IET, 2012: 123.

[20] YANG C C, PENG C K, YIEN H W, et al. Information categorization approach to literary authorship disputes[J]. Physica A, 2003, 329(3-4): 473-483.

(编辑:陈丽霞)