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

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

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

2026 Yau Hong Science Award 仅用于2026年丘成桐科学奖公示

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，采用 10 次独立重复分层随机划分（种子 100–109，每组训练/验证/测试约 191/59/23），每组划分独立训练一个模型、不做集成。MCL-Net 在 10 组测试集上取得 91.74% ± 4.10% 的分类准确率、90.78% ± 4.81% 的加权 F1 分数；平均欧氏距离误差（mean EDE）为 0.3074 ± 0.0365。上述结果在本单中心测试集上的点估计优于 ResNet-1D、Transformer Encoder、GRU、LSTM 等深度基线及随机森林等传统机器学习方法。消融实验表明，手工特征分支与跨导联注意力模块分别贡献约 32 和 20 个百分点，说明双头共享对分类与回归均有正向迁移。

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，在 10 组独立划分上提示具有潜在临床参考价值，但数据有限，尚不能与有创标测金标准直接等同，仍需多中心外部验证。

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

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

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 handcrafted 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, ten repeated stratified random splits (seeds 100–109, with approximately 191/59/23 training/validation/test cases per split) were used; each split was trained independently without ensembling. MCL-Net achieves a classification accuracy of 91.74% ± 4.10%, a weighted F1 score of 90.78% ± 4.81%, and a mean Euclidean distance error (mean EDE, computed over all test sam-

ples) of 0.3074 ± 0.0365 across the ten test sets. On this single-center test set, the point estimates outperform deep baselines including ResNet-1D, Transformer Encoder, GRU, LSTM, as well as traditional machine learning methods such as Random Forest. Ablation studies show that the handcrafted feature branch and the cross-lead attention module contribute approximately 32 and 20 percentage points, respectively, and that sharing representations between the two heads yields positive transfer for both classification and regression.

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 results across ten independent splits suggest potential clinical reference value, but the data are limited; the method is not yet equivalent to invasive mapping standards and requires multicenter external validation.

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

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

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 17

4.3.3 残差连接与层归一化 17

4.4 手工特征提取与融合模块 17

4.4.1 18 维手工特征的定义 17

4.4.2 特征提取流程 18

4.4.3 特征重要性分析 19

4.4.4 手工特征编码与融合 20

4.5 多任务学习与损失函数 21

4.5.1 多任务学习的动机 21

4.5.2 输出头设计 21

4.5.3 损失函数 22

4.6 本章小结 23

2026 S-T (Qu) High School Science Award
仅用于2026上海市中学科学奖论文公示

5 数据集与实验设置 **24**
5.1 数据集 24
5.2 实验设置 25
5.3 评估策略 25
5.4 基线方法 26
5.5 消融实验设计 26

6 实验结果与分析 **27**
6.1 模型容量分析 27
6.2 10 组独立划分的稳定性 28
6.3 与基线模型的对比 29
6.4 消融实验 30
6.5 预训练嵌入对比 31
6.6 结果讨论 32

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**

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

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或特定临床队列中发生率更高，具体比例因人群与诊断标准而异 [20]。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，并可能进一步诱发恶性心律失常 [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）是可选的重要根治性手段之一 [26]。

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 [11]。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 [22]。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 [31]。

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

2026 Tianyu High School Science Award 仅用于2026年成都市天域高级中学科学奖公示

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 [10]。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 [12]。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。消融实验表明，手工特征分支是分类性能的最大贡献来源。

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。据我们所知，目前公开文献中尚无利用临床 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，采用 10 组独立随机分层划分（种子 100–109），MCL-Net（micro 档，38.3K 参数）在测试集上平均达到 91.74% ± 4.10% 的分类准确率、90.78% ± 4.81% 的加权 F1 分数；平均 EDE 为 0.3074 ± 0.0365，优于 ResNet-1D、Transformer Encoder、GRU、LSTM、CNN+Attention 等深度基线模型以及随机森林等传统机器学习方法。

2026 STYaung High School Science Award
仅用于2026年成都中学科学奖论文公示

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；第 3 节描述问题定义和评价指标；第 4 节详细阐述 MCL-Net 的模型架构；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 节讨论临床意义和局限性；第 8 节总结全文并展望未来工作。

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

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 等较早讨论了无结构性心脏病患者室性心动过速/室早的射频消融与心电图定位思路 [6]；Betensky 等进一步提出 V2 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 [31]。Yang 等人使用卷积神经网络从 ECG 信号中识别流出道起源 [30]。Nakasone 等人利用多中心数据进行了二分类预测 [18]，Kujime 等人利用语义分割进行了三分类探索 [14]。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 [22]。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 [24]。

在 ECG 信号的特征提取方面，1D-CNN 被广泛用于提取时序特征 [13]，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 [28]，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 [12]。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 [29]。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。

2026 Tianyi High School Science Award 仅用于2026上饶市青少年科学创新大赛公示

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

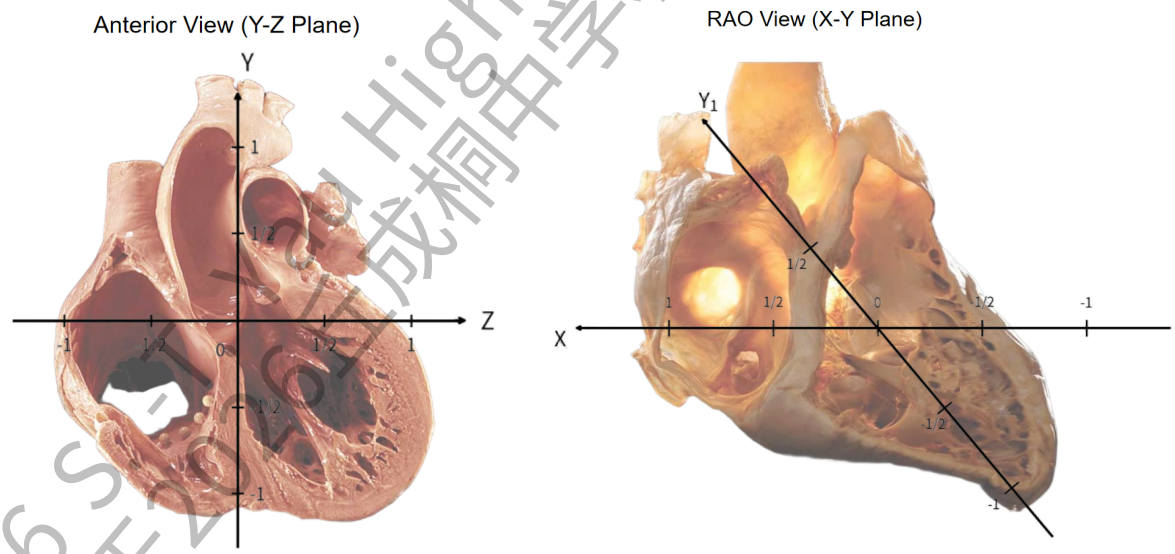
本文将 PVC 消融靶点定位建模为一个多任务学习问题：给定一个 12 导联 ECG 波形 $\boldsymbol{X} \in \mathbb{R}^{12\times T}$（其中 T 为采样点数），模型需要同时输出：（1）解剖区域类别 $y_{\text{cls}} \in \{0, 1, \ldots, 6\}$；（2）三维空间坐标 $\boldsymbol{y}_{\text{reg}} \in [-1, 1]^3$。

三维空间坐标的定义遵循参考临床标准心室解剖模板建立的空间坐标系：

- 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。



(a) 前面观 (Y-Z 平面)　　(b) 右前斜位 (X-Y 平面)

图 1: 基于标准心室模板的三维坐标映射示意图。心脏解剖底图引自 Mori S, Shivkumar K. *Atlas of Cardiac Anatomy, Volume 1: Anatomical Basis of Cardiac Interventions*. 1st ed. 2022. ISBN 978-1-942909-57-6；坐标系定义与标注由本文作者添加

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，

2026 Shi Yu High School Science Award 仅用于2026年成桐中学科学奖论文公示

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。

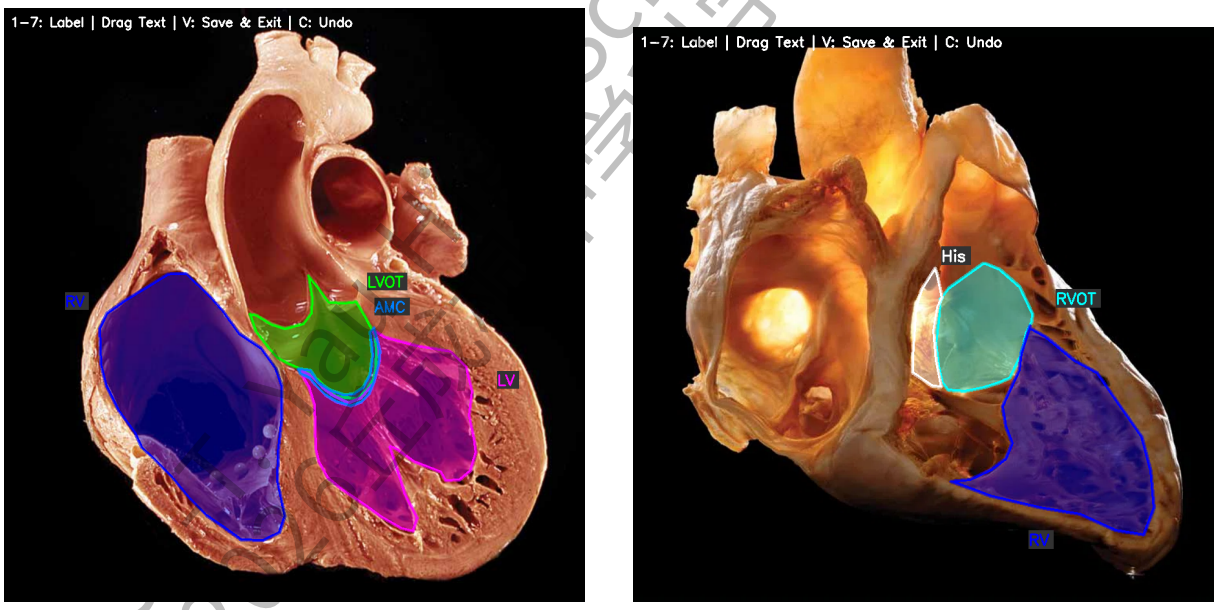
解剖区域的分类标签定义如下：

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。



(a) 前面观 (Anterior View)　　(b) 右前斜位 (RAO View)

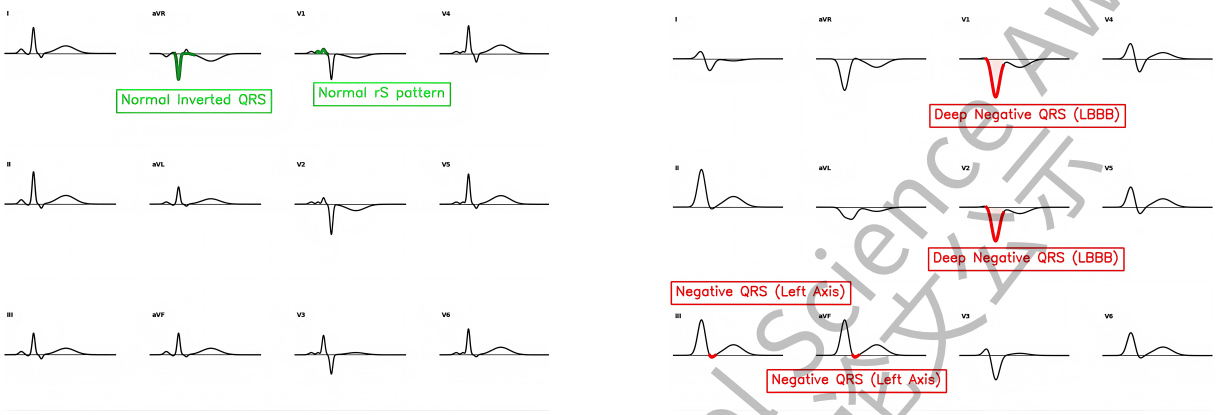
图 2: 7 个解剖区域的心脏空间位置示意图。心脏解剖底图引自 Mori S, Shivkumar K. *Atlas of Cardiac Anatomy, Volume 1: Anatomical Basis of Cardiac Interventions*. 1st ed. 2022. ISBN 978-1-942909-57-6；区域分割、坐标系与标注由本文作者添加

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 区域位于左心室心外膜顶部，解剖位置特殊；AMC 则位于左心室

2026 仅用于 Science Award 学术论文公示

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。

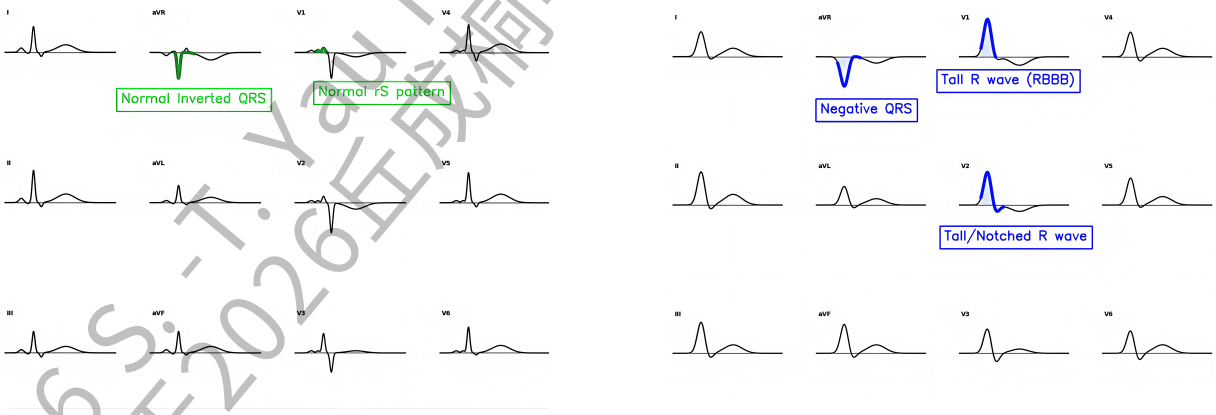
图3至图9以正常窦性心律为基准，将 7 种不同解剖起源的 PVC 波形与正常波形进行并列对比，便于观察各区域 PVC 在 12 导联上的特征性变化。图中彩色高亮标注了该起源类型最具鉴别性的导联特征。同时对每个区域给出了在本文坐标系中的大致坐标范围。



(a) 正常窦性心律 (Normal)　　(b) 右室流出道 (RVOT)

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：$x \in [0.2, 0.7]$，$y \in [0.3, 0.8]$，$z \in [-0.2, 0.4]$。



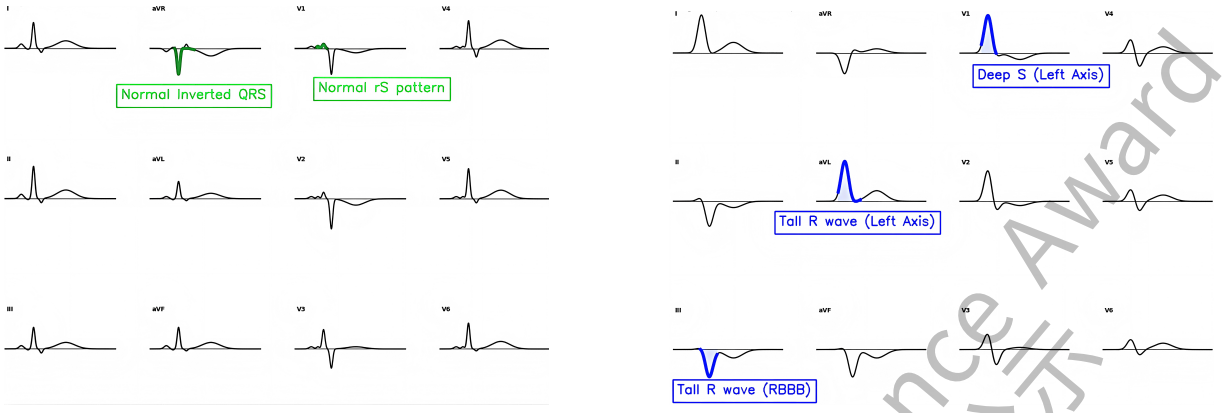
(a) 正常窦性心律 (Normal)　　(b) 左室流出道 (LVOT)

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，

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

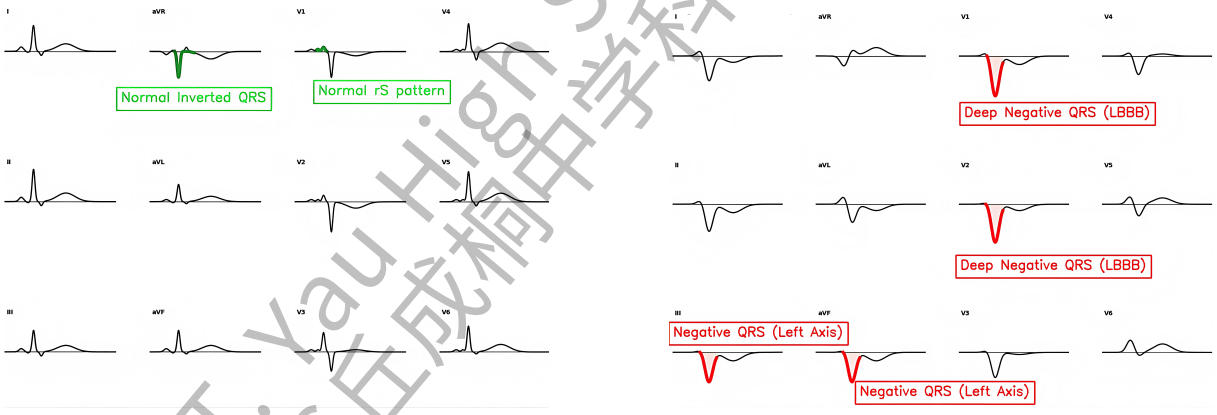
V1 导联 R 波振幅显著增高。在归一化坐标系中，LVOT 区域大致位于：$x \in [-0.6, -0.2]$，$y \in [0.2, 0.7]$，$z \in [-0.1, 0.5]$。



(a) 正常窦性心律 (Normal)　　(b) 左心室 (LV)

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：$x \in [-0.9, -0.3]$，$y \in [-0.6, 0.4]$，$z \in [-0.4, 0.6]$。

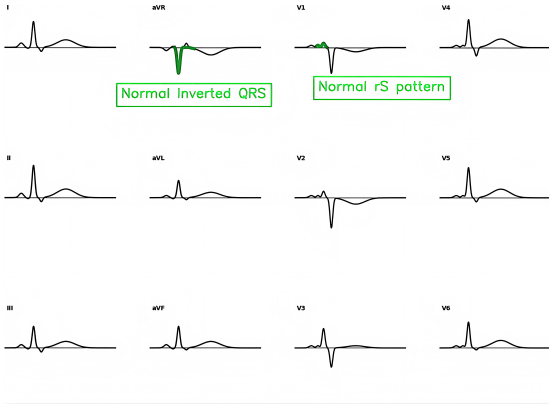


(a) 正常窦性心律 (Normal)　　(b) 右心室 (RV)

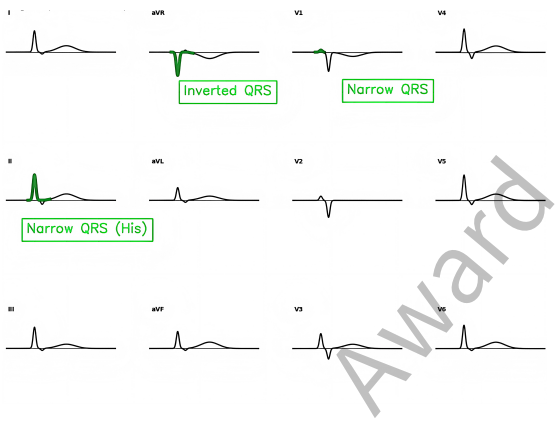
图 6: Normal 与 RV 对比

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：$x \in [0.3, 0.9]$，$y \in [-0.5, 0.5]$，$z \in [-0.5, 0.3]$。

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



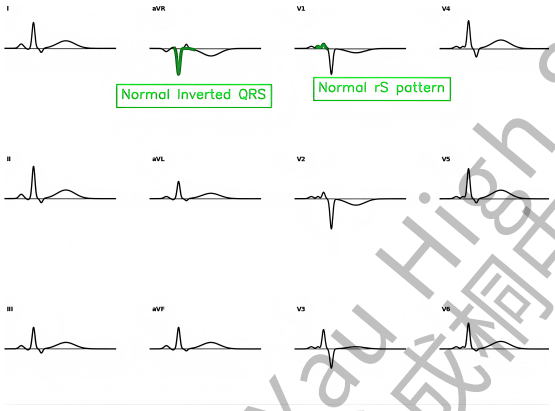
(a) 正常窦性心律 (Normal)



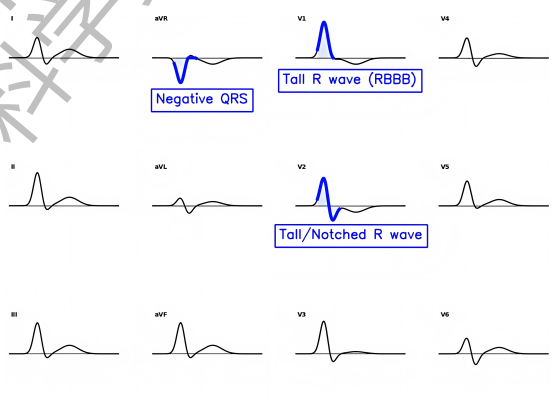
(b) 希氏束区域 (His)

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。His 束起源的 PVC 常表现为相对窄的 QRS（通常 <120–130 ms），但需与分支性起源鉴别；<120 ms 并非 His 起源独有，不能单独作为鉴别阈值。在归一化坐标系中，His 区域位于：$x \in [-0.2, 0.2]$，$y \in [0.4, 0.8]$，$z \in [-0.1, 0.3]$。



(a) 正常窦性心律 (Normal)



(b) 心顶部 (Summit)

图 8: Normal 与 Summit 对比

Summit（心顶部）的 PVC 起源于左心室心外膜顶部区域，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：$x \in [-0.4, -0.1]$，$y \in [0.5, 0.9]$，$z \in [0.1, 0.6]$。

2026 - T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示



(a) 正常窦性心律 (Normal)　　(b) 主动脉二尖瓣连接 (AMC)

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）：

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

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。

加权 F1 分数（Weighted F1 Score）：

$$\mathrm{F1}_{\mathrm{weighted}} = \sum_{c=0}^{6}\frac{n_c}{N}\cdot \mathrm{F1}_c \tag{2}$$

其中 $\mathrm{F1}_c = \frac{2\cdot \mathrm{Precision}_c \cdot \mathrm{Recall}_c}{\mathrm{Precision}_c + \mathrm{Recall}_c}$，$\mathrm{Precision}_c = \frac{\mathrm{TP}_c}{\mathrm{TP}_c+\mathrm{FP}_c}$，$\mathrm{Recall}_c = \frac{\mathrm{TP}_c}{\mathrm{TP}_c+\mathrm{FN}_c}$。$\mathrm{TP}_c$、$\mathrm{FP}_c$、$\mathrm{FN}_c$ 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，$N = \sum_c n_c$

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

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。加权 F1 计算中，当 $\text{Precision}_c + \text{Recall}_c = 0$ 时令 $\text{F1}_c = 0$。

3.3.2 回归评价指标

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

$$\text{EDE} = \frac{1}{N}\sum_{i=1}^{N} \|\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i\|_2 \tag{3}$$

其中 $\hat{\boldsymbol{y}}_i$ 和 $\boldsymbol{y}_i$ 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：$\text{EDE}_{\text{mm}} = \text{EDE} \times \frac{R}{2}$。

本文采用平均欧氏距离误差（mean EDE）作为回归主评价指标。EDE 是心脏电生理定位与 ECG 逆问题研究中被广泛采用的标准度量，用于衡量预测起源点与真实起源点之间的三维直线距离 [32, 21]；相较于测地距离，EDE 不依赖患者个体化的心脏表面网格或解剖分段，与本研究所用标准心室模板坐标系的定义一致。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 $R = 80$ mm，则 0.3074 的归一化 EDE 对应约 12.30 mm 的物理定位误差。需要说明的是，该换算基于模板尺度的近似，不是患者个体化毫米级误差。

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

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 $\boldsymbol{X} \in \mathbb{R}^{12\times T}$（本文中 $T = 500$），MCL-Net 的前向传播过程如下。本文共设计 6 个容量档位（详见 6.1 节表6），除特别说明外，正文以在 10 组独立划分上表现最优的 micro 档（38.3K 参数）为例描述各模块维度；其他档位仅缩放 hidden_dim、num_heads、lead_fc_out、feature_hidden、fusion_hidden 五个超参数，模块结构与连接方式完全一致。

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：

$$\boldsymbol{F}_{\text{lead}} = \text{LeadCNN}(\boldsymbol{X}) \in \mathbb{R}^{12\times 64} \tag{4}$$

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，输出同样维度的矩阵：

$$\boldsymbol{F}_{\text{attn}} = \text{CrossAttn}(\boldsymbol{F}_{\text{lead}}) \in \mathbb{R}^{12\times 64} \tag{5}$$

为了在降低特征维度的同时保留关键信息，我们在注意力模块后引入了一层线性映射层（Linear Projection），将各导联的 64 维特征降维至 16 维。随后，将 12 个导联的特征展平为深度特征向量：

$$\boldsymbol{f}_{\text{deep}} = \text{Flatten}(\text{Linear}(\boldsymbol{F}_{\text{attn}})) \in \mathbb{R}^{12\times 16} = \mathbb{R}^{192} \tag{6}$$

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

$$\boldsymbol{f}_{\text{fused}} = \text{Fusion}\left([\boldsymbol{f}_{\text{deep}}; \boldsymbol{f}_{\text{hand}}]\right) \in \mathbb{R}^{64}, \quad \boldsymbol{f}_{\text{hand}} = \text{MLP}(\boldsymbol{x}_{\text{hand}}) \in \mathbb{R}^{64} \tag{7}$$

其中 $[\cdot;\cdot]$ 表示向量拼接操作。融合层输入为 $192 + 64 = 256$ 维，经 $256 \to 128 \to 64$ 两层全连接后得到共享表示。

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

最后，融合特征分别经分类头和回归头产生输出：

$$\begin{aligned} \boldsymbol{y}_{\text{cls}} &= \text{Softmax}(\boldsymbol{W}_{\text{cls}}\boldsymbol{f}_{\text{fused}} + \boldsymbol{b}_{\text{cls}}) \in \mathbb{R}^7 \\ \boldsymbol{y}_{\text{reg}} &= \tanh(\boldsymbol{W}_{\text{reg}}\boldsymbol{f}_{\text{fused}} + \boldsymbol{b}_{\text{reg}}) \in [-1,1]^3 \end{aligned} \tag{8}$$

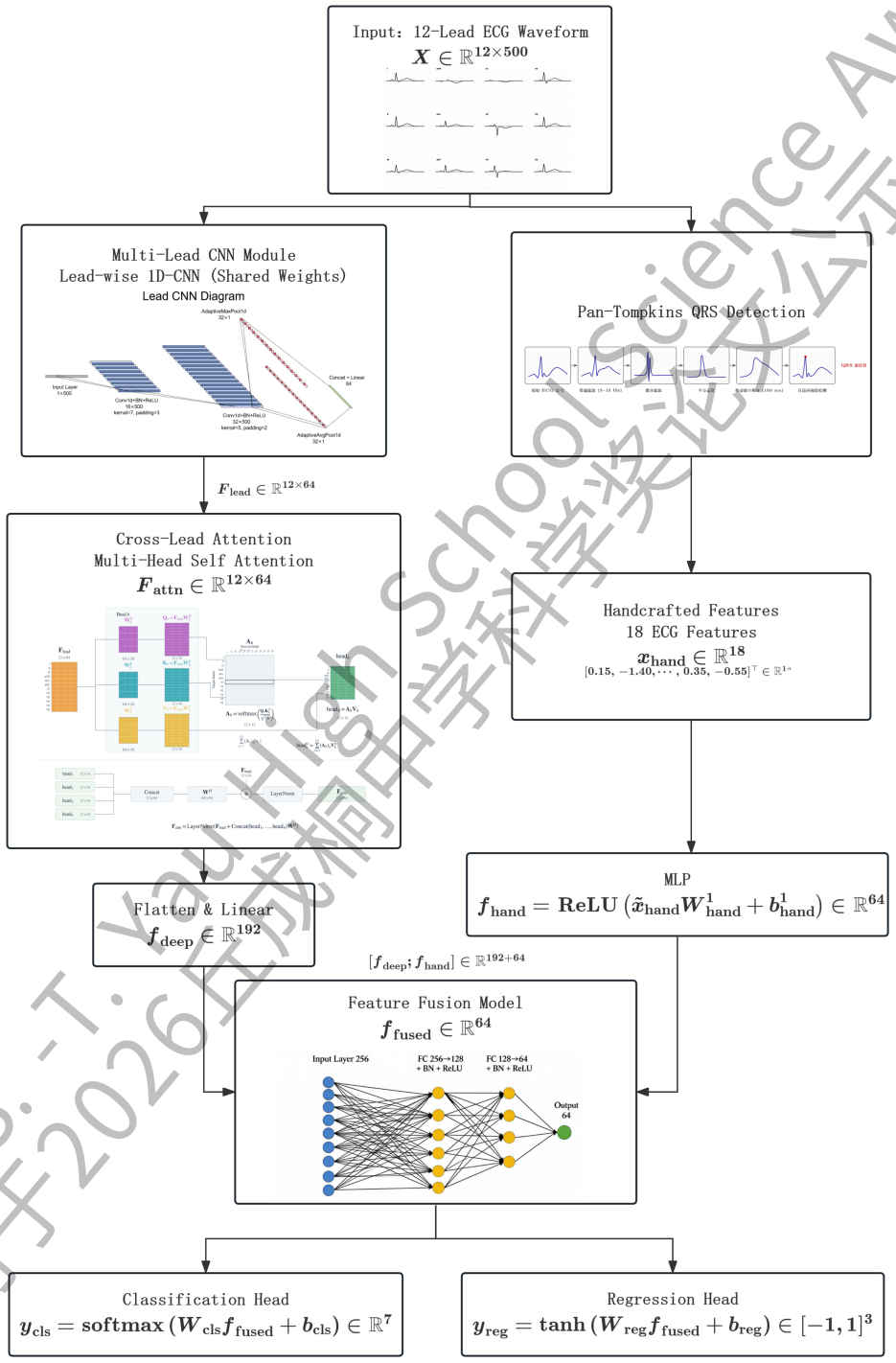


图 10: MCL-Net 整体架构图

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

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 $\boldsymbol{x}^{(i)} \in \mathbb{R}^{T}$ $(i = 1, 2, \ldots, 12)$，LeadCNN 的计算过程为：

$$\boldsymbol{h}^{(i)} = \mathrm{LeadCNN}(\boldsymbol{x}^{(i)}) \tag{9}$$

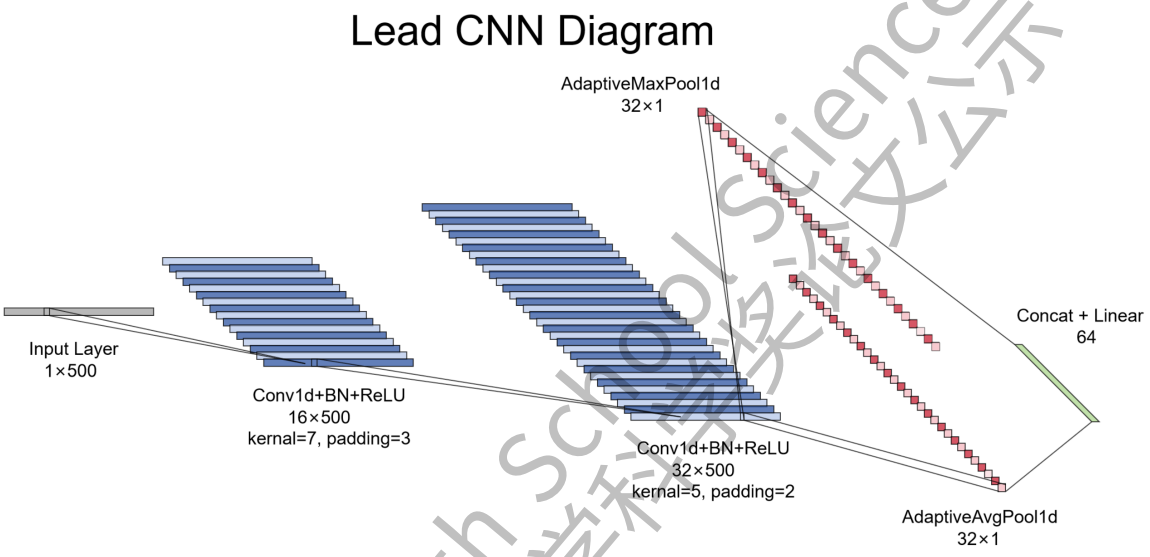


图 11: LeadCNN 网络结构示意图

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 $\boldsymbol{F}_{\mathrm{lead}} \in \mathbb{R}^{12\times 64}$ 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 [28]。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 $\boldsymbol{F}_{\mathrm{lead}} \in \mathbb{R}^{12\times 64}$，其第 i 行 $\boldsymbol{F}_{\mathrm{lead}}^{(i)}$ 表示第 i 个

2026 S.-T. Yau High School Science Award 仅用于2026江苏省中学科学奖论文公示

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。

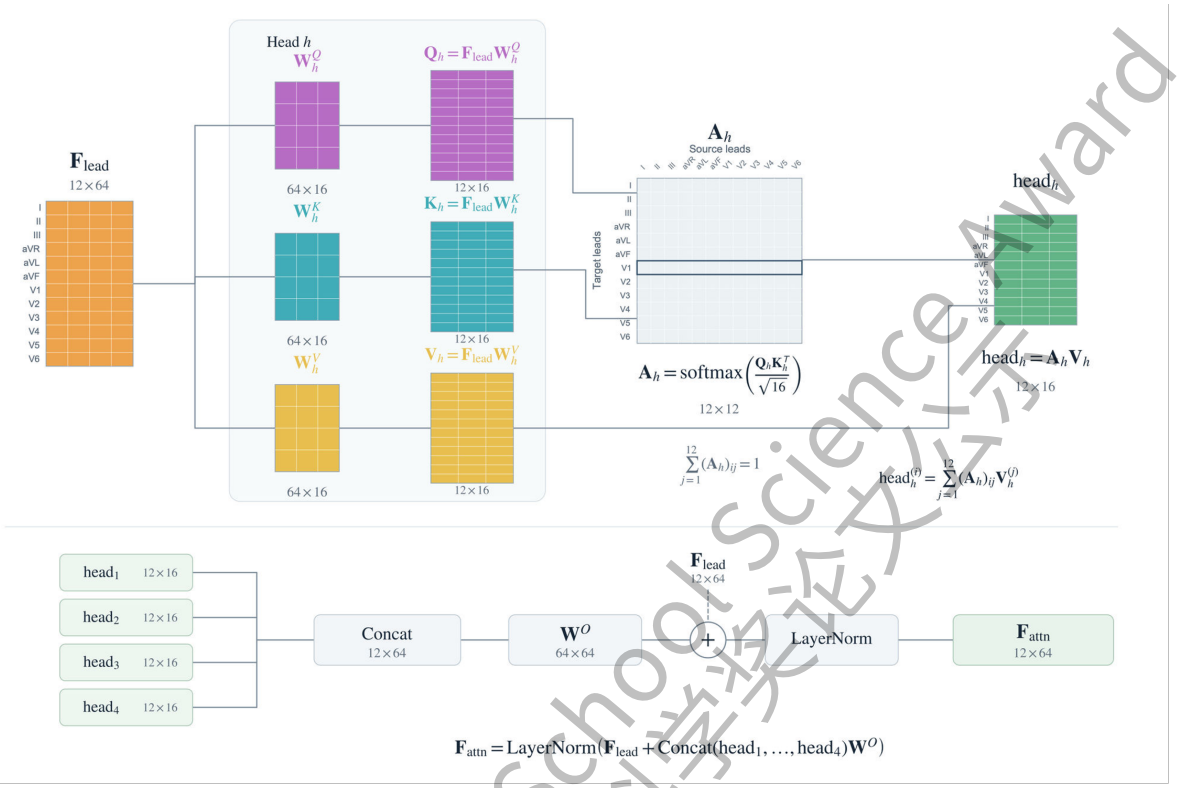


图 12: 跨导联注意力机制（Cross-Lead Attention）详细结构图。左侧为导联特征输入，中间通过 QKV 权重投影与多头自注意力计算导联间的依赖权重，右侧输出更新后的导联特征并经过残差连接与层归一化。

4.3.1 缩放点积注意力的数学形式

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

$$\boldsymbol{Q}_h = \boldsymbol{F}_{\text{lead}} \boldsymbol{W}_h^Q, \quad \boldsymbol{K}_h = \boldsymbol{F}_{\text{lead}} \boldsymbol{W}_h^K, \quad \boldsymbol{V}_h = \boldsymbol{F}_{\text{lead}} \boldsymbol{W}_h^V \tag{10}$$

其中 $\boldsymbol{W}_h^Q, \boldsymbol{W}_h^K, \boldsymbol{W}_h^V \in \mathbb{R}^{64 \times d_h}$ 为可训练投影矩阵，$d_h = 64/H = 32$ 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：

$$\boldsymbol{A}_h = \text{softmax}\left(\frac{\boldsymbol{Q}_h \boldsymbol{K}_h^{\top}}{\sqrt{d_h}}\right) \in \mathbb{R}^{12 \times 12} \tag{11}$$

其中 $\boldsymbol{Q}_h \boldsymbol{K}_h^{\top}$ 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 [28]。经过 softmax 归一化后，注意力分数满足 $\sum_{j=1}^{12} (\boldsymbol{A}_h)_{i,j} = 1$。

2026 T. Yau Hin School Science Award 仅用于2026江成利中学学术论文公示

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：

$$\mathrm{head}_h = \boldsymbol{A}_h \boldsymbol{V}_h \in \mathbb{R}^{12\times d_h} \tag{12}$$

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

$$\mathrm{head}_h^{(i)} = \sum_{j=1}^{12} (\boldsymbol{A}_h)_{i,j}\, \boldsymbol{V}_h^{(j)} \tag{13}$$

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：

$$\mathrm{MultiHead}(\boldsymbol{F}_{\mathrm{lead}}) = \mathrm{Concat}(\mathrm{head}_1, \ldots, \mathrm{head}_H)\, \boldsymbol{W}^O \in \mathbb{R}^{12\times 64} \tag{14}$$

其中 $\boldsymbol{W}^O \in \mathbb{R}^{64\times 64}$ 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：

$$\boldsymbol{F}_{\mathrm{attn}} = \mathrm{LayerNorm}(\boldsymbol{F}_{\mathrm{lead}} + \mathrm{MultiHead}(\boldsymbol{F}_{\mathrm{lead}})) \in \mathbb{R}^{12\times 64} \tag{15}$$

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。

4.4 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。

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

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置高低。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 $\boldsymbol{X} \in \mathbb{R}^{12\times T}$，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 [19]。

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，包含带通滤波、微分滤波、平方运算、移动窗口积分以及自适应阈值判决五个级联的信号处理阶段。

2026 Tsyau High School Science Award
仅用于2026年成都市学生科学奖公示

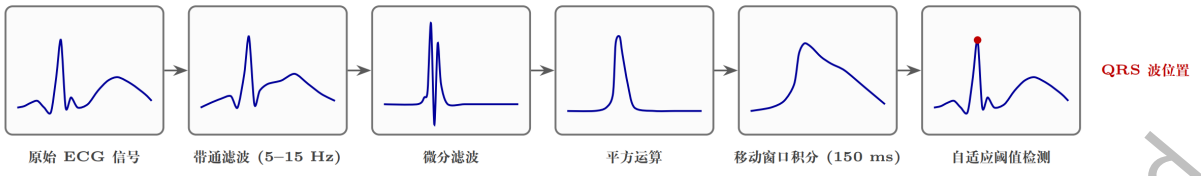


图 13: Pan-Tompkins 算法 QRS 波检测流程。图中展示了信号从原始 ECG 到最终 QRS 波位置标注的信号形态变化

经过上述步骤，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 $\boldsymbol{x}_{\text{hand}} \in \mathbb{R}^{18}$。

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。

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，与流出道起源常表现为下壁导联高 R 波的临床认识一致；其具体阈值不宜直接外推，需在独立数据中验证。III/II R 波幅值比排名第二，提示下壁电轴偏移方向对流出道起源鉴别有贡献。连

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

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，也在一定程度上映证了在回归任务中需要连续变量的信息。

4.4.4 手工特征编码与融合

得到 18 维手工特征向量 $\boldsymbol{x}_{\text{hand}} \in \mathbb{R}^{18}$ 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：

$$\boldsymbol{f}_{\text{hand}} = \text{ReLU}(\boldsymbol{x}_{\text{hand}}\boldsymbol{W}^1_{\text{hand}} + \boldsymbol{b}^1_{\text{hand}}) \in \mathbb{R}^{64} \tag{16}$$

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

将并行提取的手工特征编码与跨导联注意力模块输出的深度特征 $\boldsymbol{f}_{\text{deep}} \in \mathbb{R}^{192}$ 拼接，得到 256 维融合向量：

$$\boldsymbol{f}_{\text{concat}} = [\boldsymbol{f}_{\text{deep}}; \boldsymbol{f}_{\text{hand}}] \in \mathbb{R}^{256} \tag{17}$$

拼接操作是实现多模态特征融合最简单有效的方式之一 [23]，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：

$$\boldsymbol{f}_{\text{fused}} = \text{ReLU}(\text{BN}(\boldsymbol{f}_{\text{concat}}\boldsymbol{W}^1_{\text{fuse}} + \boldsymbol{b}^1_{\text{fuse}}))\boldsymbol{W}^2_{\text{fuse}} + \boldsymbol{b}^2_{\text{fuse}} \in \mathbb{R}^{64} \tag{18}$$

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	192 + 64	向量拼接（Concat）	256
融合层第 1 层	256	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（$256 \to 128 \to 64$），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。

2026 Siyau High School Science Award 仅用于2026年度桐中学科学奖论文公示

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 [25]。

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 $\boldsymbol{f}_{\text{deep}} \in \mathbb{R}^{192}$ 与手工特征编码 $\boldsymbol{f}_{\text{hand}} \in \mathbb{R}^{64}$ 拼接后，经融合层得到共享表示：

$$\boldsymbol{f}_{\text{fused}} = \text{Fusion}([\boldsymbol{f}_{\text{deep}}; \boldsymbol{f}_{\text{hand}}]) \in \mathbb{R}^{64} \tag{19}$$

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：

$$\boldsymbol{y}_{\text{cls}} = \text{softmax}(\boldsymbol{f}_{\text{fused}} \boldsymbol{W}_{\text{cls}} + \boldsymbol{b}_{\text{cls}}) \in \mathbb{R}^{7} \tag{20}$$

其中 $\boldsymbol{W}_{\text{cls}} \in \mathbb{R}^{64\times 7}$ 和 $\boldsymbol{b}_{\text{cls}} \in \mathbb{R}^{7}$ 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：

$$\boldsymbol{y}_{\text{reg}} = \tanh(\boldsymbol{f}_{\text{fused}} \boldsymbol{W}_{\text{reg}} + \boldsymbol{b}_{\text{reg}}) \in [-1, 1]^{3} \tag{21}$$

其中 $\boldsymbol{W}_{\text{reg}} \in \mathbb{R}^{64\times 3}$ 和 $\boldsymbol{b}_{\text{reg}} \in \mathbb{R}^{3}$ 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 $[-1, 1]$。

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

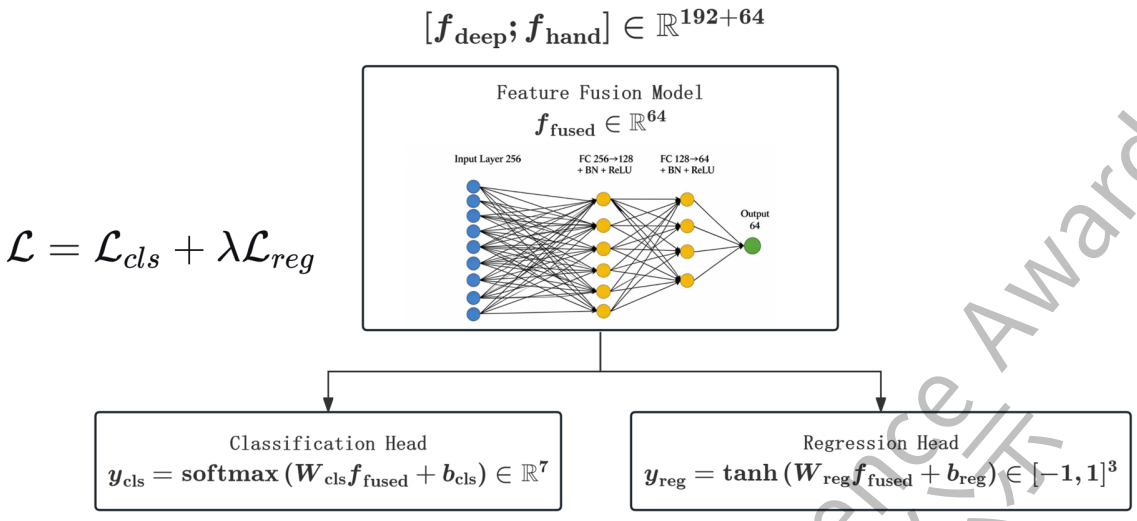


图 14: 多任务输出与特征融合结构图。融合后的 256 维特征经过全连接层降维至 64 维后，分别输入分类头与回归头，整体通过联合损失函数进行优化。

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：

$$\mathcal{L} = \mathcal{L}_{\text{cls}} + \lambda_{\text{reg}} \cdot \mathcal{L}_{\text{reg}} + \lambda_{\text{l2}} \cdot \|\boldsymbol{\Theta}_{\text{head}}\|_2^2 \tag{22}$$

其中 λ_{reg} 为回归损失权重系数，λ_{l2} 为输出头 L2 正则化系数，$\boldsymbol{\Theta}_{\text{head}}$ 为分类头与回归头的可训练参数。优化器 AdamW 的 weight_decay 作用于全部可训练参数，损失中的 λ_{l2} 仅作用于输出头参数，二者作用范围不重叠。λ_{reg} 经预实验确定为 0.9。

分类损失采用标签平滑的交叉熵损失：

$$\mathcal{L}_{\text{cls}} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=0}^{6} \tilde{y}_{i,c} \log(\hat{y}_{i,c}) \tag{23}$$

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 [27]。

回归损失采用 Smooth L1 损失函数：

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^{N} \sum_{d=1}^{3} \text{smooth}_{L1}(\hat{y}_{i,d} - y_{i,d}) \tag{24}$$

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

其中：

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2/\beta, & \text{if } |x| < \beta \\ |x| - 0.5\beta, & \text{otherwise} \end{cases} \tag{25}$$

$\beta = 0.5$ 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}。

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。

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

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 个 PVC 样本，来自 273 例患者（同一患者仅纳入一个样本，因此患者数等于样本数），所有数据均来源于合作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，以 R 峰为中心，取 R 峰前 250 个采样点、后 250 个采样点；若含端点共 501 点，则截去最后一点，最终构成 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按预先定义的心室模板分区—坐标映射表映射到标准坐标系，并由专业人员编辑审核，归一化至 $[-1, 1]$ 区间。

7 个解剖区域的真实边界为三维曲面而非规则立方体。在实际操作中，我们使用解剖模板上的轮廓点集来描述曲面边界，并依据心脏解剖常识（如 RVOT 与 LVOT 位于心室基底部两侧、His 束靠近三尖瓣环等）进行区域划分与靶点归类。

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。

数据集采用 10 组独立随机分层划分（种子 100–109），每组按 7:2:1 划分训练集、验证集和测试集，分层采样保证 7 个解剖区域在每组的三部分中均有分布。每组根据相应的种子划分，独立训练一个模型；测试集规模为 23 例。所有实验结果均报告 10 组划分的均值 ± 标准差。

7 个解剖区域的样本数如表5和图15所示。每组划分的训练/验证/测试总数分别约为 191/59/23；各类别在每组中的具体数量因分层随机略有差异。

表 5: 7 个解剖区域样本数

类别编号	解剖区域	总样本数
0	RVOT	76
1	LVOT	82
2	LV	28
3	RV	25
4	His	17
5	Summit	23
6	AMC	22
合计	–	**273**

2026 S. T. Yau High School Science Award
仅用于2026正成精品学术论文公示

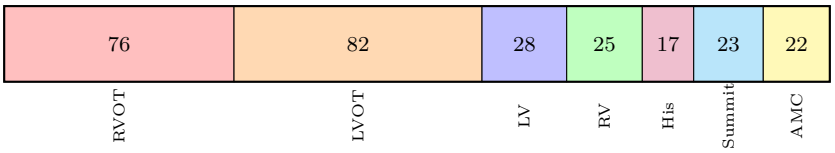


图 15: 7 个解剖区域样本数的一行堆叠可视化（合计 273 例）

5.2 实验设置

所有实验均在 PyTorch 框架下实现，运行于纯 CPU 环境（12th Gen Intel Core i7-1255U），无需依赖高成本 GPU 运算，体现了模型极高的轻量化特性与部署潜力。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^{-3}，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。

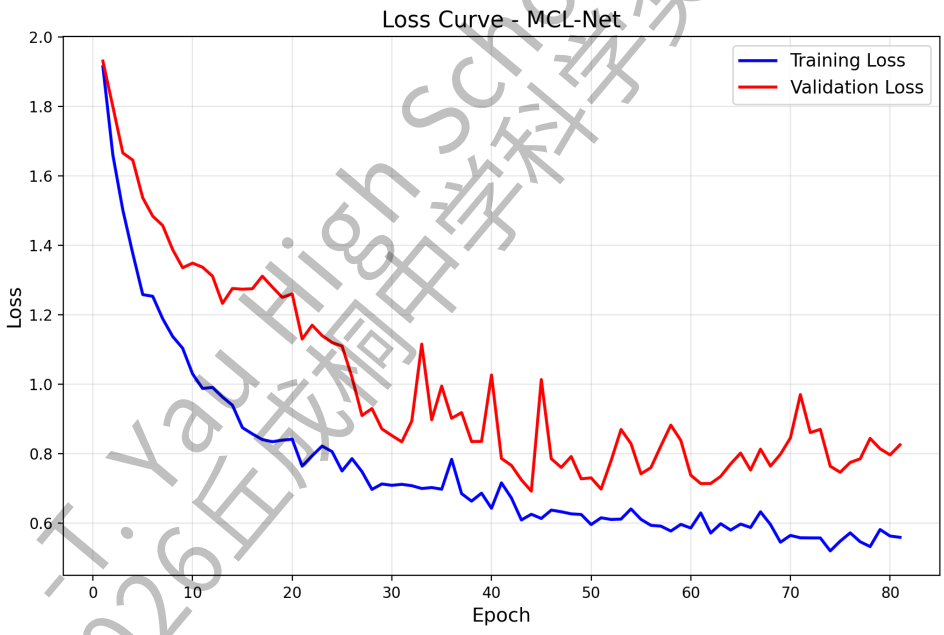


图 16: 模型训练与验证损失收敛曲线

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 10 组独立划分的均值 ± 标准差。由于测试集样本量较小（约 23 例），本文在报告百分数的同时给出正确样本数或均值口径，避免仅依赖四舍五入后的数值。

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

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 $R = 80$ mm，则 0.3074 的归一化 EDE 对应约 12.30 mm 的物理定位误差；该换算基于模板尺度近似，不是患者个体化毫米级误差。

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、深度基线使用相同训练轮数和早停策略；传统机器学习使用相同调参预算。

(1)**传统机器学习方法**：包括随机森林。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，使用 500 棵决策树。

(2)**深度基线模型**：包括 ResNet-1D、Transformer Encoder、GRU、LSTM 和 CNN+Cross-Lead Attention。ResNet-1D 采用三层残差块处理 12 导联波形，参数量约 1.2M；Transformer Encoder 将 12 导联波形展平为序列后输入 2 层标准 Transformer 编码器（$d_{\text{model}} = 128$，4 头），参数量约 2.8M；GRU 采用单层门控循环单元处理序列，参数量约 65.2K；LSTM 采用单层长短期记忆网络，参数量约 73.7K；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，参数量约 1.8M。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。

5.5 消融实验设计

为验证 MCL-Net 中各模块的贡献，本文在 micro 架构（38.3K 参数）下进行了系统的消融实验。所有消融实验均在 10 组独立划分上训练和评估，报告均值 ± 标准差。

具体地，本文设计了以下五种消融配置：（1）**完整模型**：包含多导联卷积、跨导联注意力、手工特征融合和双任务输出。（2）**no_attention**：将跨导联注意力模块替换为恒等映射，保留手工特征和双任务输出。（3）**no_handcrafted**：将手工特征输入置零，保留跨导联注意力和双任务输出。（4）**cls_only**：仅保留分类头，移除回归头。（5）**reg_only**：仅保留回归头，移除分类头。

2026 S. Tyan High School Science Award
仅用于2026全国高中学科竞赛公示

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。首先呈现模型容量分析，然后报告 10 组独立划分的稳定性，随后与多种基线方法进行对比，最后通过消融实验验证各模块的贡献。

6.1 模型容量分析

为确定小样本条件下的最优模型容量，本文在 micro 架构的基础上缩放隐藏维度、注意力头数和各层宽度，构造了 6 个容量档位：pico（18.2K）、nano（30.3K）、micro（38.3K）、small（约 65K）、base（约 162K）、large（约 314K），参数量跨度约 17 倍。所有档位使用相同的 10 组独立划分和固定超参数（dropout=0.5，lr=10^{-3}，weight decay=10^{-3}，batch=16，epochs=120，early stop patience=30，$\lambda_{\text{reg}} = 0.9$，label smoothing=0.1，SmoothL1Loss($\beta = 0.5$)，AdamW + CosineAnnealing）。各档位的具体超参数与参数量如表 6 所示，测试性能如表 7 所示。

表 6: 6 个容量档位的超参数与参数量

档位	H-dim	Heads	Lead-out	Feat-h	Fusion-h	参数量
pico	8	2	16	32	64	18.2K
nano	8	2	16	64	128	30.3K
micro	**16**	**2**	**64**	**64**	**128**	**38.3K**
small	32	2	64	64	128	∼65K
base	64	4	128	64	128	∼162K
large	96	4	192	64	128	∼314K

注：H-dim 为注意力模块内部维度（hidden_dim），Heads 为注意力头数（num_heads），Lead-out 为 LeadCNN 线性层输出维度（lead_fc_out），Feat-h 为手工特征编码维度（feature_hidden），Fusion-h 为融合层中间层维度（fusion_hidden）。参数量从 18.2K 到 314K，跨度约 17 倍。

表 7: 6 个容量档位在 10 组独立划分上的测试性能（mean ± std）

档位	参数量	Acc (%)	Weighted F1 (%)	EDE
pico	18.2K	66.09 ± 3.25	60.16 ± 4.03	0.6120 ± 0.0978
nano	30.3K	82.61 ± 5.85	76.57 ± 5.84	0.4750 ± 0.0839
micro	**38.3K**	**91.74 ± 4.10**	**90.78 ± 4.81**	**0.3074 ± 0.0365**
small	∼65K	63.04 ± 5.24	58.22 ± 6.75	0.5193 ± 0.0852
base	∼162K	61.74 ± 6.09	55.19 ± 6.62	0.6551 ± 0.0810
large	∼314K	63.04 ± 3.51	56.07 ± 4.14	0.6262 ± 0.0629

从表 7 可以得出三点结论。第一，micro（38.3K）在 Acc、Weighted F1 和 EDE 三项指标上同时达到峰值，是双任务的最优容量点。第二，分类对容量过拟合更敏感：38.3K 之后出现断崖式下降（91.74 → 63.04 → 61.74 → 63.04），65K 以上三档与 pico 的欠拟合水平相当。第三，回归对容量更宽容：small（65K）的 EDE 0.5193 仍优于 pico（0.6120）

2026 Yao High school Science Award 仅用于2026年成都中学科学论文公示

和 base/large（0.6551/0.6262），说明回归需要的是平滑连续映射，不会因容量稍大而立即崩溃。

样本/参数比可以解释这一现象：训练集约 191 例，38.3K 对应约 5.0 样本/参数，65K 对应约 2.9，162K 对应约 1.2，314K 对应约 0.6。样本/参数比降至 2 以下时分类断崖，与“每参数至少 2–10 样本”的经验法则下界一致。

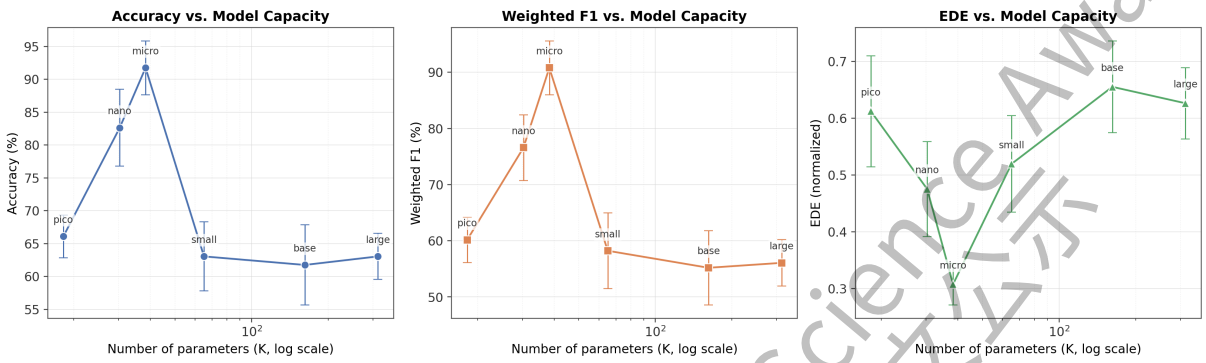


图 17: Accuracy, Weighted F1, and EDE across six capacity levels (pico to large; 18.2K–314K parameters). Mean ± std over ten independent splits. micro (38.3K) peaks simultaneously on all three metrics; above 65K, classification collapses while EDE remains in the 0.52–0.66 range.

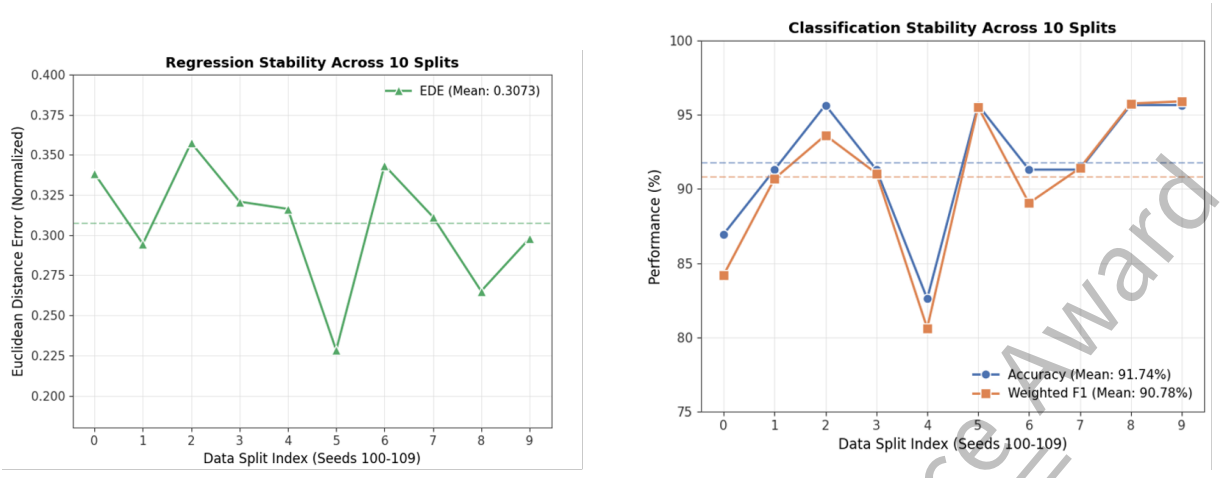
6.2 10 组独立划分的稳定性

表 8 报告了 micro 模型在 10 组独立随机分层划分（种子 100–109）上的汇总统计。10 组测试集的 Acc 均值为 91.74% ± 4.10%，Weighted F1 为 90.78% ± 4.81%，EDE 为 0.3074 ± 0.0365。

表 8: micro 模型在 10 组独立划分上的汇总统计

指标	mean ± std	min	max	极差
Acc (%)	91.74 ± 4.10	82.61	95.65	13.04
Weighted F1 (%)	90.78 ± 4.81	80.63	95.90	15.27
EDE	0.3074 ± 0.0365	0.2283	0.3575	0.1292
错分数（例/组）	1.9	1	4	3

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示



(a) EDE 波动　　(b) Acc 与 F1 波动

图 18: micro 模型在 10 组独立划分上的分类与回归稳定性波动图

从表 8 和图 18 可以看出：Acc 标准差 4.10 个百分点，对应约 23 例测试集上约 1 个样本的波动，说明模型对数据划分不敏感；10 组中 9 组 Acc $\geq$ 86.96%，8 组 $\geq$ 91.30%，最差一组（split 4）82.61% 仍高于所有基线均值；EDE 标准差 0.0365，最差 0.3575、最好 0.2283，波动幅度小于基线的均值差异。

6.3 与基线模型的对比

表 9 将 MCL-Net（micro）与多种基线方法在相同的 10 组划分上进行比较。所有深度模型从零训练，不做集成，不使用预训练权重。由于部分基线无回归能力或不报告 F1，表中以“—”标注缺失指标。

表 9: 各方法在 10 组独立划分上的性能对比（mean ± std）

方法	Acc (%)	Weighted F1 (%)	EDE
ResNet-1D	56.96 ± 7.89	49.14 ± 4.88	0.7817 ± 0.0921
Transformer	60.43 ± 4.10	53.32 ± 6.00	0.7539 ± 0.1222
GRU	73.48 ± 6.58	69.48	0.5726 ± 0.0873
LSTM	70.81 ± 12.22	68.09	0.5679 ± 0.1148
CNN+CLA（本文基线）	62.17 ± 5.16	52.78 ± 6.90	0.6827 ± 0.0628
随机森林	75.65 ± 6.51	72.82 ± 7.72	—（无回归能力）
MCL-Net (micro)	**91.74 ± 4.10**	**90.78 ± 4.81**	**0.3074 ± 0.0365**

注：“—”表示该基线未报告此指标或无此能力，不是数值为零；CNN+CLA 即 CNN+Cross-Lead Attention。

从表 9 可以看出：相对最强深度基线（CNN+Cross-Lead Attn），MCL-Net 的 Acc 提升 29.57 个百分点，EDE 降低 0.3753（降低 55.0%）；相对 Transformer，Acc 提升 31.31 个百分点，EDE 降低 0.4465；相对 ResNet-1D，Acc 提升 34.78 个百分点，EDE 降低 0.4743；相对随机森林，Acc 提升 16.09 个百分点。随机森林无回归能力，MCL-Net

2026届 T. Yau Hua School Science Award 仅用于2026届申请科学奖论文公示

的核心优势在于同时输出分类和三维坐标。值得注意的是，随机森林在纯分类任务上优于三个深度基线（75.65% vs 56.96%–62.17%），说明 18 维手工特征本身携带强判别信息；MCL-Net 在融合手工特征后进一步将 Acc 推至 91.74%。此外，LSTM 虽然在序列建模上优于 Transformer，但波动极大（Acc 标准差达 12.22%），且 EDE 误差（0.5679）依然远高于 MCL-Net。

参数量方面，MCL-Net（38.3K）仅为 ResNet-1D 的约 3.2%、Transformer 的约 1.4%、CNN+Cross-Lead Attn 的约 2.1%，但性能全面领先。这说明在小样本条件下，参数效率比绝对容量更重要。

6.4 消融实验

为验证 MCL-Net 中各模块的贡献，本文在 micro 架构下逐一移除模块，并在 10 组独立划分上训练和评估。结果如表 10 所示。

表 10: 模块消融实验结果（m10 组划分，mean ± std）

变体	Acc (%)	Weighted F1 (%)	EDE
完整模型	**91.74 ± 4.10**	**90.78 ± 4.81**	**0.3074 ± 0.0365**
no_attention	71.74 ± 5.91	65.81 ± 7.90	0.5860 ± 0.0852
no_handcrafted	59.57 ± 6.75	50.11 ± 6.63	0.7039 ± 0.0534
cls_only	69.57 ± 7.53	63.36 ± 7.60	—
reg_only	—	—	0.5946 ± 0.0719

注：完整模型包含多导联卷积、跨导联注意力、手工特征融合和双任务输出；no_attention 将跨导联注意力模块替换为恒等映射；no_handcrafted 将手工特征输入置零；cls_only 仅保留分类头，移除回归头；reg_only 仅保留回归头，移除分类头。“—”表示该变体无对应任务的输出。

从表 10 可以得出以下结论。第一，移除手工特征分支后 Acc 下降 32.17 个百分点、EDE 上升 0.3965，是最大贡献来源。第二，移除跨导联注意力后 Acc 下降 20.00 个百分点、EDE 上升 0.2786，贡献次之。第三，仅保留分类头时 Acc 下降 22.17 个百分点，说明回归任务对分类有正向迁移；仅保留回归头时 EDE 上升 0.2872，说明分类任务对回归也有正向迁移。双头共享对两个任务均有正向迁移，单头训练时另一任务显著退化。

2026 S.T. Yau High School Science Award
仅用于2026正高同年科学奖论文公示

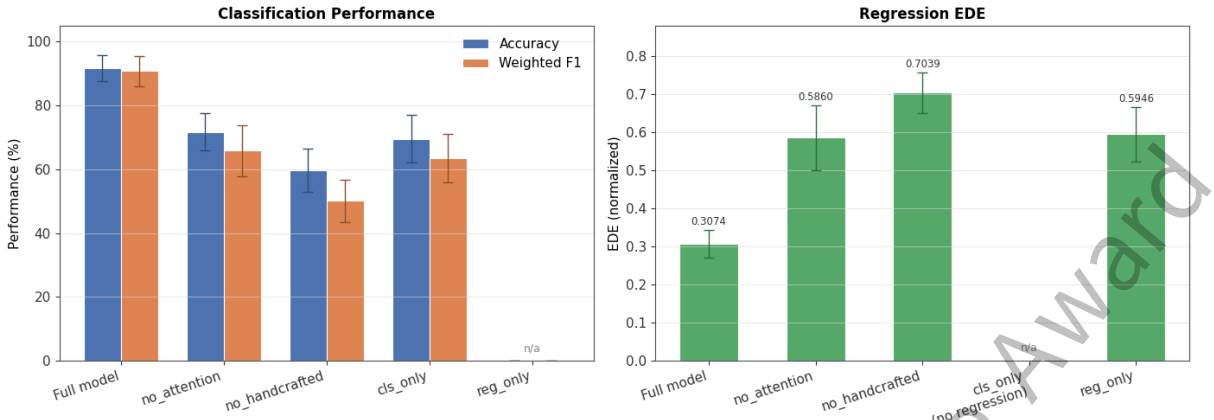


图 19: Module ablation on the micro architecture (ten splits, mean ± std). Left: Accuracy and Weighted F1; right: EDE. cls_only has no regression output and reg_only has no classification output, marked as n/a. Removing the handcrafted feature branch causes the largest accuracy drop (−32.17 pp), followed by removing cross-lead attention (−20.00 pp); sharing representations between the two heads yields positive transfer for both tasks.

6.5 预训练嵌入对比

为评估通用 ECG 预训练表示对本任务的迁移效果，本文在 micro 架构基础上替换 LeadCNN 或加载预训练权重，进行单次划分的对比实验。结果如表 11 所示。需要说明的是，本节为单次划分的探索性结果，尚未扩展到 10 组划分，结论仅作为方向性参考。

表 11: 预训练嵌入对比（单次划分，micro 架构）

变体	预训练配置	Acc (%)	F1 (%)	EDE
from_scratch	无，从零训练	**91.74**	**90.78**	**0.3074**
ssl_frozen_leadcnn	PTB-XL/本地，掩码重建，冻结	69.57	64.92	0.6064
ssl_full_finetune	PTB-XL/本地，掩码重建，全量微调	65.22	54.49	0.6740
hubert_ecg_frozen	HuBERT-ECG small，自监督，冻结	73.91	69.64	0.5160
ptbxl_cls_frozen	PTB-XL，五分类有监督，冻结	65.22	58.96	0.6050

注：本节为单次划分的探索性结果，仅作方向性参考。

从表 11 可以看出：分类任务上，从零训练最优（91.74%），所有预训练变体均低于从零训练（65.22%–73.91%），说明 6 类解剖区域的细粒度判别依赖手工特征先验和任务特定表示，通用预训练未针对该任务优化；回归任务上，HuBERT-ECG 冻结最优（EDE 0.5160），但仍高于从零训练的 0.3074。自监督掩码重建无论冻结还是全量微调，均未超越从零训练。这一分化结果提示：预训练收益是任务相关的，分类更依赖领域知识与任务特定表示，回归可能从大规模自监督表示中获益，但需要更充分的预训练轮数和更匹配的预训练任务。

2026 S.-T. Yau High School Science Award 仅用于 2026 年全国中学生科学论文公示

6.6 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。消融实验显示，移除注意力模块后 Acc 下降 20.00 个百分点，验证了该模块的独立贡献。

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。消融实验显示，移除手工特征后 Acc 下降 32.17 个百分点，是最大贡献来源，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。cls_only 和 reg_only 的消融结果分别显示 Acc 下降 22.17 个百分点、EDE 上升 0.2872，支持双头共享的双向正迁移。

第四，模型容量与样本量需要匹配。6 档容量实验表明，38.3K 的 micro 在 Acc、F1、EDE 三项同时达到峰值；65K 以上分类断崖式下降，与样本/参数比降至 2 以下一致。这提示在小样本临床任务中，参数效率比绝对容量更重要，盲目增大模型反而会损害泛化。

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，10 组独立划分的平均分类准确率为 91.74%、加权 F1 为 90.78%，平均 EDE 为 0.3074。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，仍需导联扰动实验和更大样本验证。

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 个 PVC 样本，每组测试集约 23 例。虽然 10 组独立划分在一定程度上验证了模型对数据划分的稳定性，但总样本量仍有限，稀有类别（His、Summit、AMC）的每组测试样本仅 1–3 例，逐类性能的方差较大。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**。所有数据均来源于合作医院，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**。本文采用参考标准心室模板建立的空间坐标系，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

2026 Tongji High School Science Award 仅用于2026年同济中学科学奖公示

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。

第六，预训练嵌入对比仅在单次划分上完成，尚未扩展到 10 组划分，且部分预训练配置尚未完全跑通，结论仅作为方向性参考。

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 micro 模型参数量仅 38.3K，已具备良好的部署潜力；若希望集成至医院信息系统进行实时预测，可进一步通过知识蒸馏、模型剪枝或量化等技术优化推理效率。

（6）补充分析与鲁棒性验证。后续将开展导联遮蔽与选择性预测实验，验证跨导联注意力的因果贡献，并在低置信度或分类与回归冲突时引入拒绝判断机制；同时补充逐类别混淆矩阵与召回率分析，以量化 RVOT/LVOT 等相邻区域的混淆程度。

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

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，采用 10 组独立随机分层划分（种子 100–109），MCL-Net（micro 档，38.3K 参数）在测试集上达到 91.74% ± 4.10% 的分类准确率、90.78% ± 4.81% 的加权 F1 分数；平均欧氏距离误差为 0.3074 ± 0.0365。消融实验表明，手工特征分支是最大贡献来源（移除后 Acc 下降 32.17 个百分点），跨导联注意力次之（移除后 Acc 下降 20.00 个百分点），双头共享对分类与回归均有正向迁移。6 档容量实验显示，38.3K 的 micro 在 Acc、F1、EDE 三项同时达到峰值，65K 以上分类断崖式下降，提示小样本条件下参数效率比绝对容量更重要。与 ResNet-1D、Transformer Encoder、GRU、LSTM、CNN+Cross-Lead Attention 等深度基线模型以及随机森林等传统机器学习方法相比，MCL-Net 在相同 10 组划分和调参预算下取得了更优性能，参数量仅为深度基线的 1.4%–3.2%。

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。

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

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] 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.

[11] M. E. Josephson. *Josephson's Clinical Cardiac Electrophysiology: Techniques and Interpretations*. Wolters Kluwer Health, 5th edition, 2015.

[12] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2016.

2026 Siyau High School Science Award
仅用于2026年成都中学科学奖论文公开

[13] S. Kiranyaz, O. Avci, O. Abdeljaber, et al. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151:107398, 2021.

[14] S. Kujime et al. Deep learning-based premature ventricular contraction origin classification using semantic segmentation. *Journal of Cardiovascular Electrophysiology*, 32(11):3012–3021, 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] S. Nakasone et al. Machine learning-based prediction of the origin of premature ventricular complexes from the 12-lead ecg. *Heart Rhythm*, 17(10):1746–1753, 2020.

[19] J. Pan and W. J. Tompkins. A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, 1985.

[20] 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.

[21] 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.

[22] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, et al. Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*, 2017.

[23] 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.

[24] 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.

[25] S. Ruder. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*, 2017.

[26] W. G. Stevenson and U. B. Tedrow. Catheter ablation for ventricular tachycardia. *Circulation*, 139(19):2222–2236, 2019.

2026 Si-Ti You High School Science Award 仅用于2026年成师中学科学奖论文公示

[27] 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.

[28] 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.

[29] P. Veličković, G. Cucurull, A. Casanova, et al. Graph attention networks. In *International Conference on Learning Representations (ICLR)*, 2018.

[30] T. Yang, L. Wang, X. Zhang, et al. Deep learning for outflow tract origin classification from ECG. *Journal of Electrocardiology*, 60:48–54, 2020.

[31] 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.

[32] 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. *Annals of Biomedical Engineering*, 47(2):403–412, 2019.

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

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质

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 例临床数据来源于合作医院。在数据获取过程中，合作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。

本论文图 1 和图 2 使用的心脏解剖底图来自 Mori S 与 Shivkumar K 所著 *Atlas of Cardiac Anatomy, Volume 1: Anatomical Basis of Cardiac Interventions*（2022，ISBN 978-1-942909-57-6），在此对原作者表示衷心感谢；图中的区域分割、坐标系定义与标注均为本文作者在底图基础上添加。

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。

2026 ST Youth Science Award 仅用于2026年上海青少年科学奖论文评审

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 10 组独立随机分层划分（种子 100–109）、模型容量实验、基线模型的搭建、消融实验的设计、预训练嵌入对比等，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、对 ECG 波形做固定窗口截取与幅度归一化，并对手工特征做 Z-score 标准化的方式加以解决。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！

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