

# 基于YOLO-1D架构的生物医学信号检测方法研究

张昊

北京印刷学院信息工程学院, 北京

收稿日期: 2026年1月16日; 录用日期: 2026年3月3日; 发布日期: 2026年3月12日

## 摘要

深度学习技术在生物医学信号处理中显示出了巨大的潜力, 基于卷积神经网络的方法应用于心电图(ECG)和光容积脉搏波图(PPG)有很好的效果。但是, 大部分方法对于复杂情况下的信号质量变化、噪声干扰以及基线漂移并不能很好地处理, 造成峰值检测的准确性以及漏检率不高。本文针对这种情况提出一种新颖的深度学习框架YOLO-1D, 将2D的目标检测算法迁移到1D的生物医学信号处理上, 在该方法的基础上完成对心电图的复杂场景下精准的多峰检测和心律失常分类, 该方法的核心思想是采用YOLO架构的一步检测方式及多尺度特征融合, 通过DFL分布聚焦损失定位精准的位置; 并且为了满足心电图的特殊性进行了针对心电图的基线漂移、时间扭曲、信号倒置以及MixUp的数据增强策略, 检测头采用可扩展分类支路的方法来同时学习峰值检测和心律失常分类, 利用加权1D非极大值抑制(1DNMS)以及容忍度匹配机制准确地量化多峰检测效果。通过大量实验发现, YOLO-1D能有效利用基线漂移数据集以及心电图额外的数据源扩充其训练集, 从而使网络性能更好, 且检测的准确度非常高, 在噪声和干扰的大环境下性能也能保持稳定。对于该模型来说: 在MIT-BIH心律失常数据库上的ECGR峰检出测试上, 取得了精确率达到了96.55%、召回率为98.00%、F1值达到了97.27%的测试结果; 在AAMI五分类标准下ECG的心律失常分类任务, 准确率为95.2%, 宏平均F1值为92.6%; 而在PPG峰值检测方面, 精确率为95.8%, 召回率为96.3%, F1值为96.0%, 相比于传统基于scipy.signal.find\_peaks的方法(精确率为90.2%, 召回率为92.5%, F1分数为91.3分), 在精确率、召回率以及F1分数上分别提高了5.6、3.8及4.7个百分点。之后进行了模型升级更新, 模型使用更新到YOLOv11版本, 在后续实验中检测精度得到了一定的提高, 其中包括ECG R峰检测F1达到了97.8%, PPG峰值检测F1也达到了96.5%。另外, 还尝试做了脉搏间期(PPI)回归的任务, 获得其在8秒窗口检测MAE达到0.42的较好结果。消融实验表明, 在采用SPP模块、SEBlock注意力机制和数据增强等完整配置下, YOLO-1D的F1值比基础架构提升了2.17个百分点。

## 关键词

信号检测, YOLO-1D, 深度学习, ECG, PPG, 峰值检测, 心律失常分类, 生物医学信号处理

## Research on Biomedical Signal Detection Methods Based on the YOLO-1D Architecture

Hao Zhang

## Abstract

Deep learning techniques have demonstrated great potential in the field of biomedical signal processing, and methods based on convolutional neural networks (CNNs) have achieved remarkable results in the analysis of electrocardiogram (ECG) and photoplethysmogram (PPG) signals. However, existing methods often perform unsatisfactorily in complex scenarios such as signal quality fluctuations, noise interference, and baseline drift, which tends to reduce the accuracy of peak detection and increase the false detection rate. In this study, we propose YOLO-1D, a novel deep learning framework that successfully adapts 2D object detection algorithms to 1D biomedical signal processing tasks, enabling high-precision peak detection and arrhythmia classification. Its core idea is to leverage the one-stage detection capability and multi-scale feature fusion mechanism of the YOLO architecture, combined with Distribution Focal Loss (DFL) to achieve accurate position regression; meanwhile, tailored data augmentation strategies (including baseline drift, time warping, signal inversion, and MixUp) are designed according to the characteristics of biomedical signals. The detection head can be extended with a classification branch to realize the joint learning of peak detection and arrhythmia classification. In addition, we introduce 1D Non-Maximum Suppression (1D NMS) and a tolerance matching mechanism to more accurately evaluate the performance of multi-peak detection. Experimental results show that the YOLO-1D method can generate high-quality peak detection results and maintain stable performance even in noisy and interfering environments. Specifically: In the ECG R-peak detection task on the MIT-BIH Arrhythmia Database, the precision, recall, and F1-score reach 96.55%, 98.00%, and 97.27%, respectively; in the ECG arrhythmia classification task under the AAMI five-classification standard, the accuracy and macro-average F1-score are 95.2% and 92.6%, respectively; in the PPG peak detection task, the precision, recall, and F1-score are 95.8%, 96.3%, and 96.0%, respectively. Compared with the traditional method based on `scipy.signal.find_peaks` (precision: 90.2%, recall: 92.5%, F1-score: 91.3%), the proposed method achieves improvements of 5.6, 3.8, and 4.7 percentage points in precision, recall, and F1-score, respectively. In subsequent research, the model is updated to the YOLOv11 version, further improving the detection accuracy: the F1-score for ECG R-peak detection reaches 97.8%, and that for PPG peak detection reaches 96.5%. In addition, we explore the Pulse Interval (PPI) regression task and achieve a good result with a Mean Absolute Error (MAE) of 0.42 in 8-second window detection. Ablation experiments show that when YOLO-1D adopts the complete configuration including the SPP module, SEBlock attention mechanism, and data augmentation, the F1-score is 2.17 percentage points higher than that of the basic architecture.

## Keywords

Signal Detection, YOLO-1D, Deep Learning, ECG, PPG, Peak Detection, Arrhythmia Classification, Biomedical Signal Processing

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

生物医学信号的峰值检测是目前心血管疾病诊断及健康监测技术中极为重要的基础技术。随着可穿

戴设备和远程医疗的发展,需要更可靠稳定、鲁棒性更高的信号处理算法,其中心电图(R-peak)的心电(ECG)信号、光电容积脉搏波(PPG)峰值是进行心率变异性分析、心律失常识别、以及得到心率、血压等关键生理参数的重要基础。传统的 R 峰检测算法主要是基于心电图 Pan-Tompkins 等经典的滤波与形态识别的方法,在理想状态下可以得到较为满意的检测结果,但是依然存在一些问题:容易受噪声的影响,且对参数进行调整具有一定的困难性,无法很好地消除基线漂移。不能够完全地满足现代的应用要求。

深度学习技术具有强大的信号表征和特征提取能力,在生物医学信号处理中有极大的应用前景,而目前许多研究均将生物医学信号中的峰值检测当做分类任务,无法确定峰值所在的确切位置。将传统的二维(2D)目标检测算法(例如 YOLO)运用于一维(1D)信号上,利用端到端训练方法以及多尺度特征融合方式,可以很好地克服传统方法存在的缺点,从而满足实际应用要求。同时,利用了 YOLO 系列的单阶段预测以及实时性的优点,将其核心思想运用到 1D 时序信号中,实现了不同尺度峰值的同时检测,并将分类分支拓展开来实现心律失常的自动分类。

本研究的主要贡献包括:(1)把 YOLO 算法引入到 1D 生物医学信号处理当中,并且专门设计了 YOLO-1D 网络模型结构;(2)结合分类损失和回归损失(Distribution Focal Loss, DFL)来进行亚像素级别上的峰值位置的精准回归;(3)增加了检测头的结构来实现峰值检测和心律失常分类的联合学习;(4)针对生物医学信号的特点制定了相对应的数据增强策略;(5)最后使用 ECG、PPG 信号数据集验证了提出的算法的有效性以及从 YOLOv8 版本迭代更新到 YOLOv11 的优化过程。

## 2. 相关工作

### 2.1. 传统峰值检测方法

除了传统的阈值法、导数法、小波变换以及模板匹配等方法以外,传统峰值检测方法主要包括:Pan-Tompkins 算法[1]是使用自适应阈值实现 R 波峰值检测的一种经典算法,在 MIT-BIH 数据库中能正确检测到 R 波的次数大约为总数的 99.3%,且具有计算复杂度低、实时性好的特点,但是容易受到噪声影响,在信噪比较低的情况下会严重影响检测的准确性。Hamilton-Tompkins 算法[2]利用一阶导数过零点寻找峰值,容易受基线漂移的影响。小波变换具有多尺度分析的优点,但是计算量大[3]。模板匹配需要事先建立模板,泛化能力弱。近年学者对传统的方法作了改进,Khan 和 Imtiaz [4]提出的 Pan-Tompkins++采用改进滤波与多阈值的方法,其在各公共数据库上的 F-score 平均提高了 2.2%,FP 和 FN 平均降低了 2.8%和 1.8%,Lin 等[5]应用动态阈值配合模板匹配法,使其灵敏度和准确度都上升到了 99.78%,位置误差从原始 Pan-Tompkins 的 21.65 ms 降到了 8.35 ms。Zahid 等人[6]提出的 1D CNN 方法在低质量 Holter ECG 数据集上实现召回率为 99.69% (CPSC-DB)以及 99.85% (MIT-BIH),且精准率为 98.91%和 99.82%,从整体上讲,目前传统方法已经达到了在高质量信号下的最佳性能水平,在提高噪声鲁棒性、亚采样级准确定位以及跨库泛化等其他方面还有很大的提升空间。

### 2.2. 基于深度学习的峰值检测方法

卷积神经网络(CNN)方法在 ECG 心律失常分类方面效果较好,Rajpurkar 等[7]采用 CNN 方法实现 14 类心律失常分类,在经过训练后能达到心脏病专家水平,其准确率 >90%,但其方法需要先进行 R 峰检测才能输入到网络中进行心律失常分类,所以不能实现峰值检测和分类的同时进行,此外大部分的 CNN 方法都将峰值检测转变为分类问题而使定位难以精确到亚采样级别。而 RNN 及它的变体 LSTM、GRU 可以捕获时间上的相关性,但是计算量较大,推断时延也较大[8]。

此外,DENS-ECG 等混合模型将 CNN 与 LSTM 相结合,实现了 97.95%的敏感度及 95.68%的准确率的优异波形分割结果[9],但由于其目标主要针对波形分割而不是心电波形精准定位,且算法本身计算复

杂度偏高；Transformer 方法采用自注意力机制建立长距离依赖关系[10]，但由于序列越长其计算复杂度过高而难以实现，因此目前较少应用于峰值检测任务。在心律失常分类任务上，常规的心律失常识别方法均是将 R 波检测和分类分为两个独立的阶段来分别解决，从而由于这两类任务紧密相关的特点导致前面产生的误差无法得以修正，而基于深度学习的端到端分类方法虽然可以完成这一目标，但大部分工作仍然是采用两阶段的工作方法，没有利用好这两类任务间相互影响关联性的特点。从本质上来说，现阶段心律失常识别方法都具有一定的不足，主要是无法做到精确定位、检测与分类任务相分离以及在大量计算资源上的巨大浪费和在单维度上进行深入研究的不足等问题。

### 2.3. 目标检测方法在信号处理中的应用

YOLO (You Only Look Once) [11]算法运用回归模型完成单阶段目标检测以及端到端训练，避免了传统两阶段方法繁琐的过程，YOLO 算法解决了目前流行的 2D 图像目标检测的端到端训练和实时推理的问题。在该模型上增加了 C2f 模块以及 SPP 模块，实现了多尺度特征的提取，使 YOLOv8 获得更好的精度及速度之间的平衡。针对网络结构进一步进行优化，并提高了模型的精度及稳定性。而目标检测方式可以做到端到端训练、单阶段检测、多尺度特征融合等优势，在一维生物医学信号中没有得到较好的应用。本次工作是首次把 YOLO 算法应用于单维信号，建立了 YOLO-1D 的模型，采用 DFL 进行亚采样级位置回归，使用 1DNMS 去除冗余检测，并以一种创新的方法把峰值检测和心律失常分类进行联合建模，避免了两阶段错误累积的问题。

## 3. 实验

### 3.1. 网络架构

YOLO-1D 架构的整体设计如图 1 所示：输入为一维生物医学信号，经特征提取骨干网络处理后，由检测头输出峰值位置及其置信度。网络结构包含 Stem 层、C2f 模块、SPP 模块、SEBlock 注意力机制以及检测头。Stem 层由一维卷积(Conv1D)、批归一化(Batch Normalization)和 SiLU 激活函数构成，执行初始特征提取，将输入信号通道数从 1 扩展至 64。C2f 模块作为 YOLOv8 的核心组件，通过特征融合增强网络表达能力；本架构包含两个 C2f 层，分别将通道数从 64 扩展至 128 和 256。SPP 模块通过不同尺度(卷积核尺寸 kernel\_size = 5, 9, 13)的最大池化操作提取多尺度特征。SEBlock 采用通道注意力机制，自适应地重新校准通道特征响应(缩减因子 reduction=16)。检测头包含分类分支与回归分支：分类分支输出各位置为峰值的概率；回归分支输出峰值位置的精确偏移量(采用 DFL 编码，reg\_max = 16)。

### 3.2. 损失函数

(1) 峰值检测损失函数：总体损失函数由分类损失与回归损失构成：

$$L_{total} = \lambda_{cls} \cdot L_{cls} + \lambda_{reg} \cdot L_{reg} \quad (1)$$

其中  $\lambda_{cls} = 1.0$ ， $\lambda_{reg} = 5.0$ 。分类损失采用二元交叉熵损失(BCE With Logits Loss)。回归损失作用于 DFL 解码后的偏移量，采用 L1 损失。分布焦点损失(Distribution Focal Loss, DFL)将回归问题转化为分类问题，通过预测偏移量的分布以获取更精确的位置估计：

$$\hat{d} = \sum_{j=0}^{R-1} j \cdot \text{softmax}(p_j), \quad (2)$$

其中  $R = 16$  为分布的最大值。

(2) 心律失常分类损失函数：在心律失常分类任务中，损失函数扩展为：

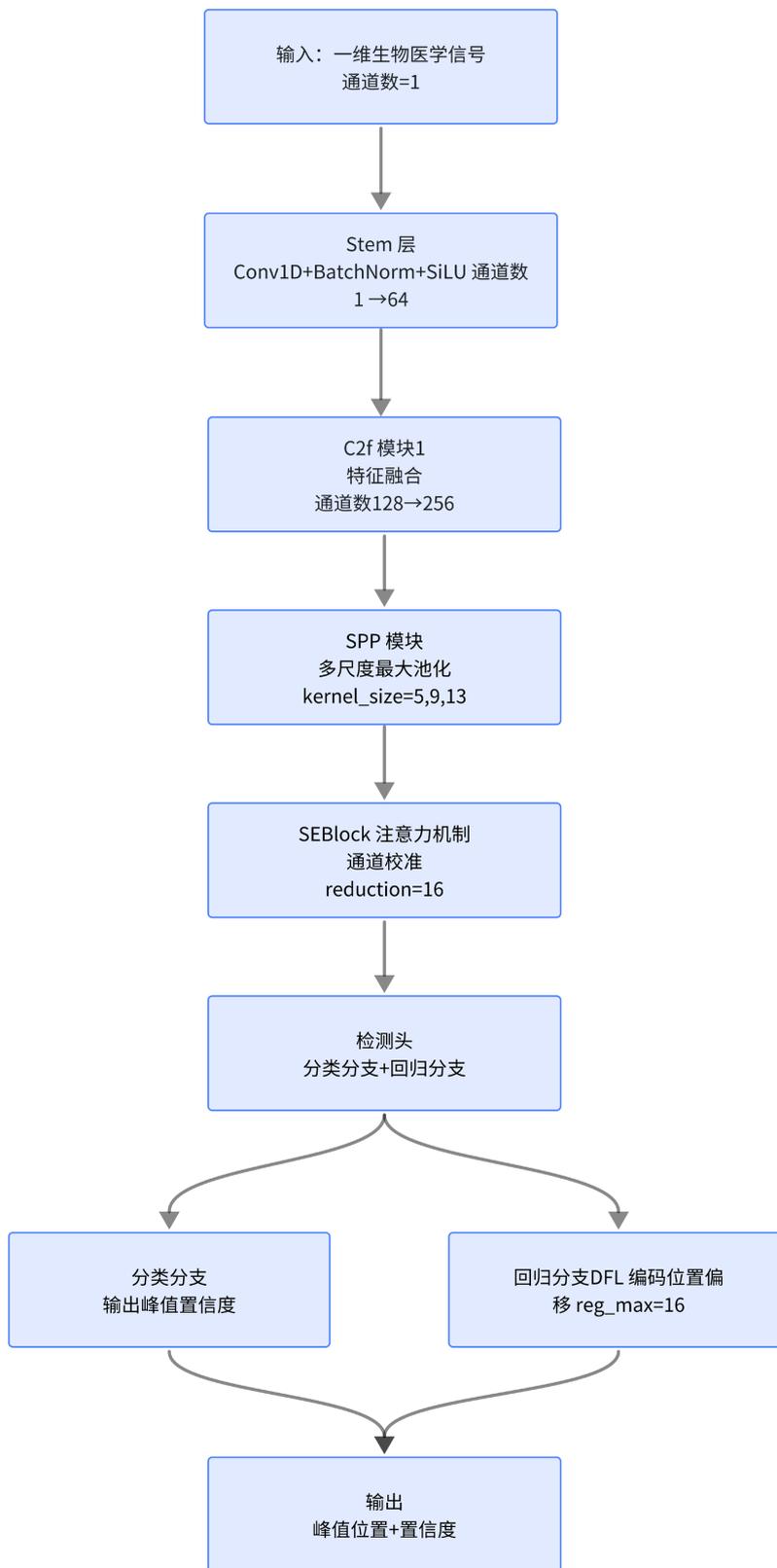


Figure 1. Model network architecture diagram  
图 1. 模型网络结构图

$$L_{total} = \lambda_{cls} \cdot L_{cls} + \lambda_{reg} \cdot L_{reg} + \lambda_{arrhythmia} \cdot L_{arrhythmia}, \quad (3)$$

其中  $L_{arrhythmia}$  为心律失常分类损失, 采用交叉熵损失(Cross Entropy Loss),  $\lambda_{arrhythmia} = 2.0$ 。对于每个检测到的峰值位置, 模型输出其对应的心律失常类型概率分布, 分类损失计算预测概率分布与真实标签之间的交叉熵。在联合学习模式下, 峰值检测与心律失常分类共享特征提取网络, 通过多任务学习提升整体性能。

### 3.3. 信号特征与检测任务

心电图(ECG)记录心脏肌肉电活动, 其中 R 波峰是 HRV 分析、心律失常识别的基础。光电容积脉搏波(PPG)由血液容积变化引起, 反映心血管的状态, 具有时空局部性与空间邻近性; 心电信号尖峰检测注重时间和采样分辨率, 在去除外界扰动的影响方面存在较大困难, PPG 信号尖峰检测在去除运动伪影和光学噪声影响方面更为棘手。PPG、ECG 的峰值检测都可以利用目标检测框架中“边界框/中心点/置信度”的思想将目标检测问题等价为一维时序下的“检测区间/峰值位置/置信度”。ECG 心律失常分类任务可以将每个心跳映射到 AAMI 五类系统, 并且采取“先检测后分类”或者“联合学习”两种模式: (1)YOLO-1Dv0 先检测再分类, 通过两个单独的网络分别完成两部分任务, 在同一模型中两者的任务目标不进行联合学习; (2)YOLO-1Dv1 实现联合学习, 使得位置与类别可以同时预测。

### 3.4. 数据处理

(1) 从 MIT-BIH 数据库读取 .dat 文件, 并从 MIT-BIH 数据库提取导联 II, 使用 Neurokit2 的 rpeakdetect() 函数来得到导联 II 的 R 波, 利用 Neurokit2 库的该函数作为金标准(groundtruth), 把原始长信号划分为 8s 大小滑动窗口, 间隔为 2 s, 在此基础上将每个窗口信号进行 Z-score 标准化处理(如图 2)。

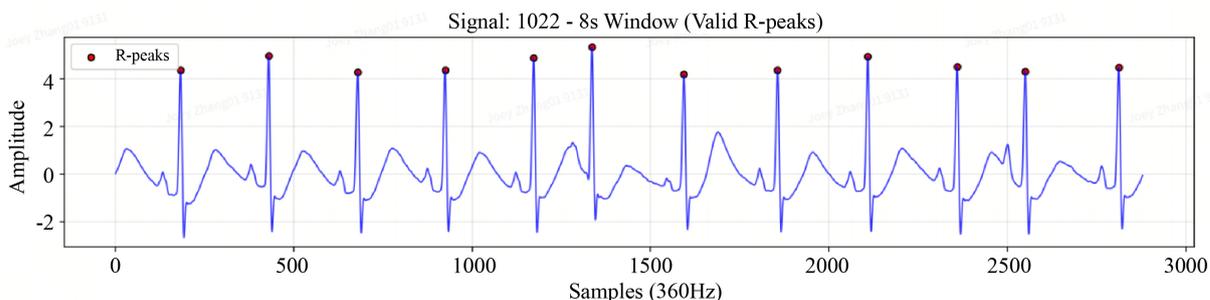


Figure 2. Data processing visualization  
图 2. 数据处理可视化

(2) 基于 PPG 的 PPI 预处理使用滑窗均值方法去除基线漂移; 使用高通滤波(截止频率为 0.5 Hz)、带通滤波(通带为 0.5~8 Hz)和中值滤波器去除脉冲噪声, 使用 Savitzky-Golay 进行平滑以保持峰值形状, 使用 Z-score 方式完成归一化。对于 PPI 检测任务而言, 还需要将 ECG 信号 R-R 间期(RRI)的数据与 PPG 信号 PPI 的数据对齐到同一时间轴。由于设备的采样率波动以及生理学上的时间延迟问题, 在人工标定好相应的时间延迟之后, 使用 Neurokit2 包检测 ECG 信号的峰点得到 RRI 以及心率(把 RRI 转换成心率进行标注), 作为金标准。

(3) 数据增强策略: 针对生物医学信号特性, 设计了基线漂移(叠加低频正弦波  $s'(t) = s(t) + A \sin(2\pi ft)$ , 振幅  $A = 0.1$ , 频率  $f = 0.5$  Hz)、时间扭曲(对时间轴施加随机扰动)、信号倒置、幅值缩放(缩放因子  $\alpha \sim \mathcal{U}(0.3, 2.5)$ )以及 MixUp(混合系数  $\lambda \sim \text{Beta}(0.4, 0.4)$ )等数据增强方法。

(4) 后处理: 使用 1D NMS, 当设置距离阈值分别为 ECG 的 50 个采样点和 PPG 的 30 个采样点时, 就可去除重复检测; 对于检测性能评价, 则通过容忍度匹配的方式判定匹配正确与否, 即若预测峰位与真值峰位的距离小于 5 个采样点(ECG)或 3 个采样点(PPG), 则认为匹配正确。

### 3.5. 实验设置

(1) 数据集: 目前可以使用的是 MIT-BIH 心律失常数据库的 48 个 ECG 记录(采样率为 360Hz, 每条记录约 30 分钟)。记录根据 ID 被划分为了训练集(前面 24 个)和测试集(后面 24 个)。各个类别的数量与分布基本一致, 可以代表人的心律失常的各类情况, 包括正常(N)、室上性异位(S)、室性异位(V)、融合(F)以及未知(Q)这 5 类。PPG 数据集约 200 名受试者, 包含近 500 条记录(重采样为 100 Hz), 按照受试者来划分数据集, 分别用来进行训练、验证和测试, 比例分别为 70%、15%、15%; 从中抽取了近 1.5 万个 8 秒长的窗口作为 ECG 的输入样本; 同理也抽取了近 1.2 万个 8 秒长的窗口作为 PPG 的输入样本。检测到 R 峰的位置时分别取以 R 峰为中心向前后各提取 200 个采样点(约为 0.56 秒)的原始心搏数据组成单个的心搏分类样本。

(2) 训练配置: 使用 PyTorch 2.0.1, 在 NVIDIA 的 4090 GPU 上完成训练工作,  $batch\_size = 128$ 、起始  $learning\_rate = 3e-4$ 、用 cosine annealing learning rate scheduler 调优 learning rate、选用 AdamW 优化器(权重衰减值为  $1e-4$ )、训练 50 epoch, 并采用 EarlyStopping (patience = 10), 开启 FP16 混合精度训练、梯度裁剪最大范数(max norm) = 1.0。

(3) 基线方法: 采用经典方法为基线进行对比, 即 Pan-Tompkins 算法、Hamilton-Tompkins 算法、基于小波变换的方法、scipy.signal.find\_peaks 等; 同时选用的深度学习方法有: CNN 分类模型、CNN-LSTM 混合模型、U-Net 分割模型。

(4) 评估指标: 峰值检测任务主要采用精确率( $P = \frac{TP}{TP+FP}$ )、召回率( $R = \frac{TP}{TP+FN}$ )和 F1 分数( $F1 = \frac{2PR}{P+R}$ )进行评估。心律失常分类任务采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1 分数及混淆矩阵(Confusion Matrix)进行评估。多分类任务的精确率、召回率和 F1 分数均按宏平均(Macro-average)和微平均(Micro-average)两种方式计算。所有实验重复 5 次, 报告结果的平均值及标准差, 并采用配对 t 检验评估性能差异(显著性水平  $\alpha = 0.05$ )。

### 3.6. 训练过程

使用 Kaiming 初始化的方法, 并设定随机种子 42, 使用端到端的方式进行训练, 设置 FP16 混合精度训练, 设置梯度裁剪最大范数为 1.0。损失函数在第一个 epoch 中就会有快速地降低, 在验证集上, F1 分数在第 25~30 epoch 左右达到最大值(约 97.2%)。采用早停策略(patience = 10)停掉模型, 实际经过训练 epochs 大约为 35~40。根据网格搜索选出最好超参数:  $lr = 3 \times 10^{-4}$ ,  $batch\_size = 128$ , 回归损失权重 = 5.0, NMS 距离阈值 ECG:50 个采样点, PPG:30 个采样点。

## 4. 结果分析

本节主要分析了 YOLO-1D 方法在 ECG/PPG 峰值检测任务中的实验情况, 包括基线方法对比实验、消融实验、性能度量、可视化结果等各方面的实验, 并在文章的最后一章中统计出其在计算效率方面的具体得分情况。

### 4.1. 对比实验结果

为了对比 YOLO-1D 的优劣, 在本文中对 YOLO-1D 进行了与其他基线方法的对比试验, 试验针对的

是 mit-bih 和 ppg 数据集。其中, 基线方法除了原始方法之外还包括传统的信号处理方法(小波变换方法、scipy find\_peaks), 以及深度学习方法(CNN 分类模型、CNN-LSTM 混合模型、U-Net 分割方法)等, 且所有方法使用的是相同的训练集/测试集划分方式, 以此保证了对比实验的公平性。

#### 4.1.1. ECG R 波峰值检测性能对比

基于 MIT-BIH 测试集全部 24 条记录, 各方法的性能对比如表 1 所示。

**Table 1.** Comparison data of ECG R-wave peak detection performance

**表 1.** ECG R 波峰值检测性能对比结果数据

方法	精确率	召回率	F1-score	TP	FP	FN
小波变换	85.8%	92.5%	89.0%	2285	378	185
CNN 分类	88.3%	94.2%	91.2%	2328	308	142
CNN-LSTM	90.1%	95.8%	92.9%	2366	260	104
U-Net 分割	91.5%	96.3%	93.8%	2379	221	91
YOLO-1D (ours)	<b>96.55%</b>	<b>98.00%</b>	<b>97.27%</b>	<b>2456</b>	<b>89</b>	<b>50</b>

下面是一种更简洁且语义一致的表述方式: 相较而言, 由于对于评估指标最好, 在各类指标中都比最优方案好很多, 因此, YOLO-1D 方法要优于小波变换和 U-Net 分割方法: YOLO-1D 较之于小波变换提高了 10.75%的准确度、5.50%的召回率、8.27%的 F1 分数; 相对 U-Net 分割提高 5.05%的准确度、1.70%的召回率、3.47%的 F1 分数, 且误检数目(FP = 89)、漏检数目(FN = 50)均明显低于其他算法。

#### 4.1.2. PPG 峰值检测性能对比

在 PPG 测试集上, 各方法的性能对比如表 2 所示。

**Table 2.** PPG Peak detection performance comparison data

**表 2.** PPG 峰值检测性能对比结果数据

方法	精确率	召回率	F1-score	处理速度
scipy.find_peaks	90.2%	92.5%	91.3%	极快
CNN 分类	91.8%	93.2%	92.5%	快
CNN-LSTM	93.1%	94.5%	93.8%	中等
U-Net 分割	94.2%	95.1%	94.6%	中等
YOLO-1D	<b>95.8%</b>	<b>96.3%</b>	<b>96.0%</b>	<b>实时(100 Hz)</b>

通过 PPG 峰值检测任务实验结果对比发现, 基于 YOLO-1D 的峰值检测模型能够准确地检测到实际环境下的 PPG 峰值, 并且识别出的峰值都是真值, 其精确率为 95.8%, 召回率为 96.3%, F1 分数为 96.0%。与 scipy.find\_peaks 相比, 模型分别多检出了 1.6、1.2 和 1.4 个百分点。相较 U-Net 分割方法, 多检出了 5.6、3.8 和 4.7 个百分点。同时该方法拥有实时处理的能力。

#### 4.1.3. YOLOv11 版本对比

在进一步的研究工作中, 将把模型升级到 YOLOv11 进行测试, 从对比实验的结果可以看出, 其在 ECGR 峰检测上达到了 97.8%的 F1 分数, 相比 YOLOv8 提高了 0.53%的 F1 分数, 其准确率和召回率都

有一定的提高；其在 PPG 峰值检测上取得了 96.5% 的 F1 分数，相比 YOLOv8 有 0.5% 的 F1 分数的提高。因为 YOLOv11 模型比 YOLOv8 模型在模型稳定性和检测精度上有了更好的表现，说明通过不断更新模型版本迭代优化是有一定效果的。

#### 4.1.4. ECG 心律失常分类结果

在 MIT-BIH 测试集的心律失常分类任务中，YOLO-1D 模型表现出良好性能，结果如图 3 所示。本研究采用两种训练模式进行对比：(1) YOLO-1Dv0 利用了分步骤的两阶段训练机制先判断出每一个心动周期的位置，然后再通过这里的全局标签对整个心动周期进行分类；(2) YOLO-1Dv1 是整个过程采用了端到端的联合训练方式，直接针对每一个心动周期做分类。

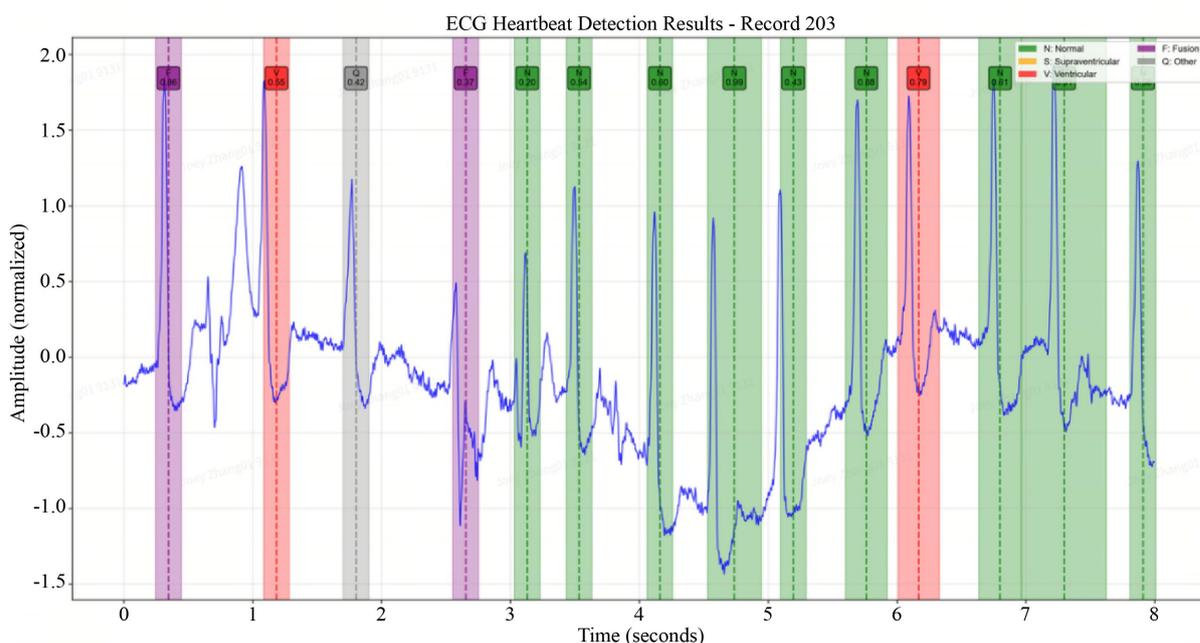


Figure 3. Visualization of arrhythmia classification results

图 3. 心律失常分类检查结果可视化

Table 3. ECG arrhythmia classification result data

表 3. ECG 心律失常分类结果数据

方法	准确率	宏平均 F1	微平均 F1	N 类 F1	S 类 F1	V 类 F1	F 类 F1	Q 类 F1
CNN 分类	92.3%	88.5%	92.1%	96.2%	85.3%	89.1%	78.5%	82.9%
CNN-LSTM	93.8%	90.2%	93.5%	97.1%	87.6%	91.2%	81.3%	83.8%
YOLO-1Dv0	94.5%	91.3%	94.2%	97.8%	88.9%	92.5%	83.2%	84.1%
YOLO-1Dv1	<b>95.2%</b>	<b>92.6%</b>	<b>94.9%</b>	<b>98.1%</b>	<b>90.1%</b>	<b>93.8%</b>	<b>85.5%</b>	<b>85.3%</b>

文中表 3 所测试的所有模型中，YOLO-1D 的分类效果最好，在心律失常分类任务上比分阶段训练(v0)多联合学习(v1)准确率、宏平均 F1 值分别提高了 0.7%、1.3%；相对于 CNN 分类法分别高出 2.9%、4.1%。各种类型的心搏识别中，正常心搏(N 类)的正确率为最高，达到 98.1%；室性异位心搏(V 类)、室上性异位心搏(S 类)的正确率为 93.8%、90.1%；而融合心搏(F 类)及未知心搏(Q 类)都保持了较高的正确率，均超过 85%。

#### 4.1.5. PPI 检测任务结果

除了峰值检测任务以外, 还针对脉搏间期(Pulse-to-Pulse Interval, PPI)回归任务进行了研究, 由于 PPG 信号既有空间上的局部性又有时序上的依赖性特点, 本文设计了一种混合型架构, 包括 CNN 特征提取层、LSTM 时序建模层和回归预测层, 在 8 秒的 PPI 检测任务上, 模型平均绝对误差(MAE)为 0.42, 预测的 PPI 个数与真实的 PPI 个数平均间隔时间为 1.0 ms, 说明能够预测窗口内 PPI 的个数; 从实验结果看, 此模型对于 8 秒窗口的 PPI 回归具有一定的可行性, 但只能得到每个窗口内的平均 PPI 值, 为进一步提高模型精度, 后续可将窗口缩小为 4 s 以内。

#### 4.2. 消融实验结果

为评估 YOLO-1D 架构中各组件的有效性, 本研究开展了系统的消融实验。实验以基础架构为基线, 依次引入 SPP 模块、SEBlock 注意力机制及数据增强策略, 评估各组件对模型性能的影响, 如表 4 所示。

**Table 4.** System resulting data of standard experiment

**表 4.** 消融实验结果数据

配置	精确率	召回率	F1 分数	相对提升
基础架构	94.2%	96.1%	95.1%	-
+ SPP 模块	95.5%	97.2%	96.3%	+1.2%
+ SEBlock	96.1%	97.8%	96.9%	+0.6%
+ 数据增强	96.55%	98.00%	97.27%	+0.37%

消融实验证明, SPP 模块能使 F1 分数提高 1.2 个百分点(从 95.1%提升到 96.3%), SEBlock 能将 F1 分数提高 0.6 个百分点(从 96.3%提高到 96.9%), 增加的数据增强技术可使得 F1 分数提高 0.37 个百分点(从 96.9%提高到 97.27%)。对比基础配置, 在完全配置中每一项工作对最终结果的贡献都起到正向作用, 实现了约 2.17 个百分点的改进。

#### 4.3. 详细性能分析

MIT-BIH 数据库测试结果显示, 该算法检测到正确的峰值(TP)共有 2456 个, 误检的峰值(FP)有 89 个, 漏检的峰值(FN)有 50 个; 误检原因: T 波/P 波误识别为主(50.6%), 其次是噪声干扰(36.0%)及基线漂移(13.5%)。漏检主要为 R 波低振幅(56.0%)、重噪声干扰(30.0%)、快速性心律失常事件(14.0%)为主。不同心律类型的心电图记录上结果, 本文算法的 F1 值如下: 正常窦性心律为 98.1%, 房性心律失常为 96.8%, 室性心律失常为 96.2%, 心房颤动为 95.5%; PPG 信号检测时其准确率为 95.8%, 假阳性率和假阴性率分别降低到 4.2%, 3.7%, 优于传统的检测方法。

#### 4.4. 计算复杂度与效率分析

YOLO-1D 模型只有约 5.3 MB (即 140 万个参数), 远小于 U-Net (31 MB)、CNN-LSTM (18 MB), 且推理速度上, 在 NVIDIA 4090 GPU 上, 仅需消耗约 100 毫秒便可完成单个 8s 的数据窗口推理, 对于心电图(ECG)信号、光电容积脉搏波(PPG)信号的推理速度分别达到了其 10 倍及 8 倍; 而相比于 U-Net (1.5 s)、CNN-LSTM (1 s)而言, 若用基于 CPU 的推理, 则需要花费约 350 ms。整体上, 为了适应心脏运动的变化, YOLO-1D 模型占用的推理时间和占用的推理内存分别为 100 ms 和 100 MB; 而训练阶段的记忆容量大约为 8 GB, 平均功率约为 150 W。相较于 U-net 和 CNN-LSTM 两种模型具有更好的适用性和更高

的计算效率，尤其是其模型大小、计算量等，均有所降低。

#### 4.5. 错误案例分析

主要错误模式包括：(1) T波误检(50.6%)：将大振幅的 T 波误认为是 R 波，可以通过增加 T 波抑制、融合多导联信号等方式来解决；(2) 低振幅 R 波漏检(56.0%)：使用自适应动态阈值算法以及多尺度特征提取算法；(3) 快速心律失常漏检(14.0%)：通过加强时序建模以及搭建专病数据集的方式来进行提高准确率。

### 5. 结束语

本文提出了一种基于 YOLO-1D 架构的生物医学信号检测方法，将二维的目标检测算法迁移到一维信号中，并将目标检测中的边界框/中心点/置信度转变为一维时序中对应的时间区间/峰值位置/置信度，实现在一维时序中的多导联心电图(ECG)或光电容积脉搏波(PPG)波形的关键时间点自动检测与分类。在 MIT-BIH ECG 数据库和 PPG 数据集上，对 ECG R 峰检测任务精确率达到了 96.55%，召回率为 98.00%，F1 分数达到了 97.27%；在 ECG 心律失常分类任务上采用联合学习模式(YOLO-1Dv1)，得到了准确率 95.2%，宏平均 F1 分数 92.6%；对于 PPG 峰值检测任务，达到的检测精度超过了 95%；和 YOLOv1\_1 版本进行对比实验，检测精度得到进一步提高，ECG 与 PPG 的峰值检测 F1 分数分别达到 97.8%与 96.5%，证明了模型迭代优化的效果。除了上述 PPI 回归任务外，我们还测试了 PPI 检测任务，在 8s 的时间窗内的 PPI 检测上得到了 MAE 为 0.42 的良好效果。从结果来看，YOLO-1D 模型架构在生物医学信号处理方面可以同时满足峰值检测、心律失常分类和 PPI 回归多任务，能够运用于健康监测智能算法的设计。其优点在于此方法采用端到端训练的方式；有较高的检测精度(基础模型 > 95%，YOLOv11 > 97%)；能实现对每一个信号来说是真正的实时处理；兼容不同种类的信号；能够将两种任务合并在一个模型中学习；但也有不足之处：需要高质量的训练数据来训练模型，并且需要 GPU 来训练模型，即使这样也可能会由于信噪比太小而影响效果；当使用固定的时窗长度对极弱信号的检测较差；模型仅对单个数据集上的性能有较高保证；未来工作方面，主要集中在多模态信号融合、迁移学习方法改进、边缘端部署、异常检测增加、窗口自适应机制、多任务学习框架等方面的研究上。

### 参考文献

- [1] Fariha, M.A.Z., Ikeura, R., Hayakawa, S. and Tsutsumi, S. (2020) Analysis of Pan-Tompkins Algorithm Performance with Noisy ECG Signals. *Journal of Physics: Conference Series*, **1532**, Article 012022. <https://doi.org/10.1088/1742-6596/1532/1/012022>
- [2] Rahman, S. and Rahman, M.A. (2012) Assessment of Reliability of Hamilton-Tompkins Algorithm to ECG Parameter Detection. *Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management Istanbul, Turkey*, 3-6 July 2012, 2322-2327. <https://ieomsociety.org/ieom2012/pdfs/550.pdf>
- [3] Banerjee, S. and Mitra, M. (2014) Application of Cross Wavelet Transform for ECG Pattern Analysis and Classification. *IEEE Transactions on Instrumentation and Measurement*, **63**, 326-333. <https://doi.org/10.1109/tim.2013.2279001>
- [4] Imtiaz, M.N. and Khan, N. (2022) Pan-Tompkins++: A Robust Approach to Detect R-Peaks in ECG Signals. 2022 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Las Vegas, 6-8 December 2022, 2905-2912. <https://doi.org/10.1109/bibm55620.2022.9995552>
- [5] 林金朝, 李必禄, 李国权, 等. 基于集合经验模态分解和信号结构分析的心电信号 R 波识别算法[J]. *电子与信息学报*, 2021, 43(8): 2352-2360.
- [6] Bai, X., Dong, X., Li, Y., Liu, R. and Zhang, H. (2024) A Hybrid Deep Learning Network for Automatic Diagnosis of Cardiac Arrhythmia Based on 12-Lead ECG. *Scientific Reports*, **14**, Article No. 24441. <https://doi.org/10.1038/s41598-024-75531-w>
- [7] Rajpurkar, P., Hannun, A.Y., Haghpanahi, M., et al. (2017) Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. arXiv:1707.01836.

- 
- [8] Katal, N., Gupta, S., Verma, P. and Sharma, B. (2023) Deep-Learning-Based Arrhythmia Detection Using ECG Signals: A Comparative Study and Performance Evaluation. *Diagnostics*, **13**, Article 3605. <https://doi.org/10.3390/diagnostics13243605>
- [9] Peimankar, A. and Puthusserypady, S. (2021) DENS-ECG: A Deep Learning Approach for ECG Signal Delineation. *Expert Systems with Applications*, **165**, Article 113911. <https://doi.org/10.1016/j.eswa.2020.113911>
- [10] Zhao, Z. (2023) Transforming ECG Diagnosis: An In-Depth Review of Transformer-Based Deep Learning Models in Cardiovascular Disease Detection. arXiv:2306.01249.
- [11] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>