



基于扩散模型的心电信号去噪方法

罗成思, 张凯旋, Abduljabbar Salem Ba-Mahel, 饶妮妮*

(电子科技大学 生命科学与技术学院, 成都 610054)

摘要 传统和深度学习的去噪技术在处理心电 (Electrocardiogram, ECG) 信号特定类型的噪声和数据泛化的验证方面存在不足。为此, 提出一种基于扩散模型的生成式 ECG 去噪模型, 该模型利用模拟数据学习干净 ECG 分布的得分函数, 基于欧拉法求解常微分方程 (ODE) 生成和分离出 ECG 和噪声。该模型在模拟数据上进行了训练, 并在独立的真实数据集上进行了验证。研究表明, 与其他相关方法比较, 该模型在去除多样性噪声以及保持 ECG 中不同振幅特征波形的一致性方面具有显著优势。

关键词 ECG; 扩散模型; 去噪; 神经网络; 信号分离

中图分类号 TP183; R318

文献标志码 A

DOI 10.12178/1001-0548.2023244

Denoising Method of ECG Based on Diffusion Model

LUO Chengsi, ZHANG Kaixuan, BA-MAHEL Abduljabbar Salem, and RAO Nini*

(School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China)

Abstract Traditional and deep learning denoising techniques exhibit shortcomings in handling specific types of noise and data generalization validation in Electrocardiogram (ECG) signals. This paper proposes a generative ECG denoising model based on diffusion models, which leverages simulated data to learn the score function of clean ECG distribution and generates and separates ECG and noise based on the Euler method for solving Ordinary Differential Equations (ODE). The model is trained on simulated data and validated on an independent real dataset. Compared to other relevant methods, the obtained results demonstrate that this model has significant advantages in removing diverse noise and maintaining consistency in ECG waveforms of different amplitude features.

Key words ECG; diffusion model; denoising; neural network; signal separation

心电图 (Electrocardiogram, ECG) 是一种常用的非侵入性检查方法, 常用于诊断心脏疾病。随着计算机技术的不断发展, 基于计算机的自动化 ECG 信号分析逐渐成为一种流行趋势。这种自动化分析不仅可以大大节省处理时间, 还能减少人力成本。然而, ECG 信号在采集和传输过程中容易受到各种类型噪声的干扰, 如基线漂移、白噪声、工频干扰和肌电伪迹等, 这些噪声会严重影响 ECG 信号的质量, 进而对自动化分析的准确性和可靠性构成威胁, 从而导致临床误诊与漏诊。

ECG 信号去噪一直是心电学领域的研究热点之一。尽管已有多多种去噪算法被提出, 如小波变

换^[1-2]、基于字典学习^[3-4]的方法和基于经验模态分解^[5]的方法等, 但这些方法都存在局限性。首先, 现有算法在去除复杂噪声, 如高斯噪声和肌电伪迹方面表现不佳; 其次, 性能优秀的传统去噪方法^[6]通常依赖于波形定位的专家标注信息。如果没有这些先验信息, 在噪声干扰下自动定位波形将带来误差, 导致这些方法的性能下降。

尽管深度学习算法在 ECG 去噪性能上表现比较出色^[6], 但过往大多数方法在模型训练和测试数据的利用上存在一些问题。如一些方法所使用的 ECG 和噪声数据通常来自单一来源的数据库; 有些甚至采用一个 ECG 记录的不同分段分别训练和测试模型^[7-9]。这样的训练测试方式难以呈现模型

收稿日期: 2023-09-26; 修回日期: 2023-11-26

基金项目: 国家自然科学基金 (62271127); 四川省自然科学基金 (2023NSFSC0627); 电子科技大学与四川大学华西医院医工合作基金 (ZYGX2022YGRH011, HXDZ22005)

作者简介: 罗成思, 主要从事生物医学信号处理方面的研究。

*通信作者 E-mail: raonn@uestc.edu.cn

的泛化能力, 影响模型的可靠性。此外, 去噪方法在性能评价中常用的指标大多是基于幅值的平方, 只能反映整体的效果, 很难判断去噪算法是否对各个波形产生了影响。

扩散模型在如语音信号的去噪任务中取得了显著效果^[10], 对于信号去噪具有很高的适用性和潜力。实际上, 在文献^[11]中, 扩散模型也已经被成功应用于 ECG 的去噪, 并取得了优于其他相关方法的性能。然而, 在这个应用中的扩散模型也存在一些局限性, 如需要 ECG 的波形定位先验信息, 也存在由于迭代次数多使得运算效率较低等问题。

为了解决这些问题, 本文提出了一种基于扩散模型的 ECG 信号去噪方法。该方法充分利用扩散模型学习 ECG 先验分布的能力, 以生成的方式对 ECG 的心房、心室和噪声 3 种信号进行分离来完成去噪。

1 数据集

生成模拟的 ECG 和噪声信号对扩散模型进行训练, 然后使用真实世界的 ECG 和噪声数据对模型进行测试。用于测试的真实 ECG 来自于 6 个数据库, 包括: MIT-BIH Arrhythmia Database (MITDB)^[12]、QT Database (QTDB)^[13]、MIT-BIH Normal Sinus Rhythm Database (NSRDB)^[14]、MIT-BIH Arrhythmia Database P-Wave Annotations (MIT PDB)^[15]、2019 China Physiological Signal Challenge Dataset (CPSC2019)^[16]、Glasgow University Database (GUDB)^[17], 真实噪声来自 MIT-BIH Noise Stress Test Database (NSTDB)^[18]。

本研究借助文献^[19]开发的程序, 生成了 500 段含 15 个导联的心房和心室 ECG 数据, 每段数据持续 1 min, 采样频率为 1 kHz。这些数据简称为 SIMU, 涵盖了正常窦性心律、房性心动过速、二联律、三联律以及室性早搏这 5 类常见心律。SIMU 有效地保留了 ECG 的主要特性, 包括各类波形形态和波群之间的间期等。文献^[19]的研究也验证了即使是经验丰富的医生也无法准确地区分出模拟 ECG 与真实 ECG, 这说明模拟 ECG 足以作为真实 ECG 数据的有效代替参与神经网络的训练。将 SIMU 中 450 段数据用于模型的训练, 剩余的 50 段用于验证。

在训练过程中, 对 SIMU 同一信号中的 15 个

导联的心房和心室数据进行随机组合, 由此将数据扩充至 101 250 段 (即 $15 \times 15 \times 450$)。同时, 将心房和心室信号的幅度进行随机正负和缩放, 缩放因子在 $\pm[0.5, 1.5]$ 范围内随机选取。此外, 还随机引入文献^[20]提供的噪声, 以及自行生成的高斯白噪声和基线漂移噪声, 作为训练阶段的 ECG 噪声数据。

MITDB、QTDB 和 NSRDB 常用作干净的 ECG 数据, 用于验证去噪算法的性能^[6], 也将它们用作干净 ECG 数据进行模型消除噪声的测试数据。需要注意的是, MITDB、QTDB 和 NSRDB 中的 ECG 信号大多不是完全干净的, 即存在背景噪声。

本研究拟消除 4 种类型的噪声, 分别为: 高斯白噪声 (White Gaussian Noise, WGN)、基线漂移 (Baseline Wander, BW)、工频干扰 (Power-line Interference, PLI) 以及肌电伪迹 (Muscle Artifact, MA)。其中, NSTDB 数据集已包含了 BW 和 MA 两种噪声。除此之外, 将自行生成 WGN, 并参照文献^[4]中的式 (5) 来生成 PLI。

因为真实数据中没有单独的心房信号, 本研究将使用 MIT PDB 数据库中的 P 波峰值位置标注信息, 并结合 P 波定位算法, 以验证本文模型在分离心房信号的准确性。此外, 还将结合 R 波^[21]和 T 波末端识别算法^[22], 利用 CPSC2019、GUDB 和 QTDB 进一步验证模型在分离心室信号的精确性。对以上所有数据, 均执行了重采样操作, 将采样频率统一设定为 250 Hz。此外, 将各个信号非重叠地划分为 10 s 一段, 将每段作为输入送入模型进行处理。

2 生成式的 ECG 信号与噪声分离方法

2.1 信号与噪声分离模型

将混入噪声的 ECG 信号视为由心房、心室和噪声 3 个源信号组成的混合信号:

$$\mathbf{y} = \mathbf{x}^a + \mathbf{x}^v + \mathbf{x}^n \quad (1)$$

式中, $\mathbf{y} \in \mathbb{R}^l$ 为长度为 l 、混入了噪声的真实 ECG; $\mathbf{x}^a \in \mathbb{R}^l$ 和 $\mathbf{x}^v \in \mathbb{R}^l$ 分别表示 ECG 中的心房和心室信号, 它们共同构成完整的干净 ECG, 表示为 $\mathbf{z} = \mathbf{x}^a + \mathbf{x}^v$; $\mathbf{x}^n \in \mathbb{R}^l$ 表示噪声信号。

采用基于得分匹配 (score matching) 的扩散模型^[23], 通过如图 1 所示的逆向常微分方程 (Ordinary Differential Equation, ODE) 对后验分布 $p(\mathbf{x}^a, \mathbf{x}^v, \mathbf{x}^n | \mathbf{y})$ 进行采样, 可以从 \mathbf{y} 中分离出 \mathbf{x}^a 、 \mathbf{x}^v 和 \mathbf{x}^n , 从而达

到噪声去除的目的。图 1 从左到右表示前向加噪过程，通过对信号逐步加入高斯噪声，直至信号完全被噪声淹没。利用这一过程训练神经网络模型学习加噪后信号分布的得分函数（score function） $\nabla_{\mathbf{x}_\sigma} p(\mathbf{x}_\sigma | \mathbf{y})$ 。图 1 从右向左表示逆向过程，将初始条件设为高斯噪声，利用学习到的 $\nabla_{\mathbf{x}_\sigma} p(\mathbf{x}_\sigma | \mathbf{y})$ 通过求解逆向 ODE 来分离信号。

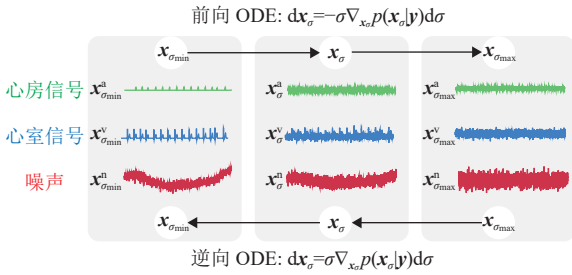


图 1 生成式的 ECG 信号与噪声分离示意图

将分离的 \mathbf{x}^a 、 \mathbf{x}^v 和 \mathbf{x}^n 依次连接起来，构成一个 $l \times 3$ 的矩阵，用 $\mathbf{x} = [\mathbf{x}^a, \mathbf{x}^v, \mathbf{x}^n] \in \mathbb{R}^{l \times 3}$ 表示。在后续的讨论中，为了表达的简洁性，也将省略 \mathbf{y} 的书写，同时，所有的得分函数和分布都是在 \mathbf{y} 的条件下进行的。因为本文没有对 ECG 作额外幅值上的任何处理，所以后续 ECG 信号幅度的单位都是 mV。

2.2 模型的训练

将数据 \mathbf{x} 的概率分布表示为 $p_{\text{data}}(\mathbf{x})$ 。上述扩散模型的训练分为两个步骤：1) 在 \mathbf{x} 中加入不同尺度的高斯噪声 $\boldsymbol{\varepsilon}_\sigma \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ ，得到带有噪声的数据 $\mathbf{x}_\sigma = \mathbf{x} + \boldsymbol{\varepsilon}_\sigma$ ，其分布记为 $p(\mathbf{x}_\sigma) = \int_{\mathbb{R}^d} \mathcal{N}(\mathbf{x}_\sigma; \mathbf{x}, \sigma^2 \mathbf{I}) p_{\text{data}}(\mathbf{x}) d\mathbf{x}$ ；2) 将 \mathbf{x}_σ 输入去噪神经网络 $D_\theta(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ ，用 $D_\theta(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ 拟合去噪后的信号 \mathbf{x} 。这一步又称为降噪得分匹配（denoising score matching）^[23]。其中， θ 表示神经网络的参数， $c_{\text{noise}}(\sigma)$ 是一个需要显式带入到神经网络中的标量项，它将噪声尺度 σ 映射为 D_θ 的条件输入。需要注意的是，训练和采样扩散模型所使用的噪声 $\boldsymbol{\varepsilon}_\sigma$ 与混入真实 ECG 中的噪声 \mathbf{x}^n 没有任何关系， $\boldsymbol{\varepsilon}_\sigma$ 是一个在模型训练和采样过程中需要的中间变量。

为了提高神经网络的训练效率，训练中需要调节高斯噪声 $\boldsymbol{\varepsilon}_\sigma$ 的尺度 σ 和各尺度下噪声在训练中出现的频率。参照文献 [23] 的做法，通过对数正态分布来安排训练中的 σ ，即 $\ln(\sigma) \sim \mathcal{N}(P_{\text{mean}}, P_{\text{std}}^2)$ 。不同的是，本文将 σ 和 c_{noise} 之间的关系定义为 $c_{\text{noise}}(\sigma) := \text{cdf}((\ln(\sigma) - P_{\text{mean}})/P_{\text{std}})$ ，其中 cdf 表示标准正态分布 $\mathcal{N}(0, 1)$ 的累积分布函数（后文用 icdf

表示 cdf 的逆函数）。训练中只需要从均匀分布 $\text{Unif}(0, 1)$ 中采样 c_{noise} ，然后通过 $\sigma = e^{\text{icdf}(c_{\text{noise}}) P_{\text{std}} + P_{\text{mean}}}$ 就可以生成服从对数正态分布的 σ 。这些参数的具体设置是根据采样生成（见 2.3 节）方法的要求进行。

在训练中沿用了文献 [23] 对神经网络 $D_\theta(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ 进行的预归一化方法：

$$D_\theta(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma)) = c_{\text{skip}}(\sigma) \mathbf{x}_\sigma + c_{\text{out}}(\sigma) F_\theta(c_{\text{in}}(\sigma) \mathbf{x}_\sigma; c_{\text{noise}}(\sigma)) \quad (2)$$

预归一化方法可以理解为将 $D_\theta(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ 分解为可学习部分和预设的部分。其中，可学习部分为 $F_\theta(c_{\text{in}}(\sigma) \mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ ，它包括了去噪神经网络所有的可学习的模块；预定义部分的所有项 $c_{\text{skip}}(\sigma)$ 、 $c_{\text{out}}(\sigma)$ 和 $c_{\text{in}}(\sigma)$ 都事先设置好，使 $F_\theta(c_{\text{in}}(\sigma) \mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ 的输入 $c_{\text{in}}(\sigma) \mathbf{x}_\sigma$ 的方差和输出的方差都约等于 1。此外，预归一化要求事先计算信号的标准差 σ_{data} 作为 $c_{\text{skip}}(\sigma)$ 、 $c_{\text{out}}(\sigma)$ 和 $c_{\text{in}}(\sigma)$ 的超参数（详细设定请参见表 1 和表 2）。

表 1 模型训练和采样生成信号中的方法设置

训练	
噪声尺度分布	$\ln(\sigma) \sim \mathcal{N}(P_{\text{mean}}, \text{diag}(P_{\text{std}}^2) \mathbf{I})$
$c_{\text{noise}}(\sigma)$	$\text{cdf}(\ln(\sigma) - P_{\text{mean}}) / P_{\text{std}}$
预归一化	
$c_{\text{in}}(\sigma)$	$1 / \sqrt{\sigma_{\text{data}}^2 + \sigma^2}$
$c_{\text{skip}}(\sigma)$	$\sigma_{\text{data}}^2 / (\sigma_{\text{data}}^2 + \sigma^2)$
$c_{\text{out}}(\sigma)$	$\sigma \sigma_{\text{data}} / \sqrt{\sigma_{\text{data}}^2 + \sigma^2}$
采样	
ODE 求解	MSDM Dirac ^[25]
时间步	$\sigma_i = \sigma_{\text{max}}^{1/p} + \frac{i}{N-1} (\sigma_{\text{min}}^{1/p} - \sigma_{\text{max}}^{1/p})^p$

表 2 模型训练和采样生成信号中的参数设置

带有注意力机制的 Unet 各类参数	取值
σ_{min}	$[1 \times 10^{-4}, 1 \times 10^{-3}, 2 \times 10^{-3}]$
σ_{max}	[10, 100, 200]
σ_{data}	[0.05, 0.5, 1]
ρ	10
P_{mean}	[-3.453, -1.151, -0.458]
P_{std}	[1.547, 1.547, 1.547]
N	5

基于噪声尺度 σ 的采样，可以得到模型的加权整体训练损失^[23]为：

$$\mathcal{L} = E_{c_{\text{noise}}(\sigma) \sim \text{Unif}(0, 1)} \left[\frac{1}{c_{\text{out}}(\sigma)^2} \mathcal{L}(D_\theta; c_{\text{noise}}(\sigma)) \right] \quad (3)$$

式中, $1/c_{\text{out}}(\sigma)^2$ 为权重; $\mathcal{L}(D_\theta; c_{\text{noise}}(\sigma))$ 为给定 σ 的降噪得分匹配损失:

$$\mathcal{L}(D_\theta; c_{\text{noise}}(\sigma)) = E_{\mathbf{x} \sim p_{\text{data}}} E_{\boldsymbol{\varepsilon}_\sigma \sim \mathcal{N}(0, \sigma^2 \mathbf{I})} \|D_\theta(\mathbf{x} + \boldsymbol{\varepsilon}_\sigma; c_{\text{noise}}(\sigma)) - \mathbf{x}\|_2^2$$

最后, 将训练得到的最优去噪神经网络记为 $D_{\theta^*}(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$, 在理想情况下它与 $p(\mathbf{x}_\sigma)$ 的得分函数 $\nabla_{\mathbf{x}_\sigma} p(\mathbf{x}_\sigma)$ 之间的关系为:

$$\nabla_{\mathbf{x}_\sigma} p(\mathbf{x}_\sigma) = (D_{\theta^*}(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma)) - \mathbf{x}) / \sigma^2 = S_{\theta^*}(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma)) \quad (4)$$

由式 (4) 可知, 可以直接使用 $S_{\theta^*}(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ 来代替 $\nabla_{\mathbf{x}_\sigma} p(\mathbf{x}_\sigma)$ 生成 \mathbf{x} 。注意, 此时条件得分函数为 $S_{\theta^*}(\mathbf{x}_\sigma; \mathbf{y}, c_{\text{noise}}(\sigma))$ 。

2.3 采样生成信号

通过求解逆向 ODE 来生成信号, 一般扩散模型的 ODE^[23] 可以表示为:

$$d\mathbf{x}_{\sigma(t)} = -\dot{\sigma}(t)\sigma(t)\nabla_{\mathbf{x}_{\sigma(t)}} \log p(\mathbf{x}_{\sigma(t)}) dt \quad (5)$$

设置 $\sigma(t) = t$, 并代入式 (4), 则式 (5) 可以表示为:

$$d\mathbf{x}_\sigma = -\sigma S_{\theta^*}(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma)) d\sigma \quad (6)$$

因此, 可以利用式 (6) 将噪声尺度 σ_a 的样本 \mathbf{x}_{σ_a} 演化为噪声尺度为 σ_b 的样本 \mathbf{x}_{σ_b} 。这里需要确定最小和最大的噪声尺度 σ_{\min} 和 σ_{\max} 作为 ODE 的边界, 让其尽量满足 $p(\mathbf{x}_{\sigma_{\min}}) \approx p_{\text{data}}(\mathbf{x})$ $p(\mathbf{x}_{\sigma_{\max}}) \approx \mathcal{N}(0, \sigma_{\max}^2 \mathbf{I})$ 的要求^[24], 以便求解 ODE 的逆向过程, 完成从纯净的高斯噪声直接生成信号的任务。

由于心房信号、心室信号和噪声的幅值存在较大差异, 如正常人 R 波 (心室信号) 的最大幅值大约为 P 波 (心房信号) 最大幅值的 10 倍^[26]。受基线漂移等干扰信号的影响, 噪声信号的幅值波动较大。经过初步计算, 并参考文献 [23] 中的基准, 本文将 3 类信号的标准差设定为 $\sigma_{\text{data}} = [0.05, 0.5, 1]$ 。鉴于此, 本文也对 \mathbf{x}^a 、 \mathbf{x}^v 和 \mathbf{x}^n 成比例地设置了不同尺度的 σ_{\min} 和 σ_{\max} , 分别为:

$$\sigma_{\min} = [\sigma_{\min}^a, \sigma_{\min}^v, \sigma_{\min}^n] = [1 \times 10^{-4}, 1 \times 10^{-3}, 2 \times 10^{-3}]$$

$$\sigma_{\max} = [\sigma_{\max}^a, \sigma_{\max}^v, \sigma_{\max}^n] = [10, 100, 200]$$

在信号与噪声分离模型的训练中, 本文也设置了与上述尺度适配的 \mathbf{P}_{mean} 和 $\mathbf{P}_{\text{std}} \in \mathbb{R}^{1 \times 3}$, 使之满足 $c_{\text{noise}}^s(\sigma_{\min}^s) = p_{\sigma_{\min}}$ 以及 $c_{\text{noise}}^s(\sigma_{\max}^s) = p_{\sigma_{\max}}$ 。其中 $s \in \{a, v, n\}$, $p_{\sigma_{\min}}$ 和 $p_{\sigma_{\max}}$ 分别对应着训练时采样的

噪声尺度小于 σ_{\min}^s 和大于 σ_{\max}^s 的概率, 将其设置为对称的关系: $p_{\sigma_{\min}} = 1 \times 10^{-4}$ 和 $p_{\sigma_{\max}} = 1 - p_{\sigma_{\min}}$ 。于是, $\mathbf{P}_{\text{mean}}^s$ 和 $\mathbf{P}_{\text{std}}^s$ 可设置为:

$$\begin{aligned} \mathbf{P}_{\text{mean}}^s &= (\ln(\sigma_{\max}^s) + \ln(\sigma_{\min}^s)) / 2 \\ \mathbf{P}_{\text{std}}^s &= (\ln(\sigma_{\max}^s) - \ln(\sigma_{\min}^s)) / (\text{icdf}(p_{\sigma_{\min}}) - \text{icdf}(p_{\sigma_{\max}})) \end{aligned} \quad (7)$$

通过以上计算得到的 \mathbf{P}_{mean} 和 \mathbf{P}_{std} 的数值, 见表 2。由于 $c_{\text{noise}}^a(\sigma^a) = c_{\text{noise}}^v(\sigma^v) = c_{\text{noise}}^n(\sigma^n)$, 所以在训练中只需要采样一个服从均匀分布的标量 c_{noise} 就可以生成 3 类信号对应的噪声尺度。

基于欧拉法来求解式 (6) 的 ODE, 以完成信号的生成。求解过程中采用了式 (8) 所示的非线性采样步长^[23]:

$$\sigma_i = \sigma_{\max}^{1/\rho} + \frac{i}{N-1} (\sigma_{\min}^{1/\rho} - \sigma_{\max}^{1/\rho})^\rho \quad (8)$$

式中, $\sigma = [\sigma^a, \sigma^v, \sigma^n]$; N 表示所划分 σ_i 的个数, 且 $0 \leq i < N$, 实际 ODE 进行迭代计算的步数是 $N-1$, 本文选取 $\rho = 10$ 。式 (8) 的步长安排同样也能满足 $c_{\text{noise}}^a(\sigma^a) = c_{\text{noise}}^v(\sigma^v) = c_{\text{noise}}^n(\sigma^n)$ 的条件。

进一步发现, 将 \mathbf{x} 初始化为 $\mathbf{0}$ 可以大大降低迭代的步数并提高性能。此外还发现, 使用文献 [25] 提出的 MSDM Dirac 技巧也可以提高生成信号的质量。因此, 本文采用了基于 MSDM Dirac 的信号生成算法, 选取 $N = 5$, 其算法伪代码如下。

```

Input:  $N, \{\sigma_i\}_{i \in \{0, \dots, N-1\}}, \mathbf{y}$ 
Define  $\mathbf{x} := [\mathbf{x}^a, \mathbf{x}^v, \mathbf{x}^n], \boldsymbol{\sigma} := [\sigma^a, \sigma^v, \sigma^n]$ 
Define  $\text{Score} := [\text{Score}^a, \text{Score}^v, \text{Score}^n]$ 
Initialize  $\mathbf{x}_0 \leftarrow [\mathbf{0}, \mathbf{0}, \mathbf{0}] \in \mathbb{R}^{I \times 3}$ 
for  $i \leftarrow 0$  to  $N-2$  do
   $[\text{Score}^a, \text{Score}^v, \text{Score}^n] \leftarrow S_{\theta^*}(\mathbf{x}_i; \mathbf{y}, c_{\text{noise}}(\sigma_i))$ 
   $\text{Score}^a \leftarrow \text{Score}^a - \text{Score}^n$ 
   $\text{Score}^v \leftarrow \text{Score}^v - \text{Score}^n$ 
   $\mathbf{x}_{i+1}^a \leftarrow \mathbf{x}_i^a - \sigma_i^a (\sigma_{i+1}^a - \sigma_i^a) \text{Score}^a$ 
   $\mathbf{x}_{i+1}^v \leftarrow \mathbf{x}_i^v - \sigma_i^v (\sigma_{i+1}^v - \sigma_i^v) \text{Score}^v$ 
   $\mathbf{x}_{i+1} \leftarrow [\mathbf{x}_{i+1}^a, \mathbf{x}_{i+1}^v, \mathbf{y} - \mathbf{x}_{i+1}^a - \mathbf{x}_{i+1}^v]$ 
Output:  $\mathbf{x}_{N-1}$ 

```

2.4 模型的主干网络架构

如图 2 所示为本文构建的信号噪声分离主干网络, 即 $D_\theta(\mathbf{x}_\sigma; c_{\text{noise}}(\sigma))$ 。该网络结构借鉴文献 [27], 采用了添加互注意力机制的 Unet 结构。网络由条件输入编码器、噪声尺度编码器和去噪编/解码器 3 个部分组成。其中, 条件输入编码器和去噪编/解

码器各有 9 层。在条件输入编码器和去噪编码器的顶部 4 层加入了互注意力机制，以融合条件输入编码器和去噪编码器分别提取的 y 和 x 的深层特征。其中，条件输入编码器的特征作为键 (key) 和值 (value)，去噪编码器的特征作为查询值 (query)，将注意力模块的输出以加法形式融合到去噪编码器的特征中。另外还将条件输入 y 与经过预归一化的 x 在通道上进行堆叠，作为去噪编码器的输入。本文网络在深层和浅层都融合了 y 和 x 的信息，网络深层的特征融合能够有效利用编码器本身的压缩去噪能力，浅层的融合可以保留原始

信号 y 中的细节。本文对 c_{noise} 进行傅里叶编码^[28]，编码长度为 128，最大频率为 64，并用黄色方框表示编码后的 c_{noise} 。最后，采用了宽度为 4、下采样率为 2 且没有偏置项的卷积核，每个卷积后的输出都以加法形式在通道上结合 c_{noise} 的编码信息 (黄色方框)，并在之后进行了组归一化 (Group norm) 和 Sigmoid 与线性函数相乘的激活函数 (Swish)^[29] 激活。条件输入编码器和去噪编/解码器的 9 层卷积核的通道数被统一设置为 [8, 8, 16, 16, 16, 16, 32, 32, 32]。

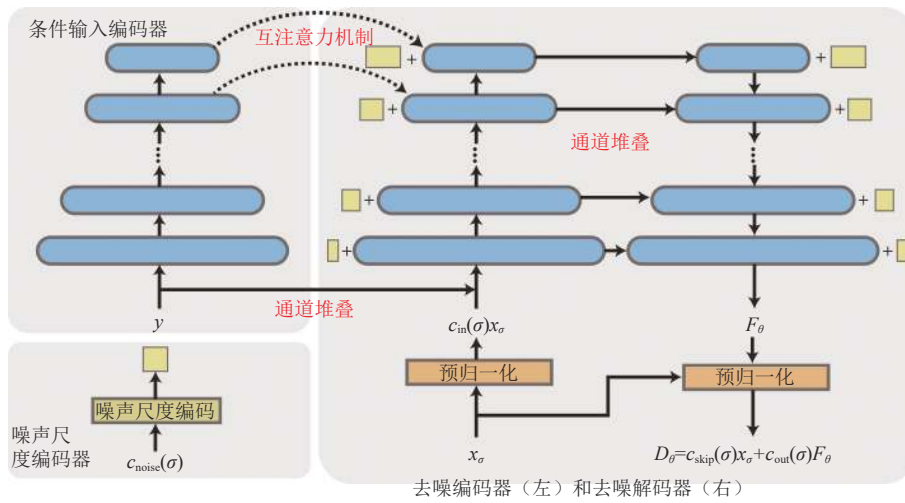


图 2 主干网络架构

3 结果和分析

3.1 模型部署与训练细节

使用 PyTorch 框架实现并部署模型，利用 A100 GPU 进行模型的训练和验证。训练期间设定批处理大小 (batch size) 为 2 048，并选择学习率为 1×10^{-4} 的 ADAM 优化器进行模型优化。此外，还采取衰减因子为 0.999 的指数滑动平均方法 (Exponential Moving Average, EMA) 对模型权重进行平滑处理。经过 20 000 个训练轮次 (epoch) 后，完成了模型的训练。

3.2 性能评价指标

以 $\hat{x} = [\hat{x}^a, \hat{x}^v, \hat{x}^n]$ 表示通过 ODE 求解得到的分离后的心房、心室以及噪声信号，以 $\hat{z} = \hat{x}^a + \hat{x}^v$ 表示去噪后的 ECG 信号。用信噪比提升 (Improvement in Signal-to-Noise Ratio, SNR_{imp})、百分比均方根差异 (Percentage-Root-Mean-Square Difference, PRD)、余弦相似度 (Cosine Similarity, CosSim)

和均方根误差 (Root-Mean-Square Error, RMSE) 来评估模型的去噪效果，分别定义为：

$$SNR_{\text{in}}(s, n) = 10 \lg \left(\frac{\sum_{i=1}^l [s(i)]^2}{\sum_{i=1}^l [n(i)]^2} \right)$$

$$SNR_{\text{out}}(s, \hat{s}) = 10 \lg \left(\frac{\sum_{i=1}^l [s(i)]^2}{\sum_{i=1}^l [s(i) - \hat{s}(i)]^2} \right)$$

$$SNR_{\text{imp}}(s, \hat{s}, n) = SNR_{\text{out}}(s, \hat{s}) - SNR_{\text{in}}(s, n)$$

$$PRD(s, \hat{s}) = \sqrt{\frac{\sum_{i=1}^l [s(i) - \hat{s}(i)]^2}{\sum_{i=1}^l [s(i)]^2}} \times 100\%$$

$$\text{CosSim}(s, \hat{s}) = \frac{\sum_{i=1}^l s(i) \hat{s}(i)}{\sqrt{\sum_{i=1}^l s_1^2(i)} \sqrt{\sum_{i=1}^l \hat{s}_2^2(i)}}$$

$$\text{RMSE}(s, \hat{s}) = \sqrt{\frac{1}{l} \sum_{i=1}^l [s(i) - \hat{s}(i)]^2} \quad (9)$$

式中, s 、 \hat{s} 和 n 分别指代真实干净信号、去噪后的信号和噪声; $\text{SNR}_{\text{in}}(z, x^n)$ 为输入信噪比, 表示真实干净 ECG 与输入噪声的信噪比; $\text{SNR}_{\text{out}}(z, \hat{z})$ 为输出信噪比, 表示真实干净 ECG 与去噪后 ECG 的信噪比; $\text{SNR}_{\text{imp}}(z, \hat{z}, x^n)$ 则表示模型去噪后 ECG 相比混入噪声的 ECG 在信噪比上的提升, 越大越好。PRD 和 RMSE 都是衡量模型预测误差的指标, 它们越小越好, 但 PRD 是相对误差, 不受信号幅度大小的影响。CosSim 是度量两个信号之间相似度的一种方法, CosSim(z, \hat{z})值越接近 1, 说明模型去噪效果越好。

为了验证本文模型是否有效地分离出了心房信号 \hat{x}^a 和心室信号 \hat{x}^v , 并且被分离的信号是否忠于原始 ECG 中波形的形态。本文也引入了敏感性 (Sensitivity, Se) 和阳性预测率 (Positive predictivity, P+) 来评价模型去噪后波形定位算法的性能。

敏感性定义为 $\text{Se} = \text{TP}/(\text{TP} + \text{FN})$, 阳性预测率定义为 $P+ = \text{TP}/(\text{TP} + \text{FP})$ 。其中, TP 表示正确定位的数量, FN 表示定位遗漏数量, FP 表示错误定位数量。计算 TP、FN 和 FP 时需要设置误差容限, 定位误差在误差容限内定义为正确定位, 否则为错误定位。根据各个数据库中的专家标注精度以及参考文献 [30] 提供的金标准来设定这些误差容限, 具体将在 3.4 节中详细描述。

3.3 模型去噪能力验证结果

3.3.1 基于模拟数据的去噪能力验证

使用了 SIMU 中未参与训练的 50 段 ECG 数据, 以及 NSTDB 中未参与训练的 4 种噪声数据 (WGN、BW、MA 和 PLI), 通过将 4 种噪声数据以不同信噪比 (-10、-5、0、5、10、15、20 以及没有噪声 +∞) 加入干净的 ECG 中构建含噪的 ECG 数据, 以验证模型的信噪分离性能。SIMU 中有 x^a 和 x^v 分离的金标准数据, 因此可以同时验证模型对 ECG 的信号分离和去噪性能。

图 3a 和 3b 分别显示在没有噪声时, 模型性能

指标 $\text{CosSim}(\hat{x}^a, x^a)$ 、 $\text{CosSim}(\hat{x}^v, x^v)$ 、 $\text{RMSE}(\hat{x}^a, x^a)$ 和 $\text{RMSE}(\hat{x}^v, x^v)$ 分别为 0.979、0.999、0.004 和 0.009, 说明模型生成的两种信号 \hat{x}^a 和 \hat{x}^v 都能很好地拟合原始 x^a 和 x^v 。

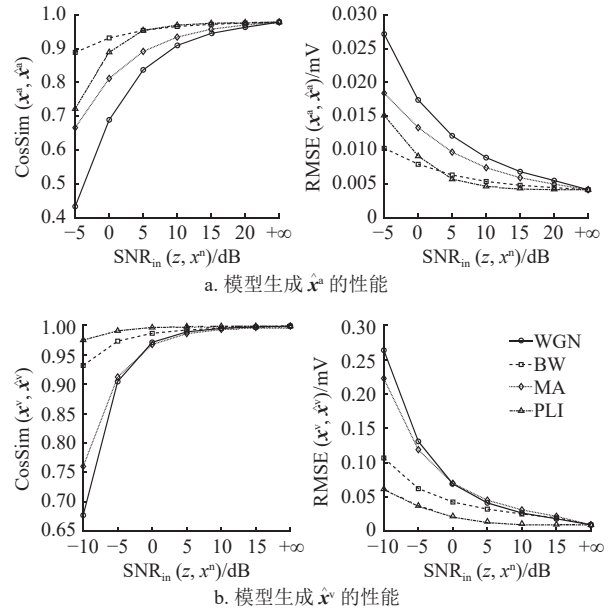


图 3 加入不同信噪比对模型生成 \hat{x}^a 和 \hat{x}^v 的影响

加入噪声后, \hat{x}^v 相比于 \hat{x}^a 明显表现出对于噪声更强的鲁棒性。以加入噪声为 WGN 的含噪 ECG 举例, 当 SNR_{in} 为 -5 dB 时, $\text{CosSim}(\hat{x}^v, x^v)$ 才接近于 0.9, 而当 SNR_{in} 仅为 10 dB 左右时, $\text{CosSim}(\hat{x}^a, x^a)$ 就已经接近 0.9。同时, 不同类型的噪声对生成 \hat{x}^v 和 \hat{x}^a 的影响也存在细微的区别。如对 \hat{x}^v 而言, BW 对它的干扰要始终大于 PLI; 但对 \hat{x}^a 而言, 在 $\text{SNR}_{\text{in}} \geq 5$ dB 时, BW 对其干扰要大于 PLI, 在小于 5 dB 时, PLI 的影响大于 BW。这符合 ECG 中波形的特性, 因为心室信号中的 ST 段的频率范围和大多数 BW 的频率范围有重合^[31], 所以 \hat{x}^v 更容易受到 BW 的干扰; 而 P 波本身幅度很小, PLI 幅度大时能轻易地掩盖 P 波, 所以在信噪比低时 \hat{x}^a 更易受到 PLI 的干扰。

接着探讨在扩散模型的训练和采样中对 x^a 、 x^v 和 x^n 设置相同尺度 (σ_{min} 和 σ_{max}) 时噪声对模型去噪性能带来的影响。模型训练和采样超参数设置为 $\sigma_{\text{min}} = [1 \times 10^{-4}, 1 \times 10^{-4}, 1 \times 10^{-4}]$ 以及 $\sigma_{\text{max}} = [100, 100, 100]$, P_{mean} 和 P_{std} 可由式 (7) 计算得出。图 4 展示了对模型设置相同的噪声尺度的结果。从图 3 和图 4 可以看出, 两个模型生成 \hat{x}^v 的性能差异很小。但图 4 中模型生成 \hat{x}^a 的性能明显不如图 3 中模

型的效果，并且求解生成图 4 数据的 ODE 需要更多的迭代步数 ($N = 10$) 才能够达到最优。

图 5 给出了一个输入混入 10 dB 的 WGN 噪声时模型生成各个信号的例子。第一行为混入了 WGN 噪声的 ECG 信号 y (即输入模型的信号)。第二、三行分别为干净的心房信号和心室信号 x^a 和 x^v (蓝色) 以及模型生成的心房信号和心室信号 \hat{x}^a 和 \hat{x}^v (橙色)。第四行分别为干净的完整 ECG $z = x^a + x^v$ 和模型生成的完整 ECG $\hat{z} = \hat{x}^a + \hat{x}^v$ 。可以看到，设置不同噪声尺度的模型 (图 5a) 生成的 \hat{x}^a 幅值只有略微的衰减，并且不含噪声；设置了相同噪声尺度的模型 (图 5b) 生成的 \hat{x}^a 幅值偏小并且噪声也较大，这说明了对于不同幅度的信号设置不同尺度噪声来生成数据，再用于对模型进行训练和采样的做法是有效的。

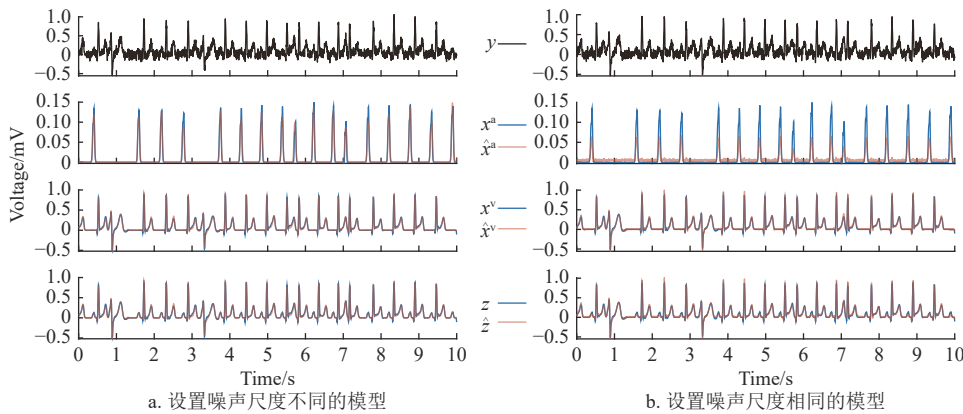


图 5 两类扩散模型对混入 WGN 噪声的 ECG 生成结果

3.3.2 基于真实数据的去噪性能验证

真实 ECG 数据来源于 MITDB、QTDB 和 NSRDB 这 3 个数据库。需要注意的是，本文没有挑选数据，而是直接采用了这 3 个数据库中第一个导联的所有信号，因此这些数据不是完全干净的 ECG 数据，而是带有不同程度的背景噪声。由于背景噪声中的基线漂移可能对性能评估结果有很大影响，且容易被去除，所以首先对这 3 个数据库的数据进行了中值滤波，以滤除基线漂移，然后再加入 WGN、BW、MA 和 PLI 这 4 种噪声生成含噪的 ECG 数据，用于验证模型的去噪性能。此外，实际的 MITDB、QTDB 和 NSRDB 数据库没有提供分离后的金标准 x^a 和 x^v 信号，因此用二者合并的 z 和 \hat{z} 来评估模型的去噪性能。

图 6 展示了模型在 MITDB、QTDB 和 NSRDB 这 3 个真实信号数据库上处理各种类型噪声的去噪

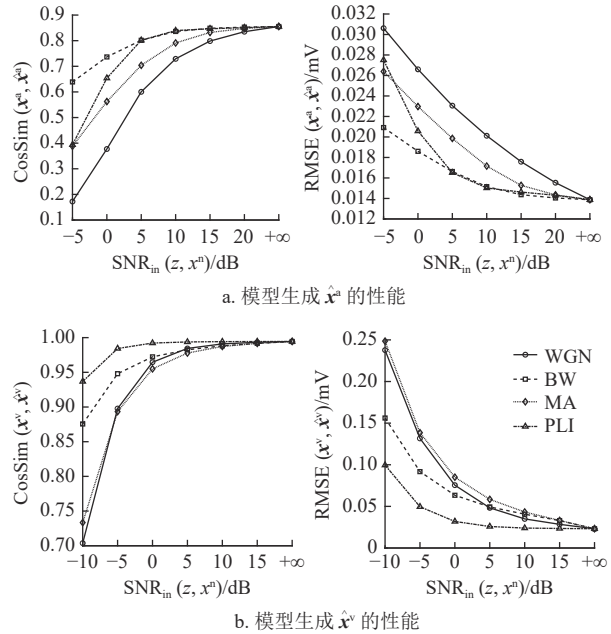


图 4 加入不同信噪比对模型生成 \hat{x}^a 和 \hat{x}^v 的影响

性能。为了比较，加入了模拟数据库 SIMU 中的结果。图 6 说明基于真实数据的结果与 3.3.1 节中模拟数据的结果一致，模型对于各类噪声的去噪性能排序为：PLI>BW>MA>WGN。值得注意的是，对于 MITDB、QTDB 和 NSRDB 这 3 个真实世界数据集，模型的去噪性能表现相差不大，而模型在 SIMU 中的性能要好于在真实世界数据集中的性能。特别对于 SNR_{imp} 指标 (图 6 第一行) 来说，当 SNR_{in} 为 10 dB 时，模型在 SIMU 和真实世界数据集上取得的性能差异最大，而此时噪声相比于信号较为微弱。本文认为导致这种差异的原因只能是模型生成的信号与真实信号之间存在固有的误差。这种固有误差可能来源于以下 3 种途径：1) 扩散模型采样导致的误差。求解 ODE 生成信号的方法本身存在一些误差^[23]，采样时设定的最小噪声尺度 σ_{min} 也会带来一些误差；2) 训练模型使用的数

数据库中 ECG 与理想无噪的 ECG 之间存在差异。这种差异会在模型生成信号时引入误差。如真实 ECG

中可能存在 U 波, 而模拟 ECG 没有将其纳入考量; 3) 真实 ECG 存在除 BW 外的其他背景噪声。

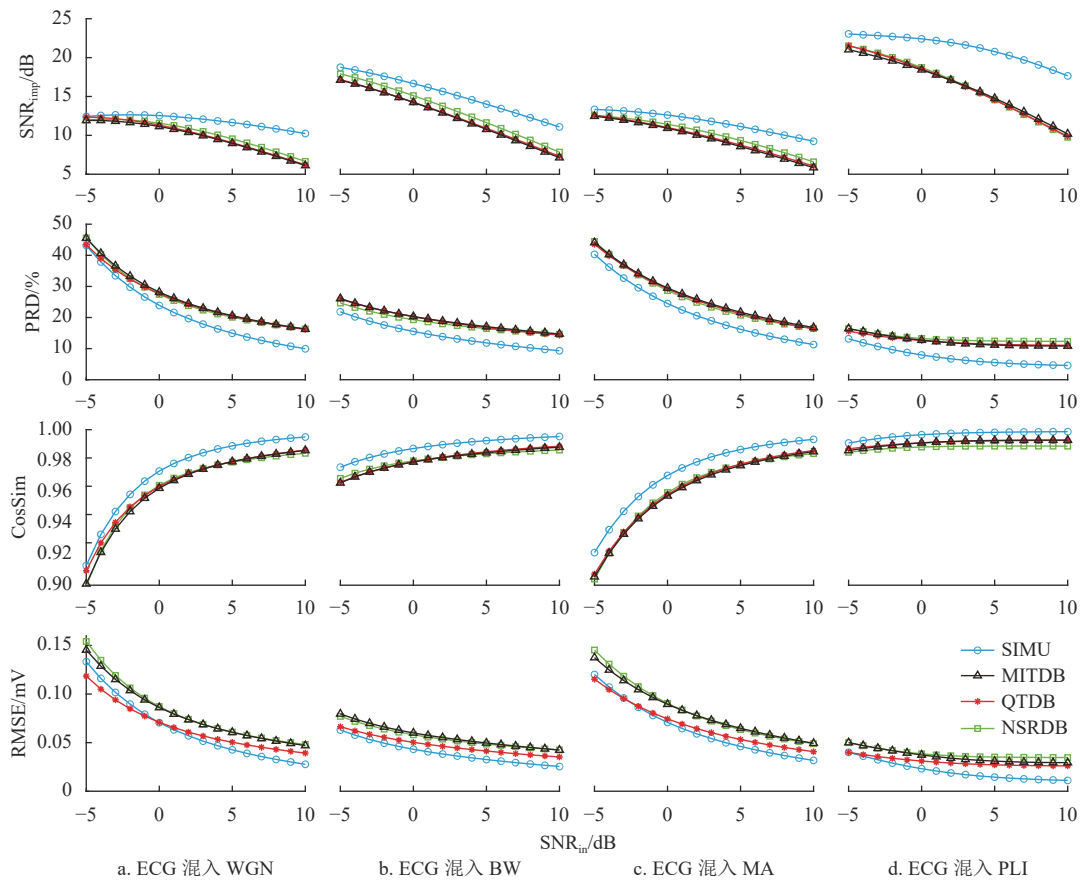


图6 模型对于 SIMU、MITDB、QTDB 和 NSRDB 中混入各类噪声的 ECG 信号的去噪效果

图7展示了模型对一个真实 ECG 数据去噪的例子, 结果表明模型对各类噪声去噪效果都很好, 去噪的 ECG (即 \hat{z}) 与真实 ECG (即 z) 之间仅仅存在细微的差别。细微差异之一是图7中箭头所指的 U 波, 模型生成的信号缺少 U 波, 将它作为噪声滤除了; 另一个细微差异是真实的 ECG 在接近 10 s 位置处中存在比较明显的背景噪声, 而模型则去除了该背景噪声。性能评估指标 SNR_{imp} 能够体现这些细微差异带来的微小误差 (详见图7描述)。

最后将本文方法与其他相关方法在不同噪声条件下进行了比较。在与这些传统去噪方法的比较中, 为了公平起见, 与文献[4]的做法一致, 结合了 QTDB 和 MITDB 数据库的数据作为测试数据。表3展示了本文方法与基于离散小波变换 (Discrete Wavelet Transform, DWT)^[1-2]、基于字典学习的稀疏表示 (Dictionary Learning based Sparse Representation, DLSR)^[4]、稀疏表示降噪 (Sparse

Representation-based Denoising, SRD)^[3] 和经验模态分解 (Empirical Mode Decomposition, EMD)^[5] 方法的比较结果。在这几种方法中, DWT 适用于去除 WGN 和 BW 噪声, SRD 适用于去除 MA 噪声, EMD 适用于去除 PLI 噪声, 而 DLSR 适用于去除所有噪声。本文方法在混入噪声为 BW 和 PLI 时逊色于 DLSR, 在混入噪声为 WGN 和 MA 且噪声较大时优于 DLSR。本文方法在混入各类型噪声条件下均分别优于基于 DWT、SRD 和 EMD^[5] 等方法。然而, DLSR 方法需要事先在 ECG 和噪声数据集中构建数据和噪声的字典, 并且还需要在去噪过程中定位 R 波, 才能够有针对性地对 QRS 波群和非 QRS 波群进行去噪, 因此存在过拟合的风险, 并且在噪声中定位 R 波本身可能会不准确, 带来二次噪声干扰。本文方法采用端到端的方式构建, 去噪前不需要对 R 波进行定位, 去噪后还可以提高 R 波定位方法的准确性 (验证请见 3.4)。

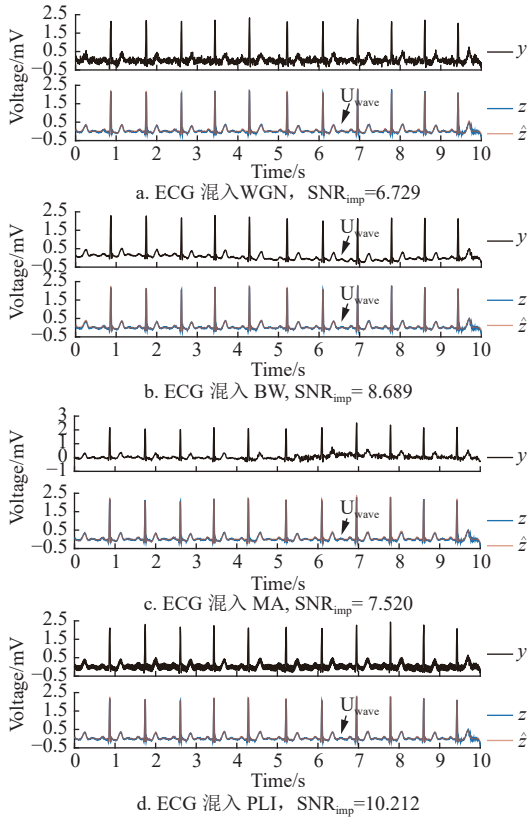


图 7 模型对 MITDB 中 ECG (序号 103 的记录) 的去噪效果

本文还与 DeepFilter^[32] 和 DeScoD^[11] 两种使用神经网络的方法进行了比较。其中, DeScoD 也使用了扩散模型, 并且两个方法都只进行了去除 BW 噪声的验证。在这个比较中选择了 QTDB 的子集作为测试数据, 结果如表 4 所示。与 3.3.1 节加入不同信噪比噪声的方式不同, 我们以信号最大幅度的比例加入 BW 噪声。从表 4 中可以看出, 除了在加入噪声比例为 1.5~2 时本文方法的 PRD 略低于 DeScoD 以外, 其他情况下它都取得了最优的性能。值得一提的是, 两种深度神经网络方法是逐心拍进行去噪的, 也就是说它们需要首先定位 R 波。虽然 DeScoD 也使用了扩散模型, 但它以高斯分布的随机数作为初始值, 通过求解 SDE (Stochastic Differential Equation) 进行 ECG 的生成。每次求解 SDE 需要迭代 50 次, 并且需要求解 10 次并取平均值以去除随机性获得的最佳性能, 运行效率比较低。本文方法是以确定的初始值 (为 0) 求解 ODE, 只需要进行 $N-1=4$ 次迭代求解一次 ODE 即可生成 ECG, 运行效率高。

表 3 与传统去噪模型的性能对比

噪声类型	方法	SNR _{in} = -5 dB		SNR _{in} = 0 dB		SNR _{in} = 5 dB		SNR _{in} = 10 dB	
		SNR _{imp} /dB	PRD /%	SNR _{imp} /dB	PRD /%	SNR _{imp} /dB	PRD /%	SNR _{imp} /dB	PRD /%
WGN	DWT ^[1]	8.063	70.714	5.974	50.550	4.380	34.208	2.907	22.815
	DLSR ^[4]	9.204	62.140	9.110	34.950	8.440	21.280	7.850	12.980
	本文方法	12.164	44.118	11.629	26.377	9.455	19.163	6.602	15.073
BW	DWT ^[2]	17.234	25.570	12.782	24.376	8.201	23.555	3.320	23.387
	DLSR ^[4]	19.890	18.018	19.700	10.380	18.610	6.660	16.130	4.992
	本文方法	17.958	23.479	15.077	18.165	11.561	15.274	7.850	13.209
MA	SRD ^[3]	10.840	51.290	7.330	43.650	2.770	41.800	2.160	41.610
	DLSR	11.690	47.190	10.550	29.680	9.530	18.770	7.230	14.020
	本文方法	12.746	42.737	11.469	27.581	9.260	19.848	6.549	15.249
PLI	EMD ^[5]	14.190	34.780	13.520	21.150	11.860	14.450	8.960	11.50
	DLSR	23.310	12.380	21.540	8.620	18.460	6.9140	16.140	5.060
	本文方法	21.873	14.628	19.434	10.996	15.303	10.075	10.393	10.027

表 4 与深度神经网络模型的性能对比 (BW)

方法	噪声幅度 0.2~0.6		噪声幅度 0.6~1.0		噪声幅度 1.0~1.5		噪声幅度 1.5~2.0	
	CosSim	PRD /%	CosSim	PRD /%	CosSim	PRD /%	CosSim	PRD /%
DeepFilter ^[32]	0.948 ± 0.042	34.026 ± 14.234	0.921 ± 0.064	43.346 ± 19.925	0.888 ± 0.093	53.777 ± 27.189	0.844 ± 0.124	66.167 ± 38.040
DeScoD ^[11]	0.962 ± 0.040	27.475 ± 14.631	0.943 ± 0.055	35.710 ± 18.718	0.920 ± 0.077	43.321 ± 22.959	0.889 ± 0.124	52.081 ± 34.681
本文方法	0.975 ± 0.012	22.170 ± 4.926	0.959 ± 0.024	28.064 ± 7.713	0.927 ± 0.050	37.423 ± 12.651	0.873 ± 0.089	50.302 ± 20.379

3.4 以去噪信号的特征波定位验证模型性能

对于 ECG 去噪方法来说, 去噪后的重要波形

(P、QRS 和 T 波) 特征与原始信号的特征保持一致是至关重要的。因为特征波包含了丰富的生理和

病理信息, 所以这些信息对于心血管疾病的临床诊断和监测至关重要。如果去噪方法改变了波形特征, 可能会导致误判或遗漏重要的诊断线索, 从而影响医生对患者心脏状况的准确评估。

3.3.2 节提供的各类性能指标从宏观上评估了去噪 ECG 与原始 ECG 在波形上的相似程度, 但我们仍然需要从波形的细节上验证去噪方法是否改变了原始波形的特征。因此, 将对比在去噪后的信号和原始信号上作波形特征定位的准确性, 以进一步验证本文模型的有效性。

对于经过模型去噪和分离出的心房信号 \hat{x}^a , 本文直接采用了 `scipy` 包中的 `find_peaks` 函数在其上完成 P 峰定位。`find_peaks` 参数设置为: `height` 为最大波峰值的 0.2 倍; `distance` 为 0.2 s; `width` 为 0.02~0.2 s。对于模型去噪和分离出的心室信号 \hat{x}^v , 使用 `Nerokit2`^[21] 中的 R 波定位方法在其上定位 R 波, 使用文献 [22] 提出的方法在其上定位 T 波末端。

选择 MIT PDB 中正常记录 (100、101、103、

117 和 122) 以及有室性心律失常的记录 (106、109、214、和 223) 来验证 \hat{x}^a 上的 P 峰定位。误差容忍被设置为 10 倍的 CSE (Common Standards in Quantitative Electrocardiography)^[30], 即 63.5 ms。表 5 展示了本文方法在 \hat{x}^a 上定位 P 峰的结果, 并与文献 [33] 方法直接在原始 ECG 中定位 P 波的结果进行了对比。结果显示出本文方法和文献 [33] 方法性能接近, 说明模型在正常和有室性心律失常的真实 ECG 中都能比较有效还原心房信号 \hat{x}^a 。

选择 CPSC2019 数据库中所有信号以及 GUDB 中肢体 II 导联信号来验证 \hat{x}^v 中进行 R 波定位的结果。CPSC2019 中信号带有多种心律失常且来源复杂。利用该数据集来验证本文模型分离 \hat{x}^v 的健壮性, 并设定了 75 ms 的误差容忍。另一方面, GUDB 数据集包括了在不同运动状态下获取的 ECG 数据, 其 R 波已经得到精确的专家标注。使用这一数据集来验证本文生成信号的精确性, 并设定了 4 ms 的误差容忍。

表 5 P 峰定位结果与对比

方法	100		101		103		117		122		%
	Se	P+	Se	P+	Se	P+	Se	P+	Se	P+	
文献[33]	95.13	99.31	98.45	99.95	99.81	100.00	99.93	99.93	100.00	99.96	
本文方法	99.91	99.29	99.52	99.40	99.27	99.32	99.80	99.87	100.00	99.96	

方法	106		109		214		223		%
	Se	P+	Se	P+	Se	P+	Se	P+	
文献[33]	99.37	94.75	98.15	97.80	99.90	94.45	99.48	92.27	
本文方法	99.07	99.47	99.54	99.42	97.94	95.64	98.23	89.46	

表 6 展示了用 \hat{x}^v 进行 R 峰定位结果, 为了进行全面的比较, 还包括了未经去噪处理的信号 (表 6 中未去噪所在的行) 以及使用 `Nerokit2` 中效果最佳的去噪方法 “`Biosppy`” 处理过的 ECG 信号 (表 6 中的第二行)。R 波定位结果表明, 在 CPSC2019 上, 用 \hat{x}^v 进行 R 波定位表现最优。同时, 未经去噪处理的 ECG 在 R 波定位上的性能甚至优于使用 `Biosppy` 处理的结果。这一方面证实了 `Nerokit2` 中 R 波定位方法本身具有一定的去噪能

力; 另一方面, 它也说明了 `Biosppy` 可能会导致 ECG 波形信息的损失, 而本文的去噪方法能够更好地保留 ECG 原始波形。在 GUDB 数据集上, 由于 R 波检测的误差容忍设置较小, 并且运动状态会在 ECG 中引入大量噪声, 未经去噪处理的 ECG-R 波定位性能弱于使用 `Biosppy` 方法去噪后的性能。综合考虑, 本文模型生成的 \hat{x}^v 上定位 R 波的性能仍然表现为最优。

表 6 R 波定位结果与对比

方法	CPSC2019		Sitting		Maths		GUDB Walking		Handbikes		Jogging		%
	Se	P+	Se	P+	Se	P+	Se	P+	Se	P+	Se	P+	
未去噪	89.38	94.27	99.30	99.48	99.13	99.72	35.88	99.66	80.70	97.93	47.23	97.64	
<code>Biosppy</code> ^[21]	89.13	94.92	99.03	99.21	98.94	98.49	90.55	98.39	94.87	97.60	89.11	98.11	
本文方法	89.86	95.09	99.41	99.56	99.44	98.65	91.41	98.83	97.91	98.88	89.71	97.78	

选择 QTDB 所有带有专家标签的信号来验证 \hat{x}^v 中的 T 波末端定位。因为 QTDB 中专家只标注了部分信号段，只能计算 Se 指标，将误差容忍设置为 10 倍的 CSE，即 153 ms。

表 7 展示了 T 波末端定位的结果。同样地，与未经去噪处理的信号和用 Biosppy 去噪后的定位结果进行了对比。结果表明，在本文模型分离的 \hat{x}^v 上定位 T 波的性能是最优的，其 Se 比未去噪信号的结果高出 6.05%。相比之下，Biosppy 的 Se 仅比未去噪信号高出 0.09%。这一结果不仅验证了本文在 3.3.2 节中关于 QTDB 数据集包含背景噪声的观点，还进一步表明传统的去噪方法通常只能去除某些特定类型的噪声，难以消除多样性噪声。

综上，本文方法能够更全面地去除影响 ECG 波形的各种噪声，同时还能保留 ECG 原始波形的关键信息。

表 7 T 波末端定位结果与对比 (Se 参数) %

未去噪	Biosppy	本文方法
91.72	91.81	97.77

4 结束语

本研究面对 ECG 信号去噪面临的挑战，特别是在自动分析算法中常见的多种噪声类型和数据泛化问题，提出了一种基于扩散模型的生成式 ECG 去噪模型。该模型在模拟数据上进行了训练，并在独立的真实数据集上得到了验证。与传统和基于深度学习的去噪技术相比，本文方法展示了一定优越性。研究结果表明，该模型不仅在去除多样性噪声方面具有显著优势，而且能够保持 ECG 关键特征波形的特征不改变。

本研究所构建的模型不仅在 ECG 去噪上展现出优越的性能，也能够有效地分离出心房和心室信号。在未来的研究中，将利用分离出的心房信号，通过提取 P 波特征，进一步进行心房颤动的自动识别研究；同样，也将利用分离出的心室信号，通过提取 T 波的特征，对 T 波的改变作自动识别研究，为后续心房和心室相关疾病的辅助诊断提供客观依据。

然而，本文的研究仍存在一些局限性。如用于训练模型的模拟数据仅包括窦性心律、房性心动过速、二联律、三联律以及室性早搏这 5 种心律类型，没有涵盖左右束支传导阻滞和房室传导阻滞等更多的心律类型。这可能影响模型在处理这些未包

含心律类型的 ECG 去噪和分离的效能。在未来的研究中，计划将更多的心律类型纳入训练数据，以提升模型的泛化能力。

总体而言，本研究不仅克服了现有 ECG 信号去噪方法的局限性，而且在实验验证方面也取得了较好的结果。

参考文献

- [1] ZHANG D H. Wavelet approach for ECG baseline wander correction and noise reduction[C]//Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference. New York: IEEE, 2005: 1212-1215.
- [2] SINGH B N, TIWARI A K. Optimal selection of wavelet basis function applied to ECG signal denoising[J]. *Digital Signal Processing*, 2006, 16(3): 275-287.
- [3] SATIJA U, RAMKUMAR B, SABARIMALAI M M. Noise-aware dictionary-learning-based sparse representation framework for detection and removal of single and combined noises from ECG signal[J]. *Healthcare Technology Letters*, 2017, 4(1): 2-12.
- [4] RAKSHIT M, DAS S. Hybrid approach for ECG signal enhancement using dictionary learning-based sparse representation[J]. *IET Science, Measurement & Technology*, 2019, 13(3): 381-391.
- [5] RAKSHIT M, DAS S. An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter[J]. *Biomedical Signal Processing and Control*, 2018, 40: 140-148.
- [6] CHATTERJEE S, THAKUR R S, YADAV R N, et al. Review of noise removal techniques in ECG signals[J]. *IET Signal Processing*, 2020, 14(9): 569-590.
- [7] CHIANG H T, HSIEH Y Y, FU S W, et al. Noise reduction in ECG signals using fully convolutional denoising autoencoders[J]. *IEEE Access*, 2019, 7: 60806-60813.
- [8] WANG J L, LI R F, LI R, et al. Adversarial de-noising of electrocardiogram[J]. *Neurocomputing*, 2019, 349: 212-224.
- [9] RASTI-MEYMANDI A, GHAFARI A. A deep learning-based framework for ECG signal denoising based on stacked cardiac cycle tensor[J]. *Biomedical Signal Processing and Control*, 2022, 71: 103275.
- [10] WELKER S, RICHTER J, GERKMANN T. Speech enhancement with score-based generative models in the complex STFT domain[C]//Proceedings of the Interspeech. [S. l.]: ISCA, 2022: 2928-2932.
- [11] LI H Y, DITZLER G, ROVEDA J, et al. DeScoD-ECG: Deep score-based diffusion model for ECG baseline wander and noise removal[J]. *IEEE Journal of Biomedical and Health Informatics*, 2024, 28(9): 5081-5091.
- [12] MOODY G B, MARK R G. The impact of the MIT-BIH arrhythmia database[J]. *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine & Biology Society*, 2001, 20(3): 45-50.

- [13] LAGUNA P, MARK R G, GOLDBERG A, et al. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG[C]//Computers in Cardiology 1997. Lund: IEEE, 1997: 673-676.
- [14] GOLDBERGER A L, AMARAL L A, GLASS L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals[J]. *Circulation*, 2000, 101(23): 215-220.
- [15] MARŠÁNOVÁ L, NĚMCOVÁ A, SMÍŠEK R, et al. Automatic detection of P wave in ECG during ventricular extrasystoles[EB/OL]. [2023-07-12]. http://link.springer.com/10.1007/978-981-10-9038-7_72.
- [16] GAO H X, LIU C Y, WANG X Y, et al. An open-access ECG database for algorithm evaluation of QRS detection and heart rate estimation[J]. *Journal of Medical Imaging and Health Informatics*, 2019, 9(9): 1853-1858.
- [17] PORR B. High precision ECG database with annotated R peaks, recorded and filmed under realistic conditions[EB/OL]. [2023-08-08]. <https://github.com/berndporr/ECG-GUDB>.
- [18] MOODY G B, MULDROW W, MARK R G. A noise stress test for arrhythmia detectors[J]. *Computers in Cardiology*, 1984, 11(3): 381-384.
- [19] BACHI L, HALVAEI H, PEREZ C, et al. ECG modeling for simulation of arrhythmias in time-varying conditions[J]. *IEEE Transactions on Bio-Medical Engineering*, 2023, 70(12): 3449-3460.
- [20] BAILON R, SORNMO L, LAGUNA P. A robust method for ECG-based estimation of the respiratory frequency during stress testing[J]. *IEEE Transactions on Biomedical Engineering*, 2006, 53(7): 1273-1285.
- [21] MAKOWSKI D, PHAM T, LAU Z J, et al. NeuroKit2: A Python toolbox for neurophysiological signal processing [J]. *Behavior Research Methods*, 2021, 53(4): 1689-1696.
- [22] ZHANG Q H, ILLANES M A, MEDIGUE C, et al. An algorithm for robust and efficient location of T-wave ends in electrocardiograms[J]. *IEEE Transactions on Biomedical Engineering*, 2006, 53(12): 2544-2552.
- [23] KARRAS T, AITTALA M, AILA T, et al. Elucidating the design space of diffusion-based generative models [EB/OL]. [2023-07-15]. <http://arxiv.org/abs/2206.00364>.
- [24] SONG Y, SOHL-DICKSTEIN J, KINGMA D P, et al. Score-based generative modeling through stochastic differential equations[EB/OL]. [2023-04-24]. <http://arxiv.org/abs/2011.13456>.
- [25] MARIANI G, TALLINI I, POSTOLACHE E, et al. Multi-source diffusion models for simultaneous music generation and separation[EB/OL]. [2023-07-06]. <http://arxiv.org/abs/2302.02257>.
- [26] BAYÉS DE LUNA A. Basic electrocardiography[M]. Malden: Wiley, 2007.
- [27] CHEN T, LI L L, SAXENA S, et al. A generalist framework for panoptic segmentation of images and videos[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2023: 909-919.
- [28] TANCIK M, SRINIVASAN P P, MILDENHALL B, et al. Fourier features let networks learn high frequency functions in low dimensional domains[EB/OL]. [2023-06-21]. <http://arxiv.org/abs/2006.10739v1>.
- [29] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions[EB/OL]. [2023-07-16]. <http://arxiv.org/abs/1710.05941>.
- [30] PARTY C W. Recommendations for measurement standards in quantitative electrocardiography. The CSE working party[J]. *European Heart Journal*, 1985, 6(10): 815-825.
- [31] LENIS G, PILIA N, LOEWE A, et al. Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: A simulation study[J]. *Computational and Mathematical Methods in Medicine*, 2017, 2017: 9295029.
- [32] ROMERO F P, PIÑOL D C, VÁZQUEZ-SEISDEDOS C R. DeepFilter: An ECG baseline wander removal filter using deep learning techniques[J]. *Biomedical Signal Processing and Control*, 2021, 70: 102992.
- [33] SACLOVA L, NEMCOVA A, SMISEK R, et al. Reliable P wave detection in pathological ECG signals[J]. *Scientific Reports*, 2022, 12(1): 6589.

编辑 叶芳