

基于多频段多任务编解码模型的心电图基准点联合检测

李磊, 廖桂鑫, 蔡瑞涵, 李珍妮, 吕俊[†]

(广东工业大学 自动化学院, 广东 广州 510006)

摘要: 基准点检测是心电图(ECG)诊断分析的基础。但是, ECG具有波形变异性, 且经常受到各种伪迹和噪声的干扰, 使得基准点检测精度受限。针对该问题, 本文首先构建概率图模型, 分析各频段ECG成分与基准点检测任务之间的推断关系。然后, 在此概率图模型的启发下提出了一种多频段多任务编解码网络。该网络先分别对不同频段的ECG成分进行一维卷积变换提取特征。然后, 通过时域卷积模组学习各频段特征的注意力掩码以抵御噪声。最后, 利用多分支关联结构, 实现多个ECG基准点的联合检测。在MIT-BIH QT和LUDB数据集上的五重交叉验证实验结果表明: 所提方法能够有效地提高ECG基准点的检测精度, 达到了当前最优的水平。

关键词: 心电图基准点检测; 编解码模型; 时域卷积神经网络

引用格式: 李磊, 廖桂鑫, 蔡瑞涵, 等. 基于多频段多任务编解码模型的心电图基准点联合检测. 控制理论与应用, 2024, 41(8): 1451–1458

DOI: 10.7641/CTA.2023.20486

Joint detection of ECG fiducial points based on multi-band and multi-task encoding and decoding model

LI Lei, LIAO Gui-xin, CAI Rui-han, LI Zhen-ni, LÜ Jun[†]

(School of Automation, Guangdong University of Technology, Guangzhou Guangdong 510006, China)

Abstract: Fiducial point detection is the basis of the electrocardiogram (ECG) diagnostic analysis. However, the ECG has waveform variability and is often disturbed by various artifacts and noises, limiting the detection accuracies of fiducial points. This paper first builds a probability graph model to analyze the inference relationships between different band ECG components and fiducial point detection tasks. Then, we propose a multi-band multi-task encoding-decoding network inspired by this probability graph model. The network first performs 1-D convolutions on each ECG component to extract features, then learns the attention masks to resist noise through temporal convolutional modules, and finally adopts the dependent multi-branch structure to realize the joint detection of ECG fiducial points. The experimental results with five-fold cross-validation on the MIT-BIH QT and LUDB databases show that the proposed method can effectively improve the detection accuracy of ECG fiducial points, comparable to the state-of-the-art level.

Key words: ECG fiducial point detection; encoding-decoding model; temporal convolutional network

Citation: LI Lei, LIAO Guixin, CAI Ruihan, et al. Joint detection of ECG fiducial points based on multi-band and multi-task encoding and decoding model. *Control Theory & Applications*, 2024, 41(8): 1451–1458

1 引言

据《中国心血管健康与疾病报告2020》^[1]统计, 我国心血管病的患病率逐年攀升, 心血管病死亡率居各种疾病之首。心电图(electrocardiogram, ECG)检查能够便捷地获取患者心脏各部位的电生理活动信息, 是心血管疾病诊断的常用手段。ECG基准点检测(P-QRS-T波的起止点和峰值点定位)是提取病理特征的重要依据。但是, 依靠医生人工判断ECG基准点, 费时

费力, 无法满足庞大患者群体的需求。

近年来, 研究人员提出了许多ECG基准点自动检测方法, 大致可以分为3类。第1类基于小波变换(wavelet transform, WT)^[2–4]、扩展卡尔曼滤波(extended Kalman filter, EKF)^[5]、互补集成经验模态分解(complete ensemble empirical mode decomposition, CEEM-D)^[6]等数字信号处理方法, 严重依赖经验参数设置, 无法适应复杂的ECG波形变化, 容易出现基准点误检

收稿日期: 2022–06–02; 录用日期: 2023–03–28.

[†]通信作者. E-mail: lujun.rylj@gmail.com; Tel.: +86 13822139845.

本文责任编辑: 夏元清.

国家自然科学基金项目(62073086, 62273106), 广东省自然科学基金项目(2022A1515011445)资助.

Supported by the National Natural Science Foundation of China (62073086, 62273106) and the Natural Science Foundation of Guangdong Province (2022A1515011445).

和漏检现象. 第2类基于经验特征和浅层分类模型, 例如: Saini等人^[7]提取ECG波形梯度特征, 利用K近邻算法(K-nearest neighbors, KNN)来定位基准点; Mehta和Lingayat^[8]提取ECG信号信息熵, 采用支持向量机定位QRS波群. 此类方法需要繁琐的特征工程, 且仍然存在适应性差、表征能力不足的问题. 第3类基于深度学习方法, 具有良好的表征能力, 能够实现端到端的ECG基准点定位, 例如: Peimankar和Puthusserypady^[9]采用卷积神经网络(convolutional neural network, CNN)与长短期记忆(long short-term memory, LSTM)网络相结合的方法, 提取ECG不同时间尺度的波形信息检测ECG基准点; Jimenez等人^[10]利用带有跳跃连接的U型网络(U-Net), 自适应提取不同层级的特征, 提高了ECG基准点的检测精度. 然而, 文献[9–10]基于通用的深度神经网络模型, 未能有效的利用ECG先验知识, 对噪声敏感, 需要大量的训练样本才能保证泛化能力. 针对该问题, Chen等人^[11]基于多视角策略提出区域聚集网络(region aggregation network, RAN), 集成不同时间片段ECG的多尺度特征, 实现了鲁棒的基准点检测; Wang等人^[12]将特征波脉宽和脉高的先验知识融入编解码网络, 增强了ECG基准点定位的鲁棒性. 但是, 文献[11–12]忽略了ECG特征波群的频域分布特点, 没有充分利用多频带ECG特征之间与多ECG基准点检测任务之间的关联推断关系.

因此, 本文首先分析了ECG特征波群的频域分布特点; 然后, 构建概率图模型刻画多频带ECG特征与多ECG基准点检测任务之间复杂的推断关系; 最后, 在该概率图模型的启发下, 提出了一种多频段多任务时域卷积神经网络(multi-band multi-task temporal convolutional network, MBMT-TCN). MBMT-TCN针对不同频段的ECG特征, 采用不同的TCN模组学习注意力掩码抵御噪声, 并通过多分支关联结构, 实现多频段ECG联合特征提取和多个基准点的联合检测.

2 方法

2.1 预处理

首先, 去除伪迹噪声: 本文采用窗宽分别为200 ms和600 ms的两个中值滤波器消除ECG基线漂移; 采用双线性变换法设计无限冲激响应(infinite impulse response, IIR)数字带阻滤波器去除工频干扰, 阻带截止频率分别是49 Hz和51 Hz; 采用截止频率为100 Hz的巴特沃斯低通滤波器去除肌电干扰. 接着, 采用基于线性预测残差的QRS波群检测算法^[13]获取各条ECG的R峰位置, 并计算平均RR间隔(RR_{ave}), 按照文献[11]设计的窗口函数 $l_i(x)$ 逐个心拍地切割ECG信号, 计算方式如下:

$$l_i(x) = \begin{cases} 1, & 0 \leq x_i - x \leq Q_{left} \times RR_{ave}, \\ 1, & 0 \leq x - x_i \leq Q_{right} \times RR_{ave}, \\ 0, & \text{其他,} \end{cases} \quad (1)$$

其中: x_i 表示第*i*个R峰位置; x 表示任意采样点位置; 随机系数 $Q_{left}, Q_{right} \in [0.6, 0.8]$ 服从均匀分布, 在各个心拍重复采样7次. 然后, 通过重采样将所有心拍的长度统一为256. 以上心拍切割方法实质是对ECG训练样本进行了平移和拉伸数据增强. 最后, 为了方便神经网络训练, 对各心拍ECG的幅值进行极差归一化, 即

$$s_{norm} = \frac{s - S_{min}}{S_{max} - S_{min}}, \quad (2)$$

其中: s 和 s_{norm} 分别为为归一化前、后的心拍信号; S_{max}, S_{min} 分别表示该心拍的最大和最小振幅.

2.2 多频段多任务概率图模型

ECG数字信号处理方法常用到如下两个先验知识, 往往被基于深度学习的ECG基准点检测方法所忽略: 1) QRS波群和P-T波的频域分布有所不同(QRS波群能量散布在0–40 Hz, P-T波能量集中于17 Hz以下, 如图1所示)^[14], 通过带通滤波能够有效提高QRS波群与P-T波成分的信噪比; 2)各特征波出现的位置具有较强的时序依赖关系(通常P波和T波分别位于QRS波群的两侧), 且相对于低矮的P-T波, QRS波群的检测难度较小, 可以先定位QRS波群, 然后检测P波和T波的基准点.

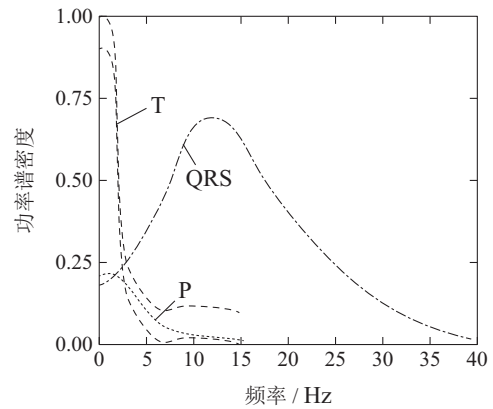


图1 心电信号特征波频谱图^[14]

Fig. 1 ECG character waves spectrogram^[14]

鉴于以上分析, 本文建立多频段ECG多基准点检测任务的概率图模型, 如图2所示. 其中: X 表示ECG信号; $\{\tilde{Y}_{QRS}, \tilde{Y}_P, \tilde{Y}_T\}$ 表示潜在的QRS波群、P波和T波的位置序列; $\{Z_{low}, Z_{high}, Z_{all}\}$ 分别表示ECG低频、高频和全频成分的特征变量; $\{F_{low}, F_{high}\}$ 分别表示ECG低频成分特征和高频成分特征的隐变量(如注意力掩码); $\{\tilde{Z}_{low}, \tilde{Z}_{high}\}$ 表示自适应增强后的特征变量; $\{U, H\}$ 表示融合特征变量. 该模型的条件概率分解如式(3)所示, 揭示了在多频带特征提取和多个基准点联合检测过程中, 各变量之间的依赖关系, 即

$$P(\tilde{Y}_{QRS}, \tilde{Y}_P, \tilde{Y}_T | X) = P(\tilde{Y}_P | H, \tilde{Y}_{QRS}, \tilde{Y}_T) P(\tilde{Y}_T | H, \tilde{Y}_{QRS}) P(\tilde{Y}_{QRS} | H) \times$$

$$\begin{aligned}
& P(H | \tilde{Z}_{\text{high}}, \tilde{Z}_{\text{high}}, Z_{\text{all}}) P(\tilde{Z}_{\text{high}} | F_{\text{high}}, U) \times \\
& P(F_{\text{high}} | U) P(U | Z_{\text{high}}, \tilde{Z}_{\text{low}}) P(F_{\text{low}} | Z_{\text{low}}) \times \\
& P(\tilde{Z}_{\text{low}} | F_{\text{low}}, Z_{\text{low}}) P(Z_{\text{high}} | X) \times \\
& P(Z_{\text{low}} | X) P(Z_{\text{all}} | X). \quad (3)
\end{aligned}$$

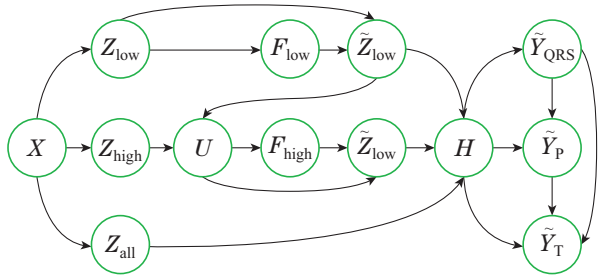


图 2 多频段ECG多基准点检测任务的概率图模型

Fig. 2 Probabilistic graphical model of multi-fiducial point detection task of multi-band ECG

2.3 多频段多任务时序卷积神经网络

受图2所示概率图模型的启发, 本文提出一种多频段多任务时序卷积神经网络(MBMT-TCN), 其网络架

构如图3所示. 该网络第1路分支用来构建低频ECG信号(P-T波主导成分)的编码映射, 通过注意力掩码估计器滤波得到提纯后的低频ECG特征; 第2路分支用来构建高频ECG信号(QRS波群主导成分)的编码映射, 并且与第1路提纯后的特征融合, 经过注意力掩码估计器滤波得到提纯后的高频ECG特征; 第3路分支单纯地构建全频段ECG的编码映射, 作为前两条支路的补充. 最后, 利用信息交互模块融合3路掩码后的特征, 送入多分支解码器子网, 联合预测3个特征波所对应的位置序列. 其中, 掩码估计器(TCN模组)由堆叠的空洞卷积组成^[15]. 相较于传统的CNN, 空洞卷积具有较大的感受野便于捕捉ECG长程的时序依赖关系. 各层空洞卷积之间的跳跃连接使得掩码器具有多尺度信息感知能力. 信息交互模块由特征矩阵拼接和 1×1 卷积构成. 所提出的MBMT-TCN可看作是单通道语音分离时序卷积网络^[15]的拓展, 通过定制化的多分支编解码结构, 利用ECG特征波的频域分布知识和基准点之间的推断关系, 增强基准点检测模型的鲁棒性和泛化能力.

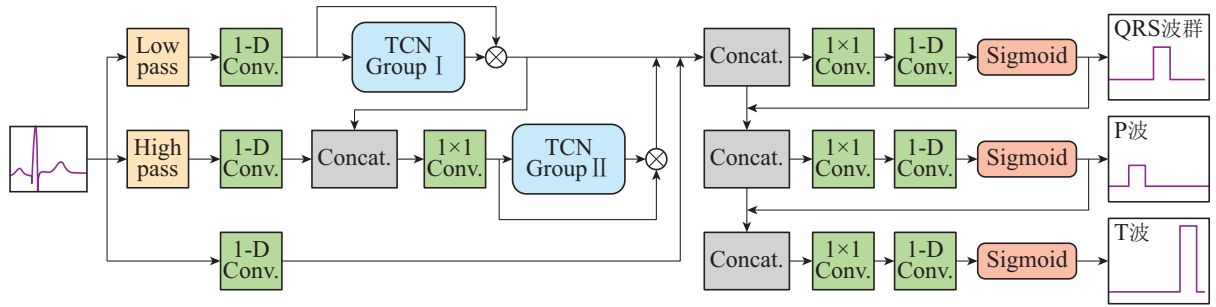


图 3 MBMT-TCN的网络结构

Fig. 3 Network structure of the MBMT-TCN

MBMT-TCN 的具体计算步骤如下: 首先, 进行一维卷积编码. 将单心拍ECG划分为 K 个长度为 L 的片段, $X \in \mathbb{R}^{K \times L}$. 输入至3个编码器子模块, 学习变换域, 分别提取低频段、高频段和全频段的特征 $Z_{\text{low}}, Z_{\text{high}}, Z_{\text{all}} \in \mathbb{R}^{K \times N}$ (N 为编码基函数的个数), 计算方式如下:

$$\begin{cases} Z_{\text{low}} = X B_{\text{low}}, \\ Z_{\text{high}} = X B_{\text{high}}, \\ Z_{\text{all}} = X B_{\text{all}}, \end{cases} \quad (4)$$

其中: $B_{\text{low}}, B_{\text{high}}, B_{\text{all}} \in \mathbb{R}^{L \times N}$ 分别表示低频段、高频段和全频段的编码基矩阵. 接着, 将编码特征矩阵 Z_{low} 输入到TCN掩码估计模块, 得到低频段的掩码特征 $F_{\text{low}} \in \mathbb{R}^{K \times N}$. 对 Z_{low} 进行掩码加权得到增强特征矩阵 $\tilde{Z}_{\text{low}} \in \mathbb{R}^{K \times N}$, 即

$$\tilde{Z}_{\text{low}} = F_{\text{low}} \odot Z_{\text{low}}, \quad (5)$$

其中 \odot 表示点乘. 然后, 将 \tilde{Z}_{low} 和低频段的编码特征矩阵 Z_{high} 输入到信息融合子模块, 通过拼接, 和 1×1

卷积得到高频段的潜在特征 $U \in \mathbb{R}^{K \times N}$. 接着, 将 U 输入TCN掩码估计器, 得到高频段的掩码特征 $F_{\text{high}} \in \mathbb{R}^{K \times N}$. 对 U 进行掩码加权得到增强特征矩阵 $\tilde{Z}_{\text{high}} \in \mathbb{R}^{K \times N}$, 即

$$\tilde{Z}_{\text{high}} = F_{\text{high}} \odot U, \quad (6)$$

将 $\tilde{Z}_{\text{low}}, \tilde{Z}_{\text{high}}$ 和 Z_{all} 拼接得到 $H \in \mathbb{R}^{K \times 3N}$. 再经过 1×1 卷积得到QRS特征矩阵 $H_{\text{QRS}} \in \mathbb{R}^{K \times N'}$. 最后, 采用由一维反卷积和Sigmoid函数组成的解码器重构出QRS波群的位置序列 $\tilde{Y}_{\text{QRS}} \in \mathbb{R}^{K \times L}$, 即

$$\tilde{Y}_{\text{QRS}} = \text{Sigmoid}(H_{\text{QRS}} D_{\text{QRS}}), \quad (7)$$

其中 $D_{\text{QRS}} \in \mathbb{R}^{N' \times L}$ 是QRS波群的解码基矩阵. 将 H 与 \tilde{Y}_{QRS} 拼接得到特征矩阵 $P = [H \ \tilde{Y}_{\text{QRS}}] \in \mathbb{R}^{K \times 4N}$, 经过 1×1 卷积得到融合后的特征矩阵 $H_P \in \mathbb{R}^{K \times N'}$. 再重构出P波位置序列 \tilde{Y}_P , 即

$$\tilde{Y}_P = \text{Sigmoid}(H_P D_P), \quad (8)$$

其中 $D_P \in \mathbb{R}^{N' \times L}$ 是心房P波的解码基矩阵. 将 H 与

\tilde{Y}_{QRS} 和 \tilde{Y}_P 拼接,并经过 1×1 卷积得到T波特征矩阵 $H_T \in \mathbb{R}^{K \times N'}$.最后重构出T波位置序列,即

$$\tilde{Y}_T = \text{Sigmoid}(H_T D_T), \quad (9)$$

其中 $D_T \in \mathbb{R}^{N' \times L}$ 是T波的解码基矩阵.将 \tilde{Y}_{QRS} , \tilde{Y}_P , \tilde{Y}_T 中 K 个长度为 L 的重构位置信息依次叠加,最终预测出整个ECG心拍所有基准点的位置序列.

2.4 损失函数

本文以均方误差作为损失函数度量基准点预测偏差,计算方式如下:

$$\begin{aligned} \min J(\omega) = & \frac{1}{2N_{tr}} \sum_{i=1}^{N_{tr}} \|\tilde{Y}_{QRS}^i - Y_{QRS}^i\|_2^2 + \\ & \frac{1}{2N_{tr}} \sum_{i=1}^{N_{tr}} \|\tilde{Y}_P^i - Y_P^i\|_2^2 + \frac{1}{2N_{tr}} \sum_{i=1}^{N_{tr}} \|\tilde{Y}_T^i - Y_T^i\|_2^2, \quad (10) \end{aligned}$$

其中: ω 为可训练的网络参数; N_{tr} 为训练样本(心拍)总数; \tilde{Y}_{QRS} , \tilde{Y}_P , \tilde{Y}_T 分别为预测的QRS波群、P波和T波的位置序列; Y_{QRS} , Y_P , Y_T 表示相应的真实位置序列,呈现为方波,由0,1构成,如图3所示.

2.5 参数设置

MBMT-TCN的网络参数设置如表1所示.在训练过程中,本文采用Adam优化器,设置学习率为0.0005,训练迭代次数为60,每次迭代都会在验证集上进行损失计算,如果模型的损失值连续3次没有下降,那么学习率减半.低通模块的频率范围为0.05~17 Hz,高通模块的频率范围为17~100 Hz.

表1 MBMT-TCN的参数设置

Table 1 Parameter configuration of MBMT-TCN

编码器(一维卷积)	
卷积核数量 N	128
低频信号的卷积核大小 L	8
高频信号的卷积核大小 L	8
原始信号的卷积核大小 L	8
卷积步长	4
TCN掩码估计器(TCN模组)	
瓶颈层卷积核数量	64
残差模块输入端的一维卷积层的卷积核数量	128
残差模块输出端的一维卷积层的卷积核数量	64
深度可分离卷积的卷积核大小	3
残差块堆叠层数	3
每层的残差块堆叠数量	8
解码器(一维反卷积)	
卷积核数量 N'	256
卷积核大小 L	8
卷积核步长	4

2.6 后处理

MBMT-TCN的输出为概率序列.本文通过阈值检测实现特征波概率序列至基准点位置序列的转化,具体方案如下:1)分别以P波概率序列 \hat{Y}_P^i 的“上升沿与阈值 $\text{Thr}_{P_{on}}$ 的交点”和“下降沿与阈值 $\text{Thr}_{P_{end}}$ 的交点”作为 P_{on} 和 P_{end} 的估计;2)以P波概率序列 \hat{Y}_P^i 超过阈值 $\text{Thr}_{P_{peak}}$ 的时段内的峰值作为 P_{peak} 的估计;3)分别以QRS波概率序列 \hat{Y}_{QRS}^i 的“上升沿与阈值 $\text{Thr}_{QRS_{on}}$ 的交点”和“下降沿与阈值 $\text{Thr}_{QRS_{end}}$ 的交点”作为 QRS_{on} 和 QRS_{end} 的估计;4)分别以T波概率序列 \hat{Y}_T^i 的“上升沿与阈值 $\text{Thr}_{T_{peak}}$ 的交点”和“下降沿与阈值 $\text{Thr}_{T_{end}}$ 的交点”作为 T_{peak} 和 T_{end} 的估计.设置不同的阈值,在验证集上计算各个基准点预测值与真实值之间的均方误差,结果如图4所示.可见,P波终点和峰值点的定位偏差对阈值变化较为敏感;不同基准点的最小均方误差所对应的阈值(虚线标识处)不尽相同.因此,文章分别选取验证集上各个基准点最小均方误差所对应的阈值作为后处理参数,将其应用于测试集的计算.因为MBMT-TCN输出的数据是重采样后固定长度的心拍片段,所以最后需要将基准点预测结果换算到原始ECG对应的位置上去,计算方式如下:

$$\hat{y}_i = \left\lfloor \frac{\tilde{y}_i r_i}{256} \right\rfloor, \quad (11)$$

其中: \tilde{y}_i 和 \hat{y}_i 分别为换算前后基准点的位置, i 为心拍序号, r_i 为重采样前心拍片段的长度, $\lfloor \cdot \rfloor$ 为取整符号.

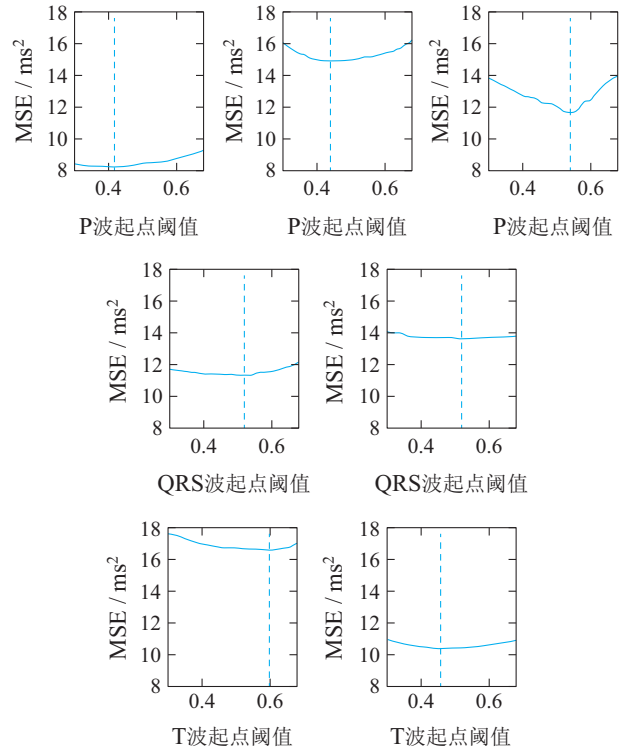


图4 在验证集上不同阈值所对应的基准点检测误差
Fig. 4 Fiducial point detection errors with different thresholds on the validation set

3 实验

3.1 数据来源

本文实验数据来自公开的麻省理工学院-波士顿心电图学会QT数据集(MIT-BIH QT) 心电数据集^[16]和LUDB(Lobachevsky university electrocardiography database)心电数据集^[17]. 其中MIT-BIH QT数据集由105条双导联的心电记录组成. 由于其样本较少, 采用第2.1节所述方式进行数据增强. 而LUDB数据集由200条12导联的心电记录组成, 数据充足不需要增强, 式(1)的 Q_{right} , Q_{left} 固定为0.7. 两套数据集都标注了QRS波群的起止点、P波的起止点和峰值点以及T波的终止点和峰值点. 采用五重交叉验证评估所提出基准点检测方法的有效性. 每一重实验按照6:2:2的比例划分训练集、验证集和测试集. 训练集、验证集和测试集的心电数据相互独立, 取自不同的心电记录.

3.2 评价指标

本文采用阳性预测值(positive predictive value, PPV)、灵敏度(sensitivity, Se), 以及时间偏差的平均值和标准差作为基准点检测性能的评价指标. PPV和Se的计算公式如下:

$$PPV = \frac{TP}{TP + FP} \times 100\%, \quad (12)$$

$$Se = \frac{TP}{TP + FN} \times 100\%, \quad (13)$$

其中: TP(true positive), FP(false positive), FN(false negative)分别为检测到的真阳性样本个数、假阳性样本个数和假阴性样本的个数. 若某个基准点的预测值与真实值的偏差小于150 ms, 则被视为真阳性样本; 否则, 被视为假阳性样本; 若在某个基准点真实值 ± 150 ms以内不存在预测值, 则该基准点被视为假阴性样本. 时间偏差的平均值(m)和标准差(σ)的计算公式如下:

$$m = \frac{1}{N_{ts}} \sum_{j=1}^{N_{ts}} (y_j - \hat{y}_j), \quad (14)$$

$$\sigma = \sqrt{\frac{1}{N_{ts}} \sum_{j=1}^{N_{ts}} (y_j - \hat{y}_j - m)^2}, \quad (15)$$

其中: N_{ts} 为测试集基准点的个数, y_j 和 \hat{y}_j 分别为基准点的真实位置和预测位置.

3.3 实验平台与基线方法

本文采用Win 10, Python 3.5安装Pytorch 1.10深度学习框架对模型进行训练和测试. 计算服务器的CPU为Intel Core i7-8700@3.2 GHz, 内存16 GB, GPU为NVIDIA Geforce RTX 3070, 显存8 GB.

在MIT-BIH QT数据集的实验中, 将所提出的MBMT-TCN与以下7种基线方法进行比较: 1) WT^[2], 将心电信号进行多层小波分解, 先对低层小波系数进行

阈值检测, 定位QRS波群, 然后对高层小波系数进行局部极值点搜索, 定位P波与T波的峰值点和起止点; 2) KNN^[7], 提取ECG信号的梯度特征, 通过最近邻分类器定位基准点; 3) EKF^[5], 采用扩展卡尔曼滤波模型估计ECG信号的隐状态(P波, T波和QRS波), 对基准点进行定位; 4) RAN^[11], 采用CNN提取ECG特征, 然后加入区域聚合网络代替传统的全连接层作为模型输出, 直接对ECG基准点位置进行预测; 5) 基于知识的深度学习(knowledge-based deep learning, KB-DL)^[12], 将ECG特征波脉高和脉宽的先验知识以及个人心拍的时长和幅值信息依次加入多级CNN编解码网络, 以增强基准点检测的鲁棒性; 6) 多频段单任务时序卷积网络(multi-band single-task temporal convolutional network, MBST-TCN), 删减MBMT-TCN的多基准点关联预测结构, 得到多通道独立任务时序卷积网络, 用于验证MBMT-TCN多任务框架的有效性; 7) 单通道多任务时序卷积网络(single-channel multi-task temporal convolutional network, SCMT-TCN), 删减MBMT-TCN的多频段分支结构, 得到单通道多任务时序卷积网络, 用于验证MBMT-TCN多频段特征提取结构的有效性.

在LUDB数据集的实验中, 将MBMT-TCN与以下4种基线方法进行比较: 1) WT^[18], 将ECG信号进行多层小波分解, 先采用高尺度系数粗略地获得各个特征波时段和峰值点位置, 然后依据低尺度系数估计特征波边界, 最后联合多电极信息对基准点位置进行细化; 2) 局部距离变换(local distance transform, LDT)^[19], 首先通过双阈值检测确定QRS位置, 接着通过自适应搜索窗定位P波和T波, 然后采用局部距离变换法确定各个电极基准点的位置, 最后联合多电极信息对基准点位置进行校准; 式(3)和式(4)分别为MBST-TCN和SCMT-TCN.

3.4 结果与分析

首先, 本文检验了在五重交叉验证中MBMT-TCN的损失函数学习曲线, 如图5所示. 可见, 经过50次epoch迭代训练后, 各重实验的训练集损失函数值均趋于收敛, 且在相应的验证集上并未出现严重的过拟合. 在验证集损失最小的模型被应用于测试集的基准点定位.

表2陈列了各种基准点定位方法在MIT-BIH QT数据集的性能指标. 由表2可见, 1) 数字信号处理方法WT^[2], EKF^[5]和KNN^[7]没有取得任何一项定位指标的最优值. WT方法对于小波基函数的选择和系数阈值的设置敏感, EKF的性能比较依赖模型初始值参数设置, KNN仅采用波形梯度特征检测基准点, 容易受到噪声的影响, 在低矮P波的起止点定位精度较差. 相较于深度学习方法(RAN, KB, SCMT-TCN, MBST-

TCN和MBMT-TCN), WT, EKF, KNN的表征能力不足, 难以适应ECG波形的变异性 and 噪声干扰的非平稳性; 2) MBMT-TCN性能指标整体优于RAN^[11], 与RAN使用传统的CNN提取ECG特征不同, MBMT-TCN引入TCN的注意力掩码估计模块, 增大了模型的感受野, 而且其多分支网络结构可以利用ECG特征波的频域分布特点和多基准点的关联推断关系, 进一步增强了检测模型的泛化能力; 3) MBMT-TCN性能指标与KB^[12]相当, 两者都是基于深度编解码架构的, 也都采用了ECG的先验知识提高基准点检测的鲁棒性. 但是, 两者有所不同. KB利用特征波时域先验知识, 而MBMT-TCN利用特征波的频域先验知识和特征波之间的关联推断关系. KB采用传统的CNN编解码网络, 在捕获ECG长程时序依赖关系上可能有所不足. 值得指出的是, KB涉及大量的与先验知识相关的超

参数, 在实验中却没有专门划分验证集选择模型, 所报告的定位性能指标可能偏高. 4) MBMT-TCN性能指标整体优于SCMT-TCN和MBST-TCN, 尤其是在 P_{peak} , QRS_{end} 和 T_{end} 的平均时间偏差仅为第2名的23%, 30%和44%. 该结果验证了多频段ECG特征提取和多基准点联合检测的优越性.

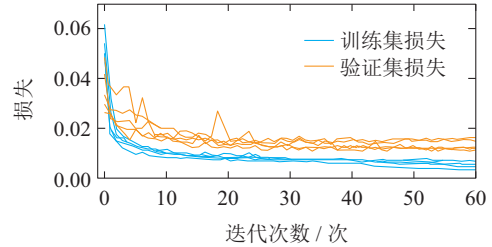


图5 MBMT-TCN模型训练和验证损失曲线

Fig. 5 Loss curves of the MBMT-TCN model in training sets and validation sets

表2 MIT-BIH QT数据集上的心电信号基准点检测性能比较

Table 2 Comparison of ECG fiducial point detection performance in MIT-BIH QT database

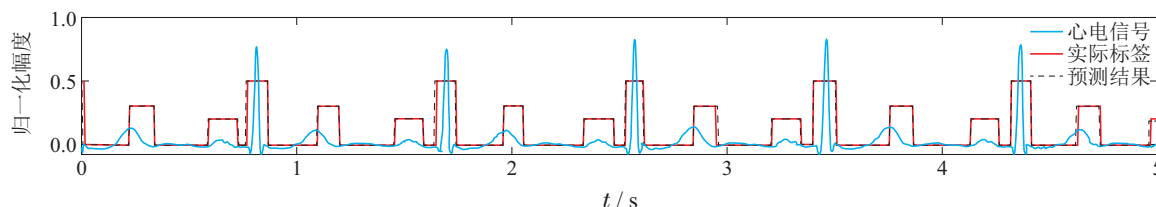
方法	指标	P_{on}	P_{peak}	P_{end}	QRS_{on}	QRS_{end}	T_{peak}	T_{end}
MBMT-TCN	Se/%	99.53	99.74	99.84	99.97	99.97	99.53	99.84
	PPV/%	99.81	99.36	99.56	98.30	98.73	99.20	99.23
	m±SD/ms	-0.9±13.4	0.3±18.8	0.7±20.2	-0.3±16.4	0.7±20.6	1.6±25.3	0.4±18.1
SCMT-TCN	Se/%	99.25	99.11	99.02	99.46	99.31	99.01	99.62
	PPV/%	99.09	98.27	98.87	96.22	96.27	98.10	98.11
	m±SD/ms	1.4±13.9	-1.6±19.0	-1.1±23.4	-1.0±18.6	-2.3±20.2	-2.3±29.8	-0.9±18.7
MBST-TCN	Se/%	99.76	99.59	99.37	99.84	99.77	99.68	99.60
	PPV/%	99.81	98.26	98.16	98.27	98.16	98.89	99.23
	m±SD/ms	-2.9±10.4	1.3±16.7	0.2±20.2	-0.3±15.9	-2.5±15.6	2.4±20.9	3.1±17.8
WT ^[2]	Se/%	98.87	98.87	98.75	99.97	99.97	N/A	99.77
	PPV/%	91.03	91.03	91.03	N/A	N/A	N/A	97.79
	m±SD/ms	2.0±14.8	3.6±13.2	1.9±12.8	4.6±7.7	0.8±8.7	N/A	-1.6±18.1
EKF ^[5]	Se/%	97.92	98.04	97.92	N/A	N/A	98.54	98.54
	PPV/%	98.61	98.63	98.61	N/A	N/A	98.81	98.81
	m±SD/ms	7.4±8.1	7.2±9.8	-9.0±6.3	N/A	N/A	3.2±13.2	-9.3±5.7
KNN ^[7]	Se/%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	PPV/%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	m±SD/ms	2.5±14.3	N/A	-3.1±13.9	2.4±8.8	4.7±9.2	N/A	1.6±18.1
RAN ^[11]	Se/%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	PPV/%	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	m±SD/ms	0.4±14.4	-0.4±10.1	-2.0±12.7	-0.7±10.9	-4.8±12.1	-3.0±10.5	-0.3±18.5
KB ^[12]	Se/%	99.40	N/A	99.40	99.98	99.99	N/A	99.15
	PPV/%	99.69	N/A	99.68	99.96	99.99	N/A	99.60
	m±SD/ms	0.3±6.4	N/A	0.1±7.9	-0.5±9.5	-0.8±7.8	N/A	-1.5±25.6

图6为MBMT-TCN, SCMT-TCN和MBST-TCN在MIT-BIH QT数据集1段ECG上的检测结果, 采用方波序列标识各心拍P波、QRS波和T波的起止位置. 由图6可见: MBMT-TCN的检测结果与真实标签更加吻合, 尤其在红色圆圈标出的 QRS_{end} , T_{peak} 和 P_{on} 点, MB-

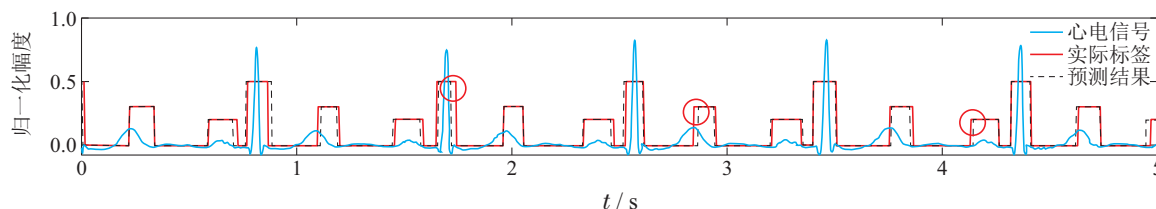
MT-TCN的定位偏差明显小于SCMT-TCN和MBST-TCN的定位偏差. 相对于SCMT-TCN的单通道多任务模型和MBST-TCN的多通道单任务模型, MBMT-TCN的多通道多任务架构具有同时利用特征波频域分布差异和时域关联信息的优势, 基准点定位更准确.

表3陈列了各种基准点定位方法在LUDB数据集的性能指标. 由表3可见, 1) 本文所提出的TCN系列方法的灵敏度、阳性预测值和平均定位误差全面优于WT和LDT, 尤其是在QRS_{on}, 所提出方法的平均定位误差仅为WT和LDT的8%~37%. 相对于数字信号处理方法, 基于深度TCN编解码模型的基准点定位方法能够更好地适应ECG的波形变化, 而且MBMT-TCN,

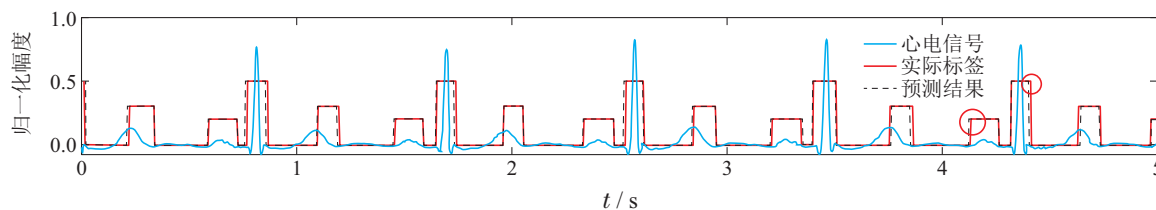
MBST-TCN和SCMT-TCN都是单电极定位方法, 无须联合多电极信息对基准点进行细化或校准; 2) MBMT-TCN的检测指标整体优于SCMT-TCN和MBST-TCN. 尤其是在 P_{peak} , T_{peak} 的平均时间偏差仅为第2名的25%, 30%. 该结果再次验证了MBMT-TCN同时利用特征波频域分布差异和时域关联关系提升检测性能的有效性.



(a) MBMT-TCN的心电信号特征波定位结果



(b) SCMT-TCN的心电信号特征波定位结果



(c) MBST-TCN的心电信号特征波定位结果

图 6 所提出的基于TCN的ECG基准点检测方法的定位结果示例

Fig. 6 Localization examples of the proposed TCN-based methods for ECG fiducial detections

表 3 LUDB数据集上的心电信号基准点检测性能比较

Table 3 Comparison of ECG fiducial point detection performance in LUDB database

方法	指标	P_{on}	P_{peak}	P_{end}	QRS _{on}	QRS _{end}	T_{peak}	T_{end}
MBMT-TCN	Se/%	99.81	99.89	99.93	100.00	100.00	99.97	99.91
	PPV/%	97.11	98.71	97.22	100.00	100.00	99.62	99.56
	m ± SD/ms	-2.3 ± 7.6	-0.2 ± 12.5	-1.3 ± 14.4	-1.2 ± 8.1	2.5 ± 14.2	-0.3 ± 13.0	0.9 ± 7.5
SBMT-TCN	Se/%	99.33	98.33	99.97	100.00	100.00	99.88	99.86
	PPV/%	92.29	99.45	92.89	99.84	99.84	99.55	99.47
	m ± SD/ms	-1.3 ± 9.5	-1.2 ± 13.0	-3.3 ± 17.6	-2.8 ± 11.5	-3.3 ± 18.8	1.5 ± 16.5	2.4 ± 8.0
MBST-TCN	Se/%	99.72	99.23	99.88	100.00	100.00	99.92	99.97
	PPV/%	96.89	99.77	96.82	99.95	99.95	99.68	99.42
	m ± SD/ms	1.9 ± 8.8	-0.8 ± 10.8	1.6 ± 14.9	3.0 ± 7.8	-2.1 ± 17.7	1.0 ± 15.1	1.2 ± 7.9
WT ^[18]	Se/%	98.46	98.46	98.46	99.61	99.61	99.03	98.03
	PPV/%	96.41	96.41	96.41	99.87	99.87	98.84	98.84
	m ± SD/ms	-2.7 ± 10.2	-0.3 ± 6.2	-0.4 ± 11.4	-8.1 ± 7.7	3.8 ± 8.8	-4.0 ± 7.4	-5.7 ± 15.5
LDT ^[19]	Se/%	98.43	98.43	98.43	99.89	99.89	99.27	99.21
	PPV/%	96.44	96.44	96.44	99.86	99.86	98.85	98.85
	m ± SD/ms	2.2 ± 7.4	-0.8 ± 5.5	-6.5 ± 10.7	15.4 ± 14.6	-3.8 ± 13.6	-0.5 ± 5.5	-1.2 ± 6.8

4 结论

现有的深度学习方法在实现心电图基准点检测时,既未直接利用特征波之间的频域分布知识,也未在网络结构上清晰地表征基准点位置之间的关联推断关系,鲁棒性和泛化能力有待进一步提高.本文提出了基于多频段多任务编解码模型的ECG基准点联合检测方法,能够有效地提高网络输入的信噪比,增强ECG特征提取的鲁棒性,同时具有多分支关联预测网络结构,能够利用基准点定位任务之间的推断关系,提高检测精度.采用MIT-BIH QT数据集和LUIDB数据集大量的临床ECG样本进行实验,结果验证了所提出方法的优越性.

参考文献:

- [1] Chinese Cardiovascular Health and Disease Report Compilation Group. China cardiovascular health and disease report 2020 summary. *Chinese Journal of Circulation*, 2021, 26(3): 209–218. (《中国心血管健康与疾病报告2020》编写组.《中国心血管健康与疾病报告2020》要点解读.中国心血管杂志, 2021, 26(3): 209–218.)
- [2] MARTINEZ J P, ALMEIDA R, OLMOS S, et al. A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE Transactions on Biomedical Engineering*, 2004, 51(4): 570–581.
- [3] LENIS G, PILIA N, OESTERLEIN T, et al. P wave detection and delineation in the ECG based on the phase free stationary wavelet transform and using intracardiac atrial electrograms as reference. *Biomedical Engineering-biomedizinische Technik*, 2016, 61(1): 37–56.
- [4] RINCIN F, RECAS J, KHALED N, et al. Development and evaluation of multilead wavelet-based ECG delineation algorithms for embedded wireless sensor nodes. *IEEE Transactions on Information Technology in Biomedicine*, 2011, 15(6): 854–863.
- [5] SAYADI O, SHAMSOLLAHI M B. A model-based Bayesian framework for ECG beat segmentation. *Physiological Measurement*, 2009, 30(3): 335–352.
- [6] HOSSAIN M B, BASHAR S K, WALKEY A J, et al. An accurate QRS complex and P wave detection in ECG signals using complete ensemble empirical mode decomposition with adaptive noise approach. *IEEE Access*, 2019, 7: 128869–128880.
- [7] SAINI I, SINGH D, KHOSLA A. K-nearest neighbour-based algorithm for P-and T-waves detection and delineation. *Journal of Medical Engineering and Technology*, 2014, 38(3): 115–124.
- [8] MEHTA S, LINGAYAT N. Comparative study of QRS detection in single lead and 12-lead ECG based on entropy and combined entropy criteria using support vector machine. *Journal of Theoretical and Applied Information Technology*, 2007, 3(2): 570–588.
- [9] PEIMANKAR A, PUTHUSSERYPADY S. DENS-ECG: A deep learning approach for ECG signal delineation. *Expert Systems with Applications*, 2021, 165: 113911.
- [10] JIMENEZ PEREZ G, ALCALINE A, CAMARA O. U-Net architecture for the automatic detection and delineation of the electrocardiogram. *Computing in Cardiology (CinC)*. Santiago de Compostela, Spain: IEEE, 2019: 1–4.
- [11] CHEN M, WANG G, XIE P, et al. Region aggregation network: Improving convolutional neural network for ecg characteristic detection. *The 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu, Hawaii: IEEE, 2018: 2559–2562.
- [12] WANG J, LI R, LI R, et al. A knowledge-based deep learning method for ECG signal delineation. *Future Generation Computer Systems*, 2020, 109: 56–66.
- [13] RAMAKRISHNAN A G, PRATHOSH A P, ANANTHAPADMAN-ABHA T V. Threshold-independent QRS detection using the dynamic plosion index. *IEEE Signal Processing Letters*, 2014, 21(5): 554–558.
- [14] ARINI P D, LIBERCZUK S, MENDIETA J G, et al. Electrocardiogram delineation in a Wistar rat experimental model. *Computational and Mathematical Methods in Medicine*, 2018, DOI: 10.1155/2018/2185378.
- [15] LUO Y, MESGARANI N. Conv-tasnet: Surpassing ideal time-frequency magnitude masking for speech separation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2019, 27(8): 1256–1266.
- [16] LAGUNA P, MARK R G, GOLDBERG A, et al. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. *Computers in Cardiology 1997*, 1997, DOI: 10.1109/CIC.1997.648140.
- [17] KALYAKULINA A I, YUSIPOV I I, MOSKALENKO V A, et al. LUIDB: A new open-access validation tool for electrocardiogram delineation algorithms. *IEEE Access*, 2020, 8: 186181–186190.
- [18] KALYAKULINA A I, YUSIPOV I I, MOSKALENKO V A, et al. Finding morphology points of electrocardiographic-signal waves using wavelet analysis. *Radiophysics and Quantum Electronics*, 2019, 61(8): 689–703.
- [19] CHEN G, CHEN M, ZHANG J, et al. A crucial wave detection and delineation method for twelve-lead ECG signals. *IEEE Access*, 2020, 8: 10707–10717.

作者简介:

李磊 硕士研究生, 目前研究方向为生物医学信号处理、机器学习, E-mail: lileihey@foxmail.com;

廖桂鑫 硕士研究生, 目前研究方向为生物医学信号处理、机器学习, E-mail: 2112104074@mail2.gdut.edu.cn;

蔡瑞涵 硕士研究生, 目前研究方向为生物医学信号处理、机器学习, E-mail: cbaen0202@gmail.com;

李珍妮 副教授, 博士, 目前研究方向为机器学习, E-mail: lizhen@gdut.edu.cn;

吕俊 副教授, 博士, 目前研究方向为生物医学信号处理、机器学习, E-mail: lujun.rylj@gmail.com.