

1. 背景

心血管疾病 (Cardiovascular Disease, CVD) 是全球致死率最高的疾病之一。心电图 (Electrocardiogram, ECG) 作为一种无创、经济且具备高时间分辨率的检测手段，长期以来被视为评估心脏电生理活动的“金标准”。在成人心脏病学领域，深度学习技术——特别是卷积神经网络 (CNN) 与 Transformer 架构——的迅猛发展，已推动自动化 ECG 诊断系统在特定任务（如心律失常分类）上的性能达到甚至超越人类专家水平。

直接将成熟的成人 AI 模型迁移至儿科场景 (Transfer Learning)，往往面临严重的**负迁移**，其根源在于成人与儿童之间存在巨大的**生理域漂移 (Domain Shift)**：

- **空间维度差异：**胎儿期及新生儿期，心脏呈现显著的生理性右室优势 (Right Ventricular Dominance)，以应对高阻力的肺循环。在心电图上，这表现为 V1 导联的高 R 波和显著的电轴右偏。对于基于成人左室优势模式训练的模型而言，V1 高 R 波是右室肥厚 (RVH) 或后壁心肌梗死的症状。因此，成人模型极易将一个健康的 3 天大新生儿误判为重症心脏病患者。
- **时间维度尺度缩放：**儿童的心率显著快于成人，且心脏传导系统尚未完全发育，导致 PR 间期、QRS 波宽和 QT 间期均显著短于成人。例如，儿童 QRS 波宽正常上限仅为 80-90ms，而成人为 120ms。现有的 CNN/Transformer 模型通常采用固定的卷积核或 Token 大小，无法适应这种非线性的时间压缩，导致对传导阻滞等疾病的判定失效。

2. 相关工作

相关工作中的一些问题：

- **解剖先验知识的缺失：**当前主流模型通常将 12 导联 ECG 信号视为 12 个独立通道，或简单地将其拼接为二维图像进行处理。这种建模方式忽视了导联系统内的空间拓扑结构。实际上，标准 12 导联是心脏电活动在额面与横面等多个解剖平面上的投影，不同导联组具有明确的生理关联——例如下壁导联 (II、III、aVF) 与侧壁导联 (I、aVL、V5、V6) 在心肌缺血定位中呈现高度协同的模式变化。忽略此类解剖-电生理先验，迫使模型在有限的儿科数据上重新学习本应已知的物理关联，极易导致过拟合。
- **长时序建模与多尺度表征能力的不足：**ECG 记录包含数千采样点（如 10 秒、500 Hz 采样率对应 5000 点）。传统 CNN 受限于局部感受野，难以有效建模远距离心律依赖；而 Transformer 虽具备全局建模能力，但缺乏显式的多尺度建模机制——它对所有时间步一视同仁，难以同时捕捉多种尺度的心电模式，如 QRS 波群的瞬时细节（毫秒级）与 RR 间期变异性等慢速节律模式（秒级）等，从而限制了其在复杂心律分析中的表现。
- **复合诊断文本导致的对比学习假阴性干扰：**为缓解标注数据稀缺的问题，对比学习被广泛用于利用未标注 ECG 数据。然而，经典 InfoNCE 损失函数默认“非正样本即为负样本”，这一假设在医学场景中存在严重缺陷：一个 batch 内的不同样本，其诊断文本可能存在相同疾病（如 A+B+C 与 A+D+E），却被强制推远至特征空间的不同区域，形成“假阴性” (False Negatives)。
- **离散化域适应范式的不适用性：**现有域适应方法（如对抗训练、最大均值差异 MMD 对齐）主要面向离散域迁移。然而，儿科 ECG 本质上构成一个由年龄驱动的**连续生理域**。若强行将发育过程离散化为若干年龄段（如 0-1 岁、1-5 岁等），不仅割裂了心脏电生理演变的连续性，还会因域边界模糊引入额外的分类歧义，显著增加模型设计与部署的复杂度。

创新点：

1. **Multi-View Cardiac Spatial Encoding** (空间维度的编码)：基于 LISA (Lateral, Inferior, Septal, Anterior) 解剖分组原则，引入空间先验，对导联组内部进行特征提取与融合，并结合通道注意力机制，学习导联组间的协同效应。

2. **Multi-Scale Shift-Adaptive Temporal Encoding** (时间维度多尺度编码)：结合Transformer的长序列建模能力与多尺度Patch Embedding策略，实现从波形级 (Waveform-level) 到节律级 (Rhythm-level) 的多尺度特征提取。
3. **DiagSim-Weighted loss**：诊断语义相似度加权的多模态对比学习损失，在训练阶段利用临床报告间的诊断相似度动态调整负样本权重，缓解因共病或相似诊断导致的假负例问题，从而提升ECG-文本跨模态表示的语义一致性与泛化能力。
4. **Age-Gated Adaptator** (年龄适应微调)：利用特征线性调制 (FiLM, Feature-wise Linear Modulation) 技术，将年龄连续生理参数作为条件变量动态调节导联组、不同尺度patch的特征权重，实现模型对不同发育阶段的平滑适应。

3. 方案

模型分为两个阶段：

1. **预训练阶段 (Pre-training)**：在MIMIC-IV-ECG上进行多模态对比学习，并在PTB-XL/CPSC2018/CSN做zero-shot验证
2. **微调阶段 (Fine-tuning)**：在Anzhen数据上，冻结或小学习率更新骨干参数，插入**Age-Gated Adaptator**，进行有监督微调。

3.1 Multi-View Cardiac Spatial Encoding (MVCSE)

针对痛点：解剖先验缺失、导联间协同关系被忽视

- **分组策略**：不将12导联视为独立通道，而是根据解剖位置分为4组：
 - I, aVL, V5, V6, aVR (侧壁，反映左室高侧及侧壁)
 - II, III, aVF (下壁，反映右冠状动脉供血区)
 - V1, V2 (间隔，反映室间隔及右室)
 - V3, V4 (前壁)
- **组内编码 (Intra-group Encoding)**：每组导联通过共享权重的时序编码 (3.2) 提取局部波形特征。
- **组间注意力 (Inter-group Channel Attention)**：使用SE-Block (Squeeze-and-Excitation) 或 ECA-Block。
 - 意义：例如在检测右室肥厚 (RVH) 时，模型会自动赋予 **S组 (V1, V2)** 更高的权重，而抑制与此无关的侧壁导联噪声。

3.2 Multi-Scale Shift-Adaptive Temporal Encoding (MS-SATE)

针对痛点：儿童心率快、QRS波窄导致的固定尺度失效

- **多尺度 Patch Embedding**：
 - 并行使用三个不同卷积核大小 (Kernel Size) 和步长 (Stride) 的卷积层作为Patch Embedding层。
 - **Scale 1 (Short)**: 捕捉高频细节 (如QRS切迹、P波形态)。对应儿童窄QRS波。
 - **Scale 2 (Medium)**: 捕捉波群特征 (如ST段形态)。
 - **Scale 3 (Long)**: 捕捉节律特征 (如RR间期)。对应心率变异性。
- **特征融合**：将三种尺度的Token输入到独立的transformer encoder中，在通道维度拼接。
- **平移适应**：引入相对位置编码 (Relative Positional Encoding)，增强模型对心跳在时间轴上的鲁棒性。

3.3 DiagSim-Weighted Loss

针对痛点：对比学习中的假阴性 (False Negatives) 问题

- 背景：**传统的InfoNCE假设只要不是同一个样本 (Augmented Pair)，就是负样本。但心电数据集中大多为复合诊断记录，两条不同的记录可能都包含诊断为“室性早搏”。强行拉远它们的距离会破坏语义空间。
- 文本编码器：** 使用预训练的临床BERT (如BioBERT或ClinicalBERT) 提取诊断文本特征 T_i 。
- 语义相似度矩阵：** 计算Batch内所有文本对的余弦相似度 $S_{ij} = \text{CosSim}(T_i, T_j)$
- 加权损失函数：**

$$\mathcal{L}_i = -\log \frac{\exp(\text{sim}(v_i, t_i)/\tau)}{\exp(\text{sim}(v_i, t_i)/\tau) + \sum_{j \neq i} \alpha_{ij} \cdot \exp(\text{sim}(v_i, t_j)/\tau)}$$



- 其中 α_{ij} 为负样本权重。
- 策略：** 当 S_{ij} 很高 (文本诊断高度相似) 时，降低 α_{ij} (甚至设为0)，避免将具有相同语义的样本作为负样本惩罚；当 S_{ij} 很低时，保持 $\alpha_{ij} = 1$ 。

3.4 Age-Gated Adaptator

针对痛点：从新生儿到青春期的连续生理域漂移

- 位置：** 插入在Transformer的每个Block之后。
- 输入：** 归一化的年龄数值 $age \in [0, 1]$ 。
- FiLM (Feature-wise Linear Modulation) 机制：**
 - 将标量 age 通过一个两层MLP映射为两个向量：缩放因子 $\gamma(age)$ 和偏移因子 $\beta(age)$ 。
 - 调制操作：对特征图 F 进行仿射变换：

$$\text{FiLM}(F|age) = \gamma(age) \odot F + \beta(age)$$



- 物理意义：**
 - Scale (γ)：** 相当于“门控”，例如对于新生儿 ($age \approx 0$)，模型可能会放大V1导联特征的权重 (因为此时右室优势是正常的)，而对于青少年，则抑制该特征以避免误判RVH。
 - Shift (β)：** 调整基准值，适应不同年龄段的波形幅度基线。

4. 实验设计

4.1 baseline

| Methods | PTBXL-Rhythm | | | PTBXL-Sub | | | PTBXL-Form | | | PTBXL-Super | | | CPS2018 | | | CSN | | |
|--|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Training ratio | 1% | 10% | 100% | 1% | 10% | 100% | 1% | 10% | 100% | 1% | 10% | 100% | 1% | 10% | 100% | 1% | 10% | 100% |
| SimCLR (Chen et al., 2020) | 51.41 | 69.44 | 77.73 | 60.84 | 68.27 | 73.39 | 54.98 | 56.97 | 62.52 | 63.41 | 69.77 | 73.53 | 59.78 | 68.52 | 76.54 | 59.02 | 67.26 | 73.20 |
| BYOL (Grill et al., 2020) | 41.99 | 74.40 | 77.17 | 57.16 | 67.44 | 71.64 | 48.73 | 61.63 | 70.82 | 71.70 | 73.83 | 76.45 | 60.88 | 74.42 | 78.75 | 54.20 | 71.92 | 74.69 |
| Barlow Twins (Zbontar et al., 2021) | 50.12 | 73.54 | 77.62 | 62.57 | 70.84 | 74.34 | 52.12 | 60.39 | 66.14 | 72.87 | 75.96 | 78.41 | 55.12 | 72.75 | 78.39 | 60.72 | 71.64 | 77.43 |
| MoCo-v3 (Chen et al., 2021) | 51.38 | 71.66 | 74.33 | 55.88 | 69.21 | 76.69 | 50.32 | 63.71 | 71.31 | 73.19 | 76.65 | 78.26 | 62.13 | 76.74 | 75.29 | 54.61 | 74.26 | 77.68 |
| SimSiam (Chen & He, 2021) | 49.30 | 69.47 | 75.92 | 62.52 | 69.31 | 76.38 | 55.16 | 62.91 | 71.31 | 73.15 | 72.70 | 75.63 | 58.35 | 72.89 | 75.31 | 58.25 | 68.61 | 77.41 |
| TS-TCC (Eldele et al., 2021) | 43.34 | 69.48 | 78.23 | 53.54 | 66.98 | 77.87 | 48.04 | 61.79 | 71.18 | 70.73 | 75.88 | 78.91 | 57.07 | 73.62 | 78.72 | 55.26 | 68.48 | 76.79 |
| CLOCS (Kiyassae et al., 2021) | 47.19 | 71.88 | 76.31 | 57.94 | 72.55 | 76.24 | 51.97 | 57.79 | 72.65 | 68.94 | 73.36 | 76.31 | 59.59 | 77.78 | 77.49 | 54.38 | 71.93 | 76.13 |
| Wav2Vec 2.0 + CMSC + RLM (Oh et al., 2022) | 76.24 | 86.34 | 92.05 | 69.10 | 80.71 | 85.01 | 52.72 | 67.81 | 80.72 | 81.15 | 84.88 | 85.53 | 75.70 | 88.16 | 92.61 | 65.65 | 78.82 | 87.87 |
| ASTCL (Wang et al., 2024) | 52.38 | 71.98 | 76.05 | 61.86 | 68.77 | 76.51 | 44.14 | 60.93 | 66.99 | 72.51 | 77.31 | 81.02 | 57.90 | 77.01 | 79.51 | 56.40 | 70.87 | 75.79 |
| CRT (Zhang et al., 2023) | 47.44 | 73.52 | 74.41 | 61.98 | 70.82 | 78.67 | 46.41 | 59.49 | 68.73 | 69.68 | 78.24 | 77.24 | 58.01 | 76.43 | 82.03 | 56.21 | 73.70 | 78.80 |
| ECGFM (McKeen et al., 2024) | 81.45 | 91.59 | 92.70 | 73.24 | 81.91 | 86.07 | 60.95 | 74.99 | 85.54 | 78.67 | 84.80 | 86.47 | 82.18 | 89.52 | 93.26 | 71.51 | 83.17 | 88.89 |
| ST-MEM (Na et al., 2024) | 51.12 | 65.44 | 74.85 | 54.12 | 57.86 | 63.59 | 55.71 | 59.99 | 66.07 | 61.12 | 66.87 | 71.36 | 56.69 | 63.32 | 70.39 | 59.77 | 66.87 | 71.36 |
| HeartLang (Jin et al., 2025) | 62.08 | 76.22 | 90.34 | 64.68 | 79.34 | 88.91 | 58.70 | 63.99 | 80.23 | 78.94 | 85.59 | 87.52 | 60.44 | 66.26 | 77.87 | 57.94 | 68.93 | 82.49 |
| MERL (Liu et al., 2024a) | 53.33 | 82.88 | 88.34 | 64.90 | 80.56 | 84.72 | 58.26 | 72.43 | 79.65 | 82.39 | 86.27 | 88.67 | 70.33 | 85.32 | 90.57 | 66.60 | 82.74 | 87.95 |
| MELP (Ours) | 88.83 | 94.65 | 96.91 | 79.22 | 84.40 | 87.46 | 63.41 | 76.71 | 83.30 | 85.82 | 87.61 | 87.87 | 88.54 | 91.75 | 94.32 | 78.25 | 84.83 | 90.17 |

4.2 消融实验

- 预训练部分

| 实验序号 | MVCSE | MS-SATE (Multi-Scale) | DiagSim Loss | 说明 |
|------|-------|--------------------------|-----------------|---------------------------------|
| A1 | × | × | × | 12导联不分组、固定切分patch、无特殊处理的文本-信号对齐 |
| A2 | × | √ | √ | 证明导联分组有效 |
| A3 | √ | × | √ | 证明多尺度时序有效 |
| A4 | √ | √ | × | 证明文本相似度加权优化了假阴性问题 |
| A5 | √ | √ | √ | 完整模型 |

- 微调部分

| 实验序号 | 内容 | 说明 |
|------|----------------------------------|-----------|
| B1 | baseline1直接微调 | 成人直接迁移不合适 |
| B2 | baseline2直接微调 | 成人直接迁移不合适 |
| B3 | my model 直接微调 | 成人直接迁移不合适 |
| B4 | my model + Age-Gated Adaptator微调 | 证明年龄门控有效性 |

4.3 可视化与可解释

- 特征空间可视化：绘制1-2个baseline的t-SNE与my model对比；绘制是否插入年龄门控的两张t-SNE，观察不同年龄、心脏病类别的聚类情况