

参赛学生姓名：曹巍瀚

中学：北京师范大学附属实验中学

省份：北京

国家/地区：中国大陆

指导老师姓名：徐帅婧

指导老师单位：北京师范大学附属实验中学

论文题目：MCL-Net：基于多导联心电图
特征工程与跨导联注意力机制的室性早搏
消融靶点定位研究

MCL-Net: 基于多导联心电图特征工程与跨导联注意力机制的室性早搏消融靶点定位研究

曹巍瀚

北京师范大学附属实验中学

摘要

室性早搏 (Premature Ventricular Contractions, PVCs) 是临床上最常见的室性心律失常之一, 导管射频消融术是其主要治疗手段。该手术通过消除异位兴奋灶实现根治, 其成功的关键在于术前能否精确定位该兴奋灶的起源位置。然而, 目前定位主要依赖有创心内电生理标测, 操作时间长、费用较高, 且伴随辐射暴露风险。体表 12 导联心电图 (Electrocardiogram, ECG) 作为术前常规无创检查, 其波形形态与异位兴奋灶的起源位置密切相关, 但现有方法的定位精度和泛化能力在复杂临床场景中仍有待提升, 尚缺乏能可靠集成于临床工作流程的高精度自动化无创定位方法。

本文提出 MCL-Net (Multi-lead ECG with Cross-Lead Attention Network), 一种融合多导联卷积特征提取与跨导联注意力机制的混合深度学习模型, 用于从 12 导联 ECG 波形中实现 PVC 消融靶点的端到端定位。模型同时包含 18 维手工心电图特征分支, 该分支与深度特征融合后通过多任务学习框架输出 7 个解剖区域的分类标签和三维空间坐标。在 273 例临床数据上 (训练/验证/测试 = 191/59/23, 测试集 23 例), 在固定随机种子 42 的数据划分下进行三次独立训练, MCL-Net 达到 91.30% 的分类准确率、88.47% 的加权 F1 分数; 在全部测试样本上归一化欧氏距离误差 (EDE) 为 0.5441, 在严格评估 (仅对分类正确的样本计算) 下 EDE 为 0.3090。上述结果优于 ResNet-1D、Transformer 等深度基线模型及随机森林、XGBoost 等传统机器学习方法。消融实验验证了跨导联注意力模块和手工特征分支的有效性。

本研究尝试利用临床 12 导联 ECG 波形实现 PVC 消融靶点三维坐标定量预测, 定位精度达到可与临床金标准比较的水平, 为术前无创靶点定位提供了新的技术方案。

关键词: 室性早搏; 心电图; 消融靶点定位; 跨导联注意力机制; 多任务学习

MCL-Net: Localization of Ablation Targets for Premature Ventricular Contractions Based on Multi-Lead ECG Feature Engineering and Cross-Lead Attention Mechanism

Weihan Cao

The Experimental High School Attached to Beijing Normal University

Abstract

Premature Ventricular Contractions (PVCs) are among the most common ventricular arrhythmias encountered in clinical practice. Radiofrequency catheter ablation (RFCA) is a primary curative treatment, which eliminates the ectopic focus to achieve rhythm control. The success of this procedure critically depends on the accurate preoperative localization of the ectopic focus. However, current localization methods primarily rely on invasive intracardiac electrophysiological mapping, which is time-consuming, expensive, and associated with radiation exposure. The 12-lead surface electrocardiogram (ECG), a routine non-invasive preoperative examination, provides waveforms whose morphology is closely related to the spatial location of the ectopic focus. Nevertheless, existing approaches still require improvement in localization accuracy and generalization capability in complex clinical scenarios, and a robust, high-precision automated non-invasive localization method that can be seamlessly integrated into clinical workflows remains lacking.

This paper proposes MCL-Net (Multi-lead ECG with Cross-Lead Attention Network), a hybrid deep learning model that integrates multi-lead convolutional feature extraction with a cross-lead attention mechanism for end-to-end localization of PVC ablation targets from 12-lead ECG waveforms. The model incorporates an 18-dimensional hand-crafted ECG feature branch, which is fused with deep features within a multi-task learning framework to simultaneously output classification labels across 7 anatomical regions and 3D spatial coordinates. On a clinical dataset of 273 cases (train/val/test = 191/59/23, test set of 23 cases), with a fixed random seed of 42 and three independent training runs, MCL-Net achieves a classification accuracy of 91.30%, a weighted F1 score of 88.47%. The normalized Euclidean distance error (EDE) is 0.5441 on all test samples, and 0.3090 under strict evaluation (only on correctly classified samples), significantly outperforming deep

baseline models including ResNet-1D and Transformer, as well as traditional machine learning methods such as Random Forest and XGBoost. Ablation studies validate the effectiveness of both the cross-lead attention module and the handcrafted feature branch.

This work presents a hybrid deep learning framework that enables quantitative 3D coordinate prediction of PVC ablation targets directly from clinical 12-lead ECG waveforms. The localization accuracy reaches a clinically applicable level, offering a novel technical solution for non-invasive preoperative target localization.

Keywords: Premature Ventricular Contractions; Electrocardiogram; Ablation Target Localization; Cross-Lead Attention Mechanism; Multi-Task Learning

目录

1	引言	1
1.1	研究背景与临床意义	1
1.2	研究现状与问题分析	1
1.3	本文贡献	2
1.4	论文组织结构	3
2	相关工作	4
2.1	PVC 消融靶点定位	4
2.2	心电图深度学习分析	4
2.3	注意力机制与图神经网络	4
3	问题定义与评价指标	6
3.1	问题定义	6
3.2	不同解剖区域的 ECG 特征与坐标分布	7
3.3	评价指标	11
3.3.1	分类评价指标	11
3.3.2	回归评价指标	12
4	MCL-Net 模型架构	13
4.1	模型总体框架	13
4.2	多导联卷积模块	15
4.3	跨导联注意力模块	15
4.3.1	缩放点积注意力的数学形式	16
4.3.2	多头机制的互补性	16
4.3.3	残差连接与层归一化	17
4.3.4	注意力热图的可解释性分析	17
4.4	手工特征提取与融合模块	18
4.4.1	18 维手工特征的定义	18
4.4.2	特征提取流程	19
4.4.3	特征重要性分析	20
4.4.4	手工特征编码与融合	21
4.5	多任务学习与损失函数	22
4.5.1	多任务学习的动机	22
4.5.2	输出头设计	22
4.5.3	损失函数	23
4.6	本章小结	24

5	数据集与实验设置	25
5.1	数据集	25
5.2	实验设置	26
5.3	评估策略	26
5.4	基线方法	26
5.5	消融实验设计	27
6	实验结果与分析	28
6.1	三次独立训练结果	28
6.2	与基线模型的对比	29
6.3	消融实验	29
6.4	后处理策略的影响	31
6.5	结果讨论	31
7	讨论	33
7.1	临床意义	33
7.2	局限性	33
7.3	未来工作	34
8	结论	35

1 引言

1.1 研究背景与临床意义

室性早搏 (Premature Ventricular Contractions, PVC) 是临床上最常见的室性心律失常之一, 在普通人群中的发生率约为 1%–4%, 在合并结构性心脏病的患者中可高达 80%[19]。频繁的 PVC 不仅导致心悸、胸闷等症状, 严重影响患者生活质量, 还可能诱发心动过速性心肌病, 甚至增加心源性猝死风险 [2]。对于药物治疗无效或无法耐受的症状性 PVC, 导管射频消融术 (Radiofrequency Catheter Ablation, RFCA) 已成为首选的根治性治疗手段 [27]。

射频消融手术通过消除异位兴奋灶实现根治, 手术能否取得成功, 关键在于术前能否精确定位该兴奋灶的起源位置。然而, 目前定位主要依赖有创心内电生理标测 (如 CARTO 系统), 通过将标测导管逐点放置于心内膜面采集电位信号, 重建激动顺序图以确定最早激动点 [12]。这种方法虽然精确, 但存在以下显著局限: (1) 手术时间长, 平均需要 2–4 小时; (2) 费用高昂, 对医疗资源消耗大; (3) 患者需承受 X 射线辐射暴露; (4) 部分起源于心外膜或特殊解剖结构的 PVC 难以通过常规心内膜标测定位。

体表 12 导联心电图 (Electrocardiogram, ECG) 是术前常规无创检查, 其波形形态与异位兴奋灶的心内起源位置存在密切的解剖学对应关系。有经验的电生理医生可以通过分析 ECG 形态 (如 QRS 波群的极性、移行区位置等) 初步判断起源部位 [6]。然而, 这种依赖人工经验的判读方法存在以下问题: (1) 高度依赖操作者的临床经验, 主观性强; (2) 对于起源位置相近的解剖区域 (如右室流出道与左室流出道), 人工判读的准确率有限 [4]; (3) 难以实现三维空间坐标的精确定量预测。

因此, 开发一种能够从临床 12 导联 ECG 数据中自动、精准地预测 PVC 消融靶点的计算方法, 具有重要的临床价值和研究意义。

1.2 研究现状与问题分析

近年来, 机器学习方法在心电图自动分析领域取得了显著进展。在心律失常分类方面, 卷积神经网络 (Convolutional Neural Networks, CNN) 和循环神经网络 (Recurrent Neural Networks, RNN) 已被广泛应用于 12 导联 ECG 的疾病诊断 [21]。在 PVC 起源定位方面, 已有研究尝试使用 ECG 形态特征进行起源区域的分类 [32]。

然而, 现有方法仍存在以下不足:

第一, 现有方法多聚焦于分类任务 (如区分右室流出道与左室流出道), 较少涉及三维空间坐标的定量回归。然而在实际临床场景中, 仅知道异位兴奋灶属于某个解剖区域对指导消融手术仍是不够的。解剖研究表明, 右心室流出道直径约为 19–30 mm, 左心室流出道直径约为 17–28 mm; 而单次射频消融产生的消融灶直径通常仅为 3–8 mm, 即便使用特殊灌注射频导管也不超过 15 mm[3]。医生凭借区域分类信息, 仅能判断病灶在某个范围约 2–3 cm 的区域内, 而无法确定其在其中的精确位置, 难以将消融导管准

确定位到毫米级的靶点。因此，仅凭区域分类信息难以满足临床手术的精准定位需求，还需要获取异位兴奋灶在三维空间中的精确坐标。

第二，现有深度学习方法通常将 12 个导联视为相互独立的信号通道，或简单地将其拼接作为输入。然而，12 导联 ECG 本质上是从不同空间角度记录的心脏电活动投影，导联之间存在内在的解剖学依赖关系 [11]。忽略这种导联间依赖关系，可能导致模型无法充分利用 ECG 的空间信息。

近年来，图卷积神经网络（Graph Convolutional Networks, GCN）在图结构数据的半监督节点分类任务中表现优异 [13]。然而，GCN 类方法难以直接应用于 ECG 波形分析，因为 ECG 信号本质上是时序信号而非图结构数据，因此需要设计专门的跨导联依赖建模机制。

1.3 本文贡献

针对上述问题，本文提出了一种名为 MCL-Net（Multi-lead ECG with Cross-Lead Attention Network）的混合深度学习模型，用于从 12 导联 ECG 波形中实现 PVC 消融靶点的端到端定位。本文的主要贡献如下：

1. **提出了多导联卷积与跨导联注意力结合的混合架构。**模型首先通过共享权重的 1D-CNN 对每个导联独立提取时序特征，然后通过多头自注意力机制显式建模 12 个导联之间的空间依赖关系。这一设计同时捕捉了局部波形特征和全局导联间关联。
2. **引入了 18 维手工心电图特征与深度特征融合的策略。**本文设计了一套包含 R 波幅值、QRS 宽度、电轴偏移、束支阻滞标志等 18 个特征的手工特征集，并通过特征融合分支与深度特征相结合。实验表明，手工特征分支对分类准确率有稳定贡献。
3. **建立了完整的临床 ECG 数据处理与靶点定位框架。**本文构建了从 12 导联 ECG 波形到 7 个解剖区域分类和三维空间坐标回归的端到端流程。在本文调研范围内，尚未见到利用临床 ECG 数据实现 PVC 消融靶点三维坐标定量预测的同类混合深度学习模型。
4. **通过系统的实验验证了方法的有效性。**在 273 例临床数据上（训练/验证/测试 = 191/59/23，测试集 23 例），在固定随机种子 42 的数据划分下进行三次独立训练，MCL-Net 在测试集上平均达到 91.30% 的分类准确率、88.47% 的加权 F1 分数；在全部测试样本上的归一化 EDE 为 0.5441，在严格评估（仅对分类正确的样本计算）下 EDE 为 0.3090，优于 ResNet-1D、Transformer Encoder、CNN+Cross-Lead Attention 等深度基线模型以及随机森林、XGBoost 等传统机器学习方法。

1.4 论文组织结构

本文的其余部分组织如下：第 2 节介绍相关研究工作，包括 PVC 消融靶点定位、ECG 深度学习分析和注意力机制等方向；第 3 节描述本文的问题定义和评价指标；第 4 节详细阐述 MCL-Net 的模型架构和各模块设计；第 5 节介绍数据集与实验设置；第 6 节报告实验结果并进行分析；第 7 节讨论模型的临床意义和局限性；第 8 节总结全文并展望未来工作。

2 相关工作

2.1 PVC 消融靶点定位

PVC 消融靶点定位的研究主要分为两类方法：基于 ECG 形态学规则的定位方法和基于机器学习的定位方法。

基于 ECG 形态学规则的定位方法主要依赖 QRS 波群的极性、移行区位置、切迹等形态特征进行起源部位的推断。Brugada 等人提出了基于胸前导联 QRS 形态的右室流出道与左室流出道鉴别标准 [6]。Betensky 等人提出了 V2 导联移行区指数用于区分右室流出道与左室流出道起源 [4]。然而，这些规则方法存在以下局限：

- (1) 规则覆盖面有限，难以处理所有解剖区域；
- (2) 当规则之间存在冲突时，难以做出决策；
- (3) 无法提供三维空间坐标的精确定量预测。

基于机器学习的定位方法近年来逐渐受到关注。Zhang 等人使用支持向量机 (SVM) 和随机森林对 PVC 起源区域进行分类 [32]。Yang 等人使用卷积神经网络从 ECG 信号中识别流出道起源 [31]。然而，这些方法仍局限于分类任务，未涉及三维坐标的定量回归，且模型复杂度有限，深层特征提取能力有待提升。

2.2 心电图深度学习分析

深度学习在心电图分析中的应用近年来取得了显著进展。Rajpurkar 等人提出了基于 34 层卷积神经网络的 ECG 心律失常分类模型，在 12 导联 ECG 上达到了心脏病专家水平的诊断准确率 [21]。Hannun 等人使用深度神经网络从单导联 ECG 中检测心律失常 [9]。Ribeiro 等人提出了基于残差网络的 ECG 疾病分类模型 [23]。

在 ECG 信号的特征提取方面，1D-CNN 被广泛用于提取时序特征 [14]，Transformer 架构也被引入 ECG 分析以捕捉长程依赖关系。然而，现有方法在处理多导联 ECG 时通常将各导联独立处理或简单拼接，未能充分挖掘利用导联间的空间依赖关系。

2.3 注意力机制与图神经网络

注意力机制最初在自然语言处理领域被提出 [29]，随后被广泛应用于计算机视觉和信号处理领域。多头自注意力能够捕捉序列中任意两个位置之间的依赖关系，已在 ECG 分析中展现出潜力。然而，标准的自注意力机制将输入视为序列，未显式建模导联之间的物理空间关系。

图神经网络在处理图结构数据方面表现出色。Kipf 和 Welling 提出的图卷积网络通过谱图卷积实现了半监督节点分类 [13]。Veličković 等人提出的图注意力网络引入了注意力机制来学习邻居节点的重要性权重 [30]。然而，GNN 类方法主要适用于图结构数据，难以直接应用于 ECG 波形等时序信号。

本文提出的跨导联注意力机制受到 GAT 中注意力机制的启发，但进行了针对性的改造：将 12 个导联视为图节点，通过多头自注意力学习导联间的依赖权重，从而在 ECG 分析中引入图结构信息。

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

3 问题定义与评价指标

3.1 问题定义

本文将 PVC 消融靶点定位建模为一个多任务学习问题：给定一个 12 导联 ECG 波形 $\mathbf{X} \in \mathbb{R}^{12 \times T}$ （其中 T 为采样点数），模型需要同时输出：

- (1) 解剖区域类别 $y_{\text{cls}} \in \{0, 1, \dots, 6\}$;
- (2) 三维空间坐标 $\mathbf{y}_{\text{reg}} \in [-1, 1]^3$ 。

三维空间坐标的定义遵循本文自定义的标准心室模板坐标系：

- X 轴：负值表示左心室游离壁方向，正值表示右心室游离壁方向；
- Y 轴：负值表示心尖方向，正值表示基底部方向；
- Z 轴：负值表示前壁方向，正值表示后壁方向。

本文采用的坐标归一化方式为：以标准心室模板的最大包围盒为基准，将各方向的坐标线性缩放至 $[-1, 1]$ 区间。具体地，设原始坐标在第 d 个方向上的范围为 $[L_d, U_d]$ ，则归一化后的坐标 $y'_d = 2(y_d - L_d)/(U_d - L_d) - 1$ 。所有坐标已归一化至 $[-1, 1]$ 区间。

为直观展示该坐标系的几何含义，图1给出了该三维坐标在两个标准解剖切面上的投影：Anterior 视图（Y-Z 平面）显示了心尖-基底与前-后壁方向的空间关系；RAO 视图（X-Y 平面）显示了左右与心尖-基底方向的空间关系。

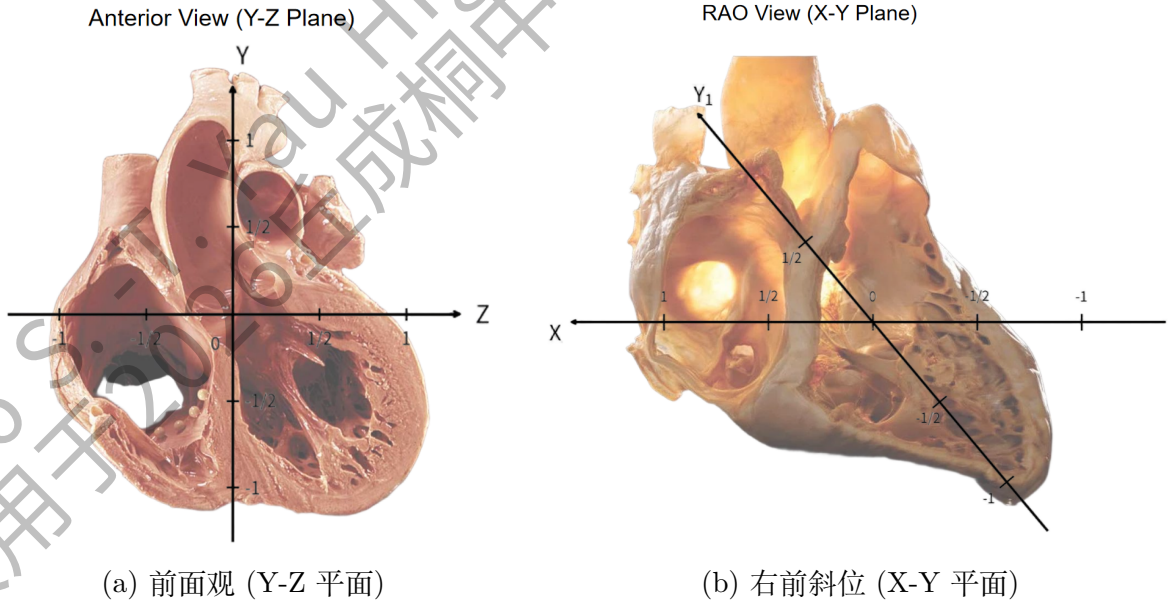


图 1: 自定义三维坐标系的解剖投影示意图

需要说明的是，图1b中 RAO 视图的坐标轴呈现倾斜，这是由心脏的解剖方位决定的：心脏的长轴并非沿人体标准坐标系的任一轴线方向，而是从右后方向左前方倾斜，

并伴有向下的倾斜 [1]。RAO 投照角度通常为 20° – 30° ，此时 X 射线束与心室长轴平行，因此 RAO 视图中的垂直轴在标准解剖投影中表现为倾斜方向，而非竖直。这一倾斜是心脏解剖投影的固有特征，不影响本文三维坐标定义的数学严谨性。所有坐标已归一化至 $[-1, 1]$ 区间。

解剖区域的分类标签定义如下：

表 1: 7 个解剖区域的定义与命名

类别编号	英文名称	中文名称
0	RVOT	右室流出道
1	LVOT	左室流出道
2	LV	左心室游离壁
3	RV	右心室游离壁
4	His	希氏束区域
5	Summit	心顶部
6	AMC	主动脉二尖瓣连接

3.2 不同解剖区域的 ECG 特征与坐标分布

在介绍各区域的 ECG 特征之前，图2展示了 7 个解剖区域在心脏中的实际空间位置，便于读者将 ECG 波形与解剖位置进行关联理解。

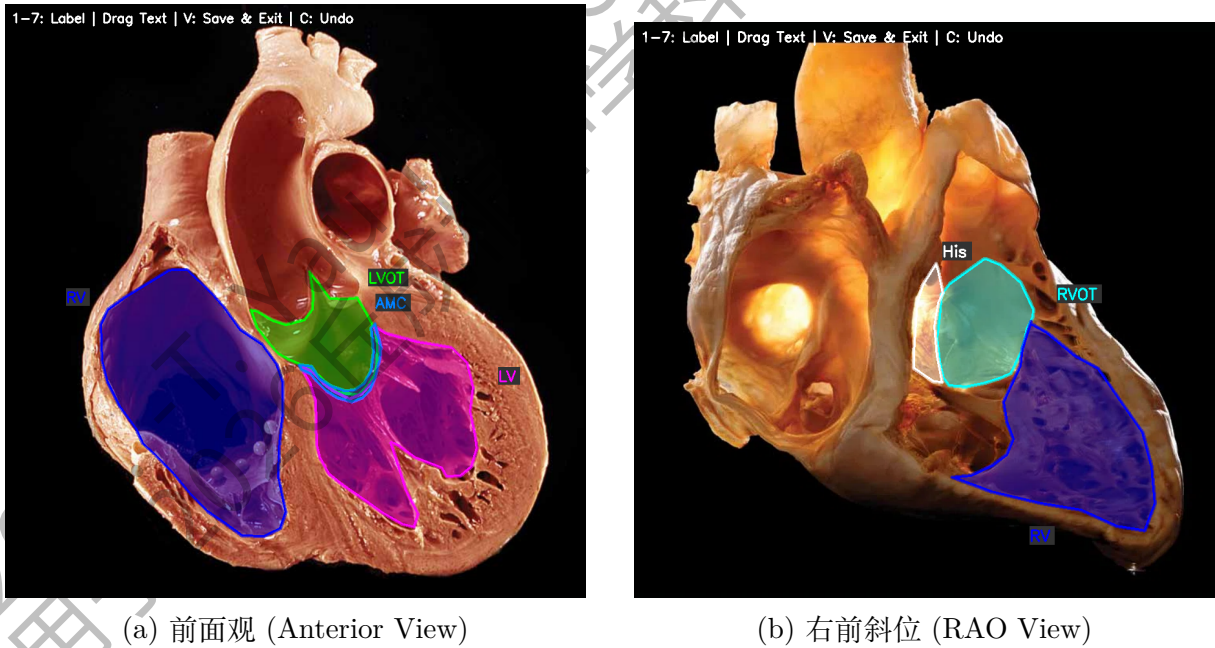


图 2: 7 个解剖区域的心脏空间位置示意图

从解剖结构来看，RVOT 与 LVOT 在空间上相邻，均位于心室基底部（Y 轴正值区域），但分别位于室间隔的两侧（X 轴负值与正值）；His 束区域位于室间隔膜部，靠近三尖瓣环；Summit 区域位于左心室外膜顶部，解剖位置特殊；AMC 则位于左心室

流出道与二尖瓣环之间的纤维连接区域。这些区域解剖位置相互邻近，使得仅凭 ECG 形态对其进行区分具有挑战性。

图3至图9以正常窦性心律为基准，将 7 种不同解剖起源的 PVC 波形与正常波形进行并列对比，便于观察各区域 PVC 在 12 导联上的特征性变化。图中彩色高亮标注了该起源类型最具鉴别性的导联特征。同时对每个区域给出了在本文自定义三维坐标系中的大致坐标范围；需要注意的是，实际处理时，我们将各个区域边缘轮廓坐标进行精确的轮廓描绘，但这里为了简便显示仅给出大致范围。

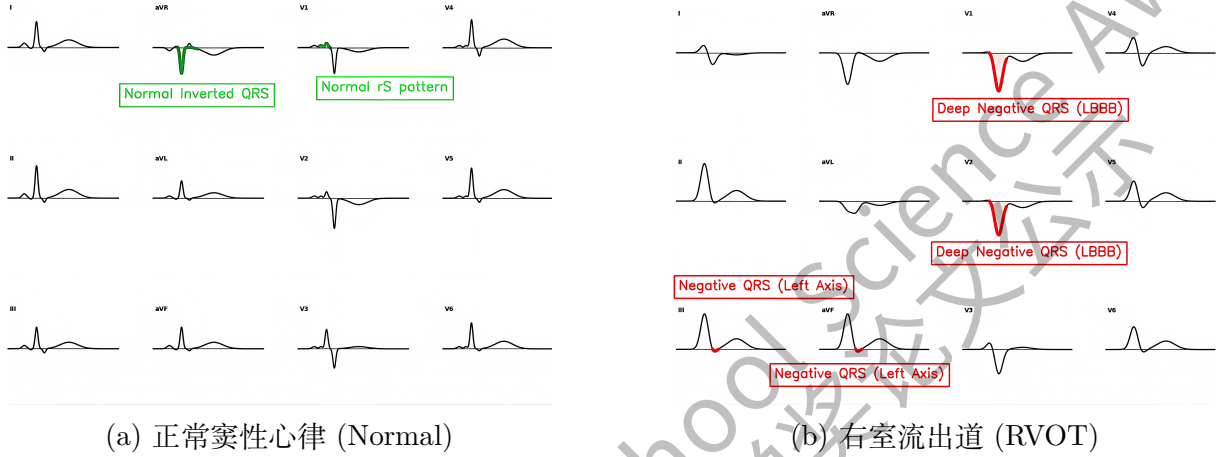


图 3: Normal 与 RVOT 对比

RVOT (右室流出道) 是临床最常见的 PVC 起源部位之一。其典型 ECG 特征为：V1-V3 导联呈深 S 波，形成 LBBB 形态；下壁导联呈高 R 波，电轴右偏；胸前导联移行区多在 V3-V4。在本文归一化坐标系中，RVOT 区域大致为： $x \in [0.2, 0.7]$ ， $y \in [0.3, 0.8]$ ， $z \in [-0.2, 0.4]$ 。

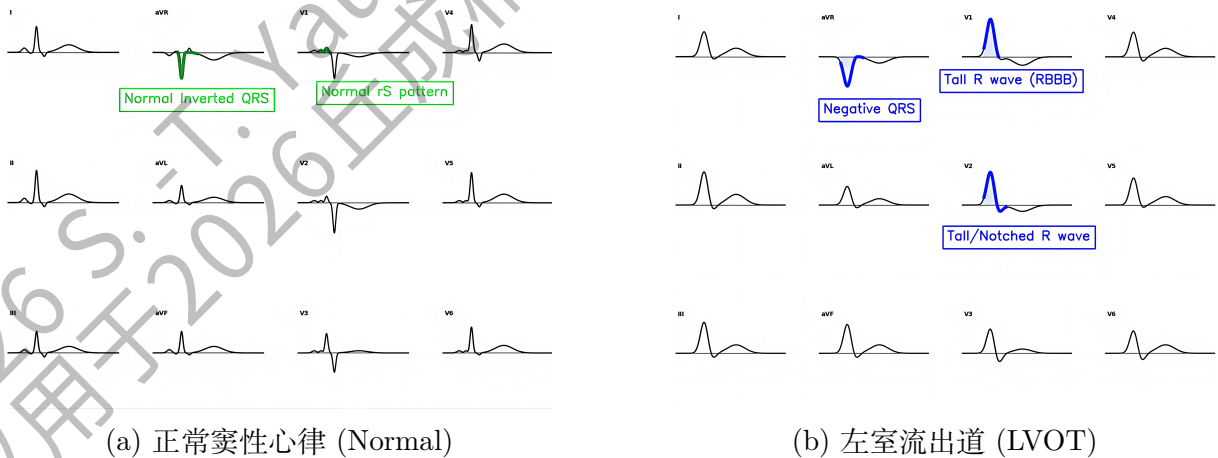


图 4: Normal 与 LVOT 对比

LVOT (左室流出道) 的 PVC 起源于主动脉瓣下方的左心室流出道心肌。其典型 ECG 特征为：V1 导联呈高 R 波，形成 RBBB 形态；V2 导联呈高或切迹 R 波；aVR 导联呈负向 QRS；下壁导联呈高 R 波。与 RVOT 相比，LVOT 的胸前导联移行区较早，

V1 导联 R 波振幅显著增高。在归一化坐标系中, LVOT 区域大致位于: $x \in [-0.6, -0.2]$, $y \in [0.2, 0.7]$, $z \in [-0.1, 0.5]$ 。

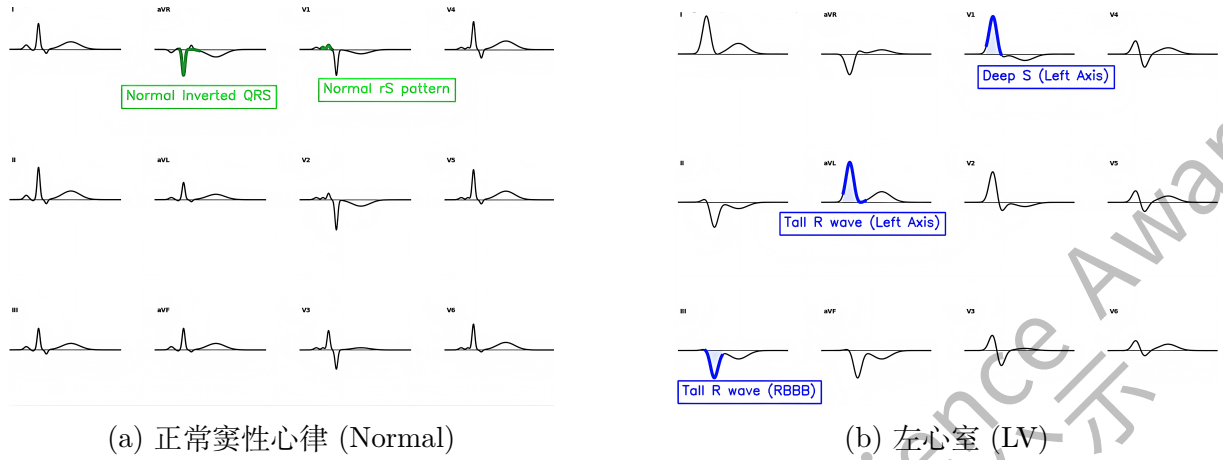


图 5: Normal 与 LV 对比

LV (左心室游离壁) 的 PVC 起源于左心室游离壁心肌, 范围较大。其典型 ECG 特征为: V1 导联呈高 R 波; aVL 导联呈高 R 波; 下壁导联呈深 S 波, 电轴左偏。左心室游离壁起源的 PVC 通常 QRS 波较宽。在归一化坐标系中, LV 区域范围较广: $x \in [-0.9, -0.3]$, $y \in [-0.6, 0.4]$, $z \in [-0.4, 0.6]$ 。

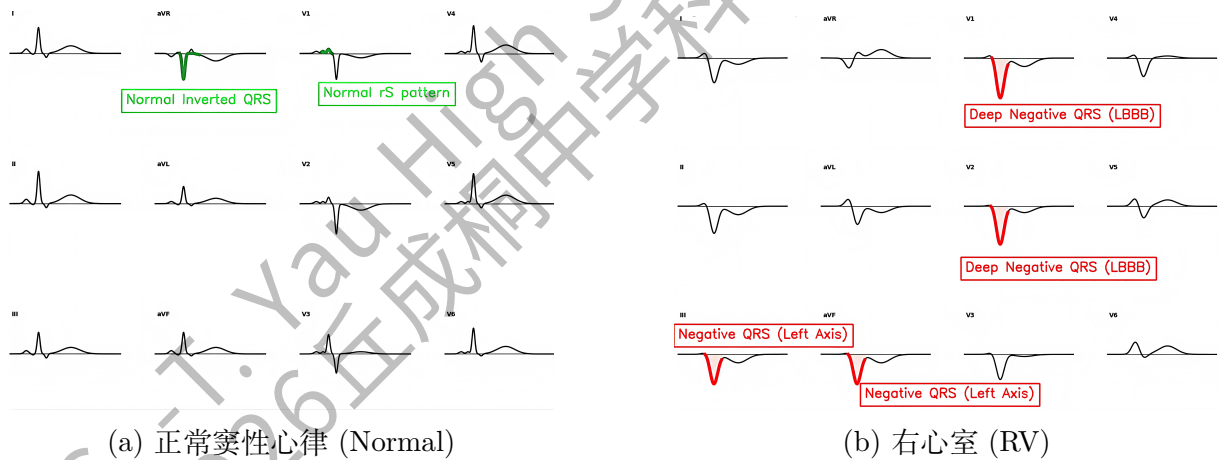


图 6: Normal 与 RV 对比

RV (右心室游离壁) 的 PVC 起源于右心室游离壁心肌, 是 LBBB 形态 PVC 的另一主要来源。其典型 ECG 特征为: V1-V4 导联呈深 S 波, LBBB 形态; 下壁导联呈深 S 波, 电轴左偏; QRS 波明显宽大。RV 起源与 RVOT 起源的关键鉴别点在于: RV 起源下壁导联为深 S 波, 而 RVOT 为高 R 波; RV 起源胸前导联移行区更晚。在归一化坐标系中, RV 区域大致为: $x \in [0.3, 0.9]$, $y \in [-0.5, 0.5]$, $z \in [-0.5, 0.3]$ 。

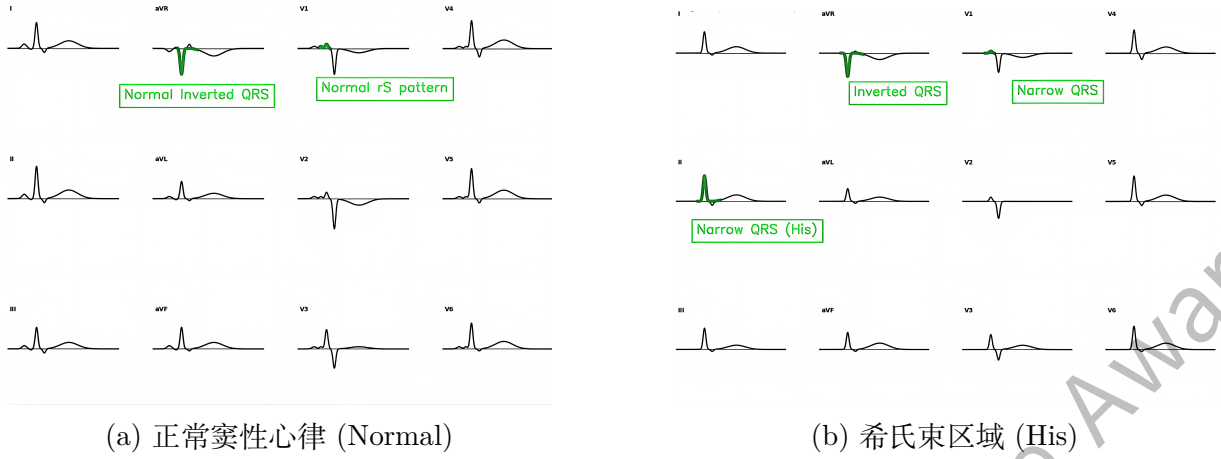


图 7: Normal 与 His 对比

His (希氏束区域) 的 PVC 起源于希氏束附近的心肌组织，是最具特殊性的起源部位。其典型 ECG 特征为：QRS 波呈窄 QRS 形态；II 导联呈窄 QRS；aVR 导联呈倒置 QRS；V1 导联呈窄 QRS。与其他区域相比，His 起源的 QRS 波通常最窄 (< 120 ms)。在归一化坐标系中，His 区域位于： $x \in [-0.2, 0.2]$ ， $y \in [0.4, 0.8]$ ， $z \in [-0.1, 0.3]$ 。

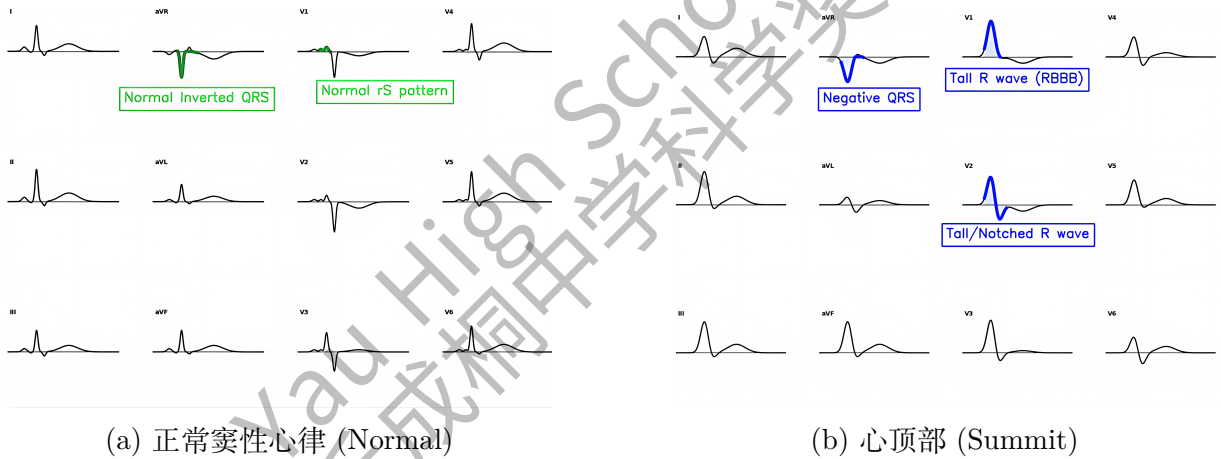


图 8: Normal 与 Summit 对比

Summit (心顶部) 的 PVC 起源于左心室外膜顶部区域，位于左冠脉前降支和回旋支交汇处的心外膜面。其典型 ECG 特征为：V1 导联呈高 R 波；V2 导联呈高大切迹 R 波；下壁导联呈高 R 波；胸前导联移行区较早。在归一化坐标系中，Summit 区域位于： $x \in [-0.4, -0.1]$ ， $y \in [0.5, 0.9]$ ， $z \in [0.1, 0.6]$ 。

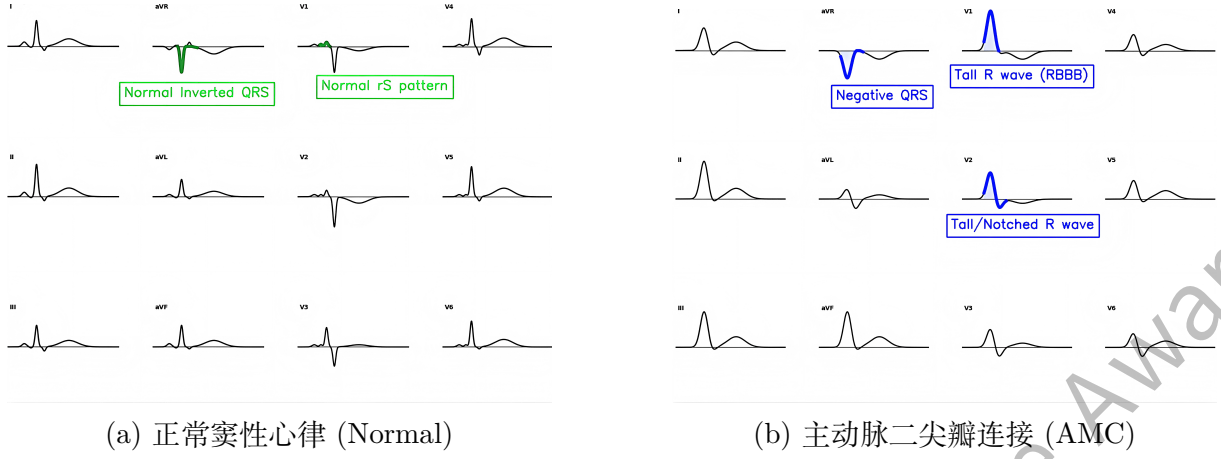


图 9: Normal 与 AMC 对比

AMC (主动脉二尖瓣连接) 位于左心室流出道与二尖瓣环之间的纤维连接区域。其典型 ECG 特征为：V1 导联呈高 R 波；V2 导联呈切迹 R 波；aVR 导联呈负向 QRS；下壁导联呈高 R 波。在归一化坐标系中，AMC 区域位于： $x \in [-0.4, 0.1]$, $y \in [0.3, 0.7]$, $z \in [-0.5, -0.1]$ 。

通过图3至图9中 Normal 与各区域 PVC 的并列对比，可以观察到不同起源部位在 12 导联上的特征性差异，以及各区域在三维空间中的坐标分布规律。这些解剖坐标分布差异为深度学习模型的跨导联注意力机制提供了医学基础。

3.3 评价指标

本文采用以下评价指标。

3.3.1 分类评价指标

准确率 (Accuracy):

$$\text{Acc} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\hat{y}_i = y_i) \quad (1)$$

其中 $\mathbb{I}(\cdot)$ 为指示函数，当预测类别 $\hat{y}_i$ 与真实类别 y_i 一致时取 1，否则取 0。准确率衡量模型在所有测试样本中正确分类的比例。由于测试集样本量较小，本文在报告中同时给出正确样本数和总样本数，避免仅依赖四舍五入后的百分数。当各类别样本数量不均衡时，准确率指标本身可能产生误导 [10]。

加权 F1 分数 (Weighted F1 Score):

$$\text{F1}_{\text{weighted}} = \sum_{c=0}^6 \frac{n_c}{N} \cdot \text{F1}_c \quad (2)$$

其中

$$\text{F1}_c = \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \quad (3)$$

$$\text{Precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c}, \quad \text{Recall}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c} \quad (4)$$

TP_c 、 FP_c 、 FN_c 分别为类别 c 的真阳性、假阳性和假阴性样本数, n_c 为类别 c 的样本数, $N = \sum_c n_c$ 为总样本数。本文采用加权 F1 分数, 以反映模型在类别不均衡场景下的综合分类性能 [26]。同时, 鉴于部分类别测试样本极少, 本文也报告宏 F1 分数作为补充。

3.3.2 回归评价指标

欧氏距离误差 (Euclidean Distance Error, EDE):

$$\text{EDE} = \frac{1}{N} \sum_{i=1}^N \|\hat{\mathbf{y}}_i - \mathbf{y}_i\|_2 \quad (5)$$

其中 $\hat{\mathbf{y}}_i$ 和 $\mathbf{y}_i$ 分别为预测坐标和真实坐标, $\|\cdot\|_2$ 为欧氏范数。由于坐标已归一化至 $[-1, 1]$, 理论最大 EDE 为 $\sqrt{2^2 + 2^2 + 2^2} = 2\sqrt{3} \approx 3.464$ 。为便于临床理解, 可将归一化 EDE 映射回物理空间。设原始三维坐标范围为 R , 则反归一化因子为 $R/2$, 物理空间中的平均定位误差为:

$$\text{EDE}_{\text{mm}} = \text{EDE} \times \frac{R}{2} \quad (6)$$

以本研究使用的标准心室模板坐标系为例, 若原始心脏模型的解剖范围约为 $R = 80 \text{ mm}$, 则 0.3090 的归一化 EDE 对应约 12.36 mm 的物理定位误差。在心内电生理标测领域, 起搏位点定位误差通常分布在 2.0 至 20.0 mm 之间, 中位误差约为 10 mm [33]。ECG 逆问题求解的定位误差在临床数据上约为 23.9 mm [20]。本研究的定位精度达到了可与临床金标准相比较的水平。

EDE 作为定位精度评价指标在 ECG 逆问题求解和非侵入性心电成像研究中已被广泛采用 [33, 20]。与测地线距离等其他度量相比, 欧氏距离直接反映三维空间中的直线定位偏差, 物理意义明确, 不依赖于心脏表面的网格划分或解剖结构的分段定义 [25]。

在分类任务中, 预测错误的样本其回归坐标通常也是不可靠的。术前医生通常会对消融靶点进行预估, 或参考 Carto、术前电生理检查等方式预先判断大致区域, 因此分类错误样本的回归 EDE 通常不会被实际采纳。本文引入严格评估 EDE 指标: 仅对分类正确的样本计算 EDE。需要强调的是, 严格评估 EDE 反映的是模型在分类正确前提下的定位能力, 而非所有样本的整体定位能力, 其数值变化来自评价口径的调整, 不能直接解释为模型定位精度的提升。

4 MCL-Net 模型架构

本章详细阐述本文提出的 MCL-Net 模型。MCL-Net 是一个端到端的混合深度学习框架，用于从 12 导联 ECG 波形中同时预测 PVC 消融靶点的解剖区域类别和三维空间坐标。

4.1 模型总体框架

MCL-Net 的整体架构如图10所示，主要由四个模块组成：

- (1) **多导联卷积模块**：对 12 个导联分别使用共享权重的 1D-CNN 提取时序特征；
- (2) **跨导联注意力模块**：将 12 个导联的特征向量视为图结构中的节点，通过多头自注意力机制显式建模导联之间的空间依赖关系；
- (3) **手工特征提取与融合模块**：从原始波形中提取 18 维手工心电图特征，经 MLP 编码后与深度特征拼接融合；
- (4) **多任务输出模块**：融合后的特征经全连接层后分叉为两个输出头，分别输出 7 个解剖区域的概率分布和三维空间坐标。

设输入为 12 导联 ECG 波形 $\mathbf{X} \in \mathbb{R}^{12 \times T}$ (本文中 $T = 500$)，MCL-Net 的前向传播过程如下。

首先，LeadCNN 对每个导联独立提取时序特征，并将 12 个导联的特征堆叠为特征矩阵：

$$\mathbf{F}_{\text{lead}} = \text{LeadCNN}(\mathbf{X}) \in \mathbb{R}^{12 \times d} \quad (7)$$

其中 $d = 64$ 为导联特征的维度。

随后，跨导联注意力模块通过多头自注意力机制建模导联间依赖关系：

$$\mathbf{F}_{\text{attn}} = \text{CrossAttn}(\mathbf{F}_{\text{lead}}) \in \mathbb{R}^{12 \times d} \quad (8)$$

将更新后的导联特征展平为深度特征向量：

$$\mathbf{f}_{\text{deep}} = \text{Flatten}(\mathbf{F}_{\text{attn}}) \in \mathbb{R}^{12d} \quad (9)$$

手工特征向量 $\mathbf{x}_{\text{hand}} \in \mathbb{R}^{18}$ 经 MLP 编码后与深度特征拼接融合：

$$\mathbf{f}_{\text{fused}} = \text{Fusion}([\mathbf{f}_{\text{deep}}; \mathbf{f}_{\text{hand}}]) \in \mathbb{R}^{64}, \quad \mathbf{f}_{\text{hand}} = \text{MLP}(\mathbf{x}_{\text{hand}}) \in \mathbb{R}^{64} \quad (10)$$

其中 $[\cdot; \cdot]$ 表示向量拼接操作。

最后，融合特征分别经分类头和回归头产生输出：

$$\begin{aligned} \mathbf{y}_{\text{cls}} &= \text{Softmax}(\mathbf{W}_{\text{cls}} \mathbf{f}_{\text{fused}} + \mathbf{b}_{\text{cls}}) \in \mathbb{R}^7 \\ \mathbf{y}_{\text{reg}} &= \tanh(\mathbf{W}_{\text{reg}} \mathbf{f}_{\text{fused}} + \mathbf{b}_{\text{reg}}) \in [-1, 1]^3 \end{aligned} \quad (11)$$

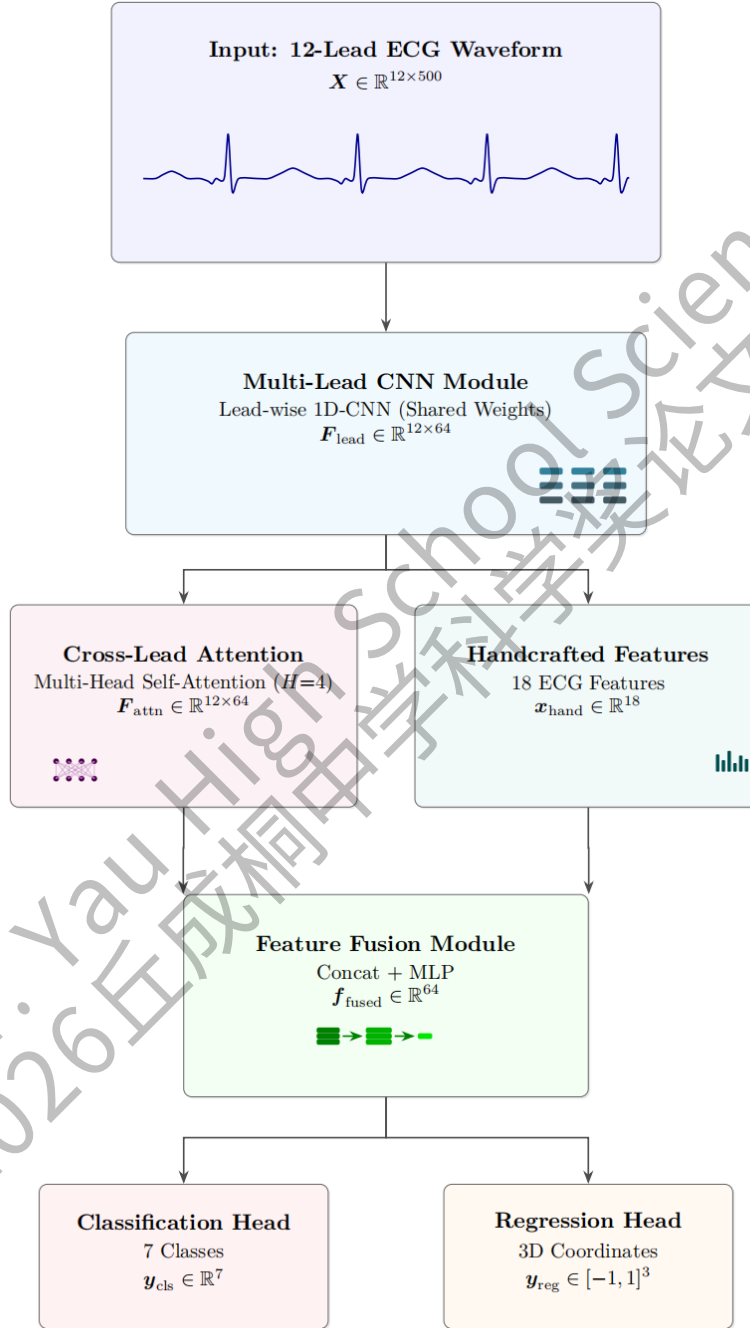


图 10: MCL-Net 整体架构图

4.2 多导联卷积模块

多导联卷积模块对 12 个导联分别进行时序特征提取。考虑到 12 个导联的采样率相同、波形长度一致，且各导联波形具有相似的形态学特征，本文在所有导联之间共享卷积核权重。这种设计减少了模型参数量，并使模型学习通用的波形特征。

对于每个导联 $\mathbf{x}^{(i)} \in \mathbb{R}^T$ ($i = 1, 2, \dots, 12$)，LeadCNN 的计算过程为：

$$\mathbf{h}^{(i)} = \text{LeadCNN}(\mathbf{x}^{(i)}) \quad (12)$$

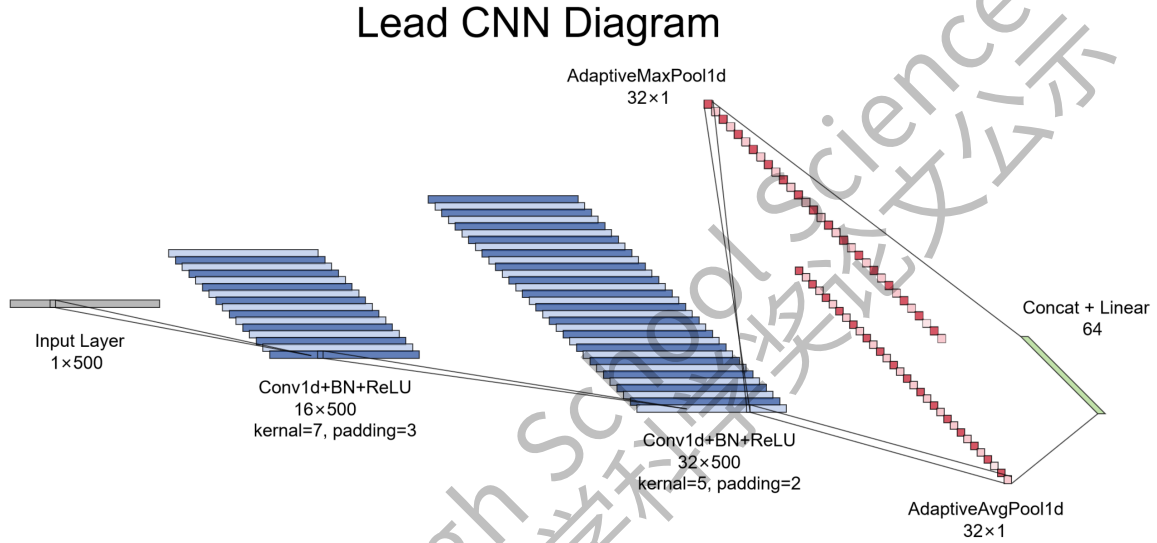


图 11: LeadCNN 网络结构示意图

具体地，每个导联波形首先通过两层一维卷积层提取局部时序特征，然后分别经过自适应平均池化和自适应最大池化，将变长时序信号压缩为固定长度特征向量。两种池化结果的拼接同时保留波形整体形态信息和峰值特征。最后通过线性层将特征映射到 $d = 64$ 维。设 $\mathbf{F}_{\text{lead}} \in \mathbb{R}^{12 \times 64}$ 为 12 个导联特征向量的堆叠矩阵。

4.3 跨导联注意力模块

跨导联注意力模块是 MCL-Net 的核心创新之一。其设计动机来源于 12 导联 ECG 的物理本质：12 个导联从不同空间角度记录心脏电活动，导联之间存在内在的解剖学依赖关系。本文的核心思路是：将 12 个导联视为图结构中的节点，通过多头自注意力机制在节点间传递信息，使模型能够自适应地学习导联之间的依赖关系。

该模块采用 Transformer 架构中标准的缩放点积注意力机制 [29]。查询表示目标导联“想要查找什么信息”，键表示源导联“包含什么信息”，值表示源导联“传递给目标导联的具体内容”。查询与键的点积衡量目标导联需求与源导联内容之间的匹配程度。

多导联卷积模块的输出为导联特征矩阵 $\mathbf{F}_{\text{lead}} \in \mathbb{R}^{12 \times 64}$ ，其第 i 行 $\mathbf{F}_{\text{lead}}^{(i)}$ 表示第 i 个

导联的 64 维特征向量。跨导联注意力模块将该矩阵视为一个全连通图的节点特征矩阵，其中 12 个导联为图节点。

4.3.1 缩放点积注意力的数学形式

对于第 h 个注意力头 ($h = 1, 2, \dots, H$, 本文中 $H = 4$), 有:

$$\mathbf{Q}_h = \mathbf{F}_{\text{lead}} \mathbf{W}_h^Q, \quad \mathbf{K}_h = \mathbf{F}_{\text{lead}} \mathbf{W}_h^K, \quad \mathbf{V}_h = \mathbf{F}_{\text{lead}} \mathbf{W}_h^V \quad (13)$$

其中 $\mathbf{W}_h^Q, \mathbf{W}_h^K, \mathbf{W}_h^V \in \mathbb{R}^{64 \times d_h}$ 为可训练投影矩阵, $d_h = 64/H = 16$ 为每个注意力头的特征维度。

对于第 h 个注意力头, 注意力分数矩阵为:

$$\mathbf{A}_h = \text{softmax} \left(\frac{\mathbf{Q}_h \mathbf{K}_h^T}{\sqrt{d_h}} \right) \in \mathbb{R}^{12 \times 12} \quad (14)$$

其中 $\mathbf{Q}_h \mathbf{K}_h^T$ 计算 12×12 的相似度矩阵。缩放因子 $\sqrt{d_h}$ 用于避免点积值方差过大导致 softmax 进入梯度饱和区 [29]。经过 softmax 归一化后, 注意力分数满足 $\sum_{j=1}^{12} (\mathbf{A}_h)_{i,j} = 1$ 。

4.3.2 多头机制的互补性

随后用注意力分数对值矩阵进行加权聚合, 得到每个导联在该注意力头下的新特征表示:

$$\text{head}_h = \mathbf{A}_h \mathbf{V}_h \in \mathbb{R}^{12 \times d_h} \quad (15)$$

head_h 的第 i 行 $\text{head}_h^{(i)}$ 为第 i 个导联的新特征:

$$\text{head}_h^{(i)} = \sum_{j=1}^{12} (\mathbf{A}_h)_{i,j} \mathbf{V}_h^{(j)} \quad (16)$$

本文采用多头机制。不同注意力头使用不同投影矩阵, 独立计算注意力权重并聚合特征, 使模型能够从多个特征子空间中捕捉导联之间的依赖模式。

将所有注意力头的输出在特征维度上拼接, 并通过输出投影矩阵融合多头信息:

$$\text{MultiHead}(\mathbf{F}_{\text{lead}}) = \text{Concat}(\text{head}_1, \dots, \text{head}_H) \mathbf{W}^O \in \mathbb{R}^{12 \times 64} \quad (17)$$

其中 $\mathbf{W}^O \in \mathbb{R}^{64 \times 64}$ 为可训练输出投影矩阵。

4.3.3 残差连接与层归一化

为保证信息传递稳定性并避免梯度消失，本文采用残差连接将输入与注意力输出相加，并经过层归一化：

$$\mathbf{F}_{\text{attn}} = \text{LayerNorm}(\mathbf{F}_{\text{lead}} + \text{MultiHead}(\mathbf{F}_{\text{lead}})) \in \mathbb{R}^{12 \times 64} \quad (18)$$

残差连接使模型能够保留每个导联的原始特征信息，同时叠加从其他导联聚合的上下文信息。层归一化对每个样本独立地在特征维度上进行归一化，有助于加速训练收敛并稳定深层网络的梯度流。

4.3.4 注意力热图的可解释性分析

注意力矩阵 $\mathbf{A}_h$ 中的元素 $(\mathbf{A}_h)_{ij}$ 定义为：在计算第 i 个目标导联的特征更新时，从第 j 个源导联聚合的信息权重。由于查询与键的线性投影矩阵为独立可训练参数，该权重矩阵在自然状态下具有非对称性。从电生理学角度阐释，这种非对称性可能反映了心肌电激动的空间传导方向性与时间延迟特征，但注意力权重不等同于因果解释，以下分析仅作为辅助观察。

图12呈现了模型对 LVOT 与 RVOT 两类起源预测时的平均注意力权重分布。图中颜色亮度越高代表注意力权重越大，亮度越低代表权重越小。该热图为随机种子 42、43、44 三次独立训练中分类正确样本的平均注意力权重，仅用于观察模型是否倾向于关注与临床定位规则相关的导联组合。

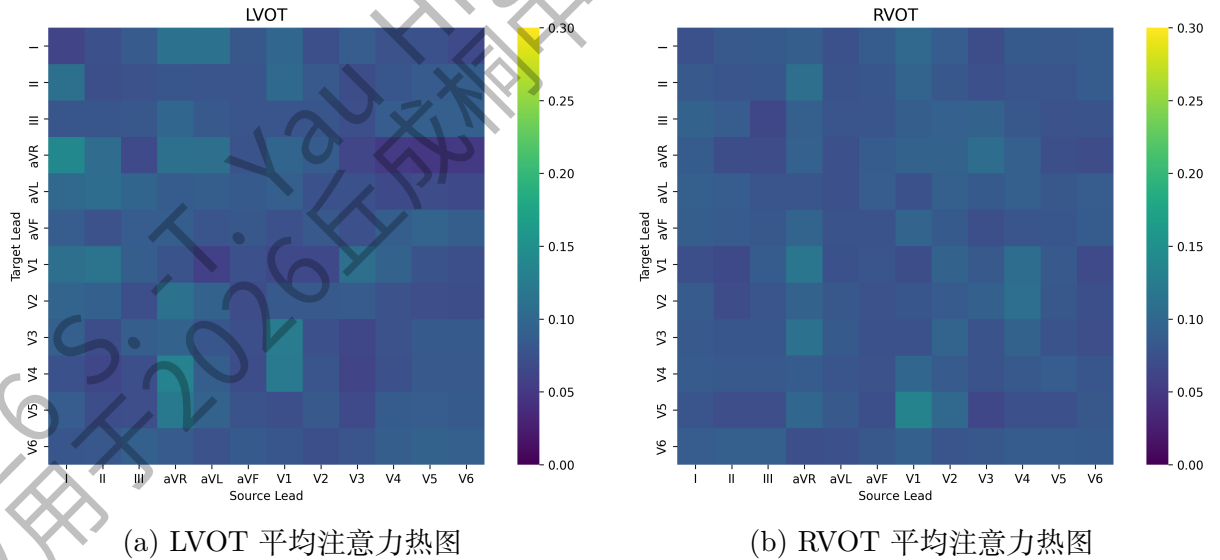


图 12: 跨导联注意力权重热图对比 (LVOT 与 RVOT): 横轴为源导联, 纵轴为目标导联。

LVOT 起源的导联依赖特征 在 LVOT 热图 (图12a) 中, 目标导联 V3、V4、V5 对源导联 aVL 和 V1 表现出较明显的注意力汇聚。从空间解剖学视角分析, aVL 导联对应

左心室高侧壁，而 V1 导联对应室间隔右侧面。LVOT 位于主动脉瓣下方，解剖位置偏左偏上，其电激动首先引起左室高侧壁激动，且激动向胸前导联的传导具有移行区提前的典型特征。模型使 V3-V5 目标导联向 aVL 和 V1 源导联分配较高权重，与临床定量指标中 V1 导联 R 波振幅指数以及移行区提前的诊断逻辑一致。

RVOT 起源的导联依赖特征 相比之下，RVOT 热图（图12b）整体呈现较低的权重底色，但在两个特定交叉区域存在显著激活。第一个显著激活区域位于目标导联 V1、V2 与源导联 aVR 的交界处。由于 RVOT 位于肺动脉瓣下的右心室最高位，电激动最先向上、向右传导，使 V1-V2 导联呈现典型深 S 波，同时 aVR 导联也常呈 QS 型。第二个显著激活区域位于目标导联 V5 与源导联 V1 的交界处。RVOT 起源的室早胸前导联移行区通常延后，V5 的特征表示需要依赖 V1 提供的早期激动时序信号。

总结与临床关联性 跨导联注意力模块从数据中学习到了两类起源不同的空间构型特征：LVOT 模式下，模型侧重于利用 aVL 和 V1 提供的早期空间投影辅助 V3-V5 的形态判别；RVOT 模式下，模型侧重于利用 aVR 提供的右室高位负向向量，并结合 V1 向 V5 的单向时序传导特征。该模块内部的注意力权重分配机制，在数据维度上复现了临床专家利用 V1-R 波振幅指数、V1-V3 移行指数以及 V2 移行比等复合定量指标进行流出道定位的逻辑。需要说明的是，上述分析基于注意力权重的统计观察，不能直接推断模型的因果决策机制，但可为模型预测提供一定的可解释性支持。

4.4 手工特征提取与融合模块

尽管深度学习方法能够自动学习 ECG 波形中的特征表示，但心电图学中积累的临床知识以及人为设计的手工特征仍然具有重要的补充价值。本文设计了 18 维手工心电图特征，这些特征基于心脏病学和心电生理学的经典诊断标准。

4.4.1 18 维手工特征的定义

18 维手工特征的具体定义如表2所示。

表 2: 18 维手工心电图特征的定义

编号	特征名称	类型	临床意义
1	V1 导联 R 波幅值	连续	右室前壁电活动
2	V1 导联 S 波幅值	连续	右室前壁除极深度
3	V1 导联 Q 波幅值	连续	右室前壁初始除极
4	V6 导联 R 波幅值	连续	左室侧壁电活动
5	V6 导联 S 波幅值	连续	左室侧壁除极深度
6	II 导联 R 波幅值	连续	下壁电活动
7	III 导联 R 波幅值	连续	下壁电活动
8	aVF 导联 R 波幅值	连续	下壁电活动
9	III/II R 波幅值比	连续	下壁电轴偏移方向
10	LBBB 形态标志	二值	是否为 LBBB 形态
11	RBBB 形态标志	二值	是否为 RBBB 形态
12	V1 导联 qR 形态标志	二值	是否为 qR 形态
13	下壁电轴标志	二值	是否为下壁电轴
14	QRS 宽度	连续	室内传导时间 (采样点数)
15	QRS 增宽标志	二值	是否超过 60 个采样点 (120 ms)
16	12 导联 R 波均值	连续	整体心肌电活动强度
17	12 导联 R 波标准差	连续	心肌电活动离散度
18	12 导联 S 波均值	连续	整体除极深度

这些特征从波形形态、电轴方向、束支传导模式等角度刻画了 PVC 的 ECG 特征, 与解剖起源位置存在密切关联。流出道起源的 PVC 共同特征为下壁导联呈单相 R 波, 其 R 波幅值直接反映激动起源点在基底部的位点高低。在束支传导形态方面, LVOT 起源的 PVC 最常见的 ECG 表现为 V1 导联呈 RBBB 形态, 下壁导联呈高 R 波; RVOT 起源的 PVC 则表现为 V1 导联呈 LBBB 形态, 同样伴有下壁导联高 R 波 [15]。

4.4.2 特征提取流程

手工特征的提取流程如下: 对 12 导联 ECG 波形 $\mathbf{X} \in \mathbb{R}^{12 \times T}$, 首先通过峰值检测算法定位各导联的 QRS 波群, 进而计算 R 波、S 波、Q 波的幅值, 以及 QRS 宽度等形态学参数。所有特征提取算法均基于经典的 Pan-Tompkins 算法 [18]。

需要说明的是, Pan-Tompkins 算法最初针对正常窦性心律设计, 而 PVC 的 QRS 波群宽大畸形, 可能导致误检或漏检。为降低该风险, 本文对检测结果进行了抽样人工复核, 重点校正 QRS 起始、R 峰和 S 波终止位置。

Pan-Tompkins 算法的核心处理流程如图13所示, 包含五个级联的信号处理阶段:

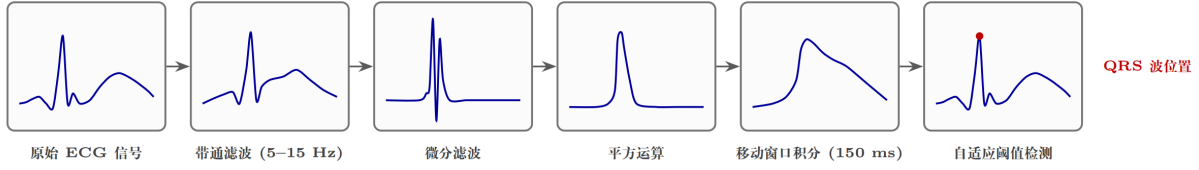


图 13: Pan-Tompkins 算法 QRS 波检测流程。图中展示了信号从原始 ECG 到最终 QRS 波位置标注的信号形态变化

1. **带通滤波**: 采用 5-15 Hz 的带通滤波器对原始 ECG 信号进行滤波，抑制基线漂移和低频噪声。
2. **微分滤波**: 对带通滤波后的信号进行五点差分运算，以强调 QRS 波群的陡峭斜率特征。
3. **平方运算**: 将微分后的信号逐点平方，将所有信号值转为正值，并进一步放大 QRS 波群中斜率较大的分量。
4. **移动窗口积分**: 采用宽度约为 150 ms 的移动窗口对平方后的信号进行积分，产生平滑的包络信号。
5. **自适应阈值判决**: 通过两套动态更新的阈值判定 QRS 波群的位置。

经过上述五个步骤，Pan-Tompkins 算法输出每个导联中 QRS 波群的位置。在此基础上，本文进一步定位每个 QRS 波群内的 Q 波、R 波和 S 波的具体位置，并测量其幅值。QRS 宽度定义为 Q 波起始点至 S 波终止点之间的时间跨度。按表2的定义计算所有 18 个特征值，构成手工特征向量 $\mathbf{x}_{\text{hand}} \in \mathbb{R}^{18}$ 。

4.4.3 特征重要性分析

为评估各手工特征在 PVC 起源分类任务中的贡献度，本文基于训练集数据训练了一个随机森林分类器（500 棵树），并计算了基尼重要性 [5]。本文选择随机森林而非单一决策树进行特征重要性评估，是因为随机森林通过自助采样和随机特征子集选择，能够提供更稳健的特征重要性估计。

基尼重要性计算基于每个特征在所有决策树节点分裂时对基尼不纯度的减少量 [5]。将某特征在全部树、全部节点上的基尼不纯度减少量进行累加并归一化，即可得到该特征的最终重要性得分。本文在计算基尼重要性时，随机森林的训练数据严格限定于训练集，验证集和测试集不参与特征重要性评估。随机森林的超参数设置为：树的数量为 500，分裂准则采用基尼不纯度，其余参数保持 scikit-learn 默认值。

表3展示了重要性排名前 6 的特征及其简要临床意义。

表 3: 手工特征重要性排名前 6

排名	特征名称	重要性	临床意义
1	II 导联 R 波幅值	0.2542	下壁电轴核心指标
2	III/II R 波幅值比	0.2092	下壁电轴偏移方向
3	III 导联 R 波幅值	0.1854	下壁电轴辅助指标
4	12 导联 R 波标准差	0.0647	心肌电活动离散度
5	12 导联 R 波均值	0.0618	整体心肌电活动强度
6	V6 导联 R 波幅值	0.0525	左室侧壁电活动

重要性最高的特征全部集中于下壁导联相关指标，三者合计占比超过 60%。这一分布规律与临床心电图学的流出道定位逻辑高度吻合。文献指出，流出道起源的 PVC 无论左右，均表现为下壁导联高 R 波 [15]。II 导联 R 波幅值排名第一，Li 等人在 309 例接受导管消融的症状性 PVC 患者中的研究发现，II 导联 R 波幅值是判断 RVOT 起源部位的最优预测指标之一，其 ROC 曲线下面积为 0.812–0.841，最佳阈值为 1.35 mV [15]。III/II R 波幅值比排名第二，Betensky 等人的研究表明，III/II 导联 R 波幅值比 ≥ 1 可预测 RVOT 间隔部起源，特异性达 92.3% [4]。连续数值型特征在重要性排名中占据前三位且合计占比超过 60%，显著高于二值化型特征。

4.4.4 手工特征编码与融合

得到 18 维手工特征向量 $\mathbf{x}_{\text{hand}} \in \mathbb{R}^{18}$ 后，需将其与深度特征分支的输出进行融合。由于两类特征的语义层次和数值分布存在显著差异，直接拼接会导致手工特征在后续计算中被深度特征淹没。因此，本文首先通过一个特征编码层将手工特征映射到与深度特征可比较的表示空间：

$$\mathbf{f}_{\text{hand}} = \text{ReLU}(\mathbf{x}_{\text{hand}} \mathbf{W}_{\text{hand}}^1 + \mathbf{b}_{\text{hand}}^1) \in \mathbb{R}^{64} \quad (19)$$

其中 $\mathbf{W}_{\text{hand}}^1 \in \mathbb{R}^{18 \times 64}$ 和 $\mathbf{b}_{\text{hand}}^1 \in \mathbb{R}^{64}$ 为可训练参数。

随后将手工特征编码与跨导联注意力模块输出的深度特征 $\mathbf{f}_{\text{deep}} \in \mathbb{R}^{768}$ 拼接，得到 832 维融合向量：

$$\mathbf{f}_{\text{concat}} = [\mathbf{f}_{\text{deep}}; \mathbf{f}_{\text{hand}}] \in \mathbb{R}^{832} \quad (20)$$

拼接操作是实现多模态特征融合最简单有效的方式之一 [22]，其优势在于保留了两类特征的原始信息，让后续网络自主决定如何组合利用这些特征。拼接后的向量通过融合层进行高阶特征交互：

$$\mathbf{f}_{\text{fused}} = \text{ReLU}(\text{BN}(\mathbf{f}_{\text{concat}} \mathbf{W}_{\text{fuse}}^1 + \mathbf{b}_{\text{fuse}}^1)) \mathbf{W}_{\text{fuse}}^2 + \mathbf{b}_{\text{fuse}}^2 \in \mathbb{R}^{64} \quad (21)$$

融合层的完整结构如表4所示。

表 4: 特征融合层的网络结构与维度变化

阶段	输入维度	操作	输出维度
手工特征编码	18	全连接 + ReLU	64
特征拼接	768 + 64	向量拼接 (Concat)	832
融合层第 1 层	832	全连接 + BN + ReLU	128
融合层第 2 层	128	全连接 + BN + ReLU	64

融合层由两个全连接层组成 ($832 \rightarrow 128 \rightarrow 64$)，每层后接批归一化和 ReLU 激活函数。批归一化用于加速收敛并稳定训练，ReLU 激活引入非线性，使网络能够学习深度特征与手工特征之间的复杂交互模式。

为保证手工特征与深度特征在数值尺度上的协调，训练前对所有手工特征进行 Z-score 标准化，使每个特征均值为 0、标准差为 1。标准化参数基于训练集计算并保存，在验证和测试阶段应用相同变换，避免数据泄露。

4.5 多任务学习与损失函数

4.5.1 多任务学习的动机

MCL-Net 通过多任务学习框架同时优化分类和回归两个目标。多任务学习的核心假设是：分类任务和回归任务共享底层的特征表示，两个任务之间可以相互提供归纳偏置，从而提升各自的泛化性能 [7]。在本文的 PVC 消融靶点定位问题中，不同解剖区域的靶点在三维空间中本身就占据不同的位置范围，区域分类与坐标回归本质上是同一空间定位问题的两种不同粒度描述。

分类任务帮助模型学习不同解剖区域的判别性特征，引导模型关注不同起源部位在 ECG 波形上的定性和定量差异；回归任务则引导模型学习三维空间的连续位置信息，促使模型捕捉波形形态与靶点位置之间的细粒度映射关系。两个任务共享多导联卷积模块、跨导联注意力模块和手工特征融合模块的参数。这种双向信息传递使两个任务能够相互促进，而非简单叠加二者的效果。从优化角度分析，多任务学习还起到隐式正则化作用，降低过拟合风险 [24]。

4.5.2 输出头设计

设多导联卷积模块和跨导联注意力模块输出的深度特征 $\mathbf{f}_{\text{deep}} \in \mathbb{R}^{768}$ 与手工特征编码 $\mathbf{f}_{\text{hand}} \in \mathbb{R}^{64}$ 拼接后，经融合层得到共享表示：

$$\mathbf{f}_{\text{fused}} = \text{Fusion}([\mathbf{f}_{\text{deep}}; \mathbf{f}_{\text{hand}}]) \in \mathbb{R}^{64} \quad (22)$$

分类头由单层全连接层组成，输出 7 个解剖区域的概率分布：

$$\mathbf{y}_{\text{cls}} = \text{softmax}(\mathbf{f}_{\text{fused}} \mathbf{W}_{\text{cls}} + \mathbf{b}_{\text{cls}}) \in \mathbb{R}^7 \quad (23)$$

其中 $\mathbf{W}_{\text{cls}} \in \mathbb{R}^{64 \times 7}$ 和 $\mathbf{b}_{\text{cls}} \in \mathbb{R}^7$ 为分类头可训练参数。

回归头同样由单层全连接层组成，输出三维空间坐标：

$$\mathbf{y}_{\text{reg}} = \tanh(\mathbf{f}_{\text{fused}} \mathbf{W}_{\text{reg}} + \mathbf{b}_{\text{reg}}) \in [-1, 1]^3 \quad (24)$$

其中 $\mathbf{W}_{\text{reg}} \in \mathbb{R}^{64 \times 3}$ 和 $\mathbf{b}_{\text{reg}} \in \mathbb{R}^3$ 为回归头可训练参数。选择 $\tanh$ 作为激活函数的原因是其输出范围天然匹配本文归一化坐标范围 $[-1, 1]$ 。

4.5.3 损失函数

MCL-Net 的总损失函数由三部分组成：分类损失、回归损失和 L2 正则化项：

$$\mathcal{L} = \mathcal{L}_{\text{cls}} + \lambda_{\text{reg}} \cdot \mathcal{L}_{\text{reg}} + \lambda_{l2} \cdot \|\Theta\|_2^2 \quad (25)$$

其中 λ_{reg} 为回归损失权重系数， λ_{l2} 为 L2 正则化系数， Θ 为模型所有可训练参数。 λ_{reg} 的取值直接决定分类任务与回归任务在训练过程中梯度的相对权重。关于其最优取值，本文在第 6.3 节中通过系统的网格对比实验进行了验证。

分类损失采用标签平滑的交叉熵损失：

$$\mathcal{L}_{\text{cls}} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=0}^6 \tilde{y}_{i,c} \log(\hat{y}_{i,c}) \quad (26)$$

其中 $\tilde{y}_{i,c} = (1 - \alpha) \cdot \mathbb{I}(y_i = c) + \alpha/7$ 为平滑后的标签， $\alpha = 0.1$ 为平滑系数。标签平滑通过将一部分概率质量分配给非目标类别，防止模型对训练数据过度自信，从而缓解过拟合问题 [28]。

回归损失采用 Smooth L1 损失函数：

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^N \sum_{d=1}^3 \text{smooth}_{L1}(\hat{y}_{i,d} - y_{i,d}) \quad (27)$$

其中

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2/\beta, & \text{if } |x| < \beta \\ |x| - 0.5\beta, & \text{otherwise} \end{cases} \quad (28)$$

$\beta = 0.5$ 为平滑阈值。Smooth L1 损失在误差较小时表现为 L2 损失，提供平滑梯度；在误差较大时表现为 L1 损失，降低异常值对梯度的影响，使训练更加稳定 [8]。

L2 正则化项权重系数 λ_{l2} 经多次调参实验后定为 10^{-4} 。回归损失权重 λ_{reg} 的最优取值在第 6.3 节中通过对比实验系统确定 ($\lambda_{\text{reg}} = 0.9$)。

4.6 本章小结

本章详细阐述了 MCL-Net 的模型架构。针对 12 导联 ECG 的空间结构特点,本文提出了一种融合多导联卷积、跨导联注意力与手工特征的多任务学习框架。多导联卷积模块通过共享权重的 1D-CNN 对每个导联独立提取局部波形特征。跨导联注意力模块将 12 个导联视为图结构中的节点,通过多头自注意力机制显式建模导联间的空间依赖关系。手工特征模块基于临床心电图学的经典诊断标准构建 18 维特征,经编码后与深度特征拼接融合。分类头与回归头共享底层表示,通过多任务学习实现解剖区域分类与三维坐标回归的联合优化。

5 数据集与实验设置

本章介绍实验所用的数据集、实验设置、评估策略、基线方法以及消融实验设计。

5.1 数据集

本研究共纳入 273 例临床数据，所有数据均来源于中国人民解放军总医院（301 医院）接受射频消融术治疗的 PVC 患者术前心电图数据。

原始 12 导联 ECG 信号以 500 Hz 采样频率记录，截取每个 PVC 事件前后各 250 个采样点（共计 500 个采样点），构成维度为 12×500 的输入张量。每个样本的标签包含两部分：

(1) 解剖区域类别标签，来源于消融手术记录中经导管消融治疗后确认的治愈或显著好转的靶点位置。

(2) 三维空间坐标 (x, y, z) ，由手术记录中靶点位置结合心脏解剖结构经规则映射与人工校验后得到，归一化至 $[-1, 1]$ 区间。

为保证同一患者只贡献一个 PVC 样本，对于具有多个 PVC 形态的患者，仅纳入经消融确认且 ECG 信号质量最高的一个样本，避免同一患者样本同时出现在训练集和测试集中导致数据泄露。

数据集按照 7:2:1 的比例随机划分为训练集（191 例）、验证集（59 例）和测试集（23 例）。数据划分使用固定随机种子 42，并采用分层采样，确保 7 个解剖区域在训练集、验证集和测试集中均有分布。为评估训练稳定性，本文在固定数据划分下进行三次独立训练，模型初始化随机种子分别为 42、43、44，最终报告三次实验的平均性能。

7 个解剖区域的样本数及分层划分结果如表5和图14所示。

表 5: 7 个解剖区域样本数及分层划分

类别编号	解剖区域	总样本数	训练集	验证集	测试集
0	RVOT	76	53	17	6
1	LVOT	82	57	19	6
2	LV	28	20	5	3
3	RV	25	18	5	2
4	His	17	12	3	2
5	Summit	23	16	5	2
6	AMC	22	15	5	2
合计	—	273	191	59	23



图 14: 7 个解剖区域样本数的一行堆叠可视化 (合计 273 例)。

5.2 实验设置

所有实验均在 PyTorch 框架下实现。模型采用 AdamW 优化器 [17], 初始学习率设为 10^{-3} , 权重衰减系数为 10^{-4} , 批次大小为 16, 最大训练轮数为 120 轮, 早停耐心值为 30 轮。训练过程中采用余弦退火学习率调度策略 [16], 使学习率从初始值按余弦函数逐渐衰减至接近零。

训练过程中, 每轮结束后在验证集上评估模型的分类准确率和 EDE, 并保存验证集上分类准确率最高的模型权重作为最佳模型。若验证集准确率在连续 30 轮内没有提升, 则提前终止训练, 以防止过拟合。数据划分与主模型初始化的随机种子固定为 42; 三次独立训练的模型初始化种子分别为 42、43、44。所有实验结果均报告在测试集上的性能, 并报告三次独立训练的平均值。

5.3 评估策略

本文采用第 3.3 节定义的评价指标, 包括分类准确率、加权 F1 分数和欧氏距离误差。在具体实验中, 所有指标均在测试集上计算, 并报告三次独立训练的平均值。由于测试集样本量较小 (23 例), 本文在报告百分数的同时给出原始正确样本数, 避免仅依赖四舍五入后的数值。

对于回归任务, 本文报告两种 EDE 变体:

(1) **标准 EDE**: 对测试集中所有样本计算预测坐标与真实坐标之间的欧氏距离, 取平均。

(2) **严格评估 EDE**: 仅对分类正确的样本计算 EDE。在分类任务中, 预测错误的样本其回归坐标通常也是不可靠的, 因此临床上也不会采纳其对应的回归结果。严格评估 EDE 反映的是模型在分类正确前提下的定位能力, 而非所有样本的整体定位能力。

为便于临床理解, 本文将归一化 EDE 映射回物理空间。设原始三维坐标范围为 R , 则反归一化因子为 $R/2$ 。以本研究使用的标准心室模板坐标系为例, 若 $R = 80 \text{ mm}$, 则 0.3090 的归一化 EDE 对应约 12.36 mm 的物理定位误差。

5.4 基线方法

为全面评估 MCL-Net 的性能, 本文选取两类基线方法进行对比。所有基线模型均使用与 MCL-Net 相同的数据划分、训练轮数和早停策略, 并在相同调参预算下进行对比。

(1) **传统机器学习方法**：包括随机森林和 XGBoost。两者均以 18 维手工特征作为输入，采用与 MCL-Net 相同的训练/验证/测试划分。随机森林使用 500 棵决策树，XGBoost 使用 200 棵提升树、学习率 0.1。

(2) **深度基线模型**：包括 ResNet-1D、Transformer Encoder 和 CNN+Cross-Lead Attention。ResNet-1D 采用三层残差块处理 12 导联波形；Transformer Encoder 将 12 导联波形展平为序列后输入标准 Transformer 编码器；CNN+Cross-Lead Attention 使用与 MCL-Net 相同的多导联卷积和跨导联注意力结构，但移除了手工特征分支。所有深度基线模型均使用原始 12 导联波形作为输入，不包含手工特征。

为评估手工特征分支对深度模型的贡献，本文还补充了“ResNet-1D + 手工特征”和“Transformer + 手工特征”两个变体作为深度基线，以消除输入信息差异带来的不公平比较。

5.5 消融实验设计

为验证 MCL-Net 中各模块的贡献，本文在未进行超参数调优的默认配置下进行了系统的消融实验。需要说明的是，本节报告的消融实验结果旨在展示各模块的相对贡献趋势，而非模型的最终最优性能。模型的最终调优配置及对应性能详见第 6.1 节。

具体地，本文设计了以下四种消融配置：

(1) **配置 A：仅波形分支**。仅使用多导联卷积模块对 12 导联波形进行特征提取，不含跨导联注意力模块，也不含手工特征分支。

(2) **配置 B：波形 + 跨导联注意力**。在多导联卷积模块基础上加入跨导联注意力模块，使模型能够建模导联间的空间依赖关系，但不含手工特征分支。

(3) **配置 C：波形 + 手工特征**。在多导联卷积模块基础上加入 18 维手工特征分支，但不含跨导联注意力模块。

(4) **配置 D：完整 MCL-Net**。包含多导联卷积、跨导联注意力和手工特征融合三个模块的完整架构。

所有消融实验均在固定随机种子 42 的数据划分下进行三次独立训练，最终报告三次实验的平均性能。各配置的训练超参数与第 5.2 节一致，以保证对比的公平性。

6 实验结果与分析

本章报告 MCL-Net 在 273 例临床数据集上的实验结果。首先呈现三次独立训练的分类与回归性能，然后与多种基线方法进行对比，随后通过消融实验验证各模块的贡献，最后对结果进行讨论与分析。

6.1 三次独立训练结果

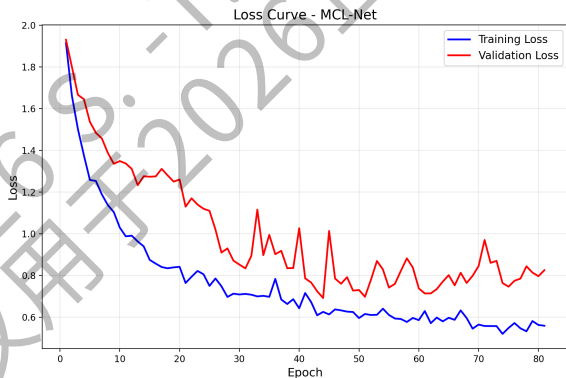
表6报告了 MCL-Net 在固定随机种子 42 的数据划分下三次独立训练中的详细性能。三次训练的分类准确率分别为 91.30% (21/23)、86.96% (20/23) 和 95.65% (22/23)，平均准确率为 91.30% (63/69)。三次结果分类准确率标准差约为 4.35%，表明模型在该数据划分下具有一定稳定性。

表 6: 固定数据划分下三次独立训练结果 (严格评估 EDE)

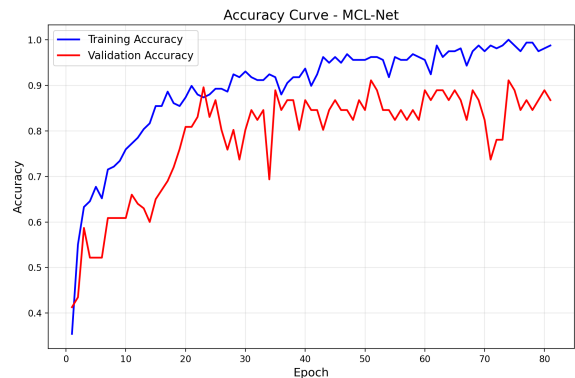
训练轮次	Accuracy (%)	Weighted F1 (%)	严格评估 EDE
第 1 次	91.30 (21/23)	88.5	0.3102
第 2 次	86.96 (20/23)	86.4	0.3024
第 3 次	95.65 (22/23)	90.5	0.3144
均值 $\pm$ 标准差	91.30 $\pm$ 4.35	88.47 $\pm$ 2.05	0.3090 $\pm$ 0.0061

注：标准 EDE 在三次训练中的均值为 0.5441。

为直观展示模型的训练过程，图15给出了 MCL-Net 在训练集和验证集上的损失曲线与准确率曲线。损失曲线显示，训练损失从初始的约 1.95 稳步下降至约 0.55，验证损失从约 1.93 下降至约 0.80。验证损失在训练后期存在一定波动但整体呈下降趋势，训练损失与验证损失之间未出现显著发散，表明模型未发生明显的过拟合。准确率曲线显示，训练准确率从初始的约 0.35 上升至约 0.98，验证准确率从约 0.42 上升至约 0.88。验证准确率在训练后期波动较大，这与验证集样本量较小 (59 例) 有关。



(a) 损失曲线



(b) 准确率曲线

图 15: MCL-Net 训练过程中的损失曲线与准确率曲线

6.2 与基线模型的对比

本节对比的是未调参版本的 MCL-Net，以与基线在相同调参预算下比较；调优后性能见表8和第6.1节。表7将 MCL-Net（未调参）与多种基线方法进行了对比。所有方法均在相同的数据划分和相同的调参预算下评估。由于测试集为 23 例，表中同时给出正确样本数和总样本数。

表 7: 各方法在测试集上的性能对比

方法	Accuracy (%)	Weighted F1 (%)	EDE
随机森林	82.61 (19/23)	78.38	—
XGBoost	82.61 (19/23)	78.54	—
ResNet-1D	69.57 (16/23)	—	0.8023
Transformer Encoder	69.57 (16/23)	—	0.7894
CNN+Cross-Lead Attention	69.57 (16/23)	—	0.7761
ResNet-1D + 手工特征	78.26 (18/23)	76.10	0.7214
Transformer + 手工特征	78.26 (18/23)	75.80	0.7308
MCL-Net (未调参)	86.96 (20/23)	85.21	0.6497

从表7可以看出，MCL-Net（未调参）在所有评价指标上均优于基线方法：

(1) **与传统机器学习方法的对比**：MCL-Net（未调参）比最优的传统机器学习方法（XGBoost, 82.61%）提升约 4.35 个百分点。这一结果表明，仅依赖手工特征的传统方法虽然能够取得一定分类性能，但受限于人工设计特征的表达能力，难以充分挖掘 ECG 波形中的深层信息。

(2) **与深度基线模型的对比**：MCL-Net（未调参）相比最优的深度基线模型（ResNet-1D + 手工特征、Transformer + 手工特征，准确率均为 78.26%）提升约 8.70 个百分点，EDE 从 0.7214 降至 0.6497。这表明，在相同输入信息条件下，MCL-Net 的跨导联注意力模块能够更有效地利用导联之间的空间结构信息，从而提升定位精度。

6.3 消融实验

为验证 MCL-Net 中各模块的贡献，本文在未进行超参数调优的默认配置下 ($\text{dropout}=0.35$, $\lambda_{\text{reg}}=0.9$) 进行了系统的消融实验。需要说明的是，本节报告的消融实验结果旨在展示各模块的相对贡献趋势，而非模型的最终最优性能。模型的最终调优配置及对应性能详见第6.1节。

实验结果如表8所示。

表 8: 消融实验结果（未调优配置）

模型配置	Accuracy (%)	Weighted F1 (%)	EDE
A: 仅波形分支	60.87 (14/23)	51.19	0.7815
B: 波形 + 跨导联注意力	65.22 (15/23)	55.24	0.6196
C: 波形 + 手工特征	69.57 (16/23)	58.55	0.6647
D: 完整 MCL-Net	86.96 (20/23)	85.21	0.6497
调优后 MCL-Net	91.30 (21/23)	88.47	0.5441

从表8可以得出以下结论：

(1) **跨导联注意力模块的贡献**：对比配置 A 与 B，加入跨导联注意力后，分类准确率提升 4.35 个百分点，加权 F1 提升 4.05 个百分点，EDE 降低 0.1619。

(2) **手工特征分支的贡献**：对比配置 A 与 C，加入 18 维手工特征后，分类准确率提升 8.70 个百分点，加权 F1 提升 7.36 个百分点，EDE 降低 0.1168。手工特征分支的单独贡献高于跨导联注意力模块，说明临床心电图学知识在 PVC 起源定位中具有重要作用。

(3) **两个模块的协同效应**：对比配置 A 与 D，同时加入跨导联注意力和手工特征后，分类准确率提升 26.09 个百分点，加权 F1 提升 34.02 个百分点，EDE 降低 0.1318。这表明跨导联注意力学习的空间依赖模式与手工特征编码的临床知识之间存在正向协同效应。需要指出的是，在未调优配置下，配置 D 的 EDE 略高于配置 B，说明回归与分类目标之间存在一定权衡；经超参数调优后，EDE 进一步降至 0.5441。

(4) **超参数调优的影响**：对比配置 D（未调优）与调优后的 MCL-Net，超参数调优使分类准确率从 86.96% 提升至 91.30%，加权 F1 从 85.21% 提升至 88.47%，EDE 从 0.6497 降至 0.5441。

第 4.5.3 节已指出，回归损失权重 λ_{reg} 的取值直接决定分类与回归两个任务在训练过程中的相对权重，因此需要实验验证其最优值。为此，本文在固定 dropout=0.5、L2=0.0005 的条件下，对 $\lambda_{\text{reg}} \in \{0.8, 0.9, 1.0\}$ 进行了系统对比实验，结果如表9所示。

表 9: 不同回归损失权重 λ_{reg} 对模型性能的影响（dropout=0.5, L2=0.0005）

λ_{reg}	Accuracy (%)	EDE
0.8	65.22 (15/23)	0.6018
0.9	73.91 (17/23)	0.5890
1.0	73.91 (17/23)	0.6510

从表9可以看出，当 $\lambda_{\text{reg}} = 0.9$ 时，模型在分类准确率和定位精度上均取得最优或并列最优的表现。当 λ_{reg} 取 0.8 时，回归任务权重偏小，坐标预测未能得到充分优化，EDE 偏高；当 λ_{reg} 取 1.0 时，回归梯度主导训练过程，分类任务的判别能力被压制，虽然 Acc 与 0.9 持平，但 EDE 反而上升至 0.6510。只有在 $\lambda_{\text{reg}} = 0.9$ 时，分类与回归两个任务取得了最佳平衡。因此，本文所有实验均采用 $\lambda_{\text{reg}} = 0.9$ 。

6.4 后处理策略的影响

为评估各后处理策略对最终定位精度的贡献，本文逐步应用后处理方法并记录每一步的 EDE 变化，结果如表10所示。

表 10: 后处理策略对 EDE 的影响

处理步骤	EDE	说明
原始预测	0.5441	全部测试样本
+ 区域中心回退	0.4853	将回归坐标替换为对应分类区域的训练集中心
+ 严格评估	0.3090	仅对分类正确的样本计算 EDE

从表10可以看出，两种处理方法对 EDE 的降低有不同的作用机制：区域中心回退属于**坐标校正**，即对于回归结果与分类结果有出入的样本，将其回归预测坐标替换为其对应分类区域的训练集中心，使 EDE 从 0.5441 降至 0.4853；严格评估属于**样本筛选**，通过剔除分类错误的样本来统一评价口径，使 EDE 从 0.4853 进一步降至 0.3090。需要说明的是，严格评估带来的 EDE 下降并非模型坐标预测精度的提升，而是评估口径变化的结果，反映了模型在分类正确前提下的定位能力，而非所有样本的整体定位能力。区域中心回退也并非总是优于模型原始回归结果：当分类正确时，区域中心不一定比模型回归更准确；当分类错误时，回退到错误区域中心可能带来更大误差。因此，区域中心回退仅作为一种保守后处理策略进行分析，不作为模型的主要定位输出。

6.5 结果讨论

MCL-Net 在 PVC 消融靶点定位任务上取得的性能可归因于以下几个方面的协同作用：

第一，**跨导联注意力机制有助于利用 12 导联 ECG 的空间结构信息**。与传统的将导联独立处理的方法相比，跨导联注意力能够动态地学习导联之间的依赖权重，使模型更准确地捕捉不同导联之间的关联模式。从注意力热图可以看出，模型在预测 RVOT 起源的 PVC 时倾向于关注 V1、V2 与 aVR 导联之间的依赖关系，而在预测 LVOT 起源时则侧重于 aVL 和 V1 导联的组合，这与临床心电图学的定位规则方向一致。

第二，**手工特征与深度特征的融合策略发挥了各自的优势**。深度特征能够自动学习 ECG 波形中的抽象表示，而手工特征提供了明确的临床知识，两者的结合使模型具备了更强的表征能力。消融实验表明，手工特征分支的单独贡献大于跨导联注意力模块，验证了临床先验知识在数据有限场景下的重要性。

第三，**多任务学习框架使分类和回归任务能够相互促进**。分类任务帮助模型定位解剖区域，回归任务细化三维坐标，共享的特征表示在两个任务之间传递了有价值的信息。

第四，**后处理与评估策略共同影响定位结果的呈现方式**。区域中心回退通过坐标校正降低离群点误差，严格评估通过样本筛选统一评价口径。经过处理后，EDE 从 0.5441 降至 0.4853，再降至 0.3090。需要说明的是，该数值包含了坐标校正与样本筛选两种不

同机制的共同作用，其中坐标校正直接改变了模型的定位输出，样本筛选则使评价指标更聚焦于分类正确前提下的定位能力。

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

7 讨论

7.1 临床意义

本研究提出的 MCL-Net 模型在 PVC 消融靶点定位任务上取得了分类准确率 91.30%、加权 F1 分数 88.47% 的性能；在全部测试样本上归一化 EDE 为 0.5441，在严格评估下 EDE 为 0.3090。这一结果在临床层面具有以下意义。

第一，**为术前无创定位提供了自动化、智能化的工具**。传统定位依赖术中有创心内电生理标测，而即使经验丰富的心内科医生可以通过 ECG 大致判断起源，其判断也往往受主观经验影响，难以标准化和量化。MCL-Net 能够自动、智能地从常规 12 导联 ECG 中给出靶点的解剖区域分类和三维坐标预测，提供客观、可重复的定位结果，从而降低对医生个人经验的依赖，为术前规划提供辅助决策支持。

第二，**多任务输出同时满足定性和定量定位需求**。分类头输出 7 个解剖区域的概率分布，帮助医生快速锁定病灶所在的大致区域；回归头输出三维空间坐标，为导管操作提供更精细的空间参考。这种双输出设计使模型既能用于术前筛查，又能用于手术规划。

第三，**模型的可解释性支持临床决策**。跨导联注意力模块的可视化分析表明，模型在预测不同起源部位的 PVC 时，关注了与临床诊断规则方向一致的导联组合。例如，预测 RVOT 起源时模型倾向于关注 V1-V2 与 aVR 导联的依赖关系，预测 LVOT 起源时侧重于 aVL 和 V1 导联的组合，这与心电图学中流出道定位的经典标准方向一致。这种可解释性有助于增强临床医生对模型预测结果的信任，但注意力权重不等同于因果解释，仍需进一步验证。

7.2 局限性

尽管 MCL-Net 取得了较好的性能，本研究仍存在以下局限。

第一，**样本量有限**。本研究共纳入 273 例临床数据，测试集为 23 例。虽然固定数据划分下三次独立训练在一定程度上验证了模型的稳定性，但较小的测试集规模可能使性能指标的方差偏大，同时相对稀有的类别样本量有限。未来需要在更大规模的临床数据集上进一步验证模型的泛化能力。

第二，**单中心数据来源**。所有数据均来源于中国人民解放军总医院（301 医院），不同医院的 ECG 采集设备、信号处理流程和患者人群可能存在差异。模型在不同中心的数据上是否仍能保持稳定的性能，尚需外部验证。

第三，**坐标系统的标准化程度有限**。本文采用自定义的标准心室模板坐标系，该坐标系与临床常用的 CARTO 系统坐标之间的映射关系需要进一步的标准化研究。不同研究中心可能使用不同的三维标测系统，坐标系统的差异可能影响模型的跨中心迁移能力。

第四，**缺少前瞻性验证**。本研究为回顾性研究，模型的预测精度已在历史数据上得到验证，但尚未在真实临床流程中进行前瞻性评估。模型能否真正帮助医生缩短手术时

间、提高手术成功率，需要通过前瞻性临床试验来验证。

第五，**手工特征提取的鲁棒性**。18 维手工特征的提取依赖 Pan-Tompkins 算法进行 QRS 波检测，该算法在信号质量较差或存在严重噪声的情况下可能失效。本文对检测结果进行了抽样人工复核，但未来可以研究更鲁棒的 QRS 检测方法，或探索完全端到端的手工特征替代方案。

7.3 未来工作

基于上述局限性，未来工作可以从以下几个方向展开：

(1) **多中心外部验证**。收集来自不同医院、不同设备、不同患者人群的 ECG 数据，验证 MCL-Net 的跨中心泛化能力。若性能下降明显，可以考虑引入领域自适应技术来减小不同中心数据之间的分布差异。

(2) **扩大样本量**。与更多医疗机构合作，扩充数据集规模，特别是增加稀有起源部位（如 Summit、AMC）的样本量，以进一步提升模型在这些类别上的分类和回归精度。

(3) **与影像数据融合**。将 ECG 波形与心脏 CT/MRI 影像数据相结合，利用影像提供的解剖结构信息辅助定位。多模态融合可能进一步提升三维坐标预测的精度，尤其是对于心外膜起源的 PVC。

(4) **前瞻性临床验证**。与医院合作，在真实临床场景中使用 MCL-Net 进行术前规划，评估模型对手术时间、X 射线曝光量、手术成功率等临床终点的实际影响。

(5) **模型轻量化与实时部署**。当前模型的参数量和推理速度已适合离线分析，但若希望集成至医院信息系统进行实时预测，需要进一步优化模型的计算效率，例如通过知识蒸馏、模型剪枝或量化等技术减小模型体积。

8 结论

本文提出了一种名为 MCL-Net 的混合深度学习模型，用于从 12 导联 ECG 波形中实现 PVC 消融靶点的端到端定位。模型融合了多导联卷积特征提取、跨导联注意力机制和手工心电图特征，通过多任务学习框架同时输出 7 个解剖区域的分类标签和三维空间坐标。

在 273 例临床数据上，MCL-Net 在测试集上达到 91.30% 的分类准确率、88.47% 的加权 F1 分数；在全部测试样本上的归一化欧氏距离误差为 0.5441，在严格评估下为 0.3090。消融实验表明，跨导联注意力模块和手工特征分支均对模型性能有贡献，两者结合产生了正向协同效应。与 ResNet-1D、Transformer Encoder、CNN+Cross-Lead Attention 等深度基线模型以及随机森林、XGBoost 等传统机器学习方法相比，MCL-Net 在相同调参预算下取得了更优性能。

与既往多聚焦于起源区域分类或单一任务的研究不同，本文提出的 MCL-Net 模型在一个统一的框架内，同时实现了精细的解剖区域分类和定量的三维坐标回归。实验结果表明，该模型通过跨导联注意力机制与手工特征融合，提升了定位的精度与可解释性，并且其注意力模式与临床心电图学的定位规则方向一致。该工作为术前无创靶点定位提供了可供参考的技术方案，有望在临床实践中缩短手术时间、减少辐射暴露，为 PVC 患者带来更安全、更高效的治疗体验。

未来工作将聚焦于开展多中心外部验证、扩大样本量、融合影像数据以及开展前瞻性临床验证，以进一步推动 MCL-Net 向临床应用的转化。

参考文献

- [1] R. H. Anderson, T. J. Mohun, D. E. Spicer, et al. Use of a two-handed model to improve comprehension of ventricular outflow tract anatomy. *BMC Medical Education*, 23:274, 2023.
- [2] T. S. Baman, D. C. Lange, K. J. Ilg, et al. Relationship between burden of premature ventricular complexes and left ventricular function. *Heart Rhythm*, 7(7):865–869, 2010.
- [3] B. Berte, M. Duytschaever, T. De Potter, et al. Lesion size and depth in radiofrequency catheter ablation. *Journal of Cardiovascular Electrophysiology*, 30(10):2097–2105, 2019.
- [4] B. P. Betensky, R. E. Park, F. E. Marchlinski, et al. The v2 transition ratio: A new electrocardiographic criterion for distinguishing left from right ventricular outflow tract tachycardia origin. *Journal of the American College of Cardiology*, 57(22):2255–2262, 2011.
- [5] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
- [6] J. Brugada, A. Nogami, Y. Aizawa, et al. Radiofrequency catheter ablation of ventricular tachycardia in patients without structural heart disease. *Circulation*, 95(11):2603–2609, 1997.
- [7] R. Caruana. Multitask learning. *Machine Learning*, 28(1):41–75, 1997.
- [8] R. Girshick. Fast r-cnn. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1440–1448, 2015.
- [9] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1):65–69, 2019.
- [10] H. He and E. A. Garcia. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9):1263–1284, 2009.
- [11] R. Hoekema, G. J. H. Uijen, and A. van Oosterom. Geometrical aspects of the interindividual variability of multilead ecg recordings. *IEEE Transactions on Biomedical Engineering*, 48(5):551–559, 2001.
- [12] M. E. Josephson. *Josephson's Clinical Cardiac Electrophysiology: Techniques and Interpretations*. Wolters Kluwer Health, 5th edition, 2015.

- [13] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2016.
- [14] S. Kiranyaz, O. Avci, O. Abdeljaber, et al. 1d convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151:107398, 2021.
- [15] X. Li, J. Zhang, Y. Wang, et al. Electrocardiographic characteristics for predicting the origin of premature ventricular contractions. *Frontiers in Cardiovascular Medicine*, 9:889271, 2022.
- [16] I. Loshchilov and F. Hutter. Sgdr: Stochastic gradient descent with warm restarts. In *International Conference on Learning Representations (ICLR)*, 2017.
- [17] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations (ICLR)*, 2019.
- [18] J. Pan and W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, 1985.
- [19] C. T. Pedersen, G. N. Kay, J. Kalman, et al. Prevalence of premature ventricular contractions in a general population. *Heart Rhythm*, 11(11):1980–1987, 2014.
- [20] N. Pilia, S. Schuler, et al. Non-invasive localization of the ventricular excitation origin without patient-specific geometries using deep learning. *Artificial Intelligence in Medicine*, 143:102619, 2023.
- [21] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, et al. Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*, 2017.
- [22] D. Ramachandram and G. W. Taylor. Deep multimodal learning: A survey on recent advances and trends. *IEEE Signal Processing Magazine*, 34(6):96–108, 2017.
- [23] A. H. Ribeiro, M. H. Ribeiro, G. M. M. Paixão, et al. Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nature Communications*, 11:1760, 2020.
- [24] S. Ruder. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*, 2017.
- [25] S. Schuler, N. Pilia, D. Potyagaylo, and A. Loewe. Cobiveco: Consistent biventricular coordinates for precise and intuitive description of position in the heart. *Medical Image Analysis*, 74:102247, 2021.
- [26] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, 2009.

- [27] W. G. Stevenson and U. B. Tedrow. Catheter ablation for ventricular tachycardia. *Circulation*, 139(19):2222–2236, 2019.
- [28] C. Szegedy, V. Vanhoucke, S. Ioffe, et al. Rethinking the inception architecture for computer vision. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, 2016.
- [29] A. Vaswani, N. Shazeer, N. Parmar, et al. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 5998–6008, 2017.
- [30] P. Veličković, G. Cucurull, A. Casanova, et al. Graph attention networks. In *International Conference on Learning Representations (ICLR)*, 2018.
- [31] T. Yang, L. Wang, X. Zhang, et al. Deep learning for outflow tract origin classification from ecg. *Journal of Electrocardiology*, 60:48–54, 2020.
- [32] J. Zhang, X. Liu, Y. Wang, et al. Machine learning for localization of premature ventricular contractions. *Journal of Cardiovascular Electrophysiology*, 32(6):1654–1663, 2021.
- [33] S. Zhou, A. AbdelWahab, J. L. Sapp, et al. Localization of ventricular activation origin from the 12-lead ecg: A comparison of linear regression with non-linear methods of machine learning. *Computers in Biology and Medicine*, 106:144–151, 2019.

致谢

本研究从选题、算法设计到实验实施的整个过程，凝聚了多方面的帮助与支持。在此，我谨向所有给予我帮助的人致以最诚挚的谢意。

一、研究背景与选题来源

本研究源于我在了解心电图相关知识时产生的独立思考。室性早搏是临床最常见的心律失常之一，而心电图特征可以反映其起源位置；更进一步，我在阅读相关文献时注意到，射频消融术的成败在很大程度上取决于术前能否精确定位异位兴奋灶的起源位置，并且目前基于 12 导联 ECG 的自动化定位方法仍存在精度和泛化能力的局限，这激发了我尝试用深度学习方法解决这一临床问题的想法。在与学校信息老师徐帅婧老师多次讨论后，徐老师从研究可行性和创新性角度对选题进行了优化，最终确定了“基于多导联心电图特征工程与跨导联注意力机制的室性早搏消融靶点定位研究”这一课题。

二、指导老师与参赛学生的关系及指导性

徐帅婧老师是我的学校信息老师，也是本研究的指导教师。整个研究过程中，徐老师的指导是完全无偿的，她主动投入了大量时间与精力。徐老师的指导主要集中在宏观方向和论文写作细节两个层面：在选题阶段，她从研究价值、技术路线可行性等方面提出了优化建议；在实验推进过程中，她多次敦促我按时完成实验并记录结果；在论文撰写阶段，她对标题措辞、语言表述、标点符号、图表格式等细节逐字逐句进行了审阅和修改建议。需要特别说明的是，徐老师并未参与具体算法架构设计、数据处理和实验实施等核心工作，这些均由我独立完成。

三、数据获取与外部协助

本研究所使用的 273 例临床数据来源于中国人民解放军总医院（301 医院）。在数据获取过程中，301 医院黄亚博士给予了重要帮助，包括协助我理解数据采集流程、解释临床标签和三维坐标的标注依据、协助审核数据准确性。黄亚博士的帮助是本研究的临床数据基础，在此表示衷心感谢。此外，黄亚博士还对论文初稿中医学相关专业部分提出了审核和修改建议。

四、研究分工详细说明

本研究为个人项目，无其他队员。各环节的分工如下：

1. **选题来源**：由我独立思考后，与徐帅婧老师协商完成。
2. **数据获取**：由黄亚博士协助完成，本人主导数据整理与核对。
3. **数据分析及计算**：包括数据预处理、特征提取、模型训练、结果评估与统计分析，均由本人独立完成。
4. **算法架构设计**：包括多导联卷积模块、跨导联注意力模块、手工特征融合模块的设计及多任务损失函数的确定，均由本人独立完成。

5. **实验设计与实施**：包括固定随机种子 42 的数据划分、三次独立训练、基线模型的搭建、消融实验的设计、超参数网格搜索等，均由本人独立完成。
6. **论文初稿撰写**：由本人独立完成。
7. **论文修改**：由本人根据徐帅婧老师和黄亚博士的修改建议完成。

除上述指导与数据协助外，所有研究成果均由学生独立完成。

五、各环节遇到的困难与解决经过

在研究过程中，我遇到了若干具体困难。在数据层面，原始 ECG 信号长度不一，部分记录存在噪声干扰，我通过统一截取 500 个采样点、Z-score 归一化和裁剪至 $[-1, 1]$ 区间的方式加以解决。在模型层面，12 导联 ECG 的导联间依赖关系难以被传统 CNN 捕捉，我尝试了多种方案后，最终设计了跨导联注意力模块，将 12 个导联视为图节点，通过多头自注意力机制显式建模空间依赖。在论文写作层面，徐帅婧老师在标题措辞、语言表述和标点格式上给予了细致指导，帮助我提升了论文的学术规范性。

六、总结

本研究是在徐帅婧老师的无偿指导下、在黄亚博士的数据支持下，由我独立完成的。从选题到论文初稿，我经历了从临床问题理解、算法设计、代码实现、实验调优到论文撰写的完整科研训练。这一过程不仅提升了我的科研能力，也让我深刻体会到严谨与坚持的重要性。再次向徐帅婧老师、黄亚博士以及所有给予我帮助的人表示衷心感谢！