

一种基于深度特征融合的可解释性12导联心电图自动诊断模型研究

卢学麒¹, 陈华元², 吴秋岑¹, 温耀棋¹, 刘国光³, 陈超敏¹

¹南方医科大学生物医学工程学院, 广东 广州 510515; ²广州南方医疗设备综合检测有限责任公司, 广东 广州 510515; ³广东省医疗器械质量监督检验所, 广东 广州 510663

摘要:目的 提升12导联心电图(ECG)自动诊断的准确性和可信度。方法 提出了一种基于深度特征融合的12导联ECG自动诊断模型(MRHL-ECGNet)。该模型包含多尺度特征提取前端、ResNet-34、全局特征混合模块及时间序列分析模块,首次将Hyena Hierarchy卷积算子应用于12导联心电图自动诊断任务中,以高效捕捉ECG中的长程依赖关系,并显著降低模型计算复杂度。同时采用基于积分梯度(IG)的可解释性分析技术,实现MRHL-ECGNet决策依据可视化。使用CPSC2018数据集对MRHL-ECGNet进行训练和测试,并采用多项定量评价指标与评估实验对MRHL-ECGNet进行全面评估。结果 在测试集上对9种类别ECG的分类任务中,MRHL-ECGNet的准确率、AUC值、F1分数、精确率和召回率分别达到0.972、0.983、0.864、0.873和0.857,均优于其他对比模型,且在GPU上对单样本输出诊断结果所需的时间为0.007s,在CPU上也仅需0.156s,内存占用为67.196MB。结论 本研究提出的MRHL-ECGNet不仅具有卓越的分类性能,还具备轻量化及可解释性的特点,在临床ECG辅助诊断中具有较高的应用价值。

关键词:心电图自动诊断;深度学习;Hyena Hierarchy卷积算子;模型可解释性

Evaluation of an interpretable 12-lead ECG automatic diagnosis model based on deep feature fusion

LU Xueqi¹, CHEN Huayuan², WU Qiucen¹, WEN Yaoqi¹, LIU Guoguang³, CHEN Chaomin¹

¹School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China; ²Guangzhou Nanfang Medical Equipment Comprehensive Testing Co., Ltd., Guangzhou 510515, China; ³Guangdong Medical Devices Quality Surveillance and Test Institute, Guangzhou 510663, China

Abstract: Objective To enhance the accuracy and reliability of 12-lead electrocardiogram (ECG) automatic diagnosis. **Methods** Herein we propose a 12-lead ECG automatic diagnosis model based on deep feature fusion (MRHL-ECGNet), which consists of a multi-scale feature extraction front-end, ResNet-34, a global feature mixing module, and a time-series analysis module. The Hyena Hierarchy Convolution Operator was applied to the 12-lead ECG automatic diagnosis task for more efficient capture of long-range dependencies while reducing computational complexity. Integrated Gradients (IG)-based interpretability analysis technology was used to achieve visualization of the decision-making basis of MRHL-ECGNet. The CPSC2018 dataset was used to train and test MRHL-ECGNet, and its performance was assessed using multiple quantitative evaluation indicators and evaluation experiments. **Results** In the 9-class ECG classification task on the test set, MRHL-ECGNet achieved an accuracy of 0.972, an AUC of 0.983, an F1 score of 0.864, a precision of 0.873, and a recall of 0.857, all surpassing other comparative models. This model only took 0.007 s to output a diagnosis for a single sample on a GPU and 0.156 s on a CPU, with a memory footprint of 67.196 MB. **Conclusion** The proposed MRHL-ECGNet model demonstrates excellent classification performance in 12-lead ECG automatic diagnosis with a lightweight design and good interpretability, and thus has great potential for clinical application in ECG-aided diagnosis.

Keywords: electrocardiogram automatic diagnosis; deep learning; Hyena Hierarchy Convolution Operator; interpretability of model

心电图(ECG)作为一种记录心脏电活动的无创检测手段,能够提供关于心脏节律、传导系统以及心肌状态等重要信息,对于多种心血管疾病(CVD)的早期筛查和监测具有不可替代的作用^[1,2]。随着医疗数据量的不

断增加以及人工智能技术的飞速发展,基于深度学习的ECG自动诊断模型已成为当前的研究热点,旨在提高诊断效率、减少人为误差并为临床决策提供有力支持^[3,4]。

近年来,众多研究者聚焦于深度学习技术在ECG自动诊断中的应用,并不断探索新方法以提升模型的准确性。2019年,Hannun等^[5]采用了一种端到端的深度神经网络对单导联ECG进行12种心律失常类别的自动分类,该方法在定量评价指标上已超过心血管医生的平均得分,但该研究仅局限于对单导联ECG进行心律失常分类,然而在实际临床诊断中,复杂的CVD往往需要综合分析12导联ECG才能做出准确判断。Yang等^[6]于

收稿日期:2025-06-24

基金项目:国家重点研发计划(2023YFC2414500,2023YFC2414502)

作者简介:卢学麒,在读硕士研究生,E-mail: 1249416727@qq.com; 陈华元,工程师,E-mail: 415081161@qq.com。卢学麒、陈华元共同为第一作者

通信作者:刘国光,高级工程师,E-mail: 67936668@qq.com; 陈超敏,教授,博士生导师,E-mail: 571611621@qq.com

2021年提出了一种基于级联卷积神经网络和专家特征的12导联ECG心律失常分类方法,其平均F1分数为0.865,但存在部分类别的分类性能较低的问题。Transformer是一种革命性的深度学习架构,此架构的设计灵感源于人类理解上下文的方式,于2017年由Vaswani等^[7]提出,该架构因其优秀的性能表现,已被广泛应用于计算机视觉、时间序列处理、自然语言处理等多个领域中^[8-10]。在心电信号的处理与分析领域,Transformer架构同样表现优异,2023年,Zhang等^[11]通过结合卷积神经网络(CNN)和Transformer架构,以及采用多尺度嵌入层和标记选择块,在12导联ECG分类任务中实现了对不同尺度特征的有效提取和信息冗余的降低,该方法在Chapman数据集^[12]中的准确率达到0.969,然而,由于Transformer架构中自注意力机制的高计算复杂度,使得模型在训练和推理过程中对计算资源的需求极为庞大,难以在计算资源受限的医疗设备中高效部署。上述研究中所提出的方法虽然在ECG自动诊断任务中取得了良好的性能,但研究的重点均集中在提升模型的准确率上,没有关注深度学习模型的可解释性,深度学习模型因其决策过程难以理解常被视为“黑盒”^[13],这在一定程度上限制了其在临床实践中的广泛应用。Zhang等^[14]提出了一种基于1D-CNN的深度神经网络,用于12导联ECG自动分类,并通过SHAP(SHapley Additive exPlanations)方法^[15]对模型进行了可解释性分析,帮助理解决策过程,但SHAP方法计算

复杂度高,对实时诊断支持不足。Reddy等^[16]提出一种名为IMLE-Net的可解释多层次多通道模型对12导联ECG进行自动诊断,通过结合CNN、双向长短期记忆网络和注意力机制,分别在心跳、节律和通道层面提取特征,并通过模型产生的注意力分数来可视化模型在不同层面上的决策依据,但这种方法容易受到模型与数据质量的影响,且缺乏因果关系分析,难以在实际临床应用场景中对深度学习模型提供快速且准确的、可解释性支持,依旧无法从根本上解决ECG自动诊断模型可解释性低的困境。

从现有的研究中可以发现,尽管基于深度学习技术的ECG自动诊断模型已具备强大的性能,但仍存在诸多问题需要解决。针对这些问题,本研究创新性地提出了一种基于深度特征融合的可解释性12导联ECG自动诊断模型,称为“MRHL-ECGNet”,该模型不仅具备强大的12导联ECG分类性能,还具有轻量化及可解释性的特点。

1 材料和方法

本研究的总体技术路线如图1所示。从公开权威的2018中国生理信号挑战赛(CPSC2018)数据集^[17]中获取ECG数据。对实验数据进行预处理,包括统一数据长度、划分数据集和设计数据增强函数。搭建12导联ECG自动诊断模型,并训练该模型;与此同时,构建模型可解释性分析方法。对搭建的12导联ECG自动诊断模型进行全方位的评估。

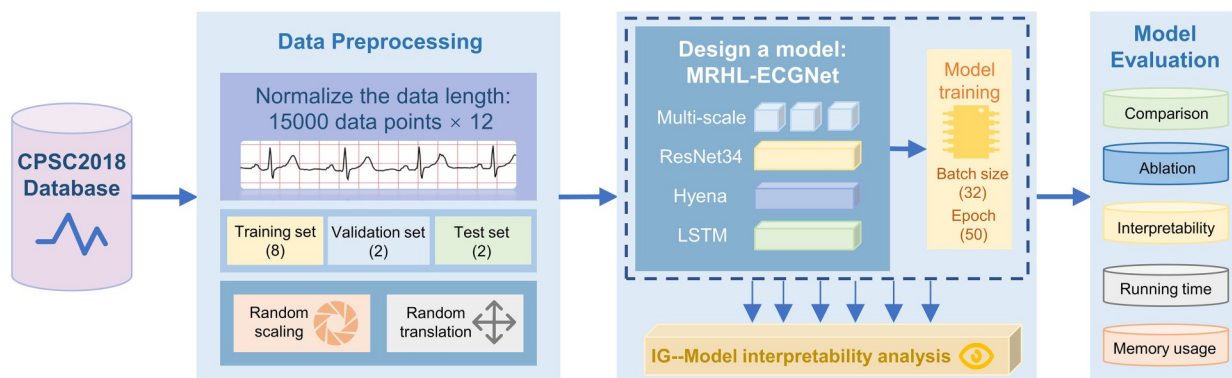


图1 总体技术路线图

Fig.1 Overall technical roadmap of the proposed model.

1.1 实验数据

本研究所使用的数据来自CPSC2018数据集^[17]。该数据集收集了来自11家医院的9831例12导联ECG记录,但该数据集公开可使用的12导联ECG记录为6877例,其中男性3699例,女性3178例。该数据集ECG持续时间为6~60 s,采样频率为500Hz。该数据集涵盖了九种诊断类别的ECG,具体包括:正常(Normal)、

心房颤动(AF)、一度房室传导阻滞(I-AVB)、左束支传导阻滞(LBBB)、右束支传导阻滞(RBBB)、房性早搏(PAC)、室性早搏(PVC)、ST段压低(STD)以及ST段抬高(STE)。

1.2 数据预处理

为了确保所有输入模型的数据具有统一的长度和格式,本研究对于数据集中超过30s的ECG记录,仅保

留其前30 s的数据,而对于信号长度不足30 s的ECG记录则使用零填充至30 s。数据的采样频率为500 Hz,因此最终每例ECG记录的形状都统一为15 000×12。按照8:2的比例分别从9种诊断类别的ECG记录中进行随机抽样,将数据集划分为训练验证集以及测试集。在后续模型训练的过程中使用五折交叉验证,因此依旧按照8:2的比例将训练验证集划分为训练集与验证集。按照上述步骤划分后的训练集样本为4404例,验证集为1101例,测试集为1372例。为了提升模型的鲁棒性,本研究设计了随机缩放和位移函数在后续模型的训练阶段对数据进行增强,训练过程中每条信号被随机缩放和位移的概率都为50%。随机缩放函数使用均值为1.0、标准差为0.05的高斯分布生成缩放因子;位移函数则将位移范围设定为-10到10个采样点之间。

1.3 ECG自动诊断模型的搭建

1.3.1 MRHL-ECGNet的整体结构 本研究所提出的

MRHL-ECGNet是一种用于12导联ECG自动诊断的深度学习模型。该模型以多阶段处理流程为框架,依次包含多尺度特征提取前端、ResNet-34、全局特征混合模块以及时间序列分析模块,最终通过特征融合与分类器输出诊断结果。图2A为MRHL-ECGNet的整体结构示意图。

多尺度特征提取前端负责接收12导联心电图信号,并利用3种不同大小的卷积核(7、15、31)提取信号在不同尺度下的特征,图2B为多尺度特征提取前端的结构示意图。这种多尺度特征提取的设计能够捕捉心电图不同频率和持续时间的信号变化,从而更加全面地捕捉心电信号中的特征信息,有助于提高模型对各种心电图异常模式的检测能力。例如,较小的卷积核可以捕捉到高频的细节特征,而较大的卷积核则能够提取到低频的全局趋势^[18]。这些特征在通道维度上进行拼接后,依次经过批归一化和ReLU激活函数,从而形成一个综合的特征表示。这一过程用公式表示为:

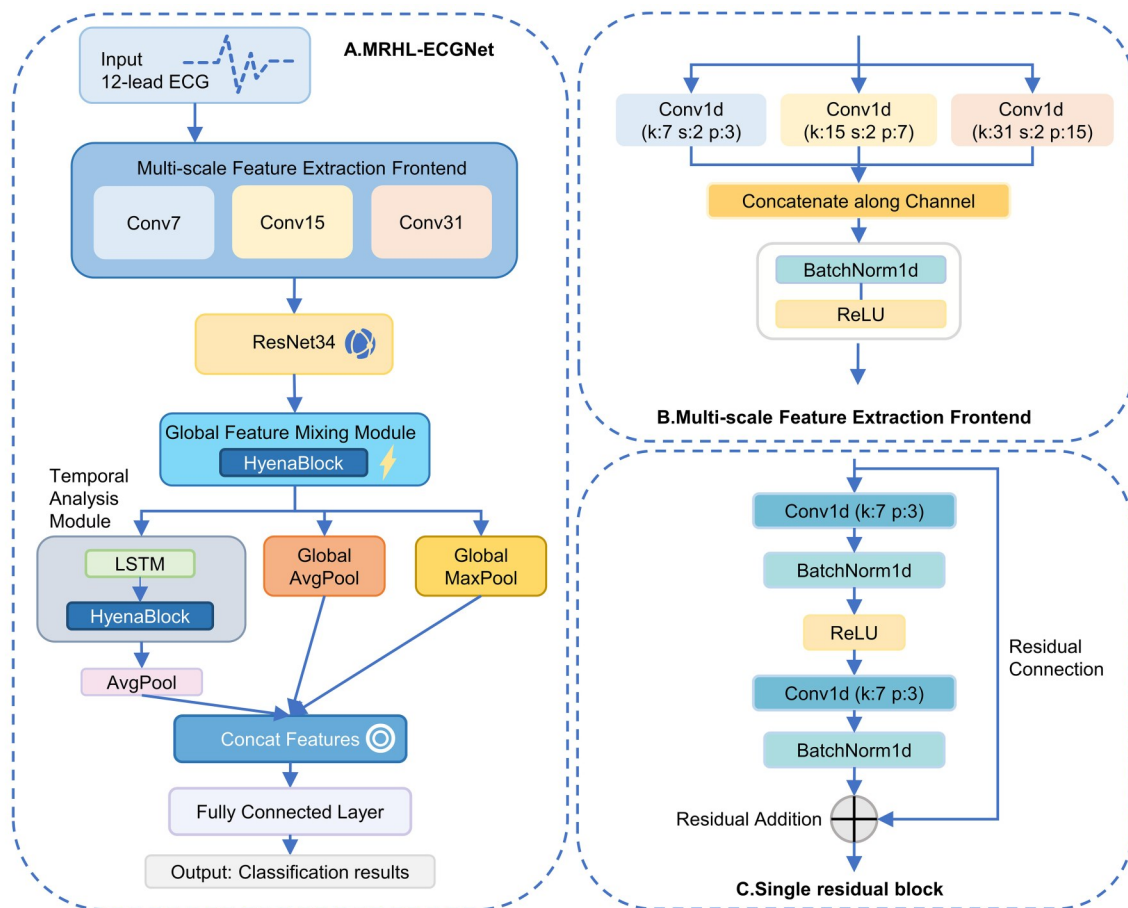


图2 模型结构示意图1

Fig.2 Schematic diagram of the model structure 1. A: Overall structure diagram of MRHL-ECGNet. B: Structure diagram of the multi-scale feature extraction front-end. C: Structure diagram of the single residual block.

$$F_{front} = ReLU$$

$$\left(\text{BatchNorm} \left(\text{Concat} \left(C_7(X), C_{15}(X), C_{31}(X) \right) \right) \right) \quad (1)$$

其中, $C_k(X)$ 表示使用卷积核大小为 k 的卷积层对输入

X 进行操作。

经过前端处理后的特征图 F_{front} 进入 ResNet-34 网络。ResNet-34 最初由 He 等^[19]于2015年提出,其核心

在于引入了残差连接机制,该机制通过将输入直接加到输出上,有效解决了深层网络的梯度消失问题。ResNet-34由4个层次的残差块组成,每个残差块包含若干卷积层和一个跳跃连接。ResNet-34的加入使得模型能够高效地提取心电图信号中的深层特征,同时避免了梯度消失问题,有助于提升模型的性能和训练效率,图2C为单个残差块的结构示意图。ResNet-34的第*l*个残差块的输出可以表示为:

$$F_{res}^l = F_{res}^{l-1} + \text{Block}^l(F_{res}^{l-1}) \quad (2)$$

其中,Block^{*l*}表示第*l*个残差块的操作。

从ResNet-34输出的特征图依次经过全局特征混合模块和时间序列分析模块。全局特征混合模块采用Hyena Hierarchy卷积算子对特征进行全局混合,其核心在于通过卷积操作和门控机制实现特征的自适应混合和增强特征之间的长程依赖关系。时间序列分析模块则结合长短期记忆网络(LSTM)和Hyena Hierarchy卷积算子,捕捉心电图信号的时间维度特征。

最后,来自全局特征混合模块输出的特征 F_{global} 经过全局平均池化和全局最大池化,时间序列分析模块输出的特征 F_{temp} 经过平均池化,这三者拼接形成融合特征

F_{fuse} 。这一过程可以表示为:

$$F_{fuse} = \text{Concat} \left(\begin{matrix} \text{GlobalAvgPool}(F_{global}), \\ \text{GlobalMaxPool}(F_{global}), \text{AvgPool}(F_{temp}) \end{matrix} \right) \quad (3)$$

其中,Concat表示特征拼接操作。

该融合特征 F_{fuse} 通过全连接层进行进一步处理,最终由分类器输出诊断结果。

在MRHL-ECGNet中各模块紧密协作,共同构建了一个强大的ECG自动诊断模型。多尺度特征提取前端能够捕捉ECG中不同频率和持续时间的信号变化,提供了丰富的特征基础。ResNet-34网络能够避免梯度消失,提取出更深层次的特征信息。全局特征混合模块和时间序列分析模块则通过Hyena Hierarchy卷积算子和LSTM网络分别从全局和时间维度增强特征之间的长程依赖关系。并且,全局特征混合模块的输出 F_{global} 经过全局平均池化和全局最大池化,这种双重池化策略是为了从不同角度提取特征信息,前者通过计算特征图的全局平均值,保留了特征的整体统计信息;后者则通过保留特征图中的最大值,突出了显著特征(表1)。

表1 MRHL-ECGNet参数设置

Tab.1 Parameter settings of the MRHL-ECGNet model

Module	In/Out & K and Other Parameters
Input Layer	In: 12×T
Multi-scale feature extraction frontend	In: 12×T; Out: 64×T/2; K: 7, 15, 31; Stride: 2
ResNet-34 Layer1	In: 64; Out: 64; K: 7; BasicBlock1d: 3; Downsample: No
ResNet-34 Layer2	In: 64; Out: 128; K: 7; BasicBlock1d: 4; Downsample: Yes
ResNet-34 Layer3	In: 128; Out: 256; K: 7; BasicBlock1d: 6; Downsample: Yes
ResNet-34 Layer4	In: 256; Out: 512; K: 7; BasicBlock1d: 3; Downsample: Yes
Global feature mixing module	In: 512; Out: 512; K: 31
Temporal analysis module (LSTM)	In: 512; Hidden state: 64
Temporal analysis module (Hyena)	In: 64; Out: 64; K: 15
Feature fusion and classifier	Concatenated in: 512×2+64; Fully connected layer; Output: Classification results

In: Input channels; Out: Output channels; K: Kernel size; T: ECG length.

1.3.2 全局特征混合模块 全局特征混合模块是MRHL-ECGNet中的关键组件,该模块通过引入Hyena Hierarchy卷积算子增强特征间的长程依赖关系。这是首次将Hyena Hierarchy卷积算子应用于12导联心电图自动诊断任务,Hyena Hierarchy卷积算子最初由Poli等^[20]于2023年提出,是一种高效的特征混合机制,在处

理长序列数据时具有显著的优势。Hyena Hierarchy卷积算子的核心在于其交织隐式参数化的长卷积(IIP-LC)和数据控制门控机制(DCGM)。交织隐式参数化的长卷积通过动态生成卷积核参数,避免了直接存储和优化超长卷积核带来的参数爆炸问题。数据控制门控机制则通过逐元素乘法对特征通道进行加权,使模型能

够自适应地增强关键特征并抑制不重要的特征。此外, Hyena Hierarchy 卷积算子利用快速傅里叶变换(FFT)和递归计算将卷积计算复杂度大幅降低,显著提高了计算效率。

与现有多数同类研究中所使用核心架构的 Transformer 相比, Hyena Hierarchy 卷积算子在多个方面展现出独特优势。首先, 对于长度为 S 的序列, Transformer 的计算复杂度为 $O(S^2)$, 而 Hyena Hierarchy 卷积算子则为 $O(\text{Slog}_2 S)$, 这种效率的提升在处理长序列时尤为明显; 例如, 当序列长度达到 8K 时, 可提速两倍^[20]。其次, Hyena Hierarchy 卷积算子在标准数据集(如 WikiText103 和 The Pile)上的语言建模任务中, 达到了与 Transformer 相当的性能, 并同时减少了约 20% 的训练计算量^[20]。最后, Hyena Hierarchy 卷积算子能够支持更长的序列长度, 能够轻松处理长达 131K 长度的序列^[20]。

全局特征混合模块接收 ResNet-34 输出的特征图后首先使用一个卷积核大小为 31 的卷积层对特征进行

处理, 这一卷积层实现了交织隐式参数化的长卷积, 通过动态生成卷积核参数, 模型能够识别到心电图中的全局模式和趋势。心电信号通常具有较长的时间跨度^[21], 交织隐式参数化的长卷积不仅解决了传统方法难以有效捕捉全局特征的问题, 还显著减少了参数量和计算成本。接下来, 通过数据控制门控机制对特征通道进行加权, 实现特征的自适应增强, 能够使模型更加关注对诊断有重要意义的特征部分。之后, 特征又依次通过投影层、丢弃层和批量归一化层。上述过程即为 Hyena Hierarchy 卷积算子, 可以通过以下核心公式表示:

$$z'_{n+1} = x'_n \cdot (h_n * z_n)^t \quad (4)$$

其中, n 表示特征处理的步骤; x'_t 表示输入特征在时间步 t 处的值; h_n 表示在第 n 个中通过隐式参数化生成的卷积核; $*$ 为卷积操作符; z_n 表示第 n 步的特征状态。

最后, 模块通过残差连接将输入特征与处理后的特征相加, 形成最终的输出特征。图 3A 为全局特征混合模块的结构示意图。

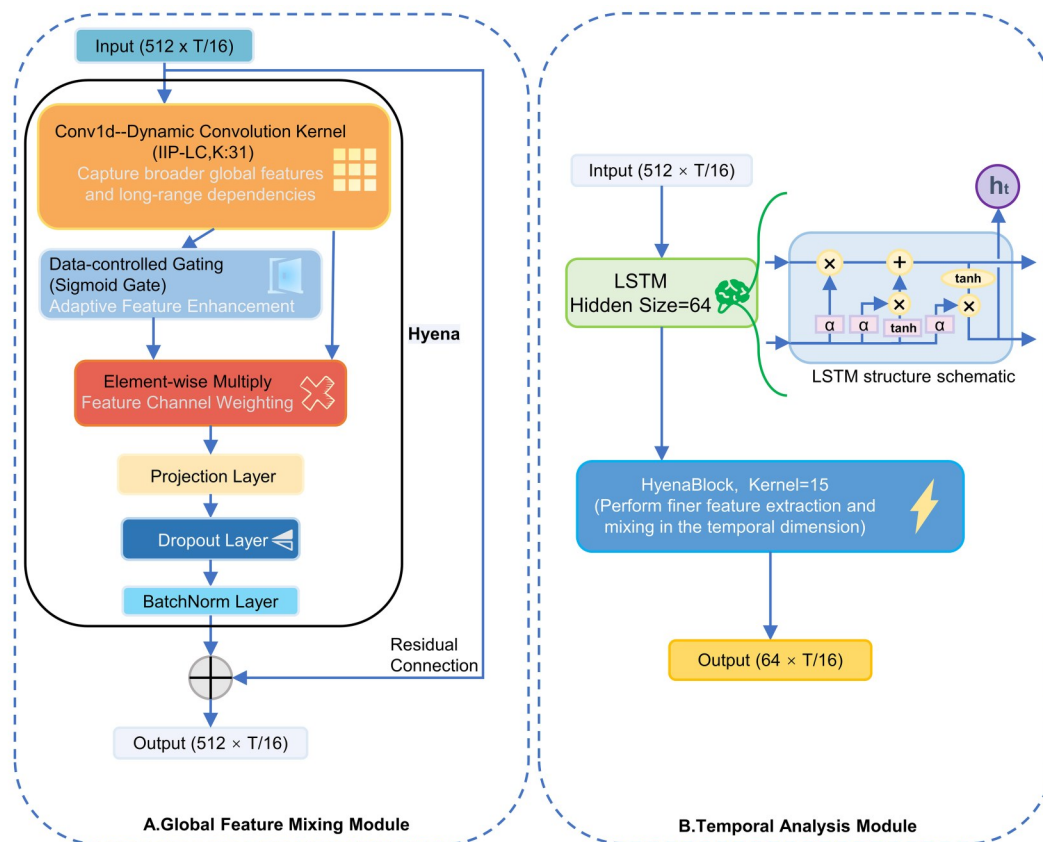


图3 模型结构示意图2

Fig. 3 Schematic diagram of model structure 2. A: Global feature merging module structure diagram; B: Temporal analysis module structure diagram.

1.3.3 时间序列分析模块 在 MRHL-ECGNet 中, 时间序列分析模块通过结合 LSTM 网络和 Hyena Hierarchy 卷积算子, 能够有效地增强模型对心电信号时间维度特

征的提取能力。LSTM 最初由 Hochreiter 等^[22] 于 1997 年提出, 是一种专为处理序列数据设计的循环神经网络 (RNN) 的变体, 其核心是引入了门控机制, 包括输入门、

遗忘门和输出门,这些门控结构能够动态调节信息的流动,使得LSTM与经典的RNN相比能够有效地捕捉序列中的长短期依赖关系,并缓解梯度消失或梯度爆炸的现象。LSTM的公式如下:

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \quad (5)$$

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \quad (6)$$

$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \quad (7)$$

$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \quad (8)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (9)$$

$$h_t = o_t \odot \tanh(c_t) \quad (10)$$

其中, i_t, f_t, g_t 分别表示输入门、遗忘门和输出门的激活值, g_t 是候选记忆单元的激活值, c_t 是记忆单元的状态, h_t 是隐藏状态。上述公式共同描述了LSTM是如何通过门控机制动态调节信息的流动。

时间序列分析模块首先使用LSTM对全局特征混合模块输出的特征序列进行处理,其隐藏状态大小为64,以捕捉心电图信号中的时间依赖关系。随后,使用IIP-LC和DCGM(即Hyena Hierarchy卷积算子)对LSTM的输出特征进行处理,进一步增强时间维度的特征提取能力(图3B)。

1.4 模型训练

在MRHL-ECGNet的训练过程中,选用Adam优化器^[23],初始学习率设定为0.0001,并通过StepLR学习率调度器来动态调整学习率,在每10个训练周期后减少到原来的10%。使用五折交叉验证,采用的批次大小为32,训练周期数为50。损失函数选用二元交叉熵损失函数,该函数结合了Sigmoid激活与二元交叉熵,其公式如下:

$$L = -\frac{1}{N} \sum_{i=1}^N \left[y_i \cdot \log(\sigma(\hat{y}_i)) + (1 - y_i) \cdot \log(1 - \sigma(\hat{y}_i)) \right] \quad (11)$$

其中, y_i 表示真实标签, \hat{y}_i 表示模型预测的概率值, σ 表示Sigmoid函数, N 表示样本数量。

研究所使用的软件开发环境和硬件配置在表2中详细列出。

1.5 深度神经网络可解释性技术

深度学习模型的透明性和可解释性在疾病自动诊断领域是不可或缺的,因为医生需要依据模型的决策依据来评估诊断的合理性和可靠性^[24],为了满足这一需求,本研究采用了基于积分梯度(IG)方法的可解释性技术,对MRHL-ECGNet模型的预测疾病时的决策依据进行解释。

IG方法是一种基于梯度的归因方法,最初在2016年由Sundararajan等^[25]提出,其核心思想是通过积分路径上的梯度来计算每个输入特征对模型输出的贡献,通

表2 实验环境

Tab.2 Experimental environment for model testing

Item	Configuration
Operation system	Windows 11
CPU	Intel Core i5-13490F
GPU	NVIDIA GeForce RTX 4060Ti
RAM	16GB
Cuda version	12.7
Pytorch version	2.3.1
Anaconda version	24.9.2

过在输入空间中沿着从基线到实际输入的路径积分梯度,来估算每个特征的重要性。该方法能够捕捉模型在非线性变化中的特征贡献,为模型的决策提供更准确的解释。其核心公式如下:

$$G(x) = \int_{\alpha=0}^1 \nabla_x f(\alpha \cdot x) d\alpha \quad (12)$$

其中, x 表示输入特征, α 是一个介于0和1之间的参数,用于定义从基线到输入的路径, f 表示模型的输出函数, $G(x)$ 表示输入特征的重要性得分。

在ECG自动诊断模型的可解释性研究方面,目前还未有研究者使用IG方法来对模型的决策过程进行解释,相较于同类型研究中所使用的模型可解释性分析方法,如Zhang等^[14]所使用的SHAP方法,IG方法具有显著优势,这是因为SHAP方法是基于合作博弈论中的Shapley值,通过枚举所有可能的特征子集来计算特征贡献,这种方法在输入的数据较长、特征较多时计算成本极高,并且在处理连续数据时存在近似问题。而IG方法则能够高效地处理长数据以及连续输入的数据,在计算复杂度上远低于SHAP方法,相比于SHAP方法较长的计算时间,IG能够快速地对决策依据进行解释,并输出归因热图,这对于ECG自动诊断模型在实际临床中的应用尤为重要。

在本研究中,使用IG的具体过程如下:首先准备了背景数据用于计算基线,以及待解释数据用于生成归因值,并获得每个心电图导联的归因值,这些归因值反映了模型在做出诊断决策时对各个导联以及导联中不同波段的关注程度。其次,为了可视化归因结果,本研究使用绘图函数绘制了ECG样本的归因热图,突出显示模型关注的关键区域。

通过上述方法,我们能够直观地理解模型在诊断过程中的决策依据,从而提高模型的透明度和可信度。

1.6 模型评价方案

1.6.1 评价指标 本研究采用了5项评价指标^[26],包括精确率(Precision)、召回率(Recall)、F1分数(F1 Score)、曲线下面积(Area Under the Curve)和准确率(Accuracy),

其中,AUC值是通过计算接收者操作特征曲线下的面积得到的,用于评估模型在不同分类阈值下的整体分类性能,AUC值越接近1,表示模型的分性能越好。其余评价指标的计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (15)$$

$$Accuracy = \frac{TP + FN}{TP + TN + FP + FN} \quad (16)$$

其中,TP、FP、TN、FN分别表示真阳性样本数、假阳性样本数、真阴性样本数以及假阴性样本数。这些评价指标的值越接近1,表示模型性能越好。

1.6.2 模型评价实验 为了全面地评估MRHL-ECGNet在ECG自动诊断任务中的性能表现,本研究进行了以下实验:将近年来同类研究中的高性能ECG自动诊断模型与MRHL-ECGNet在相同的实验条件下进行对比;MRHL-ECGNet模型消融实验;模型决策依据可视化;模型运行时间及内存占用计算。

2 结果

2.1 对比实验结果

本研究将MRHL-ECGNet与近年来同类型研究中所提出的高性能且开源的12导联ECG自动诊断模型进行对比,并且所有对比模型均采用与MRHL-ECGNet完全相同的训练及测试方案,以确保结果的可比性(表3)。

实验结果表明,MRHL-ECGNet在5项评价指标上均优于对比模型。MRHL-ECGNet在9种不同的诊断类别中均表现优秀(表4)。

2.2 消融实验结果

本研究进行一系列消融实验,依次移除了MRHL-ECGNet中的多尺度特征提取前端、ResNet-34、Hyena Hierarchy卷积算子、时间序列分析模块、全局平均池化层及全局最大池化层。移除Hyena Hierarchy卷积算子

表3 12导联ECG自动诊断模型性能比较结果

Tab.3 Comparison results of the performance of the 12-lead ECG automatic diagnosis model

Method	Accuracy	AUC	F1s	Precision	Recall
Zhang D et al. ^[14]	0.963	0.961	0.823	0.826	0.821
Reddy L et al. ^[16] (IMLE-Net)	0.895	0.913	0.810	0.812	0.809
Hwang S et al. ^[27] (ResU2D-LC)	0.905	0.928	0.797	0.808	0.787
Strodthoff N et al. ^[28] (inception1d)	0.901	0.941	0.813	0.836	0.792
MRHL-ECGNet	0.972	0.983	0.864	0.873	0.857

表4 MRHL-ECGNet在9种不同诊断类别ECG上的性能表现

Tab.4 Performance of MRHL-ECGNet model on 9 different diagnostic categories of ECG

Diagnosis category	Accuracy	AUC	F1s	Precision	Recall
Normal	0.963	0.980	0.819	0.837	0.802
AF	0.977	0.995	0.927	0.944	0.911
I-AVB	0.975	0.975	0.896	0.912	0.882
LBBB	0.994	0.998	0.920	0.986	0.862
RBBB	0.961	0.989	0.925	0.954	0.898
PAC	0.966	0.952	0.739	0.683	0.804
PVC	0.965	0.982	0.836	0.803	0.872
STD	0.951	0.976	0.793	0.844	0.747
STE	0.996	0.998	0.914	0.889	0.941
AVG	0.972	0.983	0.864	0.873	0.857

时,MRHL-ECGNet中的全局特征混合模块也随之移除。时间序列分析模块被移除时,该模块中的LSTM网络也将一并被移除。分别移除全局平均池化层和全局最大池化层的目的是为了验证双重池化策略的有效性。消融实验结果如表5中的实验1-6所示。

通过表5中实验1-6与实验8的评价指标可以看出,完整的MRHL-ECGNet在各项评价指标上均最为出色,这表明了MRHL-ECGNet中各模块对于模型性能的提升均有贡献。

此外,除上述消融实验外,本研究还将MRHL-ECGNet中的Hyena Hierarchy卷积算子替换为在人工

表5 消融实验及机制替换实验结果

Tab.5 Results of the ablation experiment and the mechanism replacement experiment

Experiment ID	Configuration	Accuracy	AUC	F1s	Precision	Recall
1	Remove multi-scale feature extraction frontend	0.951	0.968	0.846	0.844	0.851
2	Remove resnet-34	0.898	0.901	0.803	0.796	0.812
3	Remove temporal analysis module	0.912	0.919	0.820	0.821	0.819
4	Remove global avgpool	0.956	0.971	0.837	0.835	0.840
5	Remove global maxpool	0.967	0.973	0.833	0.829	0.838
6	Remove all hyenablock	0.905	0.909	0.799	0.798	0.802
7	Replace hyenablock with transformerblock	0.970	0.985	0.847	0.836	0.859
8	Full MRHL-ECGNet	0.972	0.983	0.864	0.873	0.857

智能领域被频繁使用的 Transformer 机制,以此来观察两种机制的性能和资源消耗差异。本研究将原模型中的“HyenaBlock1d”替换为基于自注意力且参数可对其的“TransformerBlock1d”,TransformerBlock1d 首先对输入的通道维度进行归一化处理,然后以“归一化后的多头自注意力+前馈网络”作为核心结构,其多头自注意力默认采用4个注意力头,每个头独立计算注意力权重,以捕捉不同子空间的全局依赖,自注意力输出后,先通过残差连接将注意力结果与归一化后的输入相加,再进入由两层全连接构成的前馈网络。机制替换的实验结果如表5中实验7所示,Hyena Hierarchy 卷积算子在性能表现上已和具有高计算复杂度的Transformer 机制持平,甚至在某些指标上微弱领先。需要注意的是,本

小节仅对比了两种机制的性能表现,他们的计算资源消耗差异将在2.4 部分进行对比。

2.3 模型决策依据可视化

本研究从8种疾病类型的ECG记录中各随机抽取了1例,通过运用基于IG方法的可解释性技术,展示了MRHL-ECGNet在对随机抽取的8例ECG记录进行诊断时所依据的关键特征(图4)。对于每例ECG记录,这里仅展示MRHL-ECGNet对于该记录进行诊断时最为关注的两个导联,每个导联显示其中2000个数据点,并将模型关注度较高的信号片段用深蓝色进行标注。MRHL-ECGNet对于ECG中的关键特征具有精准的捕捉能力。

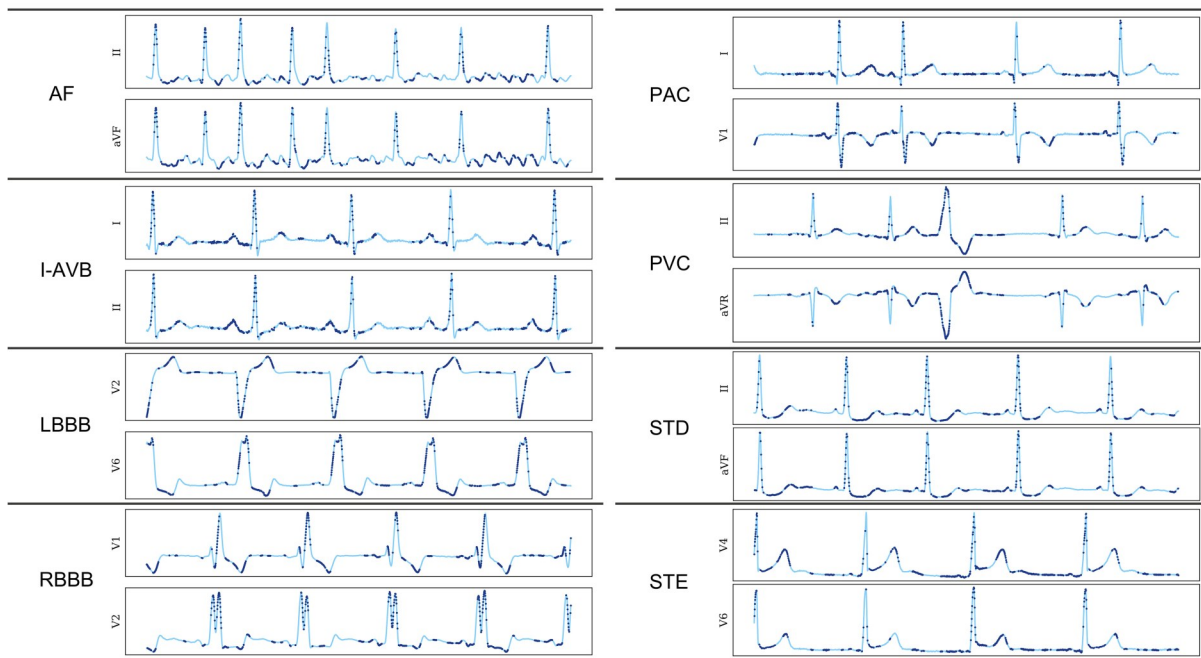


图4 使用IG方法展示MRHL-ECGNet的决策依据
Fig.4 Decision basis of MRHL-ECGNet displayed using the IG method

2.4 运行时间及内存占用

为了评估MRHL-ECGNet的运行效率和资源消耗情况,本研究分别在GPU和CPU环境下,计算了MRHL-ECGNet对1000例ECG记录进行自动诊断的平均时间,及运用基于IG方法的可解释性技术生成决策依据图的平均时间,并计算了MRHL-ECGNet及其主要模块的参数量和内存占用情况。最后,将MRHL-ECGNet中的Hyena Hierarchy卷积算子替换为Transformer机制后的再次重复上述实验,以此来对比两种机制对模型运行效率的影响以及内存占用情况(图5、表6)。

在运行效率方面,无论是在高性能的GPU环境还

是在计算资源受限的CPU环境下,MRHL-ECGNet的自动诊断耗时均能保持在1s以内,且基于IG方法的可解释性分析耗时也较短,能够充分满足临床应用中的实时性需求(图5)。相比之下,将Hyena Hierarchy卷积算子替换为Transformer机制后,模型的自动诊断时间和可解释性分析时间均显著增加,尤其在CPU环境下,生成决策依据图的耗时长达215.098s。

在内存占用方面,MRHL-ECGNet的优势同样显著。MRHL-ECGNet的内存占用仅为67.196MB,能够满足模型实际部署时的轻量化需求(表6)。然而,当使用Transformer机制替换Hyena Hierarchy卷积算子后,模型的内存占用上升约20%。

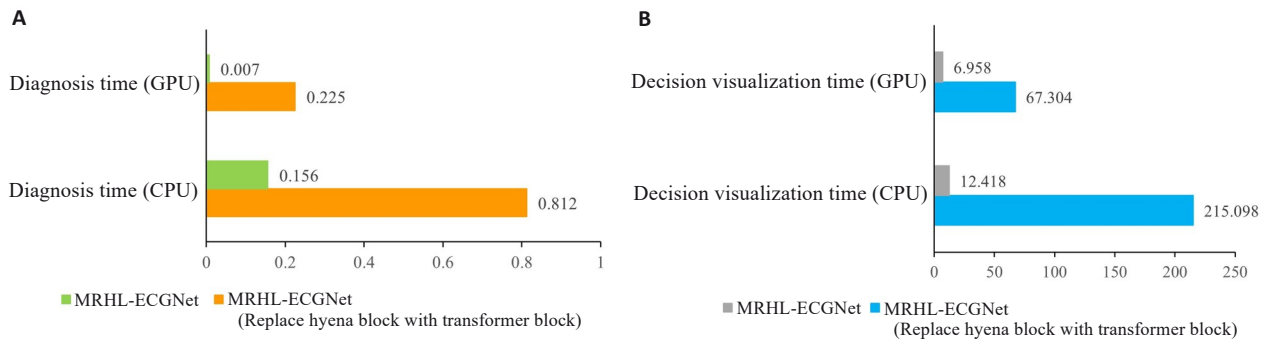


图5 模型运行时间

Fig.5 Model running time. A: Time required for diagnostic results output under different conditions. B: Time for generating the decision-making basis diagram under different conditions.

表6 模型参数量及内存占用

Tab.6 Model parameter quantity and memory usage

Module	Parameters	Memory (MB)
Multi-scale feature extraction frontend	13920	0.053
ResNet-34	16589312	63.283
Global feature mixing module	295424	1.127
Temporal analysis module	154176	0.588
Hyenablock	8256	0.031
Mrhl-ecgnet	17615017	67.196
Mrhl-ecgnet (replace hyenablock with transformerblock)	20746921	80.226

3 讨论

本研究提出的MRHL-ECGNet在12导联ECG自动诊断任务中展现出了卓越的性能。从对比实验结果来看,MRHL-ECGNet在5项评价指标上均优于其他对比模型^[14, 16, 27, 28]。这主要归功于其创新性的模型架构设计,多层次特征融合策略使得模型能够从不同层次对ECG特征进行全面提取与分析。通过进一步分析MRHL-ECGNet在9种不同诊断类别ECG上的性能表现,可以发现对于不同的诊断类别,模型的表现较为均衡,但在PAC这一诊断类别中,MRHL-ECGNet的表现相对较差,这可能是由于各别PAC记录在某些特征与其他正常或异常心律的特征存在一定的相似性,例如,PAC的P波形态与正常的窦性P波不同,但当异位心房起搏点靠近窦房结时,P波形态可能与窦性P波相似^[29],这增加了模型区分的难度。

在消融实验中,当移除ResNet-34时,MRHL-ECGNet的准确率下降得尤为严重,这是由于ResNet-34通过其残差连接机制有效解决了深层网络的梯度消失问题^[19],使得MRHL-ECGNet能够高效地提取ECG中的深层特征信息。与此同时,当Hyena Hierarchy卷积算子被移除时,MRHL-ECGNet的准确率也出现了明显的下降,而将其替换为Transformer机制后,性能虽与原模型接近但在某些指标上仍略逊一筹,这两种结果均

证明了Hyena Hierarchy卷积算子通过其交织隐式参数化的长卷积和数据控制门控机制能够高效地捕捉ECG中的全局模式和趋势,在ECG自动诊断任务中具有独特的优势,在性能上已经可以媲美其他高性能12导联ECG自动诊断模型^[11]所依赖的Transformer机制。

从基于IG方法的可解释性技术生成的模型决策依据可视化图中,可以发现MRHL-ECGNet能够精准地捕捉到ECG中对于疾病诊断最为重要的特征,如AF在ECG上的特点包括不规则的RR间期、P波消失、呈锯齿状的小f波等^[30],对于RBBB,其ECG特点包括V1/V2导联中的rSR'波形(“M”型或兔耳样)、QRS波群宽大等^[31],说明了模型的决策依据与临床诊断思维相吻合。并且这种可解释性方法使得使用者能够快速查看MRHL-ECGNet的决策依据,大幅提高了模型的透明度和可信度。

此外,MRHL-ECGNet在具备高性能的同时还满足了实时性和轻量化的需求。从实验结果中可以发现,无论在GPU还是CPU环境下,MRHL-ECGNet进行自动诊断的时间及基于IG方法的可解释性分析时间均较短,能够充分满足临床应用中的实时性需求,这主要得益于Hyena Hierarchy卷积算子对计算复杂度的优化,其利用快速傅里叶变换和递归计算大幅降低了卷积计算复杂度,使得模型在处理长序列数据时仍能保持高效

性。相比之下,将 Hyena Hierarchy 卷积算子替换为 Transformer 机制后,模型的运行时间显著增加,尤其在计算资源受限的 CPU 环境下,替换为 Transformer 机制后,模型的自动诊断时间增加约 420%,可解释性分析时间更是增加至原先的 17 倍,无法满足实时性的需求。在内存占用方面,MRHL-ECGNet 的内存占用仅 67.196 MB,而使用 Transformer 机制替换 Hyena Hierarchy 卷积算子后,模型的内存占用上升约 20%。这些情况的出现是由于在 Transformer 架构中注意力的计算和存储复杂度呈二次级放大,易在超长序列或大批量输入下触发显存瓶颈;而 Hyena Hierarchy 卷积算子则通过隐式卷积与门控设计,将时间复杂度降至亚二次级,从而在保证分类性能的同时,实现更优秀的内存和时间效率表现。

综上所述,本研究针对现有 ECG 自动诊断模型所存在的问题,提出了一种全新的 12 导联 ECG 自动诊断模型—MRHL-ECGNet。主要创新点总结如下:(1)多层次特征融合:MRHL-ECGNet 以多阶段处理流程为框架,依次包含多尺度特征提取前端、ResNet-34、全新构建的全局特征混合模块以及时间序列分析模块,并加入双重池化策略,这种架构设计使得模型能够从不同层次对 ECG 信号的特征进行提取和分析,更加全面地挖掘 ECG 中的关键信息,提升模型对各种 ECG 异常模式的检测能力。(2)首次将 Hyena Hierarchy 卷积算子应用于 12 导联心电图自动诊断任务中:引入 Hyena Hierarchy 卷积算子不仅大幅增强了模型对于长程依赖的捕捉能力,还显著降低了模型在处理长序列数据时的计算复杂度。替代了目前高性能 12 导联 ECG 自动诊断模型^[1]所依赖的 Transformer 架构,使得 MRHL-ECGNet 在保持高性能的同时,还具备极快的运行速度和极低的内存占用。(3)采用准确且高效的模型可解释性分析方法:首次将基于 IG 方法的可解释性技术应用于 ECG 自动诊断模型,对模型的决策依据进行可视化。与 SHAP 或注意力分数方法相比,该方法具有计算准确且速度快的优势,有效提高了 MRHL-ECGNet 的透明度和可信度。

这推动了 ECG 自动诊断领域的技术进步,辅助临床医生更加快速、准确地进行 CVD 的诊断;更为构建高效、可靠、可解释的医疗诊断模型提供了新思路与方法,有望在临床辅助诊断和智能医疗发展方面发挥重要价值。

但与此同时,本研究仍存在一些局限性。首先,研究所使用的数据集无法完全涵盖所有类型的心血管疾病和临床场景的多样性,缺乏某些罕见病理类型的样本。其次,当前所使用的可解释性分析方法虽已在解释的准确性和计算速度上表现优异,但该方法主要基于输入特征的归因,对于模型内部更深层次的特征学习逻辑

的解释还不够深入;与此同时,由于临床心电图医师对上千条 12 导联 ECG 进行独立逐像素点特征标注的工作量巨大,本研究尚未将模型决策依据可视化结果与临床心电图医师所标注的关键特征进行一致性对比分析,临床可解释性仍具有进一步验证的空间。此外,本研究尚未在实际临床环境中部署和验证模型,缺乏对模型在现实医疗场景中表现的评估。

未来的研究将集中在以下几个方面:采集来自不同地区、不同人群的 ECG 数据,尤其是罕见 CVD 的 ECG 记录,对模型进行更进一步的训练与评估,以提升模型的泛化能力。探索更先进的深度学习模型可解释性技术,精准解析深层特征学习逻辑,并邀请多名临床心电图医师对研究所使用的 12 导联 ECG 记录进行独立逐像素点特征标注,将模型决策依据可视化结果与医师标注的“金标准”进行一致性分析,力求完全穿透模型内部“黑箱”。尝试将 MRHL-ECGNet 部署于真实的医疗场景,通过与专业医师诊断结果的对比及前瞻性研究,全方位评估模型的实际应用效能与可靠性。

Declaration of interests: The authors declare no competing interests.

参考文献:

- [1] Berkaya SK, Uysal AK, Gunal ES, et al. A survey on ECG analysis [J]. Biomed Sig Proc Control, 2018, 43: 216-35.
- [2] Chou's Electrocardiography in Clinical Practice[M]. Elsevier Science Health Science div, 2008.
- [3] Hong SD, Zhou YX, Shang JY, et al. Opportunities and challenges of deep learning methods for electrocardiogram data: a systematic review [J]. Comput Biol Med, 2020, 122: 103801.
- [4] Priori SG, Blomström-Lundqvist C, Mazzanti A, et al. 2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: The Task Force for the Management of Patients with Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death of the European Society of Cardiology (ESC). Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC) [J]. Eur Heart J, 2015, 36(41): 2793-867.
- [5] Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network [J]. Nat Med, 2019, 25(1): 65-9.
- [6] Yang XZ, Zhang XY, Yang MY, et al. 12-Lead ECG arrhythmia classification using cascaded convolutional neural network and expert feature [J]. J Electrocardiol, 2021, 67: 56-62.
- [7] Vaswani A, Shazeer NM, Parmar N, et al. Attention is all you need [C]//Neural Information Processing Systems., 2017
- [8] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale [J]. arxiv preprint arxiv: 2010. 11929, 2020.
- [9] Zhou HY, Zhang SH, Peng JQ, et al. Informer: beyond efficient transformer for long sequence time-series forecasting [J]. Proc

- AAAI Conf Artif Intell, 2021, 35(12): 11106-15.
- [10] Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining[J]. *Bioinformatics*, 2020, 36(4): 1234-40.
- [11] Zhang SY, Lian C, Xu BR, et al. A token selection-based multi-scale dual-branch CNN-transformer network for 12-lead ECG signal classification[J]. *Knowl Based Syst*, 2023, 280: 111006.
- [12] Zheng JW, Zhang JM, Danioko S, et al. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients[J]. *Sci Data*, 2020, 7(1): 48.
- [13] Petch J, Di S, Nelson W. Opening the black box: the promise and limitations of explainable machine learning in cardiology[J]. *Can J Cardiol*, 2022, 38(2): 204-13.
- [14] Zhang DD, Yang S, Yuan XH, et al. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram[J]. *iScience*, 2021, 24(4): 102373.
- [15] Lundberg SM, Lee SI. A unified approach to interpreting model predictions[C]//*Neural Information Processing Systems*, 2017
- [16] Reddy L, Talwar V, Alle S, et al. IMLE-net: an interpretable multi-level multi-channel model for ECG classification[C]//2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC). October 17-20, 2021. Melbourne, Australia. IEEE, 2021: 1068-1074.
- [17] Liu FF, Liu CY, Zhao LN, et al. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection[J]. *J Med Imaging Hlth Inform*, 2018, 8(7): 1368-73.
- [18] Zhang HZ, Zhao W, Liu S. SE-ECGNet: a multi-scale deep residual network with squeeze-and-excitation module for ECG signal classification[C]//2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). December 16-19, 2020. Seoul, Korea. IEEE, 2020: 2685-2691.
- [19] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-8.
- [20] Poli M, Massaroli S, Nguyen E, et al. Hyena hierarchy: Towards larger convolutional language models[C]//*International Conference on Machine Learning*. PMLR, 2023: 28043-78.
- [21] Yang W, Deo R, Guo WS. Functional feature extraction and validation from twelve-lead electrocardiograms to identify atrial fibrillation[J]. *Commun Med (Lond)*, 2025, 5(1): 32.
- [22] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Comput*, 1997, 9(8): 1735-80.
- [23] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arxiv preprint arxiv:1412.6980, 2014.
- [24] Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI[J]. *Inf Fusion*, 2020, 58: 82-115.
- [25] Sundararajan M, Taly A, Yan Q. Gradients of counterfactuals[J]. arxiv preprint arxiv:1611.02639, 2016.
- [26] Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks[J]. *Inf Process Manag*, 2009, 45(4): 427-37.
- [27] Hwang S, Cha J, Heo J, et al. Multi-label abnormality classification from 12-lead ECG using a 2D residual U-net[C]//*ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. April 14-19, 2024, Seoul, Korea, Republic of. IEEE, 2024: 2265-9.
- [28] Strodthoff N, Wagner P, Schaeffter T, et al. Deep learning for ECG analysis: benchmarks and insights from PTB-XL[J]. *IEEE J Biomed Health Inform*, 2021, 25(5): 1519-28.
- [29] Conen D, Adam M, Roche F, et al. Premature atrial contractions in the general population: frequency and risk factors[J]. *Circulation*, 2012, 126(19): 2302-8.
- [30] Joglar JA, Chung MK, Armbruster AL, et al. 2023 ACC/AHA/ACCP/HRS guideline for the diagnosis and management of atrial fibrillation: a report of the American college of cardiology/American heart association joint committee on clinical practice guidelines[J]. *Circulation*, 2024, 149(1): e1-156.
- [31] Ikeda T. Right bundle branch block: current considerations[J]. *Curr Cardiol Rev*, 2021, 17(1): 24-30.

(编辑:经 媛)