

本文引用格式：倪娟,王剑卓.基于 Mixup 的心电图多标签异常心律检测方法[J].自动化与信息工程,2024,45(3):51-55.

NI Juan, WANG Jianzhuo. Multi label abnormal heart rate detection method for ECG based on Mixup[J]. Automation & Information Engineering, 2024,45(3):51-55.

## 基于 Mixup 的心电图多标签异常心律检测方法\*

倪娟 王剑卓

(广东工业大学, 广东 广州 510006)

**摘要:** 针对心电图不同样本间的高变异性, 以及深度学习模型泛化能力不足的问题, 提出一种基于 Mixup 的心电图多标签异常心律检测方法。首先, 通过 Mixup 方法将心电图与白噪声混合; 然后, 利用混合样本训练深度学习模型; 最后, 在 CPSC 2018 数据集上进行实验, 该方法的 F1 分数相较于 Inception-ResNet-v2、MLC-CNN、STA-CRNN 分别提升了 0.014、0.031、0.023。

**关键词:** Mixup 方法; 心电图; 多标签; 异常心律检测; 深度学习模型

**中图分类号:** TP391

**文献标志码:** A

**文章编号:** 1674-2605(2024)03-0008-05

**DOI:** 10.3969/j.issn.1674-2605.2024.03.008

### Multi Label Abnormal Heart Rate Detection Method for ECG Based on Mixup

NI Juan WANG Jianzhuo

(Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** Multi label abnormal heart rate detection method based on Mixup is proposed to address the high variability of electrocardiograms between different samples and the insufficient generalization ability of deep learning models. Firstly, the electrocardiogram is mixed with white noise using the Mixup method; Then, use mixed samples to train deep learning models; Finally, the experiment was conducted on the CPSC 2018 dataset, compared to Inception-ResNet-v2, MLC-CNN, STA-CRNN, the F1 scores of this method have increased by 0.014, 0.031, and 0.023, respectively.

**Keywords:** Mixup method; electrocardiogram; multi label; abnormal heart rate detection; deep learning models

#### 0 引言

心血管疾病是导致人类死亡的首要原因<sup>[1]</sup>。在临床实践中, 心电图 (electrocardiogram, ECG) 是诊断心血管疾病的主要手段之一。但 ECG 心律种类繁多, 波形变化复杂, 易受测量噪声和伪迹的干扰。人工分析 ECG 效率较低, 且对医生临床经验和专业能力要求较高。因此, ECG 多标签心律检测成为信号处理、机器学习和生物医学工程等多学科交叉领域的研究热点<sup>[2-5]</sup>。

近年来, 学者们提出了许多具有较强表征能力的深度学习方法, 提升了心律检测的精度。ATTALLAH 等<sup>[6]</sup>通过集成 5 种不同结构的深度学习模型, 分类

COVID-19 与其他心脏病例; 实验结果显示, 该方法的分类准确率分别为 98.80% 和 91.73%。YANG 等<sup>[7]</sup>提出一种由逐点卷积和深度可分离卷积模块组成的卷积神经网络 (convolutional neural network, CNN), 能有效地识别不同类别的异常心律。YAO 等<sup>[8]</sup>设计一种基于注意力的时间增量卷积神经网络, 将 ECG 的空间信息和时间特征融合; 在 2018 年中国生理信号挑战赛上, 分类准确率达到了 81.2%。JIANG 等<sup>[9]</sup>针对多标签异常心律检测任务忽略了类别相关性的问题, 提出一种基于图卷积神经网络的异常心律检测方法, 将不同类别之间的关联性融入到检测网络, 以提高异常心律的检测性能。CAI 等<sup>[10]</sup>提出一种多模块多

\* 基金项目: 广东省自然科学基金 (2022A1515011445); 广东省药品监督管理局科技项目 (2022YDZ06)。

标签心电分类模型，结合了一维卷积、挤压块等，可同时检测 55 种心脏病症状，其宏观 F1 分数为 0.863。WANG 等<sup>[11]</sup>提出一种基于生成对抗网络模型的异常心律分类方法，提高了小样本分类的精度。但因异常心律之间的特征空间分布较为复杂，上述方法面临过拟合的风险，只能学习到常见特征，而无法学习到罕见特征。

为此，本文通过 Mixup 方法在 ECG 中引入一定强度的白噪声来训练多标签异常心律检测模型，降低了模型在决策边界附近的震荡，提高了模型泛化能力及多标签异常心律检测的性能。

## 1 算法设计

### 1.1 问题描述

深度学习模型训练一般采用经验风险最小化 (empirical risk minimization, ERM) 原则<sup>[12]</sup>。对于给定的损失函数，预测一个模型函数  $f \in F$ ，使数据在其分布  $P(X, Y)$  上的预期风险最小。模型训练过程可表示为  $\min_{\theta} L(f_{\theta}(X_i), Y_i)$ ，其中， $f$  为模型结构， $\theta$  为待优化的参数， $L$  为损失函数。

由于数据分布  $P(X, Y)$  是未知的，深度学习模型只能求解其近似分布。在模型训练过程中，从  $P(X, Y)$  中随机采样  $(X_i, Y_i)$ ， $1 \leq i \leq n$  作为训练样本，根据  $\min_{\theta} L(f_{\theta}(X_i), Y_i)$  求解深度学习模型的函数。但采用 ERM 原则训练模型时，因过度依赖训练集并追求优异的表现，易导致过拟合，影响测试集的泛化能力。

针对深度学习模型易受噪声影响，出现过拟合甚至梯度爆炸的现象，导致其泛化性能较差的问题，本文引入邻域风险最小化原则<sup>[13]</sup>，在 ECG 中加入比例为  $\alpha$  的白噪声  $\varepsilon$ ，实现数据的真实分布建模。ECG 混合样本表示为

$$X'_i = X_i + \alpha \times \varepsilon_i \quad (1)$$

由于白噪声不属于待检信号，因此  $Y_{\varepsilon_i} = 0$ ，即 ECG 混合样本的标签可表示为

$$Y'_i = Y_i + \alpha \times Y_{\varepsilon_i} = Y_i \quad (2)$$

其近似数据分布可表示为

$$P_v(X, Y) = \frac{1}{n} \sum_{i=1}^n P_{X_i + \alpha \times \varepsilon_i, Y_i}(X, Y) \quad (3)$$

采用最小化损失函数来确保 ECG 混合样本能够准确地映射到相应的标签，并获得接近实际数据分布的结果，即

$$\min_{\theta} L(f_{\theta}(X_i + \alpha \times \varepsilon_i), Y_i) \quad (4)$$

### 1.2 损失函数及其优化

本文采用二元叉熵损失函数 (BCELoss) 来度量异常心律的分类差，计算公式为

$$L_{\text{BCE}}(f_{\theta}(X), Y) = \frac{1}{n} \sum_{i=1}^n \left( Y_i \times \log(f_{\theta}(X_i)) + (1 - Y_i) \times \log(1 - f_{\theta}(X_i)) \right) \quad (5)$$

式中： $f_{\theta}(X_i)$  和  $Y_i$  分别为模型的预测值和数据的标签值。

采用 Mixup 方法对原始 ECG 进行数据增强，根据  $\min_{\theta} L(f_{\theta}(X_i), Y_i)$  数学模型，深度学习模型参数优化过程的损失函数为

$$\begin{aligned} L'_{\text{BCE}}(f_{\theta}(X'), Y') &= \frac{1}{n} \sum_{i=1}^n (Y'_i \times \log(f_{\theta}(X'_i)) + (1 - Y'_i) \times \log(1 - f_{\theta}(X'_i))) \\ &= \frac{1}{n} \sum_{i=1}^n (Y_i \times \log(f_{\theta}(X'_i)) + (1 - Y_i) \times \log(1 - f_{\theta}(X'_i))) \\ &= L_{\text{BCE}}(f_{\theta}(X'), Y) = L_{\text{BCE}}(f_{\theta}(X + \alpha \times \varepsilon), Y) \end{aligned} \quad (6)$$

本文采用 AdamW 优化器，初始学习率设置为  $2e-4$ ，训练迭代次数为 100 次，训练批次大小为 64。

### 1.3 深度学习模型结构

深度学习模型在训练时容易出现过拟合，导致模型虽然在训练样本上拟合得很好，但在未见过的数据上表现不佳。本文通过引入正则化方法——Mixup 来减小模型过拟合风险。首先，采用文献[14]提出的 Jitter 方法给原始 ECG 添加时间步长为  $t$  的白噪声  $\varepsilon$ ，以  $\alpha$  为样本混合策略，扩展训练样本范围，如此模型能

够在噪声干扰的环境下学习和训练，从而缓解深度学习模型的过拟合；然后，利用 ECG 混合样本训练深度学习模型。本文采用一维残差神经网络 (residual network, ResNet)<sup>[15]</sup>作为多标签异常心律的检测网络。基于 Mixup 的 ECG 多标签心律分类模型框架如图 1 所示。

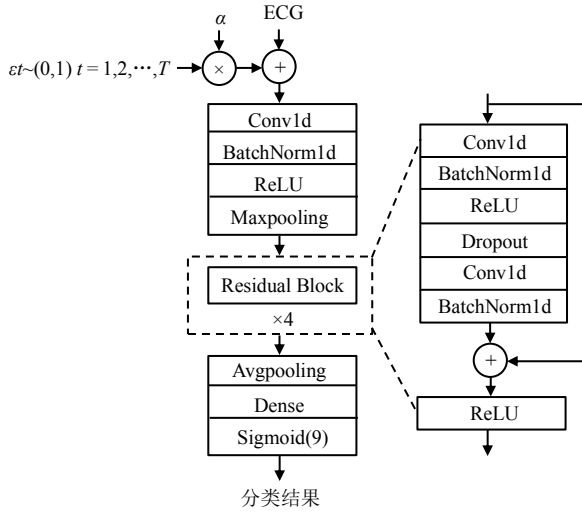


图 1 基于 Mixup 的 ECG 多标签心律分类模型框架

- 1) 利用一个一维卷积 (Conv1d) 层 (卷积核大小为  $15 \times 1$ , 步长为 1) 初步提取 ECG 特征;
- 2) 通过批量归一化层 (BatchNorm1d) 和激活函数 (ReLU) 来减轻梯度消失和梯度爆炸的问题;
- 3) 采用最大池化 (Maxpooling) 层 (窗口大小为  $3 \times 1$ , 步长为 2) 进一步提取 ECG 特征;
- 4) 利用 4 个堆叠的残差块 (Residual Block) 提取 ECG 的深层特征, 每个残差块包括 2 个 Conv1d 层 (卷积核大小为  $7 \times 1$ , 步长为 1), 用于捕获 ECG 局部关联关系; 2 个 BatchNorm1d 层, 用于标准化卷积层输出, 从而加速训练并减轻梯度消失的问题; 1 个 Dropout 层, 用于减少过拟合风险; 2 个 ReLU 层, 用于引入非线性激活, 增强特征表征能力;
- 5) 通过平均池化层 (Avgpooling), 对最后一个残差块的输出特征降维;
- 6) 采用全连接层 (Dense) 将提取的特征向量降维, 并作为分类器的输入;
- 7) 利用一个激活函数 (Sigmoid) 得到最终的分类结果

类结果  $\hat{Y} \in R^{9 \times n}$ 。

## 2 数据

### 2.1 数据集

为验证本文方法的有效性和准确性, 选用第一届中国生理信号挑战赛采用的数据集 (CPSC 2018)<sup>[16]</sup>进行实验。该数据集的 ECG 采样频率为 500 Hz, 包括正常类别 (N) 和 8 种异常心律 (AF、I-AVB、LBBB、RBBB、PAC、PVC、STD、STE), 其中, 每个数据样本由 12 个导联组成, 数据样本时长为 6~60 s。

### 2.2 数据预处理

首先, 通过中值滤波器消除 ECG 的基线漂移; 然后, 利用双线性变换法设计数字带阻滤波器, 去除工频干扰; 接着, 采用 5 级小波变换, 选用 bior4.4 小波基函数, 消除肌电干扰; 最后, 对各个导联的 ECG 进行极差归一化, 去除数据样本幅值分布差异的影响。

由于 CPSC 2018 数据集的 ECG 样本时长不一致, 本文将该数据集的所有 ECG 样本重采样为 30 s。对于长度不足 15 000 个采样点的样本, 采用补零操作至 15 000 个采样点; 对于长度超过 15 000 个采样点的样本, 截取前 15 000 个采样点。将该数据集按 7:2:1 的比例划分为训练集、验证集和测试集, 并确保各子集中的类别比例一致。

## 3 实验结果与分析

### 3.1 评价指标

本文采用准确率 (precision rate, P)、召回率 (recall rate, R) 和 F1 分数 (F1-score, F1) 作为异常心律分类性能的评价指标, 计算公式分别为

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (9)$$

式中:  $TP$ 、 $FP$ 、 $FN$  分别为预测正确存在的异常

心律总数、预测错误实际不存在的异常心律总数和预测错误实际存在的异常心律总数。

### 3.2 实验对比与分析

本实验运行环境为 Windows 10 操作系统, Python 3.8, 利用 PyTorch 1.11 深度学习框架进行模型训练和测试。计算机的 CPU 为 Intel Core i9-13900K@ 5.80 GHz, 内存为 128 GB, GPU 为 NVIDIA Geforce RTX4090, 显存为 24 GB。

为验证本文方法的有效性, 与如下 3 种方法进行对比实验, 实验结果如表 1 所示。

方法 1 (Inception-ResNet-v2<sup>[17]</sup>), 利用递归图和 Inception-ResNet-v2 架构, 将异常心律分类问题建模为二维图像分类任务;

方法 2 (MLC-CNN<sup>[18]</sup>), 利用数据分布的先验知识, 增强心电异常的融合特征;

方法 3 (STA-CRNN<sup>[19]</sup>), 利用时空注意力的卷积递归神经网络, 重点关注 ECG 时空轴上的代表性特征。

表 1 4 种方法的异常心律分类性能比较

心律类别	F1 分数			
	方法 1	方法 2	方法 3	本文方法
N	0.812	0.792	0.819	<b>0.886</b>
AF	0.875	0.936	0.936	<b>0.971</b>
I-AVB	<b>0.923</b>	0.889	0.866	0.853
LBBB	<b>0.929</b>	0.779	0.862	0.903
RBBB	0.776	0.744	0.926	<b>0.947</b>
PAC	0.753	<b>0.866</b>	0.789	0.742
PVC	0.793	<b>0.876</b>	0.865	0.873
STD	0.837	<b>0.876</b>	0.812	0.817
STE	<b>0.900</b>	0.683	0.640	0.730
平均值	0.844	0.827	0.835	<b>0.858</b>

由表 1 可以看出, 本文方法在 CPSC 2018 数据集上的分类性能最好。

针对  $\alpha$  的不同取值, 设置 2 组实验方案:

实验组 (1), 将  $\alpha$  设置为 0, 即不采用 Mixup 方法;

实验组 (2), 将  $\alpha$  依次取值为 0.2、0.4、0.6、0.8, 实验结果如表 2 所示。

表 2 消融实验 F1 分数

心律类别	F1 分数				
	实验组 (1)		实验组 (2)		
	$\alpha=0$	$\alpha=0.2$	$\alpha=0.4$	$\alpha=0.6$	$\alpha=0.8$
N	0.863	0.852	<b>0.886</b>	0.836	0.836
AF	0.971	<b>0.975</b>	0.971	<b>0.975</b>	0.966
I-AVB	0.863	<b>0.865</b>	0.853	0.845	0.811
LBBB	0.915	<b>0.918</b>	0.903	0.913	0.897
RBBB	0.941	<b>0.952</b>	0.947	0.948	0.940
PAC	0.737	0.747	0.742	0.729	<b>0.756</b>
PVC	0.852	0.857	<b>0.873</b>	0.860	0.862
STD	0.791	<b>0.828</b>	0.817	0.822	0.782
STE	0.641	0.714	0.730	0.780	<b>0.809</b>
平均值	0.841	0.827	<b>0.858</b>	0.857	0.851

对比实验组 (1)、(2) 的平均 F1 分数可知, 原始 ECG 加入适量的噪声, 有助于减轻模型的过拟合, 同时提高了模型的泛化能力。

此外, 通过 CPSC 2018 多标签异常心律分类混淆矩阵 (见图 2), 解释该模型在预测异常心律时的有效性及原因。

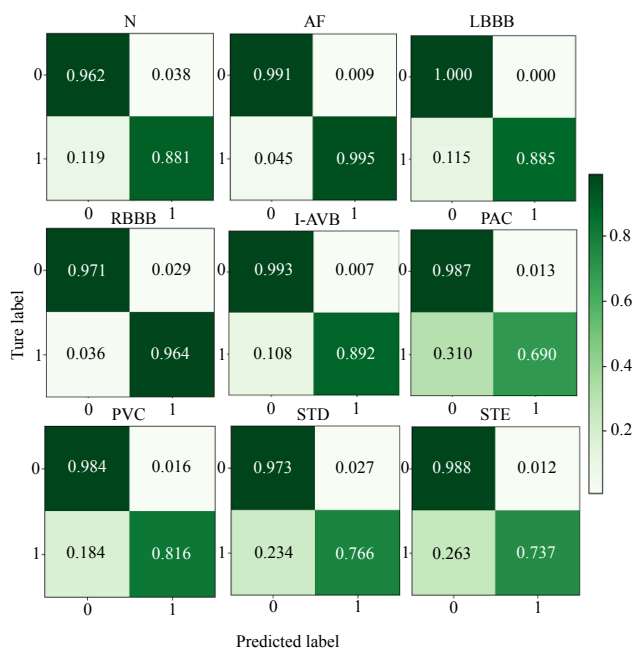


图 2 CPSC 2018 多标签异常心律分类混淆矩阵

由图 2 可以观察到, 9 类患者的异常心律预测错误率普遍较低, 正确预测率相对较高; 然而, 相较于

其他类别, PAC、STD 和 STE 存在较高的错误率, 主要因为这些类别的样本数量较少, 模型难以充分学习异常心律的相关特征。

#### 4 结论

针对多标签异常心律检测任务存在难以准确提取代表性特征的问题, 本文采用 Mixup 方法在原始 ECG 中引入一定强度的白噪声, 不仅有助于增加数据的多样性, 还可以提高模型的泛化能力, 从而更准确地捕捉异常心律特征。然而, 一些小样本类别的检测精度仍不尽如人意, 影响了检测性能。之后考虑利用集成学习的方法, 将多个模型的预测结果进行融合, 进一步提高小样本类别的检测性能。

#### 参考文献

- [1] 张晓东,白晓君.防治心脑血管疾病进入脑心同治时代[N].健康报,2024-04-17(4).
- [2] YANG S, LIAN C, ZENG Z, et al. A multi-view multi-scale neural network for multi-label ECG classification[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2023.
- [3] RAN S, LI X, ZHAO B, et al. Label correlation embedding guided network for multi-label ECG arrhythmia diagnosis[J]. Knowledge-Based Systems, 2023,270:110545.
- [4] SRIVASTAVA A, PRATIHER S, ALAM S, et al. A deep residual inception network with channel attention modules for multi-label cardiac abnormality detection from reduced-lead ECG[J]. Physiological Measurement, 2022,43(6):064005.
- [5] LIU H, CHEN D, CHEN D, et al. A large-scale multi-label 12-lead electrocardiogram database with standardized diagnostic statements[J]. Scientific Data, 2022, 9(1):272.
- [6] ATTALLAH O. ECG-BiCoNet: An ECG-based pipeline for COVID-19 diagnosis using Bi-Layers of deep features integration[J]. Computers in Biology and Medicine, 2022,142:105210.
- [7] YANG F, ZHANG X, ZHU Y. PDNet: A convolutional neural network has potential to be deployed on small intelligent devices for arrhythmia diagnosis[J]. CMES-Computer Modeling in Engineering & Sciences, 2020, 125(1):365-382.
- [8] YAO Q, WANG R, FAN X, et al. Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network[J]. Information Fusion, 2020, 53: 174-182.
- [9] JIANG Z, ALMEIDA T P, SCHLINDWEIN F S, et al. Diagnostic of multiple cardiac disorders from 12-lead ECGs using graph convolutional network based multi-label classification[C]//2020 Computing in Cardiology. IEEE, 2020: 1-4.
- [10] CAI J, SUN W, GUAN J, et al. Multi-ECGNet for ECG arrhythmia multi-label classification[J]. IEEE Access, 2020, 8: 110848-110858.
- [11] WANG Y, SUN L, SUBRAMANI S. Cab: classifying arrhythmias based on imbalanced sensor data[J]. KSII Trans Internet Inf Syst (TIIS), 2021,15(7):2304-2320.
- [12] SUGIYAMA M, BAO H, ISHIDA T, et al. Machine learning from weak supervision: An empirical risk minimization approach[J]. MIT Press, 2022.
- [13] PINTO F, YANG H, LIM S N, et al. Using Mixup as a regularizer can surprisingly improve accuracy & out-of-distribution robustness[J]. Advances in Neural Information Processing Systems, 2022, 35: 14608- 14622.
- [14] LI M P. Jitter, noise, and signal integrity at high-speed[M]. Pearson Education, 2007.
- [15] ALLEN-ZHU Z, LI Y. What can resnet learn efficiently, going beyond kernels[J]. Advances in Neural Information Processing Systems, 2019,32.
- [16] LIU F, LIU C, ZHAO L, 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; 6.
- [17] ZHANG H, LIU C, ZHANG Z, et al. Recurrence plot-based approach for cardiac arrhythmia classification using inception-ResNet-v2[J]. Frontiers in Physiology, 2021,12:648950.
- [18] GE Z, JIANG X, TONG Z, et al. Multi-label correlation guided feature fusion network for abnormal ECG diagnosis[J]. Knowledge-Based Systems, 2021, 233:107508.
- [19] ZHANG J, LIU A, GAO M, et al. ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network[J]. Artificial Intelligence in Medicine, 2020,106:101856.

#### 作者简介:

倪娟, 女, 1997 年生, 硕士研究生, 主要研究方向: 模式识别、生物信号处理。E-mail: 1431679754@qq.com

王剑卓, 男, 1998 年生, 硕士研究生, 主要研究方向: 模式识别、生物信号处理。E-mail: 2913725799@qq.com