分阶段聆听心跳：基于生理学的相位感知心电图生物特征识别

Jintao Huang、Lu Leng（IEEE会员）、Yi Zhang（IEEE高级会员）和Ziyuan Yang（IEEE会员）

arXiv:2601.00170v1 [eess.IV] 1 Jan 2026

摘要—由于心电图（ECG）具有个体特异性特征和内在活体性，因此被用于可穿戴设备中的身份认证。然而，现有方法通常将心搏视为同质信号，忽略了心动周期内各相位的特异性特征。为此，我们提出一种分层相位感知融合（HPAF）框架，通过三阶段设计明确避免跨特征纠缠。第一阶段，相位内表征（IPR）分别提取每个心动相位的表征，确保保留相位特异的形态学特征和变化线索，同时不受其他相位干扰。第二阶段，分阶段分组层次融合（PGHF）以结构化方式聚合生理相关相位，从而可靠地整合互补的相位信息。最终阶段，全局表征融合（GRF）进一步融合分组后的表征，并自适应地平衡其贡献，以生成统一且具有区分性的身份表征。此外，考虑到心电图信号是连续采集的，每个人都可以采集多个心搏。我们提出一种心搏感知多原型（HAM）注册策略，构建多原型图库模板集，以降低心搏特异性噪声和变异性的影响。在三个公开数据集上进行的大量实验表明，与其他方法相比，HPAF在闭集和开放集设置下均取得了最先进的结果。

索引词—心电图生物特征识别、多视图表征、图神经网络、身份验证。

I. 引言

长期以来，心电图（ECG）在临床诊断中发挥着基础性作用。随着可穿戴设备的兴起，心电图日益被视为一种具有生理学依据的生物特征信号。与其他与面部或指纹等模态不同，心电图采集直接源于心脏活动，因此天然具备活体检测能力 [1]–[3]。这些特性使心电图成为实现隐私保护和持续身份认证的有前景模态 [4]–[7]。

近年来，越来越多的研究开始探索使用心电图进行身份识别。这些方法已从手工特征方法发展到基于深度学习的方法 [8]，[9]。大多数现有方法使用单一模型对不同的心动周期阶段进行编码。例如，Ibtehaz 等人 [10] 将多分辨率孪生网络应用于完整心搏，以实现心电图身份认证。Chu 等人 [11] 采用并行多尺度残差网络，并结合中心与边界损失来提取判别性特征。Wang 等人 [12] 提出了一种自监督卷积神经网络（CNN）框架，用于比较不同的心电图分割策略。Allam 等人 [13] 采用并行卷积神经网络-长短期记忆网络架构，从 R 波峰对齐心搏中学习表征。尽管这些方法在理想条件下表现出良好前景，但它们都将包含 P 波、QRS 波群和 ST/T(U) 段的心搏视为同质序列，并通过共享编码器进行处理。这些方法基于一个基本假设，即不同生理阶段具有同质的信号模式，忽略了其固有的结构和生理异质性。这导致判别性特征学习效果不佳，并使识别任务的性能未达最佳。

具体而言，心电图包含不同的波形阶段，分别源自心脏不同的电生理活动，包括 P 波、QRS 波群和 T 波。由于电生理机制各不相同，这些阶段在形态和频谱特征方面存在显著差异，如图 1 所示。P 波反映心房去极化，其形态相对平滑、振幅较低；QRS 波群则对应快速的心室去极化，由于心肌质量较大且传导速度快，因此能量最高、转变最陡峭。T 波代表心室复极，其特征是电活动恢复较慢，波形模式更为平滑。因此，若将这些生理特性迥异的阶段视为同质的信号模式，将严重限制模型捕捉稳定且具有区分性的身份特征的能力。

受心动周期内在生理结构的启发，我们提出了一种多粒度心动周期阶段

本文已提交至 IEEE，拟供发表。版权可能在未另行通知的情况下转让，届时此版本可能不再可用。

本研究部分受国家自然科学基金委员会资助，项目编号为 62466038；部分受江西省图像处理与模式识别重点实验室资助，项目编号为 2024SSY03111 和 ET202404437；另部分受南昌航空大学信息中心高性能计算服务资助。（通讯作者：Lu Leng 和 Ziyuan Yang）

Jintao Huang 任职于江西省图像处理与模式识别重点实验室、南昌航空大学，邮编 330063，中国（电子邮箱：2520083500018@stu.nchu.edu.cn）。

Lu Leng 任职于中国南昌航空大学江西省图像处理与模式识别重点实验室，邮编 330063（电子邮箱：leng@nchu.edu.cn）。

Yi Zhang 和 Ziyuan Yang 就职于中国成都 610065 的四川大学网络空间安全学院（yzhang@scu.edu.cn 和 cziyuanyang@gmail.com）。

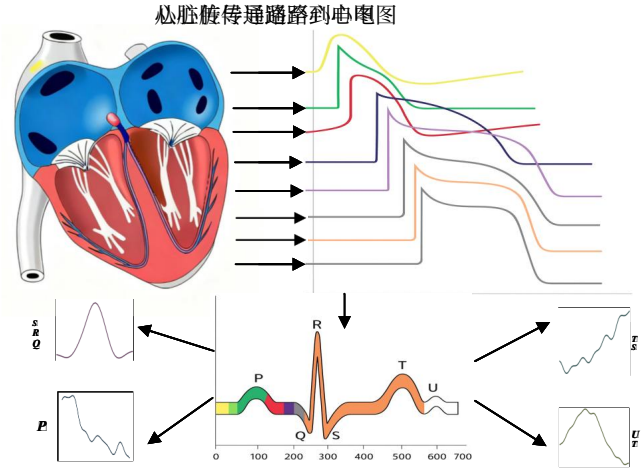


图 1：心动周期不同生理阶段与相应心电图信号之间关系的示意图。

表示框架，以阶段感知的方式对异质阶段模式进行建模。具体而言，我们首先提出一种心动周期分割（CPS）方法，将每个心搏分割为四个不同的阶段，包括 P波、QRS波群、ST段和T/U波，并以 R 峰为锚点。通过这种方式，我们的方法获得了具有生理意义的时间动态阶段，并保留了阶段特异性形态特征，为后续的阶段特异性建模提供了可靠的基础。

如前所述，采用单一共享编码器会忽略不同生理阶段之间的生理异质性，从而妨碍稳定且具有判别性的表征提取。为解决这一问题，我们提出一种层次化阶段感知融合策略（HPAF），包含阶段内表示（IPR）、分阶段分组层次融合（PGHF）和全局表征融合（GRF）模块。在 IPR 中，为避免不同阶段的特征纠缠，我们设计了特定于各阶段的形态—变化感知特征提取器（MVFE），由各自独立的 MVFE 处理每个阶段，以提取具有判别性的特征。MVFE 由用于形态提取的标准卷积分支和用于变化建模的可学习 Gabor 分支组成。标准卷积分支能够有效捕捉波形形态和时间结构，而可学习的 Gabor 滤波器则能够对细粒度的局部变化进行建模，例如细微的振荡和突变。

随后，在 PGHF 中，我们将各生理阶段划分为两个具有生理一致性的组，并分别执行相位分组层次融合：一个 PGHF 模块用于 P波和 T/U波特征，另一个用于 QRS 波群和 ST段特征。具体而言，P波和 T/U波特征被归为一组，因为二者都与复极化相关，并呈现平滑、低振幅的波形。相比之下，QRS波群和 ST段特征被归为一组，是因为它们分别与心室去极化和早期复极化相关，其特点是振幅更高、转变更陡峭，且形态更为复杂。随后，我们设计了一种跨相位门控融合器，通过门控校准而非直接融合来选择性地耦合各生理阶段，从而实现对不同生理阶段组的稳定校准。

最后，PGHF模块融合后的两个特征被送入GRF模块，再次通过跨相位门控融合器进行自适应融合，得到最终表征。具体而言，GRF首先将每个融合特征投影到共享评分空间，并计算一个标量类注意力门控，以衡量QRS–ST组与ST–T/U组的相对重要性。随后，该门控控制两个融合特征的加权混合，并通过线性投影得到最终全局特征。由此，模型能够根据不同的噪声条件和节律模式，动态判断应以快速瞬态线索（QRS–ST）还是缓慢复极化线索（ST–T/U）为主。

为了优化网络并提取判别性特征，我们引入了对比损失。该损失在特征空间中约束全局融合嵌入，以促使阶段特异性身份线索与跨阶段一致性达到平衡。在实现阶段，由于心电图信号是连续采集的，因此可以为每位个体采集多个心搏。我们提出了一种心搏感知多原型（HAM）注册策略。不同于以往使用单个特征向量表示每位个体的方法，HAM使用一组原型表示每位受试者。通过这种方式，我们的方法能够构建一个图库集，减少心搏特异性噪声和变异性的影响，从而提高匹配可靠性和性能。我们的主要贡献总结如下：

- 我们提出 HPAF，这是一种以生理机制为基础的阶段感知心电图身份识别框架，通过三阶段设计避免跨阶段特征纠缠。
- 我们设计了双分支 MVFE，以提取形态特征和变化特征，并引入 PR-GAT 对二者的交互进行建模，从而实现显式的可靠性感知细化。
- 我们提出了一种新颖的注册策略 HAM，利用心电图信号中的多个心搏，为每个人构建稳健的多原图库模板。
- 在公开数据集上开展的大量实验，在闭集和开放集协议下全面验证了该方法的有效性。

II. 相关工作

A. 心电图身份识别的表征学习

1) 基于基准点的特征学习：基于基准点的心电图身份识别方法依赖于对显著标志点（P 波、QRS 波群、T 波）的精确检测，并构建时域特征。例如，Biel 等人 [2]提出从静息12导联心电图中提取振幅、持续时间和波间期，并使用SIMCA分类器进行受试者验证。Israel 等人 [3] 进一步

从高采样率心电图中经过仔细标注的 P/R/T 点提取归一化时间特征，并结合LDA与多数投票，在不同电极位置和焦虑状态下实现稳健识别。因此，准确的波形 delineation 是基于基准点流程的关键前提；例如，Martínez 等人 [14] 开发了一种基于小波变换的心电图波形 delineator，用于在标准数据库上定位QRS波群，并描绘P/QRS/T峰值及边界。这些基于标志点的方案具有物理可解释性，但对边界检测误差、运动伪影和EMG噪声高度敏感，正如Fratini 等人 [15], 所总结的那样；而 Wang 等人 [16]采用基于小波的时频表示与堆叠稀疏自动编码器，无需显式检测基准点，凸显了传统基于基准点的方法在噪声环境和跨条件场景下的局限性。

2）非基准点表征学习：非基准点特征提取绕过对 P/QRS/T 标志点的显式检测，而是从完整心电图或其变换域/重构域中提取描述符。代表性方法包括：Jung 等人 [17], 将预处理后的单导联心电图分割为固定窗口，并结合窗口移除策略和经典分类器，计算基于自相关和离散余弦变换（DCT）的特征；Fang 等人[18], 重构心电图相空间，并量化不同图谱之间的相似性/差异性；Srivastva 等人 [19], 将自相关与 DCT/离散傅里叶变换（DFT）/沃尔什–哈达玛变换（WHT）以及主成分分析（PCA）/线性判别分析（LDA）相结合；Goshvarpour 等人 [20], 使用基于匹配追踪的统计特征和非线性特征，并结合概率神经网络（PNN）/k近邻（kNN）及特征选择。Gutta 等人 [21], 将多类别心电图身份识别重新表述为一对多多任务学习问题，并采用特征缩放的概率核分类器，通过稀疏–低秩权重分解联合进行特征选择与分类。这些方法避免了因基准点检测不准确而导致的错误，但通常会产生高维表征，因此需要进行特征选择或降维。

3) 基于深度学习的特征提取：近年来，深度学习已成为基于心电图进行身份识别的主流方案，能够端到端地学习判别性特征。大多数方法依赖于作用于以 R 峰为锚点的心搏或节律片段上的时域卷积神经网络，同时通过固定的时频变换（STFT、CWT）或手工设计的 Gabor 描述符来加入频率信息，并在后续进行融合；这可能引入时频串扰，使性能对窗口长度和采样率较为敏感。代表性方法包括围绕 R 峰进行的单心搏时频建模 [22], 2基于 STFT 频谱图的 D 维卷积神经网络，用于节律分类 [23], 用于身份验证的多尺度残差卷积神经网络和孪生架构 [10], [11], [24], 以及能够捕捉多个心搏之间长程时间依赖关系的基于LSTM的模型。

B. 心电图生物识别注意力机制

近年来，心电图生物识别系统越来越多地利用注意力来突出信息丰富的时间模式和心搏。Chee 等人 [25] 在心电图序列对上采用 Transformer 风格自注意力编码器，以学习灵活的身份嵌入，从而支持跨多个数据库的身份识别和身份验证。Jyotishi 等人 [26] 设计了一种带有内部注意力模块的分层长短期记忆网络，为承载更丰富生物识别信息的心电图波群分配更大的权重；在体内和体外数据集上，其性能优于传统的基准点基线方法和非基准点基线方法。在卷积层面，Hammad 等人 [27] 将注意力块集成到残差卷积神经网络（ResNet-Attention）中，从而在特征图中突出判别性心电图片段；而 Pan 等人 [28]提出了一种多分支域自适应融合网络，其中加权自适应注意力机制增强了 ECGID、PTB、CYBHi 和 Heartprint 上的会话鲁棒心跳特征。

除了纯时间注意力之外，一些研究还将注意力应用于通道、特征或图空间。Sun 等人 [29] 提出了 RandSaD，这是一种采用分裂注意力的双路径残差模型，可从小波去噪心电图窗口中选择性地聚合多尺度通道特征，识别准确率达到 $\approx$ 99.6%。Wang 等人 [30] 在协同嵌入框架中融合一维和二维心电图表示，并利用维度注意力提高判别性潜在特征的权重。Ma 等人 [31] 提出 ORGNNFL，该模型将多特征注意力模块与图神经网络融合相结合，以便在分布偏移的情况下自适应地聚合拓扑感知心电图特征。Al Alfi 等人 [32]构建了一个 12 导联心电图图，其边的权重由基于互信息的“类似注意力”分数确定。

III. 所提方法

A. 概述

本文提出了一种用于心电图身份识别的新型多粒度心动周期阶段表示框架，所提方法的概述见图2。采用单一共享编码器提取不同心动相位的特征，会忽略生理异质性，从而可能妨碍稳定且具有判别性的特征提取。为解决这一问题，我们提出心动相位分割（CPS）方法，将心电图分割为四个心动周期阶段，以避免后续特征提取中的相位失真。随后，我们提出分层相位感知融合（HPAF）策略。HPAF由相位内表示（IPR）、分阶段分组层次融合（PGHF）和全局表征融合（GRF）模块组成。

具体而言，在 IPR 中，每个相位都由专属的形态变化感知特征提取器（MVFE）处理，以避免相位间特征纠缠。每个 MVFE 都包含一个用于波形形态的卷积分支，以及一个可学习的 Gabor 分支，用于细粒度

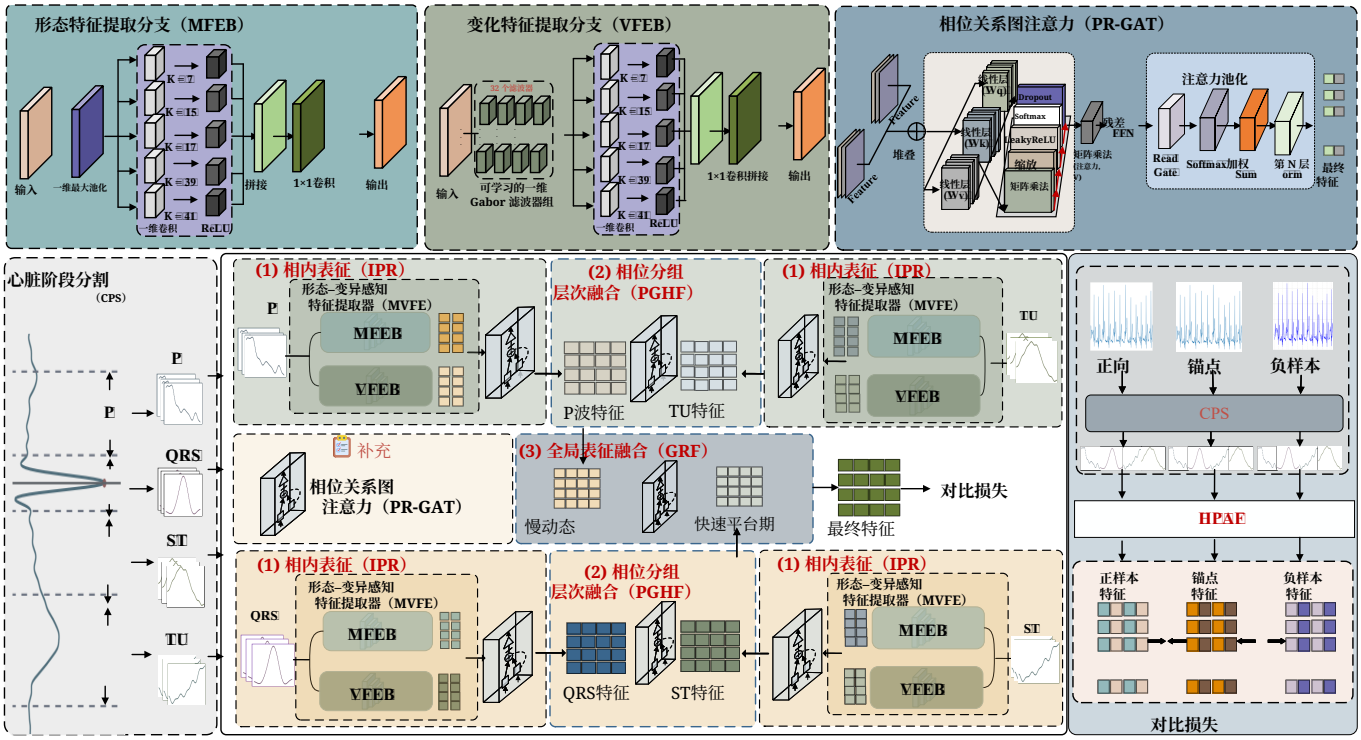


图 2：我们提出了用于相位感知心电图身份识别的 HPAF 框架。CPS 检测 R峰，并将每个心搏划分为四个相位对齐的片段，即 P波、QRS波群、ST段和 T/U波。每个相位由配备 MFEB 和 VFEB 的 MVFE 进行编码，再由 PR-GAT 融合以形成相位表示，随后 PGHF 和 GRF 将其汇总为单一的心搏级身份嵌入。编码器采用对比损失进行训练，以提高受试者间可分性。

局部变化。随后，PGHF 将各生理阶段分组为 P–T/U 和 QRS–ST，并通过跨相位门控融合器在每组内进行层次化融合。接着，GRF 将两组的表示汇总为统一的心搏级表征，作为最终特征。此外，我们还引入对比损失，以增强跨阶段一致性，同时避免过度依赖任何单一阶段。下文将详细介绍这些组成部分。

B. 分层相位感知融合（HPAF）

1) 心动周期分割（CPS）： 为缓解不同心电图相位之间的异质性，我们提出CPS方法进行显式的相位分割，从而有效防止后续特征提取过程中不同相位之间的特征纠缠。具体而言，CPS将R峰作为锚点，将每个心搏显式划分为不同的生理相位，而不是采用传统方法对整个心电图信号进行平均分段。每个心搏被划分为四个阶段，包括心房去极化、心室去极化和早期平台期，以及心室复极，分别对应P波、QRS波群、ST段和T/U波。

具体而言，CPS 首先在预处理心电图信号中检测 R 峰，然后围绕每个 R 峰划定四个时间窗。这些时间窗根据预定义的相对偏移量和持续时间，大致对应 P 波、QRS波群、ST段和 T/U相位。这四个定长片段通过选取预定义偏移窗口内的样本提取，偏移量以相对于 R 峰的样本数计：P 波为 $[-80, -20]$ 至 $[-20, +20]$、QRS波群为 $[+20, +80]$、TU波为 $[+80, +160]$。这样，CPS 生成的四相位片段具有明确的生理意义和一致的时间尺度，可有效防止后续特征提取过程中出现特征纠缠。每个分段相位均采用专用的相位特定 MVFE 独立编码。

2）相位内表征（IPR）： 在 CPS 之后，每个分割后的相位都由专属的 MVFE 独立编码。尽管常用的卷积滤波器能有效捕捉粗略的形态信息，却往往无法保留心电图信号中的细微变化。Gabor 滤波器能够有效捕捉心电图信号中的细微变化 [?], ，因为它们着重于短时间区域内的局部波形变化。基于这些互补特性，MVFE 被设计为一种双分支架构，包括一个基于卷积神经网络的形态特征提取分支（MFEB），以及一个基于可学习 Gabor 滤波器的变化特征提取分支（VFEB）。此外，为了建模形态与变化表征之间的关系，我们引入了相位关系图注意力（PR-GAT）模块。通过将形态表征和变化表征视为图节点，PR-GAT 能够显式捕捉

它们之间的关系依赖，从而获得全面且具有判别性的特征。

VFEB旨在提取细粒度变化。与需要手动指定超参数的传统 Gabor 滤波器不同，我们引入了可学习的 Gabor 卷积（LGC）层，其滤波器参数直接从数据中学习，无需预先掌握领域知识。一维Gabor核 g在第 t个相对位置处的第 k个通道定义如下：

$$g_k(t) = \exp\left(-\frac{t^2}{2\sigma_k^2}\right) \cdot \cos(2\pi f_k t + \psi_k), \tag{1}$$

其中， σ_k、 f_k和 ψ_k 均为可学习参数，分别表示 Gabor 核的缩放、中心频率和相移。为避免模型受到缓慢变化的信号偏移的影响，并突出局部变化，我们施加如下零均值约束：

$$\hat{g}_k(t) = g_k(t) - \frac{1}{|T|}\sum_{\tau=1}^{T} g_k(\tau), \tag{2}$$

其中， T 为滤波器长度，$\hat{}g_k$ 表示 g_k的零均值版本。

接下来，为了同时提取短期和长期特征，我们进一步设计了一个多尺度特征块（MSFB），由具有不同卷积核大小 $k = \{7, 15, 17, 39, 41\}$ 的并行卷积层组成。然后，沿通道维度拼接提取的多尺度特征，接着使用 a1 $\times$ 1 卷积进行特征融合，并进一步下采样。最后，后续特征层将该表征映射为紧凑的嵌入 $\mathbf{z}_v$。

此外，MFEB 专注于从心电图信号中提取形态特征。MFEB 与 VFEB 的整体架构相似，主要区别在于第一个特征提取阶段：MFEB 使用传统卷积层，而 VFEB 采用 LGC 层。随后，通过 MFEB 将提取的形态特征表示为 $\mathbf{z}_m$。

不同于将 $\mathbf{z}_v$ 和 $\mathbf{z}_m$ 一视同仁的静态拼接，PR-GAT 通过图注意力显式建模形态与变化表征之间的交互。

具体而言，我们首先将 $\mathbf{z}_v$ 和 $\mathbf{z}_m$ 堆叠为 $\mathbf{H} \in \mathbb{R}^{2\times d}$，其中 d 表示特征维度。然后，我们使用可学习矩阵 $\mathbf{W}^Q, \mathbf{W}_K, \mathbf{W}_V$ 将 $\mathbf{H}$ 投影为查询、键和值嵌入，如下所示：

$$\mathbf{Q} = \mathbf{H}\mathbf{W}_Q, \quad \mathbf{K} = \mathbf{H}\mathbf{W}_K, \quad \mathbf{V} = \mathbf{H}\mathbf{W}_V, \tag{3}$$

其中，$\mathbf{Q}$、$\mathbf{K}$ 和 $\mathbf{V}$ 分别表示查询、键和值表示。

然后，通过基于注意力的聚合获得最终注意力特征$\mathbf{A}$，如下所示：

$$\mathbf{A} = \mathrm{softmax}\left(\mathrm{LeakyReLU}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right)\right), \tag{4}$$

其中，LeakyReLU(·)表示Leaky ReLU激活函数。 d_k 是缩放因子。

然后，将$\mathbf{A}$ 输入由残差连接和层归一化层构成的标准注意力块，得到精炼特征$\mathbf{H}' = [\mathbf{h}'_v;\ \mathbf{h}'_m]$。$\mathbf{H}'$的两个通道分别表示精炼后的变化特征和形态特征。

为融合$\mathbf{h}'_v$ 和$\mathbf{h}'_m$，我们设计了一个注意力池化模块。对于 $k \in \{v, m\}$，使用轻量级多层感知机预测标量分数 s_k，并将其归一化为权重。然后，该过程可表述为：

$$\alpha_k = \frac{\exp(s_k)}{\sum_{j\in\{v,m\}}\exp(s_j)}, \quad k \in \{v, m\}, \tag{5}$$

其中， α 表示归一化后的特征。

然后，可按如下方式获得最终嵌入：

$$\mathbf{h}_{\text{phase}} = \mathrm{LN}(\alpha_v\mathbf{h}'_v + \alpha_m\mathbf{h}'_m), \tag{6}$$

其中，LN(·)表示层归一化操作。$\mathbf{h}_{\text{phase}} \in \mathbb{R}^D$ 是融合后的相位表示。在实践中，该机制会动态地为不同特征分配不同权重，从而增强鲁棒性。

3）相位分组层次融合（PGHF）：从 IPR 模块获得解耦后的特定相位表示后，PGHF 旨在基于生理驱动的融合策略融合这些特征。我们首先根据形态特征，将四个阶段的嵌入划分为两个连贯的组。具体而言，P波和T/U相位被归为慢波组，因为它们呈现出相对平滑、振幅较低的波形轮廓，与去极化-复极化转变相关。相反，QRS波群和ST段阶段构成快波组，其特点是由心室活动驱动的高振幅、尖锐瞬态变化。随后，这些组由PR-GAT模块进行处理，通过门控校准实现选择性的跨相位耦合，确保不同生理模式下的特征适应保持稳定。

4）全局表征融合（GRF）：GRF作为层次架构的最终阶段，将慢波组和快波组合并为具有判别性的嵌入。与前一步类似，我们将全局融合构建为高层图交互任务，而非简单的特征拼接。我们提出的PR-GAT模块将两组的嵌入视为图节点，并通过成对交互评估其相对重要性。这使我们的方法能够动态建模慢动态与快动态之间的关系，自适应地合成整体搏动级嵌入，从而有效应对搏动间变异性。

C. 对比损失

身份验证与分类有着本质区别：前者需要将查询代码与图库模板进行匹配，而非直接分配类别标签。因此，传统分类损失并不足以胜任验证任务，因为它主要促使特征聚集在类别中心附近。此外，这种损失容易导致过拟合，使模型难以泛化到未见身份，这是验证场景中常见的挑战

。这种损失无法充分约束特征空间，因此不适用于心电图生物特征识别：在这类任务中，每个受试者仅由少量注册心搏表示，却必须与众多高度相似的负样本个体区分开来。

为缓解这一问题，我们使用基于边界的对比损失而非分类损失来优化编码器。对于小批量中的第 i 个样本，我们使用其配对的正样本 P（同一受试者），并从由小批量中所有锚点和正样本组成的候选池中挖掘负样本 N。N 根据最高余弦相似度选出（排除锚点及其配对正样本）。令 $s(\cdot,\ \cdot)$ 表示余弦相似度公式。于是，我们的损失可表示如下：

$$\mathcal{L}_{\mathrm{CL}} = \frac{1}{B}\sum_{i=1}^{B}\left[\, m + s\big(\mathbf{u}_i, \mathbf{u}_i^{N}\big) - s\big(\mathbf{u}_i, \mathbf{u}_i^{P}\big)\,\right]_{+}, \qquad (7)$$

其中，m 是边界，$[\cdot]_+ = \max(\cdot,\ 0)$。而 $\mathbf{u}_i$ 表示 GRF 生成的心搏级特征。

D. 实现

由于心电图信号是连续采集的，因此每位受试者都可以采集到多个心搏。随后，我们采用心搏感知多原型（HAM）注册策略，以确保身份验证的稳健性。与以往的工作不同，以往方法用单个特征向量表示每位受试者，而 HAM 则用一组原型表示每位受试者。

第 s 个个体的嵌入集合可表示为 $\mathcal{E}_s = \{\mathbf{u}_{s,1},\ldots,\mathbf{u}_{s,N_s}\}$，其中 N_s 表示心搏数量。为构建多原型表示，我们将 $\mathcal{E}_s$ 聚为 K 组，并用一个原型表示每组。

具体而言，第 k 个原型 $\mathbf{p}_{s,k}$ 是分配至第 k 个聚类的嵌入的均值：

$$\mathbf{p}_{s,k} = \frac{\sum_{n=1}^{N_s}\mathbb{I}(c_{s,n}=k)\,\mathbf{u}_{s,n}}{\sum_{n=1}^{N_s}\mathbb{I}(c_{s,n}=k)}, \quad k=1,\ldots,K, \qquad (8)$$

其中，$c_{s,n}\in\{1,\ldots,K\}$ 表示第 n 个嵌入的聚类分配，$\mathbb{I}(\cdot)$ 是示性函数。

在身份验证过程中，每个查询嵌入都会与图库中存储的所有原型进行匹配。将与查询距离最小的原型所对应的身份指定为最终匹配结果。这种设计有助于减轻单个心搏产生的噪声和偏差的影响，从而提高验证过程的稳健性。

IV. 实验

A. 数据集与预处理

1）数据集：我们在三个公开心电图数据集上，采用不同的实验设置评估所提出的方法。

ECGID 数据集：ECGID 数据集提供了90名受试者的310次会话。每次会话包含一条20秒的单导联I记录，分为两列（原始/滤波后分别作为第0/1列），采样率为500赫兹，分辨率为12位。会话持续时间从1天到6个月不等，均在非受限条件下进行，并带有粗略的R/T标注 [34]。

MIT-BIH 数据集：MIT-BIH 数据集包含来自 47 名受试者的 48 段半小时动态心电图片段，通常采用 MLII 和 V1/V2 导联，采样率为 360 赫兹（11 位，量程 10 mV），并附有人工核验的心搏标签（~110k 个心搏）[35]。

PTB 心电图数据集：PTB 数据集包含来自 290 名受试者的 549 条记录（每名受试者 1–5 条），包含 15 个同步导联（12 个标准导联、 + Frank XYZ），采样率为 1000 赫兹/16 位，并附有临床元数据 [36]。

2）预处理详情：为应对不同心电图数据集在导联配置、采样率和信号质量方面的差异，我们采用统一的预处理流程，以标准化模型输入。

原始信号经 0.5–40 赫兹带通滤波器去噪，以去除基线漂移和高频噪声，随后统一重采样至 200 赫兹。采用 CPS 检测 R 峰，并将每个心搏分割为四个生理阶段，包括 P波、QRS波群、ST段和 TU 波阶段。在统一的 200 赫兹采样率下，每个样本代表 0.005 秒。因此，分配给 P波、QRS波群、ST段和 T/U相位的分段长度分别为 60、40、60 和 80 个样本，对应的时长分别为 0.30 秒、0.20 秒、0.30 秒和 0.40 秒。

B. 实验设置与比较方法

1）训练与评估设置：我们在闭集和开放集设置下评估所提出的方法。两种情况下，数据均按受试者划分，而非随机拆分所有样本。

所有方法均使用 PyTorch 实现，并采用动量为 0.9 的随机梯度下降（SGD）训练 40 个周期，批量大小为 32，初始学习率为 1×10^{-4}，并使用余弦退火学习率调度。实验使用 AMD Ryzen 7 5800X CPU 和单块 NVIDIA RTX 3060 Ti GPU 进行。

与开放集协议类似，每条查询心电图都会与所有注册心电图进行匹配，并依据最小距离准则确定预测身份。

2）评估指标与比较方法：具体而言，我们在两项任务中评估所提出的方法，分别是身份验证任务和识别任务。对于身份验证任务，我们采用等错误率（EER），其定义为错误接受率与错误拒绝率相等时的值。此外，我们还使用受试者工作特征（ROC）曲线评估身份验证性能。对于识别任务，我们使用 Top-1准确率（准确率）、ROC曲线下面积（曲线下面积）和累积匹配特性（CMC）曲线评估性能。其中，CMC 用于衡量正确身份排名进入前-k 个预测结果的概率。

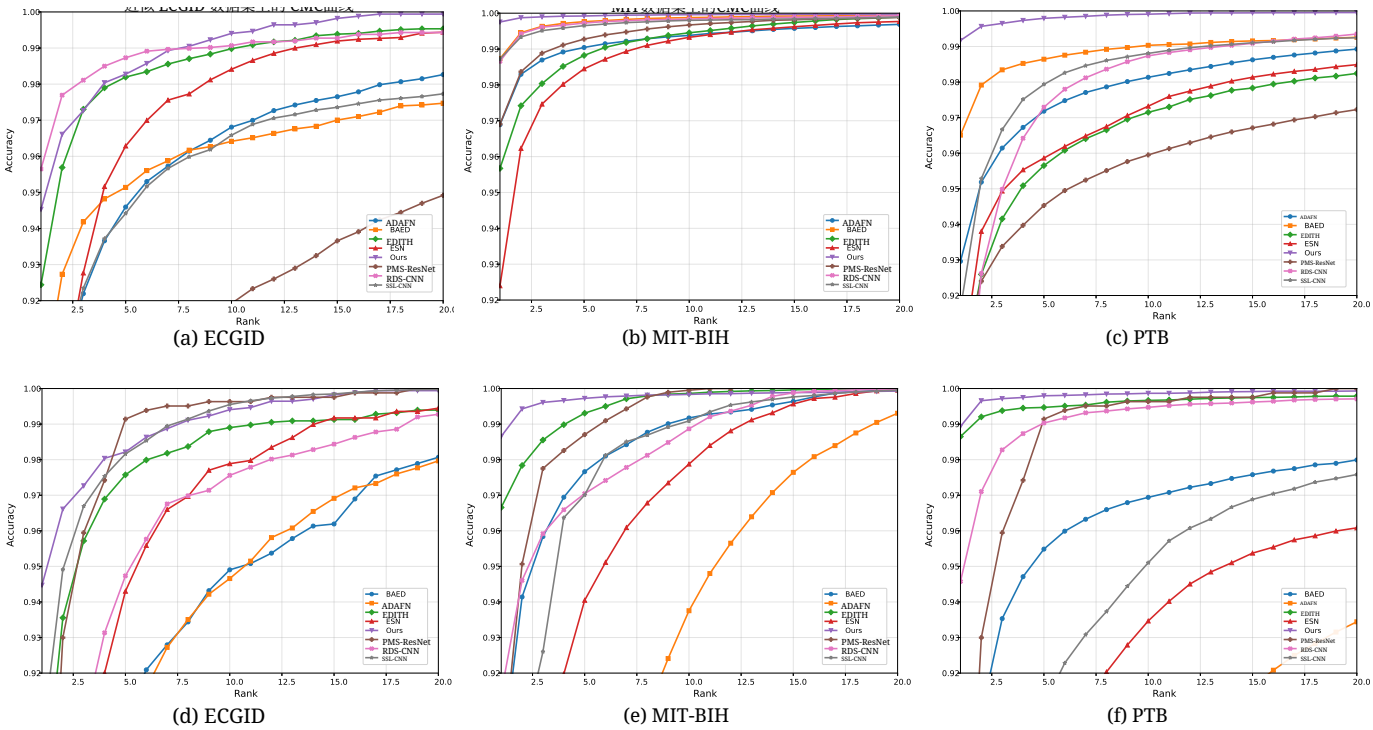
为确保比较公平且全面，我们将七种基线方法进行比较，包括 BAED [13], RDS- 卷积神经网络 [24], EDITH [10], PMS-ResNet [11], ADAFFN [28],

表I：三个数据集上的闭集分段级识别。我们的方法始终优于最先进的基线方法，展现出在受试者内条件下更优越的表征学习能力。

方法	ECGID			MIT-BIH			PTB		
	准确率（%）↑	曲线下面积（%）↑	等错误率（%）↓	准确率（%）↑	曲线下面积（%）↑	等错误率（%）↓	准确率（%）↑	曲线下面积（%）↑	等错误率（%）↓
PMS_残差网络	96.8700%	93.1958%	12.4092%	96.9625%	99.8303%	1.2697%	94.3336%	93.1958%	12.4092%
ESN	82.9200%	85.0162%	21.2550%	95.4948%	98.5032%	4.9168%	92.1900%	92.1200%	14.3600%
BAED	89.2708%	90.1200%	13.8700%	98.7302%	99.9106%	0.7120%	98.0616%	93.6600%	11.9300%
ADAFN	83.6000%	90.6300%	13.4800%	96.8900%	99.7700%	0.9200%	92.9600%	96.5300%	9.7000%
EDITH	91.8700%	90.1500%	14.5600%	95.6700%	94.0900%	10.8000%	98.8500%	96.1100%	8.3800%
SSL_CNN	88.4347%	94.2343%	12.1726%	98.7169%	12.1726%	0.4241%	98.7069%	92.9626%	13.5638%
RDS_CNN	88.3745%	90.5300%	14.9500%	98.6600%	99.1208%	4.0864%	92.5735%	90.8400%	15.8600%
Ours	97.4829%	95.4247%	9.9058%	99.7553%	99.9692%	0.1114%	99.1678%	96.8652%	8.6857%

表II：ECGID、MIT-BIH 和 PTB 上的开放集片段级识别。我们的模型在具有多样病理状况和采集设置的数据集上，对未见受试者展现出稳健的泛化能力。

方法	ECGID			MIT-BIH			PTB		
	准确率 (%) ↑	曲线下面积 (%) ↑	等错误率 (%) ↓	准确率 (%) ↑	曲线下面积 (%) ↑	等错误率 (%) ↓	准确率 (%) ↑	曲线下面积 (%) ↑	等错误率 (%) ↓
PMS_残差网络	82.4300%	84.6029%	22.7084%	89.2250%	85.0718%	23.7836%	80.7700%	82.3452%	24.5576%
ESN	71.7800%	71.2400%	34.3700%	69.4326%	66.0120%	39.5100%	70.4300%	69.5938%	35.2144%
BAED	71.2361%	69.8344%	34.1535%	89.3586%	87.3643%	19.0133%	85.4745%	94.7807%	9.6553%
ADAFN	64.6900%	65.3800%	34.9100%	57.7700%	66.6600%	37.9600%	60.5900%	63.6300%	38.0400%
EDITH	87.8700%	92.2100%	12.4300%	96.6600%	97.4600%	6.0400%	98.6500%	95.1100%	10.2800%
SSL_CNN	57.1344%	66.9551%	38.8534%	90.0811%	90.0811%	17.8188%	87.8999%	87.0397%	20.0727%
RDS_CNN	89.3680%	83.1503%	22.1800%	91.4315%	91.0700%	15.9090%	94.5760%	91.6559%	15.5920%
Ours	94.4709%	92.5548%	15.1465%	98.2179%	97.6833%	8.0360%	98.9339%	96.2483%	9.6234%



(a) ECGID　(b) MIT-BIH　(c) PTB

(d) ECGID　(e) MIT-BIH　(f) PTB

图3：不同方法在三个数据集上的 CMC 对比。子图 (a)–(c) 分别表示不同方法在 MIT-BIH、ECGID 和 PTB 数据集上的闭集设置表现，(d)–(f) 则展示了相应的开集设置结果。

SSL-CNN [12], 和 ESN [37]。所有基线方法都与我们的方法采用相同的预处理流程、输入配置和训练计划。

C. 闭集实验

首先，我们在闭集设置下评估所提出的方法，此时训练集和测试集中的身份完全相同。对于每位受试者，一半样本用于训练，另一半用于测试。

表 III：基于多个数据集的双分支基础分割消融实验。

协议	方法	ECGID			MIT-BIH			PTB		
		准确率(%)↑	曲线下面积(%)↑	等错误率(%)↓	准确率(%)↑	曲线下面积(%)↑	等错误率(%)↓	准确率(%)↑	曲线下面积(%)↑	等错误率(%)↓
闭集	仅使用CNN	88.4347	94.2343	12.1726	90.9963	99.8905	2.0421	98.7069	92.9626	13.5638
	仅使用Gabor	87.9113	94.7270	11.5600	99.6345	99.6345	1.9013	98.0266	94.6215	11.7848
	Ours	97.4829	95.4247	9.9058	99.7553	99.9692	0.1114	99.1678	96.8952	8.6857
开放集	仅使用CNN	57.1344	66.9551	38.8534	90.0811	90.0811	17.8188	87.8999	87.0397	20.0727
	仅使用Gabor	67.5981	79.7807	27.7885	85.3700	88.8950	19.3803	88.7787	89.0093	19.3619
	Ours	94.4709	92.5548	15.1465	98.2179	97.6833	8.0360	98.9339	96.2483	9.6234

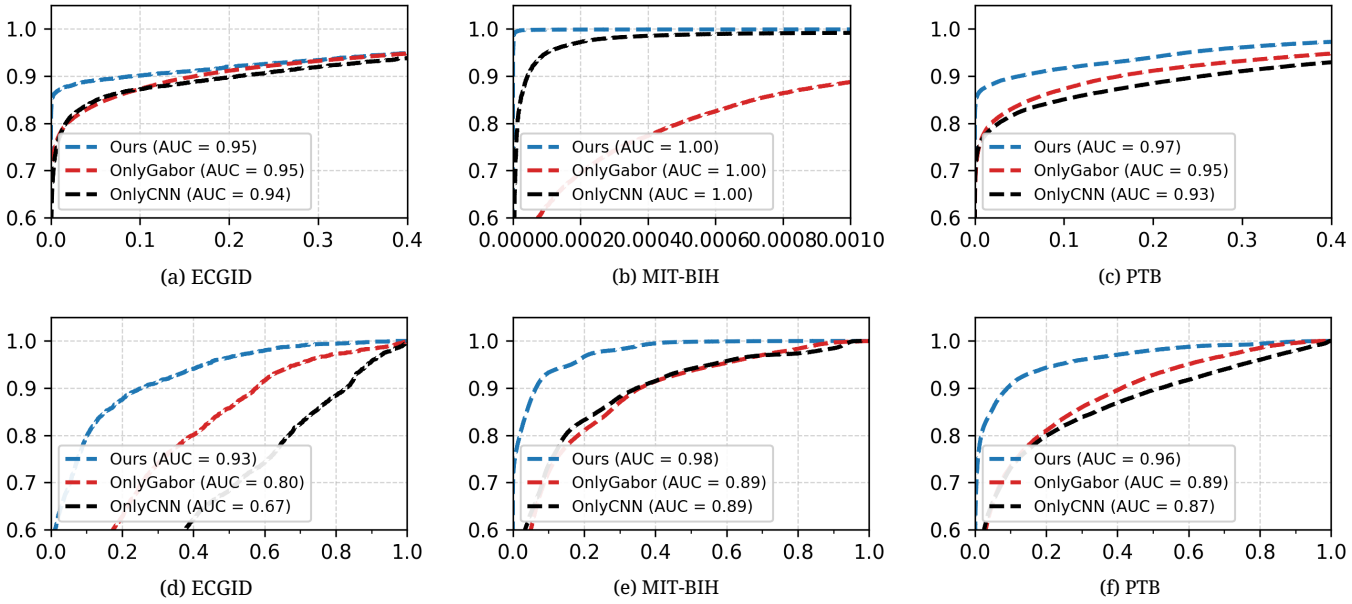


图4：不同变体的 ROC 对比。子图 (a)–(c) 分别展示了闭集设置下 ECGID、MIT-BIH 和 PTB 上的 ROC 曲线，而 (d)–(f) 则展示了开放集设置下的对应结果（数据集顺序相同）。

表IV：分阶段模块堆叠的消融实验（始终启用 IPR）。

数据集	CPS	PGHF	GRF	曲线下面积(%)↑	等错误率(%)↓
ECGID	✓	✗	✗	92.4538	15.4915
	✓	✓	✓	92.5548	15.1465
MIT-BIH	✓	✗	✗	97.0113	9.7294
	✓	✓	✓	97.6833	8.0360
PTB	✓	✗	✗	95.7800	10.3812
	✓	✓	✓	96.2483	9.6234

定量结果列于表I，CMC曲线绘于图3 (a)–(c)。在全部三个公开数据集上，我们的方法均取得了最高的身份识别准确率和最低的等错误率。在 ECGID 上，我们的方法准确率达到 97.48%，优于 PMS-ResNet（96.87%）和 BAED（89.27%）。在 MIT-BIH 数据集上，提升最为明显：我们的模型取得了 99.75% 的 Top-1 准确率和 0.11% 的等错误率。值得注意的是，该等错误率显著低于最接近的竞争方法 BAED（0.71%）和 PMS-ResNet（1.27%）。

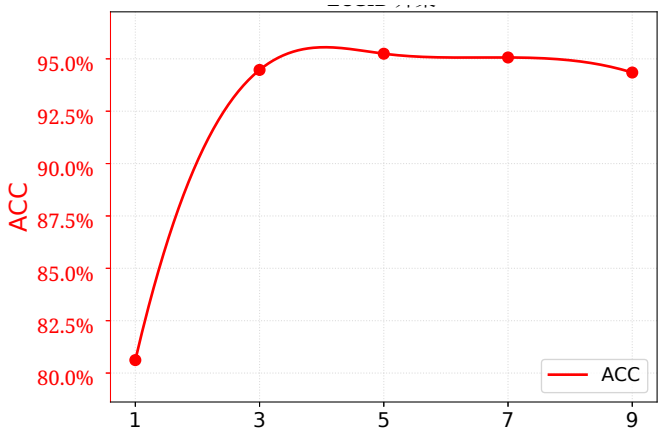


图5：HAM 中注册原型数量的消融实验。

D. 开放集实验

此外，我们还在更具挑战性的开放集设置下评估了所提出的方法。该设置假设测试受试者在训练期间从未出现过。所有受试者都被随机划分为互不重叠的训练集和

测试集，模型参数仅使用训练受试者进行训练。

对于测试集中的每位受试者，一半样本用于注册，其余一半用作查询样本。将每个查询心电图的特征与所有注册心电图的特征进行匹配，并将对应于最小距离的身份作为最终预测结果。

表I和图3(d)–(f)总结了结果。由于未见受试者存在分布偏移，各基线方法的整体性能均有所下降。SSL-CNN和ADAFN受到的影响尤为明显，在ECGID上的准确率分别仅为57.13%和64.69%。可以看出，我们的方法在这一问题上表现稳健。在所有公开数据集上，我们的方法均取得最佳结果：在ECGID上的准确率为94.47%（比排名第二的方法高5.11%），在MIT-BIH上为98.21%，在PTB上为98.93%。此外，我们可以观察到，我们的方法在几乎所有设置下都能实现最低的等错误率。CMC曲线也体现了这一优势：与其他方法相比，我们的方法在Rank-1时的身份识别率更高。

这些结果验证了我们策略的有效性和稳健性。以往方法在闭集设置下能够取得令人满意的性能，但通常无法将这种性能保持在未见过的个体上。相比之下，我们的方法通过有效建模相位特异性信息，保持了稳健的性能。

E. 消融实验

如表 III 和图4所示，我们评估了 MVFE 中各分支的有效性。所有变体均在相同的训练协议下采用相同的NPD分割协议。可以发现，基于单分支的变体在开放集协议下表现出明显局限性，而融合形态特征和变化特征能够提升跨数据集的判别能力，在 ECGID 上尤为明显。具体而言，曲线下面积从 66.96%（最佳单分支）提升至 92.55%。这些结果表明，两个分支能够捕获仅凭单一视角难以获得的互补线索。

此外，为验证 HAM 中原型数量这一超参数设置的合理性，我们在 ECGID 数据集的开放集设置下评估了性能，结果见图5。我们观察到，随着原型数量从 1 增加到 3，准确率持续提升，表明所提出的多原型设计能够有效减少单个心搏引入的噪声，从而带来更稳健的性能。然而，当原型数量进一步增加时，性能趋于饱和，且没有显著变化。因此，从经验上看，我们建议将注册原型的数量设为 3，以平衡性能和存储成本。

最后，关于层次融合策略的消融实验结果见表IV。可以注意到，将 PGHF 和 GRF 添加到 CPS 和 IPR 中能够带来稳定的性能提升，这表明每个模块都有助于细化全局嵌入并收紧特征空间结构。重要的是，在相同协议下，与所比较的其他变体相比，组合模块取得了最强的整体性能，这支持了将相位引导的层次融合与全局细化相结合有助于实现稳健的开放集识别这一结论。

V. 结论

在本研究中，我们提出了一种用于心电图身份识别的多粒度心脏相位表示框架，将每个心电图信号划分为四个具有生理意义的相位。随后，我们提出了分层相位感知特征提取框架 HPAF，由 IPR、PGHF 和 GRF 模块组成。具体而言，IPR 为每个独立相位分配特定相位的 MVFE，以提取特定相位表示；随后，PGHF 将相似相位分组，GRF 则逐步提取全局特征。我们的方法在公开数据集上取得了具有竞争力的表现。结果表明，将心电图信号表示为由特定相位线索构成的组合，而非同质的完整搏动模式，有助于提升基于心电图的身份识别鲁棒性。尽管我们的方法表现出色，但目前仍存在局限：该方法依赖准确的 R 峰定位，在严重噪声或出现异位搏动时，定位性能可能下降。未来，我们将探索自适应相位定位和无 R 峰分段方法，以实现更稳健的性能。

VI. 参考文献

[1] M. Komeili、N. Armanfard 和 D. Hatzinakos，“融合心电图与指纹进行活体检测和自动模板更新”，IEEE 信息取证与安全汇刊，第13卷，第7期，第1810–1822页，2018年7月。[2] L. Biel、O. Pettersson、L. Philipson 和 P. Wide，“心电图分析：一种人体身份识别新方法”，IEEE 仪器与测量汇刊，第50卷，第3期，第808–812页，2001年6月。[3] S. A. Israel 等人，“利用心电图识别个体”，模式识别，第38卷，第3期，第133–142页，2005年。[4] K. Weimann 和 T. O. F. Conrad，“基于深度神经网络的联邦学习：一种用于增强心电图分类的隐私保护方法”，IEEE 生物医学与健康信息学汇刊，第28卷，第11期，第6931–6943页，2024年11月。[5] J. Ma、T. Zhang 和 M. Dong，“一种采用自适应傅里叶分解并具有安全保障的电子健康应用心电图数据压缩新方法”，IEEE 生物医学与健康信息学汇刊，第19卷，第3期，第986–994页，2015年5月，doi：10.1109/JBHI.2014.2357841。[6] S.-C. Wu 等人，“心电图生物特征识别：不可关联性、不可逆性与安全性”，IEEE 物联网汇刊，第8卷，第6期，第487–500页，2020年。[7] M. A. Serhani 等人，“心电图监测系统：综述、架构、流程与关键挑战”，Sensors，第20卷，第6期，文章编号1796，2020年。[8] D. Jyotishi 和 S. Dandapat，“一种利用心电图信号进行人员识别的基于LSTM的模型”，IEEE 传感器快报，第4卷，第8期，第1–4页，2020年8月。[9] G. Petmezas 等人，“心电图数据领域最先进的深度学习方法：系统综述”，JMIR 医学信息学，第10卷，第8期，文章编号e38454，2022年。

[10] N. Ibtehaz 等人，“EDITH：借助深度学习实现可靠个体身份认证的心电图生物特征识别”，IEEE 新兴计算智能主题汇刊，第6卷，第3期，第928–940页，2021年。[11] Y. Chu 等人，“基于并行多尺度一维残差网络与中心和边界损失的心电图身份认证方法”，IEEE Access，第7卷，第51598–51607页，2019年。[12] G. Wang、S. Shanker、A. Nag、Y. Lian 和 D. John，“面向物联网边缘传感器的基于自监督学习的心电图生物特征身份认证”，IEEE 生物医学与健康信息学汇刊，第28卷，第11期，第6606–6618页，2024年，doi：10.1109/JBHI.2024.3455803。[13] A. J. Prakash 等人，“BAED：一种基于深度学习技术、采用心电图信号的安全生物特征认证系统”，生物控制论与生物医学工程，第42卷，第4期，第1081–1093页，2022年。[14] J. P. Martinez 等人，“基于小波的心电图波形界定器：在标准数据库上的评估”，IEEE 生物医学工程汇刊，第51卷，第4期，第570–581页，2004年4月。[15] A. Fratini 等人，“通过心电图分析进行个体识别”，生物医学工程在线，第14卷，第1期，第78页，2015年。[16] D. Wang 等人，“一种基于时频自编码的新型心电图生物特征身份识别方法”，Electronics，第8卷，第6期，第667页，2019年。[17] W. H. Jung 和 S. G. Lee，“基于窗口移除法进行非基准点特征提取的心电图身份识别”，应用科学，第7卷，第11期，第1205页，2017年。[18] S. C. Fang 和 H. L. Chan，“通过量化心电图相空间中的相似性与差异性进行人体身份识别”，模式识别，第42卷，第9期，第1824–1831页，2009年。[19] R. Srivastva 和 Y. N. Singh，“采用非基准点方法进行人体识别的心电图分析”，IET 生物特征识别，第8卷，第5期，第295–305页，2019年9月。[20] A. Goshvarpour 和 A. Goshvarpour，“利用新型基于匹配追踪的心电图特征集进行人体身份识别”，生物医学计算机方法与程序，第172卷，第87–94页，2019年4月。[21] S. Gutta 和 Q. Cheng，“面向基于心电图的生物特征识别的联合特征提取与分类器设计”，IEEE 生物医学与健康信息学汇刊，第20卷，第2期，第460–468页，2016年3月。[22] D. A. AlDuwaile 和 M. S. Islam，“利用卷积神经网络和单次心跳进行心电图生物特征识别”，熵，第23卷，第6期，文章编号733，2021年。[23] J. Huang 等人，“利用基于STFT的频谱图和卷积神经网络进行心电图心律失常分类”，IEEE Access，第7卷，第92871–92880页，2019年。[24] E. Ihsanto 等人，“一种采用残差深度可分离卷积神经网络的快速精准心电图身份认证算法”，应用科学，第10卷，文章编号3304，2020年。[25] K. J. Chee 和 D. A. Ramli，“利用Transformer自注意力机制作为序列对特征提取器，并实现灵活注册范围识别的心电图生物特征识别”，Sensors，第22卷，第9期，文章编号3446，2022年。[26] D. Jyotishi 和 S. Dandapat，“一种采用带注意力机制的分层长短期记忆网络的心电图生物特征识别系统”，IEEE 传感器期刊，第22卷，第6期，第6052–6061页，2022年。[27] M. Hammad 等人，“用于基于心电图信号进行人体身份认证的ResNet注意力模型”，专家系统，第38卷，第6期，文章编号e12547，2021年。[28] P. Yi 等人，“基于注意力增强型域自适应特征融合网络的心电图生物特征识别”，IEEE Access，第12卷，第1291–1307页，2023年。[29] L. Sun 等人，“用于心电图身份识别的随机注意力与双路径系统”，工程应用中的人工智能，第132卷，文章编号107883，2024年。[30] K. Wang 和 N. Wang，“利用双层级特征、协同嵌入和维度注意力权重学习进行心电图生物特征识别”，Sensors，第25卷，第17期，文章编号5343，2025年。[31] T. Ma 等人，“面向心电图生物特征识别的分布外表征与图神经网络融合学习”，IEEE 生物特征、行为与身份科学汇刊，第7卷，233，2025年2，225第。–期[32] M. Al Alfi 等人，“利用12导联心电图信号和图卷积网络增强生物特征识别”，数字健康前沿，第7卷，文章编号1547208，2025年。[33] T. S. Kumar 和 V. Kanhangad，“基于Gabor滤波器的一维局部相位描述符：利用单导联心电图检测阻塞性睡眠呼吸暂停”，IEEE 传感器快报，第2卷，第1期

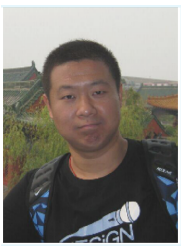
，第 1–4 页，2018年。

[34] T. Lugovaya，《基于心电图的人体生物特征识别》，理学硕士论文，圣彼得堡国立理工大学，俄罗斯圣彼得堡，2005年。[35] G. B. Moody 和 R. G. Mark，《MIT-BIH心律失常数据库的影响》，IEEE Eng. Med. Biol. Mag.，第20卷，第3期，第45–50页，2001年。[36] R. Bousseljot 等，《通过互联网使用 PTB 的 CAR-DIODAT 心电图信号数据库》，Biomed. Tech.，第40卷，第1期，第317–318页，1995年。[37] M. Hazratifard 等，《在智慧医疗系统中使用心电图信号进行人体身份认证的集成孪生网络（ESN）》，Sensors，第23卷，第10期，文章编号4727，2023年。



第一作者 JINTAO HUANG 于 2025年 获得南昌航空大学软件学院工学学士学位。目前正在中国南昌航空大学攻读硕士学位。他的

研究 h i兴趣 includ计算机视觉 i -化、生物特征识别，以及生成式人工智能的应用。



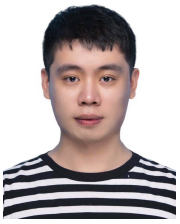
第二位，Lu Leng 于2012年在中国成都西南交通大学获得博士学位。他曾在韩国首尔延世大学和中国南京南京航空航天大学从事博士后研究。他曾是美国西弗吉尼亚大学和韩国延世大学的访问学者。目前，他是南昌航空大学计算机视觉研究院的正教授兼院长。

Leng教授已发表100余篇国际期刊和会议论文。他曾获得多项奖学金和科研项目资助，其中包括六项国家自然科学基金委员会（NSFC）资助的项目。他担任100余种国际期刊和50个会议的审稿人。他的研究兴趣包括计算机视觉、生物特征模板保护、生物特征识别、医学图像处理、信息隐藏等。



第三位 Yi Zhang 于2005年、2008年和2012年分别获得四川大学计算机学院计算机科学与技术专业的理学学士、理学硕士和博士学位。

2014年至2015年，他在美国纽约州特洛伊的伦斯勒理工学院生物医学工程系担任博士后研究员。他目前是四川大学网络空间安全学院正教授，并且还担任深度成像组（DIG）主任。他在图像处理领域发表了140多篇论文，刊载于或合著于多家顶尖期刊和会议，包括IEEE医学影像汇刊、IEEE信息取证与安全汇刊、MedIA、IJCV和CVPR；相关成果还曾由物理学会（IOP）报道，并在林岛诺贝尔奖得主大会上展示。他的研究兴趣包括医学影像、压缩感知和深度学习。他曾获得国家重点研发计划、国家自然科学基金委员会以及中国四川省科技支撑计划的重大资助。他还担任《国际生物医学成像杂志》和《传感与成像》的客座编辑，以及IEEE医学影像汇刊和IEEE辐射与等离子体医学科学汇刊的副编辑。



第四作者 Ziyuan Yang 目前是香港中文大学电子工程系的博士后研究员。他于2021年在中国南昌的南昌大学信息工程学院获得计算机科学理学硕士学位，并在中国四川大学计算机学院获得博士学位。

他曾在新加坡科技研究局（A*STAR）前沿人工智能研究中心担任研究实习生。技术与研究。他近几年在机器学习领域的顶级会议和期刊上发表了50余篇论文，包括 CVPR、AAAI、IJCV、IEEE T-IFS、IEEE T-NNLS、IEEE T-SMCS 和 IEEE T-AI。他曾为顶级期刊或会议担任审稿人，例如 IEEE T-PAMI、IEEE T-TIP、IEEE T-IFS、IEEE T-MI、CVPR 和 ICCV。