

# 基于混合时频域特征的卷积神经网络心律失常分类方法的研究

吕 杭<sup>1</sup>, 蒋明峰<sup>1</sup>, 李 杨<sup>1</sup>, 张鞠成<sup>2</sup>, 王志康<sup>2</sup>

(1. 浙江理工大学计算机科学与技术学院, 浙江杭州 310018; 2. 浙江大学医学院附属第二医院, 浙江杭州 310009)

**摘 要:** 心律失常是常见的心血管疾病之一, 目前很多方法通过计算机辅助系统对心电图进行分析以识别心律失常, 但由于大多数心律失常数据样本较少, 计算机辅助系统识别心律失常效果不佳. 本文提出了一种基于混合时频域分析特征提取的卷积神经网络方法, 该方法提取心电图的RR间期时域特征、希尔伯特-黄变换提取的频域特征和连续小波变换提取的时频域联合特征, 经过特征融合后输入卷积神经网络训练分类模型, 并采用Focal Loss作为网络的损失函数, 实现对心律失常的分类. 本文使用MIT-BIH (Massachusetts Institute of Technology-Boston's Beth Israel Hospital) 心律失常数据库验证本文提出方法对4类心电数据分类的结果, 实验结果表明, 与现有的分类算法相比, 本文所提出的混合时频域特征方法能有效提升心律失常分类的准确性.

**关键词:** 时频域分析; 连续小波变换; 希尔伯特-黄变换; 心律失常分类; Focal Loss; 卷积神经网络

**基金项目:** 浙江省科技厅重点研发项目(No.2020C03060); 国家自然科学基金(No.62101497); 浙江省自然科学基金-数理医学学会联合基金重点项目(No.LSZ19F010001)

**中图分类号:** TP391.41

**文献标识码:** A

**文章编号:** 0372-2112(2023)03-0701-11

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

**DOI:** 10.12263/DZXB.20211181

## Research on Arrhythmia Classification by Using Convolutional Neural Network with Mixed Time-Frequency Domain Features

LÜ Hang<sup>1</sup>, JIANG Ming-feng<sup>1</sup>, LI Yang<sup>1</sup>, ZHANG Ju-cheng<sup>2</sup>, WANG Zhi-kang<sup>2</sup>

(1. School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou, Zhejiang 310018, China;

2. The Second Affiliated Hospital, School of Medicine Zhejiang University, Hangzhou, Zhejiang 310009, China)

**Abstract:** Arrhythmia is one of cardiovascular diseases, and many methods are used to analyze electrocardiogram by computer-aided system to identify arrhythmia. However, most of the data samples of arrhythmia are small, and the computer-aided system is not effective in identifying arrhythmia. In this paper, a mixed time-frequency domain feature extraction method is proposed for arrhythmia classification by using convolution neural network method. The fused features consist of the time domain characteristics from the RR interval, frequency domain characteristics from hilbert-huang transform, and joint time-frequency domain features extracted from continuous wavelet transform. Then the fused features are used as an input to the convolution neural network for training classification model, and the focal loss is used as the loss function of the training model, so as to realize the arrhythmias classification. In addition, the MIT-BIH (Massachusetts Institute of Technology-Boston's Beth Israel Hospital) arrhythmia database is used to verify the performances of the proposed method for arrhythmias classification of four types of ECG (Electrocardiograph) data. Experimental results show that compared with the existing classification algorithms, the proposed method improves the F1 of class obviously.

**Key words:** time-frequency domain analysis; continuous wavelet transform; hilbert-huang transform; arrhythmia classification; focal loss; convolutional neural network

**Foundation Item(s):** Key Research and Development Program of Zhejiang Province (No.2020C03060); National Natural Science Foundation of China (No.62101497); Joint Fund of Zhejiang Provincial Natural Science Foundation (No.LSZ19F010001).

## 1 引言

近年来,心血管疾病(CardioVascular Diseases, CVDs)的发病率在全球范围内大幅增加,其中心律失常是CVDs<sup>[1,2]</sup>死亡的主要原因之一. 大多数心律失常并不严重,但有些是有害的,甚至危及生命. 早期发现心律失常并有效干涉可以预防CVDs的发生,在临床中,心律失常通常是通过分析心电图(ElectroCardioGraph, ECG)<sup>[3]</sup>来诊断的. 心电图是一种无创的测量心脏电活动的方法,通常用于监测和评估病人的心脏状态. 发生心律失常时心脏会出现异常电活动,心电图的P波、QRS波和T波的形态学发生变化<sup>[4]</sup>,因此心电图可作为鉴别诊断的一部分. 但是监测这一过程既耗时又费力,而在计算机辅助智能诊断下<sup>[5]</sup>,可以很大程度提高心电图的解释和帮助心血管病的处理. 因此,基于深度学习的计算机辅助工具有很大的潜力来帮助医生提供更好和更快速的心律失常诊断.

心律失常的分类研究主要分为传统的特征提取与机器学习相结合的方法和深度学习方法. 传统的分析方法将机器学习,如支持向量机(Support Vector Machine, SVM)和随机森林(Random Forest, RF),与根据经验提取的时间和频率特征等多种特征相结合. Daqrouq等人<sup>[6]</sup>从心电记录中提取小波特征,然后输入概率神经网络识别房颤,是机器学习和特征提取的典型结合. Wang等人<sup>[7]</sup>通过计算RR区间特征、基于小波包分解的形态学特征和统计特征输入RF进行心律失常分类. 深度学习方面的方法,一维卷积神经网络(Convolutional Neural Network, CNN)被广泛用于心律失常的分类,在识别心房颤动等心律失常方面优于循环神经网络<sup>[8-10]</sup>.

一维深度学习网络在心律失常分类中具有较好的快速效果,但只分析了心电记录的形态特征,忽略了其频率和能量分布信息. Çinar等人<sup>[11]</sup>采用短时傅立叶变换将心电记录转换为二维谱图,获得心电记录中的频率和能量信息,用于对正常窦性节律、异常心律失常进行分类. Zhang等人<sup>[12]</sup>提出了一种基于ResNet-101的混合时频分析和迁移学习的心跳分类方法,该方法采用希尔伯特变换(Hilbert Transform, HT)和Wigner-Ville分布的混合时频分析,将一维心电记录转换为二维时频图,然后将这些时频图输入基于ResNet-101的迁移学习分类器,进行分类任务. Sabut等人<sup>[13]</sup>提出了一种基于混合时频特征的深度神经网络方法的室性心动过速和室颤分类方案,采用小波变换、经验模态分解和变模态分解方法对信号进行分解. Xu等人<sup>[14]</sup>采用基于两个心拍的R峰区段(RR间期)、离散小波变换(Discrete Wavelet Transform, DWT)分解心电信号的混合时域和小波时频特征,利用极限学习机作为分类器,他们还证实了不同种类的心律失常的RR间期也会随之不同. Wang等人<sup>[15]</sup>提出了一种基于连续小波变换和卷积神经网络的ECG

自动分类方法,利用连续小波变换(Continuous Wavelet Transform, CWT)对ECG信号进行分解,得到不同的时频分量,再用卷积神经网络对上述时频分量组成的二维尺度图进行特征提取,也提取4个RR间期特征并结合卷积神经网络特征输入到全连接层进行心电图分类.

目前使用较多的时频技术<sup>[16]</sup>为CWT和短时傅里叶变换(Short Time Fourier Transform, STFT)<sup>[17,18]</sup>. CWT继承并发展了STFT的思想,但与STFT不同的是,CWT不仅可以在低频处提供高频分辨率和低时间分辨率,而且在高频处具有高时间分辨率和低频分辨率. 还有一种希尔伯特黄变换(Hilbert-Huang Transform, HHT)的时频分析方法<sup>[19]</sup>,HHT是数据驱动和后验的自适应模式分解,可以自适应提取ECG信号的单组分量性质,并将信号表示为几个单组分量震荡和趋势的叠加,这种单分量分解特性能够精确估计每个组成部分的瞬时频率和瞬时幅度,这些瞬时的参数可以深入了解信号的频率成分和时间变化. 由于心律失常分类的复杂性,单一的RR间期特征、频域特征、时频特征的心律失常分类准确性不高,所以本文提出了一个基于混合时频域特征的卷积神经网络心律失常识别的方法.

## 2 方法

本文研究的方法是从一维时域ECG信号序列中获取RR间期特征、时域信号经HHT后得到的不同频率段的瞬时频率特征和经CWT后得到的二维时频能量特征,最后将特征级联后输入CNN分类模型. 该方法如图1所示,包括预处理、特征提取和CNN分类模型组成. 预处理包括ECG去噪、心拍分割和RR间期提取. 特征提取包括HHT提取的瞬时频率特征和CWT的时频能量特征,CNN分类模型由三个卷积层组成,下面将逐一介绍.

### 2.1 数据集

本文中用于训练模型的ECG数据来自MIT-BIH (Massachusetts Institute of Technology-Boston's Beth Israel Hospital)的心律失常数据库,该数据库是以采样率360 Hz从4 000份24小时动态心电图记录中,随机选取23份记录临床常规的代表性样本和包含罕见但临床显著的心律失常的25份记录. 数据集共包含47名受试者收集的48段带注释的半小时双导ECG记录<sup>[20]</sup>. MIT-BIH数据的每条记录由MLII(Modified Lead II)导联(上部信号:改良的肢体导联II)和V5导联(下部信号:改良的导联V5)信号组成,MLII导联经过修正,所以本文使用MLII导联的ECG数据进行实验.

MIT-BIH的心律失常数据库的记录还包含由两名或两名以上的专业医生分别标记的15种心律失常类型,根据ANSI/AAMI EC57-2012的标准<sup>[21]</sup>建议,我们进

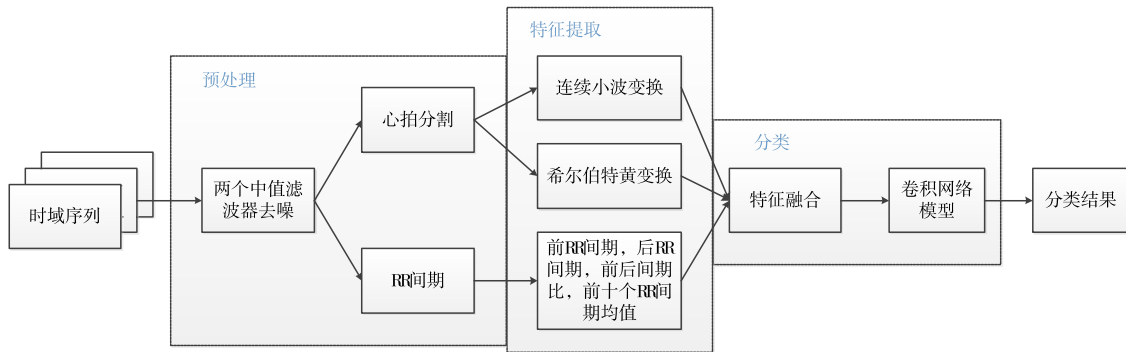


图1 方法的总体流程

一步将这些心律失常分为五类,并且还去除了四项记录(即 102、104、107 和 217),具体如表 1 所示.此外,由于 Q 类没有应用价值,在实验过程中将其忽略,不参与本文方法的评价<sup>[22,23]</sup>.

表1 MIT-BIH数据集根据AAMI标准划分的5种心拍

N	S	V	F	Q
正常搏动(N)	房性早搏(A)	室性早搏(V)	心室和正常搏动的融合(F)	起搏心跳
左束支传导阻滞搏(L)	异常房性早搏(a)	室性逸搏(E)		起搏和正常搏动的融合(f)
右束支传导阻滞搏(R)	交界性早搏(J)			未分类搏动(Q)
心房逸搏(e)	室上性早搏或异位性搏动(S)			
交界性逸搏(j)				

为了便于与现有工作进行直接比较,将 MIT-BIH 心律失常数据库分为训练集(DS1)和测试集(DS2)数据集<sup>[24]</sup>,DS1 和 DS2 都由 22 条记录组成,这些记录的心拍类型比例相似. DS1 用于训练模型,DS2 用于测试方法的性能,DS1 和 DS2 的具体样本数如表 2 所示.

表2 训练集(DS1)和测试集(DS2)ECG样本

数据集/标签	N	S	V	F	总
DS1	45 824	3 788	943	414	50 969
DS2	44 218	3 219	1 836	388	49 661

## 2.2 特征提取方法研究

ECG 分类中存在一个问题,由于 ECG 信号是由不同的频率分量组成的,增加了基于深度学习的方法提取特征的难度.本研究计算 RR 间期特征信息,利用 CWT 将 ECG 信号变换为时频域特征,以及 HHT 获取 ECG 信号不同频率成分的瞬时频率.临床采集的 ECG 信号通常会受到基线漂移、肌电干扰、工频干扰等各种噪声的影响.通常在提取特征之前做一个预处理,通过滤波处理后再提取特征,因为过度滤波会导致有用信息的丢失,所以本文只去除基线漂移<sup>[25]</sup>.本文采用了两个中值滤波器(一个宽度为 200 ms 的中值滤波器和一个宽度为 600 ms 的中值滤波器)来模拟基线漂移,然后将其从原始信号中减去,得到基线校正后的 ECG 信号<sup>[15]</sup>.首先将 ECG 序列分割成单个心拍的形式,以单个心拍作为样本输入.本文是以 MIT-BIH 心律失常数据库已经标注好的 R 峰位置为基准点,采集 R 峰前 100

个样本点和 R 峰后 150 个样本点,得到固定大小为 250 个样本点的样本,将 ECG 信号分割成样本序列.

### 2.2.1 RR 间期特征

本文提取了四种较常用的 RR 间期特征<sup>[26]</sup>,分别为当前心拍和前一个心拍的前 RR 间期,当前心拍和下一个心拍的后 RR 间期,前 RR 间期和后 RR 间期的比值和前 10 个前 RR 间期平均值的局部 RR 间期.同时前 RR 间期、后 RR 间期和局部 RR 间期都减去了平均 RR 间期,用于消除患者之间的差异.

### 2.2.2 基于连续小波变换的时频域特征

CWT 是一种广泛应用于信号时频分析的方法,它克服了傅里叶变换的局限性,它与 STFT 不同的是,CWT 通过调整尺度和平移参数<sup>[27]</sup>分离信号成分,产生信号的时频表示,比传统的使用固定长度的滑动时间窗口的 STFT 更精确.

本文中用于 ECG 信号小波变换的原型小波定义为:

$$C_a(b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \varphi\left(\frac{t-b}{a}\right) dt \quad (1)$$

其中  $a$  为尺度参数, $b$  为平移参数, $\varphi(t)$  为小波函数(也称为母小波).尺度可通过式(2)转换为频率, $F_c$  是母小波的中心频率, $f_s$  为样本  $x(t)$  的采样频率<sup>[28]</sup>.

$$F = \frac{F_c \cdot f_s}{a} \quad (2)$$

母小波的选择往往是影响时频分析效果的关键.

本文参考<sup>[15]</sup>选择小波 Mexican hat (mexh) 作为母小波, 如下式:

$$\varphi(t) = \frac{2}{\sqrt{3} \sqrt[4]{\pi}} \exp\left(-\frac{t^2}{2}\right) (1-t^2) \quad (3)$$

因为它与 QRS 波形状相近, 在 ECG 信号分析中得到了广泛的应用. 利用小波变换的不同尺度因子, 得到了信号在不同尺度下的小波系数. 这些小波系数可以看作是 ECG 信号在时频域的二维尺度.

### 2.2.3 基于希尔伯特-黄变换的瞬时频率特征

希尔伯特-黄变换 (Hilbert-Huang Transform, HHT)<sup>[29]</sup> 作为一种强大的时频技术在非平稳信号处理中得到了广泛的应用. HHT 执行数据自适应分解, 将信号分解为一组接近正交的分量, 它分为两个步骤, 即经验模态分解 (Empirical Mode Decomposition, EMD) 和希尔伯特变换. EMD 将信号分解为一系列本征模态函数 (Intrinsic Mode Function, IMF) 的振荡模态和一个残差. 对  $x(t)$  应用 EMD 分解后得到  $n$  个 IMF (记作  $c_i$ ) 和残差 (记作  $r_n$ ) 如下:

$$x(t) = \sum_{i=1}^n c_i + r_n \quad (4)$$

得到所有的 IMF 后, 将每个 IMF 进行 HT, 即把信号  $x(t)$  与函数  $h(t) = \frac{1}{\pi t}$ , 进行卷积运算, 公式如下:

$$H(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau \quad (5)$$

因此解析的信号  $z(t)$  可以由  $x(t)$  和  $H(t)$  表示为:

$$z(t) = x(t) + iH(t) = a(t)e^{i\phi(t)} \quad (6)$$

其中:

$$\begin{aligned} a(t) &= [x^2(t) + H^2(t)]^{1/2}, \\ \phi(t) &= \arctan(H(t)/x(t)) \end{aligned} \quad (7)$$

$a(t)$  和  $\phi(t)$  分别是  $x(t)$  的瞬时幅值和瞬时相位. 当  $x(t)$  为单分量, 得到瞬时频率  $\omega(t)$  如式 (8) 所示:

$$\omega(t) = \frac{d\phi(t)}{dt} \quad (8)$$

应用 EMD 生成的  $x(t)$  的 IMF 是单分量的, 所以  $x(t)$  可以表示为:

$$x(t) = \sum_{k=1}^n a_k(t) \exp\left(i \int \omega_k(t) dt\right) \quad (9)$$

每个心拍样本  $x(t)$  应用 EMD 后, 得到每个 IMF 对应的瞬时频率特征.

### 2.3 基于卷积神经网络的 ECG 分类模型

卷积神经网络 (CNN) 使用卷积运算来代替神经网络中的一般乘法, 其中卷积层一般通过卷积核的权重来提取特征, 可以通过增加卷积层数来提取相关的高层特征; ReLU 层作为激活函数实现非线性功能; 批归一化层 (Batch Normalization layer, BN) 通常位于卷积层和

ReLU 层之间, 对每个通道的特征图进行归一化处理, 减少训练时间和网络初始化的敏感性; 池化层, 用于降低特征维数, 加快训练过程<sup>[28,30]</sup>. CNN 的最后一层通常是全连接层 (Fully Connected layer, FC), 用于将 CNN 提取的特征表示整合到一起. 在本研究中, 我们搭建了一个基于 CNN 的 ECG 分类模型, 如图 2 所示, 将卷积层、批归一化层、ReLU 层和池化层组合成一个卷积单元, 采用连续三次卷积操作提取特征. CNN 分类模型详细信息如表 3 所示, 输入层 1 为 CWT 得到的特征, 输入层 2 为 RR 间期特征, 输入层 3 为 HHT 得到频域分解的不同频率成分的瞬时频率特征. 输入层 1 和 3 经过卷积处理展开后分别为 64 和 80 个神经元, 再和输入层 2 的 4 个神经元级联得到 148 个神经元输入全连接层分类器, 经过 softmax 得到最后结果.

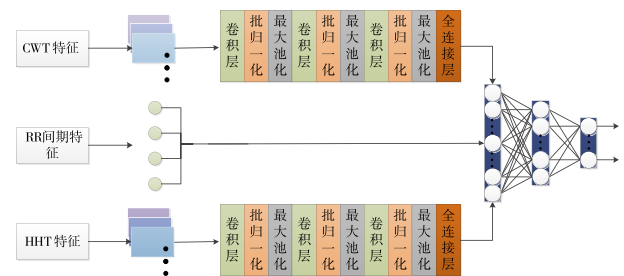


图2 CNN分类模型

CNN 模型训练的损失函数采用 Focal Loss 函数, 通过调整多数样本和少数样本的权重系数, 可解决类别之间极度不平衡问题. 对于多数样本, 取较小的值来减少权重, 反之, 对于少数样本取较大的值. 因此通过控制, 控制不同数量样本的共享权重占总损失的比. 还通过添加控制因子, 控制易分类样本和不易分类样本的权重. 降低易分类样本的权重, 使模型在训练时更加关注不易分类的样本. 所以 Focal Loss 函数定义为:

$$\text{focal loss}(p_t) = -\alpha_t (1-p_t)^\gamma \log(p_t) \quad (10)$$

其中  $p_t$  为类别  $t$  的概率, 分别将控制因子  $\gamma$  和  $\alpha$  设为 2 和 0.25<sup>[31]</sup>.

### 2.4 实验环境及评价标准

#### 2.4.1 实验环境及超参数设置

本文中 CNN 分类模型采用的深度学习平台为 keras, 后端采用 TensorFlow, 实验使用的硬件平台显卡为 NVIDIA GeForce RTX 3090, 操作系统是 Ubuntu 18.04.5. 其中损失函数分别采用交叉熵和 Focal Loss 进行对比, 优化器都设置为 Adam. 数据分批输入网络, batch size 设为 32, 开始训练前, 随机初始化权重. 学习率设置为 0.001, 每 2 个 epoch 降低 0.1 倍. 最大 epoch 设置为 30, 如果训练 10 个 epoch 后, 训练效果没有提升就提前终止训练.

表 3 CNN 分类模型的参数

网络层名	核尺寸	滤波器	填充	步长	输出形状	参数
输入层 1	-	-	-	-	100×250×1	-
卷积层	7×7	16	vaild	1	94×244×16	800
批归一化	-	-	-	-	94×244×16	64
最大池化	5×5	-	-	-	18×48×16	-
卷积层	3×3	32	vaild	1	16×46×32	4 640
批归一化	-	-	-	-	16×46×32	128
最大池化	3×3	-	-	-	5×15×32	-
卷积层	3×3	64	vaild	1	3×13×64	18 496
批归一化	-	-	-	-	3×13×64	256
最大池化	3×3	-	-	-	1×4×64	-
全连接层	-	-	-	-	1×4×16	1 040
展开层	-	-	-	-	64	-
输入层 2	-	-	-	-	4	-
输入层 3	-	-	-	-	8×248×1	-
卷积层	7×7	16	same	1	8×248×16	800
批归一化	-	-	-	-	8×248×16	64
最大池化	2×5	-	-	-	4×49×16	-
卷积层	3×3	32	same	1	4×49×32	4 640
批归一化	-	-	-	-	4×49×32	128
最大池化	2×3	-	-	-	2×16×32	-
卷积层	3×3	64	same	1	2×16×64	18 496
批归一化	-	-	-	-	2×16×64	256
最大池化	2×3	-	-	-	1×5×64	-
全连接层	-	-	-	-	1×5×16	1 040
展开层	-	-	-	-	80	-
融合层	-	-	-	-	148	-
全连接层	-	-	-	-	32	4 768
softmax 层	-	-	-	-	4	132

### 2.4.2 评价标准

采用准确度(accuracy)、精度(precision)、召回率(recall)和 F1 分数等评价标准,定义如下:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (11)$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (12)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (13)$$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (14)$$

其中 TP 为阳性样本的正确预测数, TN 为阴性样本的正确预测数, FN 为阳性样本的错误预测数, FP 为阴性样本的错误预测数<sup>[32]</sup>.

## 3 实验结果

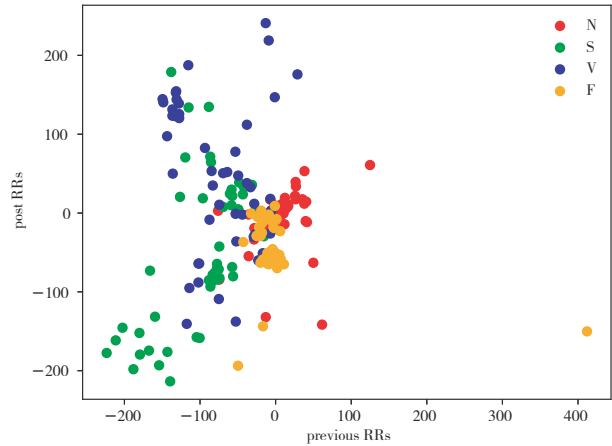
### 3.1 特征提取结果

本研究总共提取三组特征,分别是 RR 间期特征、

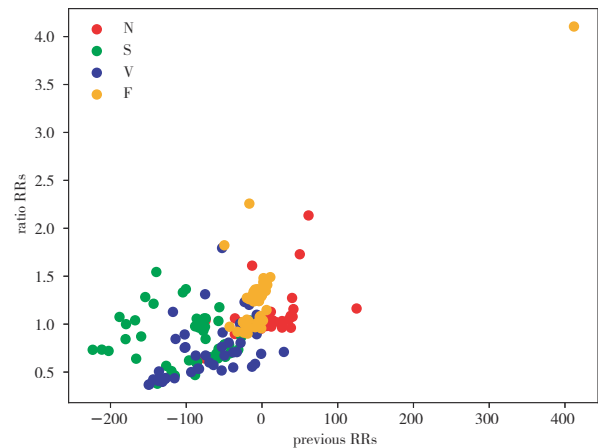
CWT 时频特征和 HHT 不同频率段的瞬时频率.

#### 3.1.1 RR 间期特征

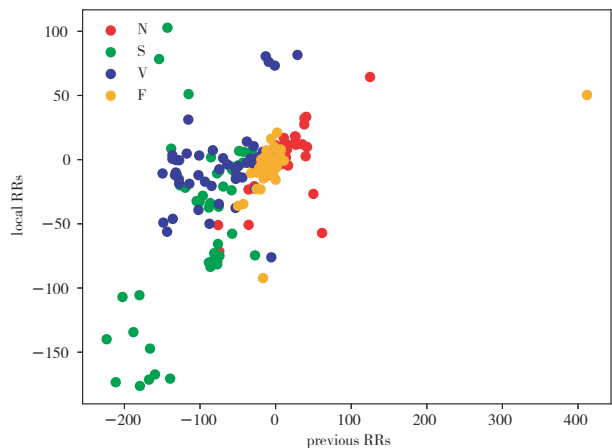
为了说明 RR 间期特征的有效性,图 3 为从四类样本中各随机抽取的 50 个样本,绘制的散点分布图. 三个子



(a) 前 RR 间期和后 RR 间期



(b) 前 RR 间期和前后 RR 间期比



(c) 前 RR 间期和 10 个 RR 间期平均值

图 3 不同心拍的四个 RR 间期分布

图为前RR间期(当前心拍R峰与前一个心拍R峰的区域,previous RRs)分别与后RR间期(当前心拍R峰与后一个心拍R峰的区域,post RRs)、前后RR间期比(前RR间期与后RR间期的比,ratio RRs)和前10个RR间期的平均值(当前心拍R峰与前十个心拍R峰的区域,local RRs)的分布.可知N和F类样本分布比较集中,并且能明显看出来它们之间的分布差异;而S和V类样本虽然分布较为分散,但不同样本之间分布的区域还是有明显的不同.

### 3.1.2 CWT 频谱特征

如图4所示,图4(a)、(b)分别为N的原始ECG信号图、CWT时频特征图,图4(c)、(d)分别为V的原始信号图、CWT时频特征图,由图可知,N类别心拍的时频特征图在R峰处能量最高,然后逐渐降低,R峰两侧能量会突变为零;而V类的心拍时频能量相较于N类更平均,R峰能量很少,但在R峰两侧能量最多,可见不同类别心拍的时频能量图有明显的不同.

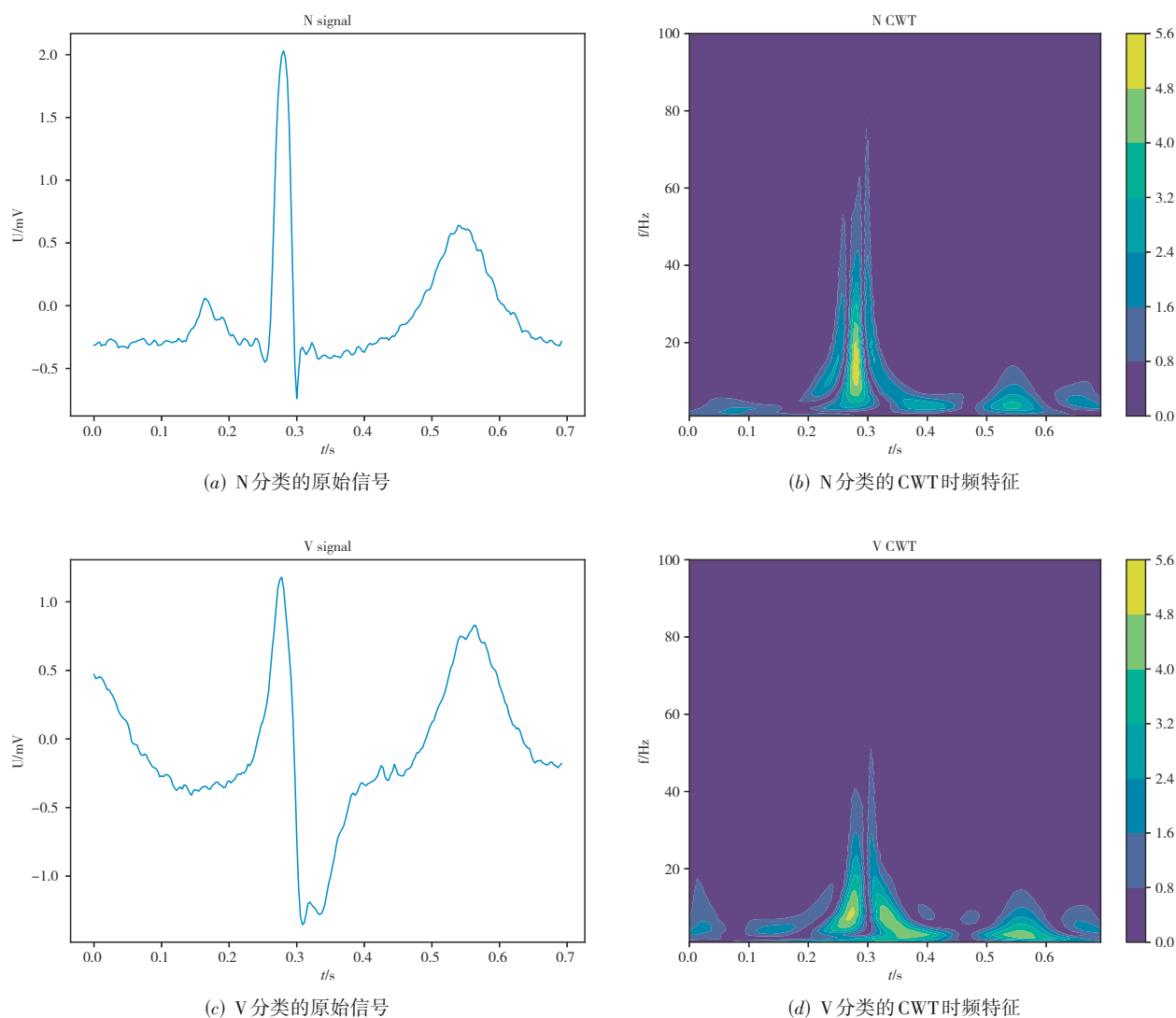


图4 原始信号图及CWT时频特征

### 3.1.3 HHT 瞬时频率特征

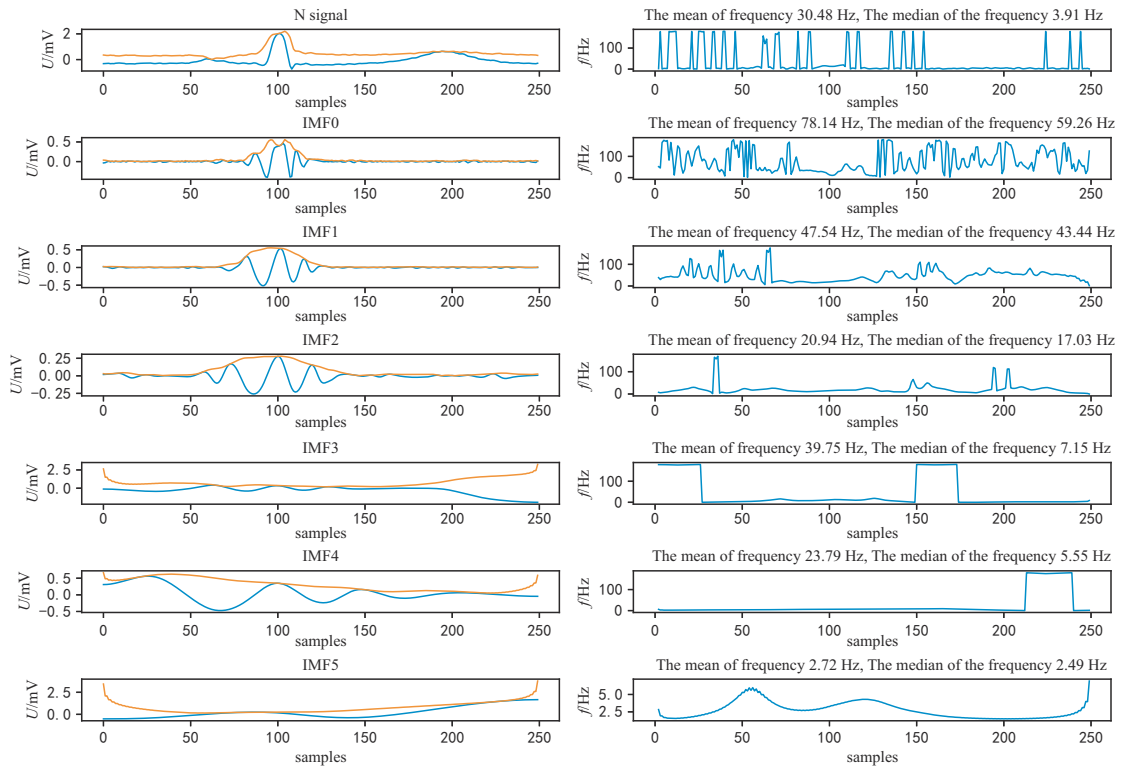
图5为ECG的N和V类原始样本经EMD得到的不同IMF及对应的瞬时频率特征.左侧蓝色曲线为IMF,其中第一组为原始信号,最后一组IMF为残差,橙色曲线为IMF对应的HHT变换图;右侧为IMF对应的瞬时频率特征图.可以看到N和V类样本的IMF数目不同,

N类IMF数更多,其IMF瞬时频率值也更大一些.

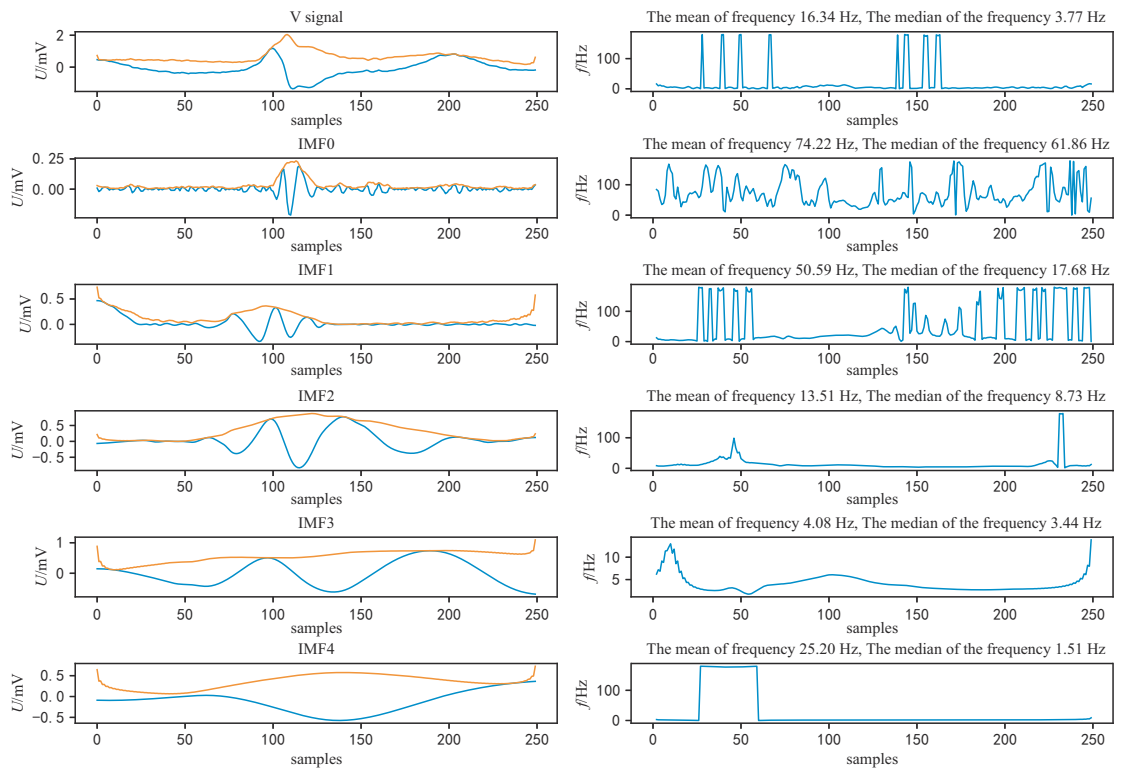
由图3~图5特征图可知,不同类别的心拍CWT时频特征、HHT瞬时频率特征以及RR间期特征有明显的不同,所以可以将上述三个特征作为心拍分类的输入特征.

## 3.2 实验结果

混淆矩阵说明了预测标签(predicted label)与真实



(a) N 分类



(b) V 分类

图5 原始信号及其对应的HHT变换图和瞬时频率

标签(true label)之间的一致,行标签表示每行中的真实标签,列标签表示每列中的预测标签,颜色表示上述记录占同一行所有记录的比例.分别采用RR间期与CWT时频特征为输入(简称RR+CWT);RR间期与CWT时频特征为输入,损失函数采用Focal Loss(简称RR+CWT+Focal Loss);RR间期与HHT瞬时频率特征为输入(简称RR+HHT);RR间期与HHT瞬时频率特征为输入,损失函数采用Focal Loss(简称RR+HHT+Focal Loss),RR间期、CWT时频特征与HHT瞬时频率特征联

合为输入,(简称RR+CWT+HHT);RR间期、CWT时频特征与HHT瞬时频率特征联合为输入,损失函数采用Focal Loss(简称RR+CWT+HHT+Focal Loss)对四类心律失常进行分类,其分类结果的混淆矩阵如图6所示.可以发现,由于F的样本很少,六种方法的预测F标签很容易受到其他标签的影响.图6(f)混淆矩阵相较于其他5个混淆矩阵,显示RR+CWT+HHT+Focal Loss方法具有较高的准确性,对于样本数很少的F类也有一定的效果.

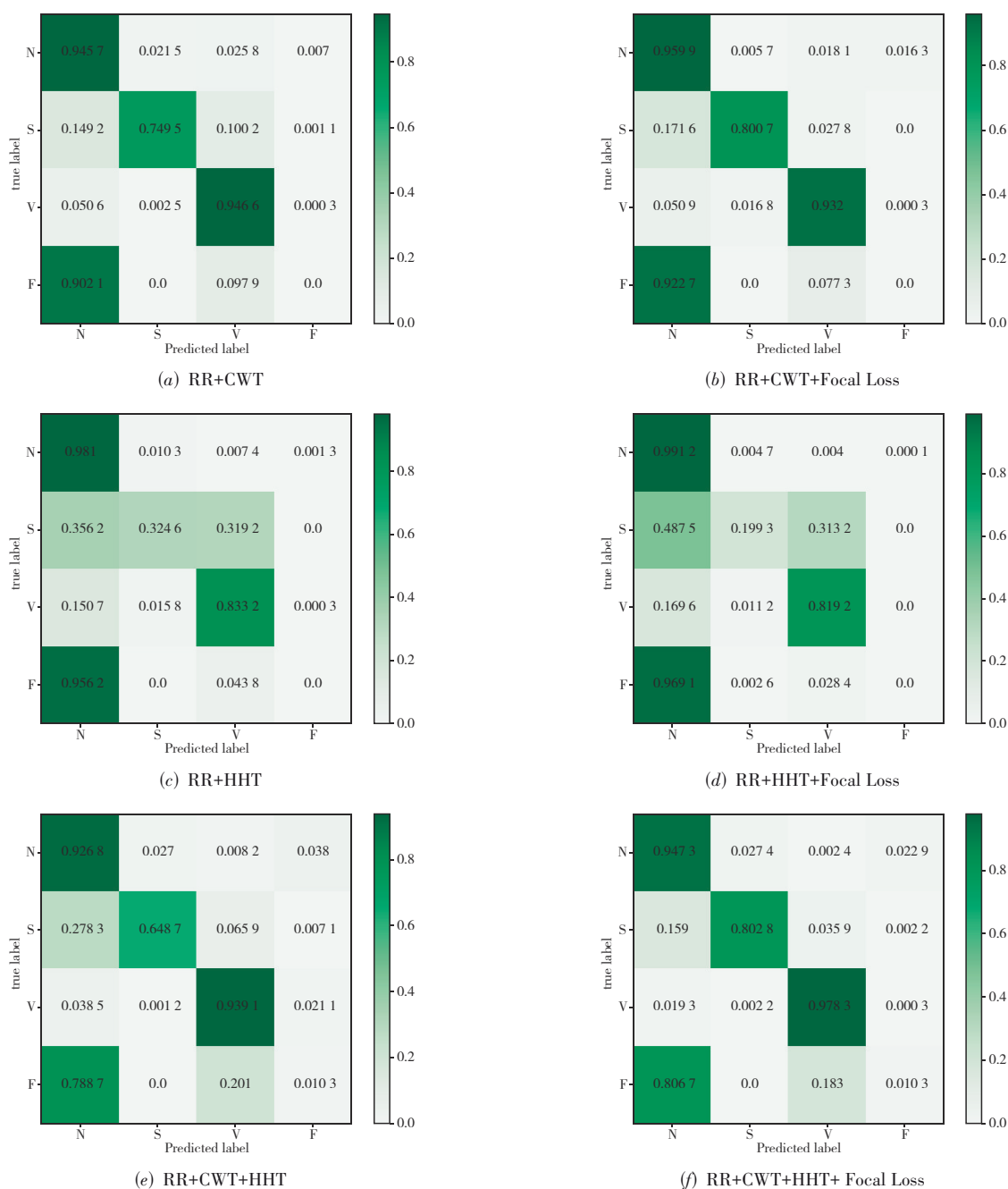


图6 各类算法分类结果的混淆矩阵

### 3.3 消融实验结果及分析

为了验证实验的有效性,我们采用消融实验的策略.如表4所示,分别验证了损失函数为交叉熵和将损失函数改为 Focal Loss 的三种方法:文献[15]中的方法(RR+CWT),RR间期和HHT提取的不同频率段的瞬时频率(RR+HHT)以及本文提出的RR间期、CWT提取的时频域特征和HHT提取的不同频率段的瞬时频率(RR+CWT+HHT).

通过与文献[15]中方法复现的结果比较,可以发现,在文献[15]的方法中加上 Focal Loss,整体性能有明显的提升,少数样本类也获得了不错的效果, $F1$ 分别从0.963 3,0.660 1,0.798 7提升到0.970 2,0.814 0,0.845 1,证明了 Focal Loss 的有效性.然后我们又讨论了HHT生成的不同频率段的瞬时频率对模型的影响,实验结果表明RR+HHT效果不理想,但是V类样本的精度提升了0.052,再加上 Focal Loss 后,RR+HHT+Focal Loss 比RR+HHT多数评价指标好一些,N、S、V的 $F1$ 分别提高了0.01、0.09、0.004,再次验证了 Focal Loss 的有效性,同时少数类样本V的 $F1$ 也比RR+HHT提高了0.001.

最后我们讨论了RR+CWT+HHT级联的融合特征

表4 不同方法下的评价指标

方法	类别	Precision	Recall	$F1$	Accuracy
RR CWT	N	0.981 5	0.945 7	0.963 3	0.931 1
	S	0.589 8	0.749 5	0.660 1	
	V	0.690 8	0.946 6	0.798 7	
	F	0.000 0	0.000 0	0.000 0	
RR CWT Focal Loss	N	0.980 7	<b>0.959 9</b>	<b>0.970 2</b>	0.944 7
	S	<b>0.827 7</b>	<b>0.800 7</b>	<b>0.814 0</b>	
	V	<b>0.773 0</b>	0.932 0	<b>0.845 1</b>	
	F	0.000 0	0.000 0	0.000 0	
RR HHT	N	0.973 2	<b>0.960 8</b>	<b>0.966 9</b>	0.921 6
	S	0.539 9	0.324 6	0.405 4	
	V	<b>0.742 5</b>	0.833 2	0.785 2	
	F	0.000 0	0.000 0	0.000 0	
RR HHT Focal Loss	N	0.960 2	<b>0.959 8</b>	<b>0.975 5</b>	0.933 0
	S	0.601 0	0.377 5	0.499 4	
	V	<b>0.775 4</b>	0.819 1	<b>0.789 7</b>	
	F	0.000 0	0.000 0	0.000 0	
RR CWT HHT	N	0.977 6	0.926 8	0.951 5	0.936 5
	S	0.498 3	0.648 7	0.563 7	
	V	<b>0.843 5</b>	<b>0.939 1</b>	<b>0.888 7</b>	
	F	<b>0.002 3</b>	<b>0.010 3</b>	<b>0.003 7</b>	
RR CWT HHT Focal Loss	N	<b>0.984 3</b>	<b>0.947 3</b>	<b>0.965 5</b>	<b>0.946 7</b>
	S	0.547 3	<b>0.802 8</b>	0.650 9	
	V	<b>0.928 9</b>	<b>0.978 3</b>	<b>0.952 9</b>	
	F	<b>0.003 9</b>	<b>0.010 3</b>	<b>0.005 7</b>	

分类结果的情况,由表可知,样本较多的N和S类整体评价指标不如RR+HHT,但样本较少的V和F类 $F1$ 从0.798 7、0.000 0提升到0.888 7、0.003 7,我们给RR+CWT+HHT加上 Focal Loss 后整体性能提升明显,N、V、F类的 $F1$ 分别提升了0.002 2、0.154 2、0.005 7,但S类 $F1$ 值稍微降低了0.009 2.

## 4 结论

本文提出了一种基于时频域混合特征的 Focal Loss CNN 分类方法,CNN 分类模型以RR间期的时域特征、HHT分解不同频率组成的频域特征和CWT的时频域联合特征为输入.时域分析的优点在于计算简单且有很高的时间精度和准确性,频域分析可以反映出ECG信号中隐藏的信息,相比较单独时域分析、单独频域分析与时频域分析提取的特征训练模型,时频混合特征取得了更好的分类效果.实验结果表明,本文提出的方法不平衡的样本也有很好的效果,特别对于室性早搏和室性逸搏的识别效果有明显的提升.

心律失常诊断中,ECG数据总是不平衡的,未来工作中,将采集大量的临床实测ECG数据进一步去验证本方法的有效性,并希望继续改进方法,进一步解决数据不平衡问题和提高算法的准确性,为心律失常早发现、早诊断,提供新的思路,以推动智慧医疗的发展.

## 参考文献

- [1] BARMPOUTIS P, DIMITROPOULOS K, APOSTOLIDIS A, et al. Multi-lead ECG signal analysis for myocardial infarction detection and localization through the mapping of Grassmannian and Euclidean features into a common Hilbert space[J]. Biomedical Signal Processing and Control, 2019, 52: 111-119.
- [2] LIN X J, GREEN J C, XIAN H, et al. Holiday and weekend effects on mortality for acute myocardial infarction in Shanxi, China: A cross-sectional study[J]. International Journal of Public Health, 2020, 65(6): 847-857.
- [3] YE C, KUMAR B V, COIMBRA M T. Heartbeat classification using morphological and dynamic features of ECG signals[J]. IEEE Transactions on Biomedical Engineering, 2012, 59(10): 2930-2941.
- [4] KUTLU Y, KUNTALP D. A multi-stage automatic arrhythmia recognition and classification system[J]. Computers in Biology and Medicine, 2011, 41(1): 37-45.
- [5] JOHNSON K W, TORRES SOTO J, GLICKSBERG B S, et al. Artificial intelligence in cardiology[J]. Journal of the American College of Cardiology, 2018, 71(23): 2668-2679.
- [6] DAQROUQ K, ALKHATEEB A, AJOUR M N, et al. Neu-

- ral network and wavelet average framing percentage energy for atrial fibrillation classification[J]. *Computer Methods and Programs in Biomedicine*, 2014, 113(3): 919-926.
- [7] WANG Z Z, LI H Y, HAN C, et al. Arrhythmia classification based on multiple features fusion and random forest using ECG[J]. *Journal of Medical Imaging and Health Informatics*, 2019, 9(8): 1645-1654.
- [8] CHEIKHROUHOU O, MAHMUD R, ZOUARI R, et al. One-dimensional CNN approach for ECG arrhythmia analysis in fog-cloud environments[J]. *IEEE Access*, 2021, 9: 103513-103523.
- [9] POURBABAEI B, ROSHTKHARI M J, KHORASANI K. Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients [J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, 48(12): 2095-2104.
- [10] NURMAINI S, TONDAS A E, DARMAWAHYUNI A, et al. Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks[J]. *Future Generation Computer Systems*, 2020, 113: 304-317.
- [11] ÇINAR A, TUNCER S A. Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks[J]. *Computer Methods in Biomechanics and Biomedical Engineering*, 2021, 24(2): 203-214.
- [12] ZHANG Y T, LI J Y, WEI S S, et al. Heartbeats classification using hybrid time-frequency analysis and transfer learning based on ResNet[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(11): 4175-4184.
- [13] SABUT S, PANDEY O, MISHRA B S P, et al. Detection of ventricular arrhythmia using hybrid time-frequency-based features and deep neural network[J]. *Physical and Engineering Sciences in Medicine*, 2021, 44(1): 135-145.
- [14] XU Y, ZHANG S, CAO Z, et al. Extreme learning machine for heartbeat classification with hybrid time-domain and wavelet time-frequency features[J]. *Journal of Healthcare Engineering*, 2021, 2021: 6674695.
- [15] WANG T, LU C, SUN Y, et al. Automatic ECG classification using continuous wavelet transform and convolutional neural network[J]. *Entropy (Basel, Switzerland)*, 2021, 23(1): E119.
- [16] 郑近德, 潘海洋, 戚晓利, 等. 基于改进经验小波变换的时频分析方法及其在滚动轴承故障诊断中的应用[J]. *电子学报*, 2018, 46(2): 358-364.
- ZHENG J D, PAN H Y, QI X L, et al. Enhanced empirical wavelet transform based time-frequency analysis and its application to rolling bearing fault diagnosis[J]. *Acta Electronica Sinica*, 2018, 46(2): 358-364. (in Chinese)
- [17] HUANG J S, CHEN B Q, YAO B, et al. ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network[J]. *IEEE Access*, 2019, 7: 92871-92880.
- [18] GRAMATIKOV B, GEORGIEV I. Wavelets as alternative to short-time Fourier transform in signal-averaged electrocardiography[J]. *Medical and Biological Engineering and Computing*, 1995, 33(3): 482-487.
- [19] 杨向林, 严洪, 许志, 等. 基于 Hilbert-Huang 变换的 ECG 消噪[J]. *电子学报*, 2011, 39(4): 819-824.
- YANG X L, YAN H, XU Z, et al. ECG de-noising based on Hilbert-Huang transform[J]. *Acta Electronica Sinica*, 2011, 39(4): 819-824. (in Chinese)
- [20] MOODY G B, MARK R G. The impact of the MIT-BIH arrhythmia database[J]. *IEEE Engineering in Medicine and Biology Magazine*, 2001, 20(3): 45-50.
- [21] LIU F F, LIU C Y, ZHAO L N, et al. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection[J]. *Journal of Medical Imaging and Health Informatics*, 2018, 8(7): 1368-1373.
- [22] MAR T, ZAUNSEDER S, MARTÍNEZ J P, et al. Optimization of ECG classification by means of feature selection [J]. *IEEE Transactions on Biomedical Engineering*, 2011, 58(8): 2168-2177.
- [23] MONDÉJAR-GUERRA V, NOVO J, ROUCO J, et al. Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers[J]. *Biomedical Signal Processing and Control*, 2019, 47: 41-48.
- [24] DE CHAZAL P, O'DWYER M, REILLY R B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features[J]. *IEEE Transactions on Biomedical Engineering*, 2004, 51(7): 1196-1206.
- [25] HE J Y, RONG J, SUN L, et al. A framework for cardiac arrhythmia detection from IoT-based ECGs[J]. *World Wide Web*, 2020, 23(5): 2835-2850.
- [26] ZHANG Z C, DONG J, LUO X Q, et al. Heartbeat classification using disease-specific feature selection[J]. *Computers in Biology and Medicine*, 2014, 46: 79-89.
- [27] GUO M F, ZENG X D, CHEN D Y, et al. Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant

grounding distribution systems[J]. IEEE Sensors Journal, 2018, 18(3): 1291-1300.

- [28] WU Z Q, LAN T J, YANG C W, et al. A novel method to detect multiple arrhythmias based on time-frequency analysis and convolutional neural networks[J]. IEEE Access, 2019, 7: 170820-170830.
- [29] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-995.
- [30] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.  
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251. (in Chinese)
- [31] LU Y, JIANG M F, WEI L Y, et al. Automated arrhythmia classification using depthwise separable convolutional neural network with focal loss[J]. Biomedical Signal Processing and Control, 2021, 69: 102843.
- [32] KAPLAN BERKAYA S, UYSAL A K, SORA GUNAL E, et al. A survey on ECG analysis[J]. Biomedical Signal Processing and Control, 2018, 43: 216-235.



**张鞠成** 男, 1989年10月出生, 山东潍坊人, 工程师, 主要研究方向为MRI关键技术和心脏电生理研究.

E-mail: jucheng@zju.edu.cn



**王志康** 男, 1970年5月出生, 浙江杭州人, 研究员, 现为浙江大学校医院副院长, 主要研究方向为生物医学信号处理.

E-mail: 2192009@zju.edu.cn

### 作者简介



**吕 杭** 男, 1997年8月出生于湖北省黄冈市, 现为浙江理工大学研究生. 研究领域方向为医学影像处理.

E-mail: hanker.lv@foxmail.com



**蒋明峰(通讯作者)** 男, 1977年5月生于江西丰城, 现为浙江理工大学计算机科学与技术学院教授、博士生导师, 主要研究方向为计算机医学图像处理、生物医学信号处理.

E-mail: m.jiang@zstu.edu.cn



**李 杨** 男, 1986年05月出生于辽宁抚顺, 现为浙江理工大学计算机科学与技术学院讲师、硕士生导师, 主要研究方向为计算机医学图像处理、计算机视觉.

E-mail: dr.yangli@outlook.com