

DMAM-ECG: A Diffusion Model with Self-Attention Module for ECG Signal Denoising

Zheng-Dong Hu, Yang Hong, Jia-Yan Huang

New Engineering Industry College
Putian University, Putian 351100, China
zdong_hu@163.com, luuvletter@gmail.com, jyan_huang@163.com

Kai-Hong Chen

School of Mathematics and Data Science
Minjiang University, Fuzhou 350108, China
kaihong_chenc@163.com

Wan-Qi Zhao

The University of Sheffield, Sheffield S10 2TN, United Kingdom
wanqizhao0110@163.com

Antoni Grau, Edmundo Guerra, Chuan-Sheng Wang*

Department of Automatic Control Technical
Polytechnic University of Catalonia, Barcelona 08034, Spain
antoni.grau@upc.edu, edmundo.guerra@upc.edu, wangcleaner@gmail.com

Fu-Quan Zhang*

Fujian Provincial Key Laboratory of Information Processing and Intelligent Control
Minjiang University, Fuzhou 350108, China
zfq@mju.edu.cn

*Corresponding author: Chuan-Sheng Wang and Fu-Quan Zhang

Received August 23, 2023, revised December 9, 2023, accepted March 4, 2024.

ABSTRACT. *Electrocardiogram (ECG) is the most common and non-invasive auxiliary diagnostic technique for heart disease. However, the ECG signal is susceptible to various noises during the collection process, which will impact subsequent analysis and diagnosis. Therefore, ECG signal denoising has become the primary task after signal acquisition. ECG signal denoising aims to remove noise interference from the collected noisy signals for clarify abnormal waveforms in it, and thus diagnose diseases more accurately. In this paper, to obtain high-quality and high-fidelity denoised ECG signals, we propose a novel diffusion model with self-attention module for ECG signal denoising (DMAM-ECG). Specifically, the DMAM-ECG contains two main branches used to extract features of input noise signal observations and potential variables, respectively. Each branch consists of 6 designed Convolution-Instancenormal-LeakyReLU Filter Blocks (CIL-Filter Blocks). Meanwhile, DMAM-ECG introduces the self-Attention Module Fusion Blocks (SAM-Fusion Blocks) to further extract features for the first branch, and then fuse the output features with the second branch. Besides, we use a multi-shot reconstruction strategy to further improve signal denoising. Quantitative and qualitative results on the QT database show that the proposed DMAM-ECG performs better than the existing 5 state-of-the-art ECG signal denoising methods, especially for strong noise interference ECG signal removal. To sum up, our DMAM-ECG introduces self-attention module into the diffusion model to improve the grasp of global information and achieve high-quality information preservation for extremely noise removal of ECG signal. In particular, it uses only 3-shot reconstruction gets much better denoised results than that of the best baseline with 10-shot. The code of the proposed DMAM-ECG is available at <https://github.com/lwvletterh/DMAM-ECG>.*

Keywords: Heart Disease, ECG Signal Denoising, Baseline Wander, Diffusion Model, self-Attention Module

1. Introduction. Heart disease seriously threatens people's lives, and the auxiliary diagnostic tools and technologies for early detection of it are in great demand. Electrocardiogram (ECG) can reflect the electrical activity process of cardiac excitation, and it is a common non-invasive auxiliary diagnostic technique in clinical practice [1, 2, 3]. However, ECG signals have the characteristics of strong randomness and poor anti-reference ability, and thus they are prone to various noise interference from the human body's internal or external environment, such as baseline wander caused by human respiration, power frequency interference caused by the power system, Electromyographic (EMG) interference caused by muscle tremors, and motion artifacts. As shown in Figure 1, it gives a clean ECG signal with different noise interference. The noisy ECG signal will impact the subsequent analysis and diagnosis. Therefore the removal of ECG signal noise has become the primary task after signal acquisition and is of decisive significance for ECG signal diagnosis, especially automatic diagnosis.

High-quality and high-fidelity ECG signal denoising can improve the availability of collected ECG signals and reduce the cost of clinical trials. The existing ECG denoising methods can be divided into traditional filter-based methods and deep learning-based methods. Traditional filter-based methods often denoise ECG signals from signal processing perspective. They usually use different filters to preliminary remove noises first and then use some transform, such as wavelet transform [4, 5] or S-transform [6], to obtain the refined ECG signals. However, traditional filter-based methods are difficult to balance high-frequency and low-frequency noise removal for ECG signals with extreme noise interference.

To solve the shortcomings of traditional methods, some more effective deep learning-based ECG signal denoising methods have been proposed [7, 8]. Antczak et al. [9] proposed an ECG signal denoising method based on recurrent neural network, which can

effectively denoise for extreme noise interference ECG signal. Nichol et al. [10] improved the denoising autoencoder using wavelet transform to construct a deep neural network for ECG signal enhancement. Based on the fully convolution neural networks (CNNs), Chiang et al. [11] designed a denoising automatic encoder, which can remove high noise reference with low signal distortion. Romero et al. [12] proposed a deep-learning-based filter to remove the most common uncontrollable baseline wander. By combining generative adversarial networks (GANs) and residual networks (ResNet), Xu et al. [13] presented an ECG denoising method to remove noise while preserving the effective information of the original signal. However, deep learning-based methods are limited by the strength of noise interference, resulting in limited availability of denoised ECG signals.

In this paper, to recover high-quality and high-fidelity clean signal from an ECG signal with extreme noise interference, we propose a diffusion model with self-attention module for ECG signal denoising (DMAM-ECG). Specifically, the DMAM-ECG contains two main branches, and each consists of six designed Covolutional-InstanceNormal-LeakyReLU Filter Blocks (CIL-Filter Block). It first uses the two branches to extract features of input noise signal observations and potential variables, respectively. Then, considering that the self-attention module is better at extracting internal correlations of features, the DMAM-ECG introduces it to further extract the observation features from the first branch and fusion with the second branch to obtain denoised result. To sum up, our main contributions are as follows:

(1) We proposed a diffusion model with self-attention module for ECG signal denoising (DMAM-ECG), which introduces self-attention module into the diffusion model to improve the grasp of global information and achieve high-quality information preservation of denoised ECG signal. In other words, the DMAM-ECG can gradually improve its denoising effects by multi-shot signal reconstruction, especially for extreme noise removal.

(2) We introduced the self-attention module in DMAM-ECG to extract the correlation of different positions in the input sequence, i.e., the better extraction of internal correlations of features, for improving the denoising effectiveness of ECG signal with extreme noise interference.

(3) Experiments on the ECG records from the QT database with noise from the MIT-BIH NST database show that the DMAM-ECG outperforms five existing state-of-the-art ECG denoising methods in ECG denoising with five similarity-based metrics. In particular, the proposed DMAM-ECG with only 3-shot gets much better denoising results than the best baseline with 10-shot.

2. Related Work. This section will introduce the existing ECG signal denoising methods and denoising diffusion model.

2.1. ECG signal denoising. The common ECG signal interference includes baseline wander, power frequency interference, and EMG. Baseline wander is usually caused by poor contact between the electrode and the body surface, and it belongs to low-frequency signal. Power frequency interference is caused by electromagnetic radiation, it is relatively stable. EMG is caused by involuntary muscle contraction of the subject, its frequency range is wide. EMG is relatively irregular and manifests as high-frequency sawtooth waves on the ECG. The most of existing ECG signal denoising methods mainly focus on baseline wander removal and can be roughly divided into traditional digital filter-based methods and deep learning-based methods.

2.1.1. Traditional filter-based methods. Early traditional methods usually process ECG signals from a signal processing perspective, they designed targeted filters for noise removal. Based on the assumption of different statics between noise and the original ECG

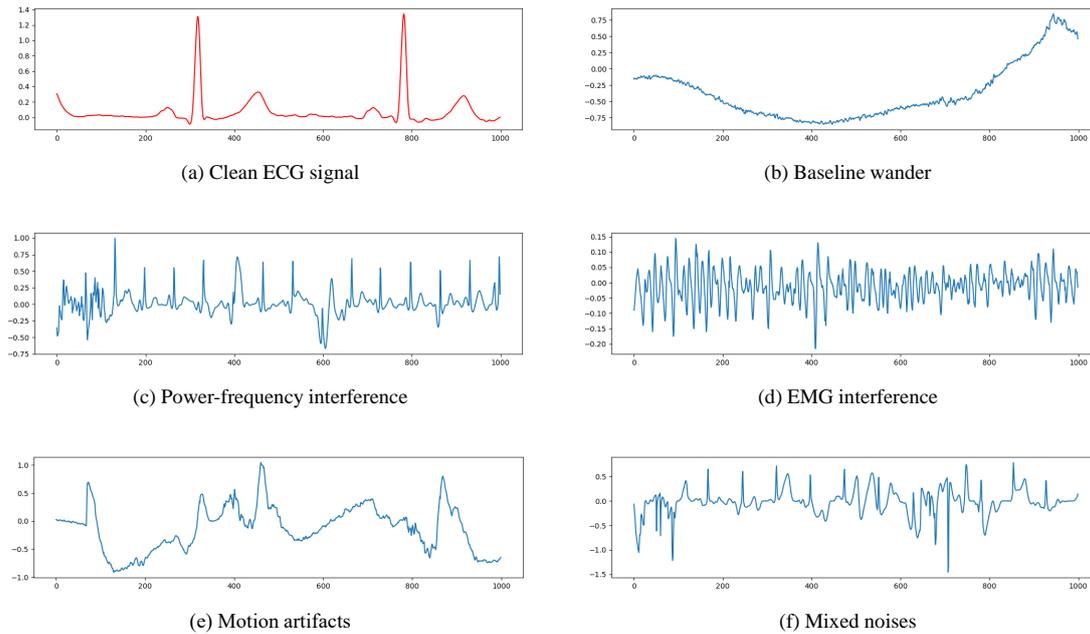


FIGURE 1. Examples of ECG signal under different noise interference.

signal, Barati et al. [14] performed independent component analysis on the ECG signal channel to remove baseline wander. Kumar et al. [15] combined finite pulse recovery (FIR) and infinite pulse recovery (IIR) digital filters for ECG signal denoising. Taouli et al. [16] used morphological filters to suppress ECG signal noise and baseline wander. Unfortunately, despite FIR-based methods have good effects on high-frequency noise, they easily cause distortion of the low-frequency signal. On the contrary, morphological filter-based methods have an ideal effect on baseline wander removal, but they are not suitable for high-frequency noise removal.

2.1.2. Deep learning-based methods. Recently, deep learning-based methods have attracted increasing attention [17, 18, 19, 20, 21, 22]. Compared to shallow models that rely on rich information inherent in data, deep models based on data-driven show better performance. Representative deep learning-based methods include CNN-based [9], Denoising Autoencoder (DAE-based) [10, 11], and GAN-based [13]. For instance, DRNN [9] first uses deep Res-Net to pre-train on synthesized data, and then uses transfer learning to fine-tune in reality for ECG signal denoising. Considering that DAE can reconstruct clean signals from noisy signals, Nichol et al. constructed a deep neural network using wavelet transfer to improve DAE. Similarly, FCN-DAE [11] constructs a DAE using a fully convolutional neural network. Based on the reconstruction of GANs, CGAN [13] proposes an ECG signal denoising method, which uses residual blocks and skip connecting in the generator to extract deeper information and ResNet framework in the discriminator. Unfortunately, deep learning-based methods show good effects in weak noisy corrupted ECG signals but limited denoising performances in extreme noise conditions.

2.1.3. Denoising diffusion model. The GAN has potential instability and lack of generation diversity due to adversarial training. And the VAE has poor generation quality due to its reliance on substitution loss. Compared to GAN and VAE, the diffusion model does not require adversarial training and can obtain high-quality generated data with better analyticity and flexibility. In other words, a well-trained diffusion model [23, 24] can generate data that conforms to the data distribution from a randomly sampled noise

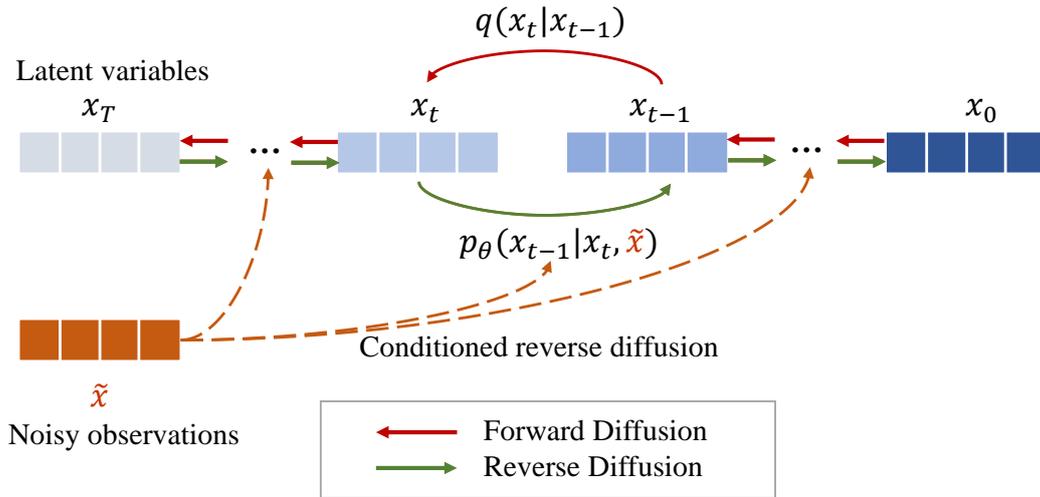


FIGURE 2. Denoising diffusion model of the DMAM-ECG.

through the learned denoising process. Therefore, we apply the denoising diffusion model to ECG signal denoising and demonstrate its effectiveness in this paper.

The denoising diffusion model is a latent variable model that has been widely used in image generation. As shown in Figure 2, it first adds noise gradually to the input data using a fixed Markov chain to obtain pure Gaussian noise data (Forward Diffusion). Then, it trains and learns the reverse process to convert the noise back into a target distribution sample for generating new data (Reverse Diffusion). In other words, the forward diffusion $q(\cdot|\cdot)$ gradually adds noise for clean ECG signal, and the reverse diffusion $p_\theta(\cdot|\cdot)$, which is conditioned by the corresponding noisy ECG observations \tilde{x} , aims to learn the denoising function θ to generate the clean signal from random Gaussian noise. Specifically, given a data distribution $q(x_0)$, it expects to obtain an approximate distribution $p_\theta(x_0)$.

Forward diffusion gradually adds noise to the data to approximate the posterior distribution $q(x_{1:T} | x_0)$, where the latent variables x_1, \dots, x_T has the same dimension with x_0 . In [24], the diffusion process is set as a simple parameterized fixed Markov chain, with each step using conditional Gaussian translation,

$$q(x_{1:T} | x_0) := \prod_{t=1}^T q(x_t | x_{t-1}), \tag{1}$$

$$q(x_t | x_{t-1}) := \mathcal{N}\left(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t\mathbf{I}\right),$$

where β_1, \dots, β_T denotes a variance scheduling, whose value can be constantly changing by learning or fixed. $\mathcal{N}(x; \mu, \sigma)$ is a Gaussian probability density function with parameters μ and σ .

Noted that, on any time step t for x_t , the closed form for sampling can be represented as,

$$q(x_t | x_{t-1}) := \mathcal{N}\left(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)\mathbf{I}\right), \tag{2}$$

where $\alpha_t := 1 - \beta$ and $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$. So x_t can be further represented as,

$$x_t = \sqrt{\bar{\alpha}_t}x_0 + (1 - \alpha_t)\epsilon \tag{3}$$

where ϵ is a parameter of the Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{I})$.

Reverse diffusion aims to restore the initial value of x_0 from x_t . Starting from pure Gaussian noise $p(x_T) := \mathcal{N}(x_T, \mathbf{0}, \mathbf{I})$, the inverse diffusion process is defined by the following Markov chain,

$$p_\theta(x_{0:T}) := p(x_T) \prod_{t=1}^T p_\theta(x_{t-1} | x_t), x_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad (4)$$

$$p_\theta(x_{t-1} | x_t) := \mathcal{N}(x_{t-1}; \boldsymbol{\mu}_\theta(x_t, t), \sigma_\theta(x_t, t) \mathbf{I})$$

Noted that the time-related parameters of transformation are obtained through learning [24]. The specific parameterization method for $p_\theta(x_{t-1} | x_t)$ is described as,

$$\boldsymbol{\mu}_\theta(x_t, t) = \frac{1}{\alpha_t} \left(x_t - \frac{\beta_t}{\sqrt{1 - \alpha_t}} \boldsymbol{\epsilon}_\theta(x_t, t) \right), \quad (5)$$

$$\sigma_\theta(x_t, t) = \sqrt{\tilde{\beta}_t}, \text{ where } \tilde{\beta}_t = \begin{cases} \frac{1 - \alpha_{t-1}}{1 - \alpha_t} \beta_t & t > 1 \\ \beta_1 & t = 1 \end{cases},$$

where $\boldsymbol{\epsilon}_\theta(\cdot, \cdot)$ is a learnable denoising function to estimate the added noise vector $\boldsymbol{\epsilon}$. Such parameterization results in the following alternative loss function,

$$\mathcal{L}_{\text{simple}}(\theta) := \mathbb{E}_{t, x_0, \boldsymbol{\epsilon}} \left[\left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, t) \right\|^2 \right], \quad (6)$$

Equation (6) can be regarded as a weighted combination of denoising score matching. Compared to the original score matching loss [25, 26], it introduces the learnable parameter $\boldsymbol{\epsilon}$. That way, the model can flexibly adjust the noise estimation strategy to adapt to different data distributions and noise types, thus achieving more stable training and getting better results.

Nevertheless, the diffusion model requires a large amount of computing resources and time to handle complex denoising tasks. Specifically, its computational complexity experiences a sharp increase in dealing with high-dimensional data. It is challenge to select appropriate parameters for achieving better denoising results. It depends on certain noise types, but for non specific types of noise and complex noise distributions, diffusion model may perform poorly. Besides, it is difficult in boundary processing, which easily leads to poor denoising effect at he boundary. Furthermore, diffusion model may exsist the excessive smoothing problem in some case. In other words, even the noise can be effectively removed, it may lead to a decrease in signal details at the same times.

3. Methodology. This section first introduces the diffusion model for ECG signal denoising. Then, it introduces the overall structure of the DMAM-ECG, the details of the designed CIL-Filter Block, and the SAM-Fusion Block. Finally, it introduces the used loss function.

3.1. Diffusion model for ECG signal denoising. Owing to the signal smooth processing ability, diffusion model has achieved significant results in image processing tasks, and it also has the great potential to be applied to ECG signal denoising. In fact, the probability of each state transition in a Markov chain only depends on the current state instead of past states. In signal processing, we can consider noise as a random state change, while diffusion model simulate the propagation and diffusion process of noise using the idea of Markov chain. Therefore, we can update the signal value by applying the transition probability of Markov chains. And these transition probabilities define the possibility of signal propagation from one location to the adjacent locations. During the

propagation process, noise gradually spreads to adjacent positions, resulting in a smoothing effect throughout the entire signal. By iteratively applying the transfer process of Markov chains, noise propagates from high-energy regions to low-energy regions, resulting in a gradual reduction of noise and an improvement in signal clarity. Therefore, we utilize the diffusion model for ECG signal denoising in this paper.

For the ECG signal denoising task, it assumes that the clean denoised ECG signal is x_0 and the noisy ECG signal is \tilde{x} . Inspired by DeSocD-ECG [27], we model the conditional distribution $p_\theta(x_0 | \tilde{x})$ based on a diffusion model to estimate the true conditional distribution $q(x_0 | x)$. Thus, the reverse diffusion process (Equation 4) can be extended to the following conditional form,

$$p_\theta(x_{0:T} | \tilde{x}) := p(x_T) \prod_{t=1}^T p_\theta(x_{t-1} | x_t, \tilde{x}), x_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \quad (7)$$

$$p_\theta(x_{t-1} | x_t, \tilde{x}) := \mathcal{N}(x_{t-1}; \boldsymbol{\mu}_\theta(x_t, t | \tilde{x}), \sigma_\theta(x_t, t | \tilde{x}) \mathbf{I}).$$

With the noisy observations \tilde{x} as conditions, the reverse diffusion process iteratively reconstructs the clean signal x_0 from an original Gaussian distribution x_T by denoising on every small step, as shown in Figure 2. Besides, to improve synthetic quality with fewer forward/reverse diffusion steps, we adapt the same noise schedule β_i initial strategy in [23, 28], i.e., the $(\beta_1, \dots, \beta_n)$ are given by,

$$\beta_t = \left(\frac{T-t}{T-1} \sqrt{\beta_1} + \frac{t-1}{T-1} \sqrt{\beta_t} \right)^2. \quad (8)$$

where the initial and the final noise schedule values are set to $\beta_1 = 0.0001$, $\beta_T = 0.5$, respectively.

Moreover, to improve the model effectiveness, we further use the specific noise schedule for the training process [27], which can avoid the noise function ϵ_θ by conditioning on the step index t . Specifically, with the definitions $\alpha_t := 1 - \beta_t$ and $\bar{\alpha}_t := \prod_{i=1}^t \alpha_i$, we first use the predefined noise schedule $S = \{1, \sqrt{\bar{\alpha}_0}, \dots, \sqrt{\bar{\alpha}_T}\}$ in training process. Then, we add the denoising function condition on the continuous noise level $\bar{\alpha} \sim \text{Uniform}(S_{t-1}, S_t)$. Therefore, the denoising function ϵ_θ receives the noisy observations \tilde{x} as the condition, and the loss function is defined as,

$$\mathbb{E}_{x_0, \bar{\alpha}, \epsilon} \left[\left\| \epsilon - \epsilon_\theta(\sqrt{\bar{\alpha}}x_0 + \sqrt{1 - \bar{\alpha}}\epsilon, \tilde{x}, \bar{\alpha}) \right\|^2 \right] \quad (9)$$

3.2. The structure of the DMAM-ECG. The structure of the DMAM-ECG is shown in Figure 3, it contains two main branches, each consisting of six *Convolutional InstanceNormal LeakyReLU Filter Blocks* (**CIL-Filter Blocks**), which is motivated by the half normalized filter proposed by DeSocD-ECG [27]. Specifically, the two branches are first used to extract features of input noisy ECG signal observation \tilde{x} and potential variable x_t . Then, *self-attention module fusion blocks* (**SAM-Fusion Blocks**) are introduced to further extract features from the first branch and to fuse the feature information with those of another branch.

CIL-Filter Block. The structure detail of the designed CIL-Filter Block is shown in Figure 4. Firstly, the input is sent into two improved convolutional blocks with kernel sizes 3×3 and 5×5 for preliminary feature extraction, respectively. Notably, the improved convolutional block is used to improve training stability while preserving natural statistical characters of features, and it is composed of one common convolution, an instance-normalization layer, and a LeakyReLU layer, as shown in the sub-figure below Figure 4. Then, those features are fed into multi-scale filters for further extraction under

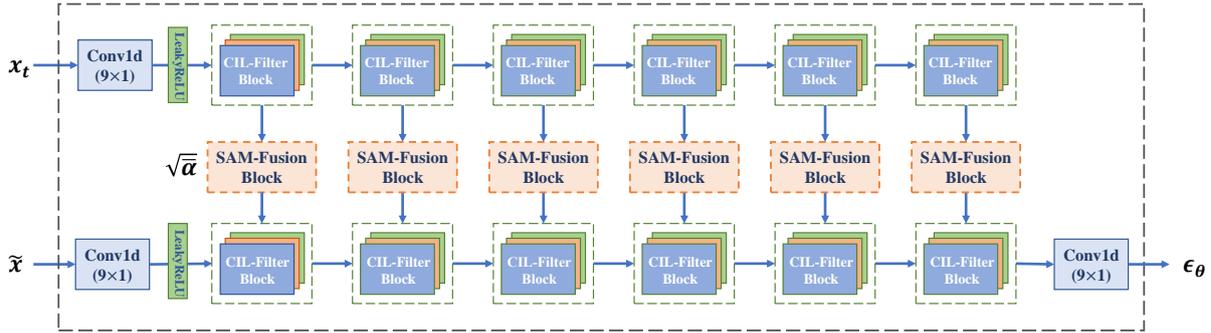


FIGURE 3. The structure of the proposed DMAM-ECG.

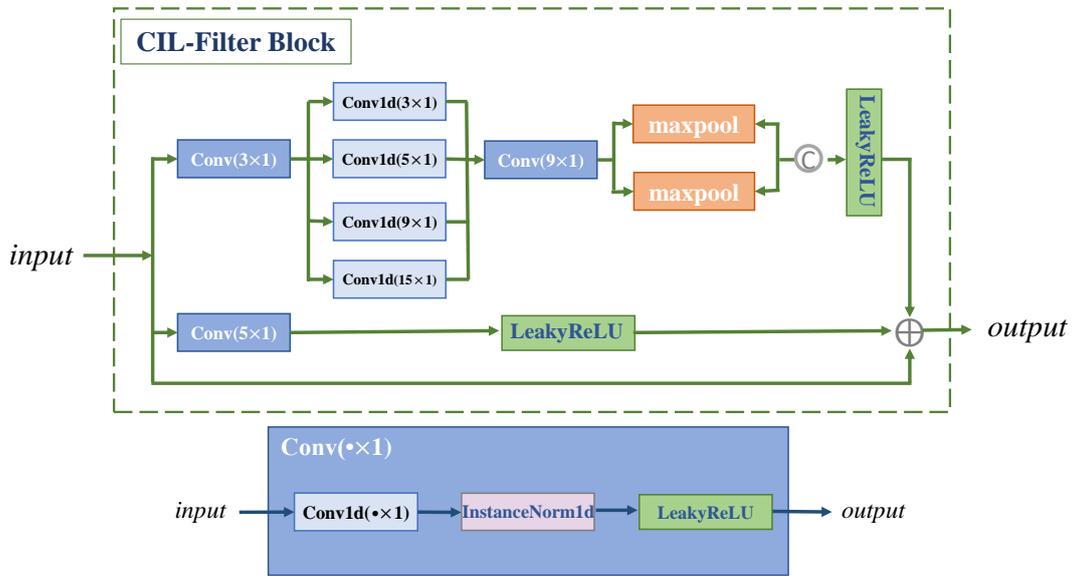


FIGURE 4. The structure of CIL-Filter Block.

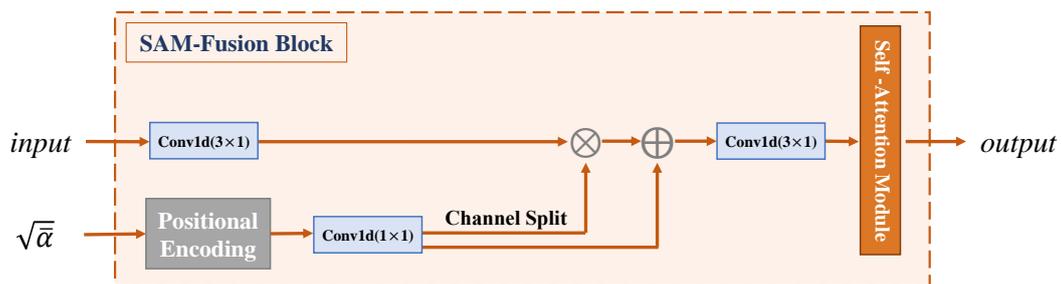


FIGURE 5. The structure of the SAM-Fusion Block.

different receptive fields. Next, they are concatenated in channel-wise and then aggregated by a 9×9 improved convolutional block. Finally, those features are through LeakyReLU layer, and the following add to the extracted features of the second convolutional block after the LeakyReLU layer and the input.

SAM-Fusion Block. Considering that the self-attention module can reduce the dependence on external information and is better at capturing internal correlations of features, we introduced *self-attention module fusion blocks* (SAM-Fusion Blocks) to the

further extraction of features and fusion of two branches. Motivated by the DeScoD-ECG [27], the SAM-Fusion Block is designed based on the feature-wise linear modulation (FiLM) [29], and its specific structure is shown in Figure 5. Firstly, the input features are input into a 3×3 convolutional layer while the noise level $\sqrt{\alpha}$ is encoded by a sinusoidal positional encoder [30]. Then, the $\sqrt{\alpha}$ is further encoded by a 1×1 convolutional layer to obtain the learnable scaling and shift vectors. Next, the vectors multiply and then add to the input feature. Finally, they are fed into a 3×3 convolutional layer followed by a self-attention module for better capture of internal correlations.

3.3. Loss function. Since the Smooth L1 loss can solve the un-smooth problem of L1 loss, and the gradient explosion problem at outliers of L2 loss, we used the Smooth L1 loss for the proposed DMAM-ECG. Using x and y to represent the input and the output denoised ECG signal respectively, the loss function $\mathcal{L}_{1_{smooth}}$ of DMAM-ECG can be defined as,

$$\mathcal{L}_{1_{smooth}}(x, y) = \frac{1}{N} \sum_{n=1}^N \begin{cases} 0.5 * [x(n) - y(n)]^2, & \text{if } |x(n) - y(n)| < 1 \\ |x(n) - y(n)| - 0.5, & \text{otherwise.} \end{cases} \quad (10)$$

4. Experiments.

4.1. Training Setting, Database, and Metrics.

4.1.1. Training Setting. All the experiments were executed on Intel(R) Xeon(R) Platinum 8375C CPU @ 2.90GHz CPU and NVIDIA GeForce RTX 3090 GPU. The proposed DMAM-ECG was coded with the PyTorch framework [31], the training epochs are 600, the initial learning rate is 0.001, and the batch size is 128. We compared our DMAM-ECG with 5 the-art-of-state ECG denoising methods, including DRNN [9], FCN-DAE [11], CGAN [13], DeepFilter [12], and DescoD-ECG [27], and we kept their default parameter values to obtain their best results. Specifically, the specific parameter settings for each algorithm were shown in Table 1. Notably, to encourage more researchers to validate and build upon this work, we publish our code of the proposed DMAM-ECG at <https://github.com/luvletterh/DMAM-ECG>.

TABLE 1. Training setting details of different ECG denoising methods.

Methods	DRNN [9]	FCN-DAE [11]	CGAN [13]	DeepFilter [12]	DeScoD-ECG [27]	Ours
Learning_rate	0.001	0.001	0.1	0.001	0.001	0.001
Batch_size	64	128	64	128	96	128
Epochs	10	25	200	60	400	600
Framework	TensorFlow	PyTorch	PyTorch	TensorFlow	PyTorch	PyTorch

4.1.2. Database. We use the QT database [32] and the Massachusetts Institute of Technology-Beth Israel Hospital Noise Stress Test (MIT-BIH NST) database for all experiments. The QT database [32] is composed of 105 fifteen-minute excerpts of two-channel ECG Holter recordings, with various QRS and ST-T morphologies. The MIT-BIH NST database [33] contains 3 types of typical noise recordings (including baseline wander, EMG artifacts, and motion artifacts) that are often present in stress tests caused by motion-related interference. The noises are collected by placing the electrodes on the patient’s limbs without the presence of ECG signals. To obtain different noising ECG signals, we corrupt the normal ECG signal from QT with the noise recordings from MIT-BIH NST using different random interference factors ranging from 0.2 to 2.5. Notably, the specific pre-processing of the dataset in our work is following that of DeepFilter [12].

4.1.3. *Evaluation Metrics.* To quantitatively evaluate different ECG denoising methods, we use four similarity-based metrics [34, 35], including the Sum of the Square of the Distances (SSD), Percentage Root-mean-square Difference (PRD), Absolute Maximum Distance (MAD), and Cosine Similarity (Cos_Sim). Noted that the lower (\downarrow) the values of *SSD*, *PRD*, and *MAD* and the higher (\uparrow) the *Cos_Sim* value, the better the denoising effect. By giving an input noisy ECG signal $x(n)$ and an output denoising ECG signal $y(n)$, the calculation formulas of the above metrics are as follows,

$$SSD = \sum_{n=1}^N [x(n) - y(n)]^2 \quad (11)$$

$$PRD = \sqrt{\frac{\sum_{n=1}^N [x(n) - y(n)]^2}{\sum_{n=1}^N [x(n) - \frac{1}{N} \sum_{n=1}^N x(n)]^2}} \times 100\% \quad (12)$$

$$MAD = \max |x(n) - y(n)|, \text{ for } 1 \leq n \leq N \quad (13)$$

$$\text{cos_Sim} = \frac{\langle x, y \rangle}{\|x\| \|y\|} \quad (14)$$

4.2. **ECG Signal Denoising Results.** We compared the DMAM-ECG with five existing representative ECG denoising methods qualitatively and quantitatively, including DRNN [9], FCN-DAE [11], CGAN [13], DeepFilter [12], and DeSocD-ECG [27]. Meanwhile, to better observe the denoising abilities of all methods in different degrees of noise interference, we divided the noise into different segments according to the random factor, i.e. 0.2 to 0.6, 0.6 to 1.0, 1.0 to 1.5, 1.5 to 2.0, and an extreme noise interference (2.0 to 2.5).

4.2.1. *Qualitative Results.* To observe intuitively the denoising effect of the proposed DMAM-ECG, we provide the visual denoised results of it with different shot reconstructions on each noise segment. It is worth noting that, we use the same signal reconstruction strategy of DeSocD-ECG [27] to improve reconstruction accuracy, i.e., we test 3, 5, and 10-shot averages for ECG signal denoising in this paper. As shown in Figure 6, the clean, noisy, and denoised ECG signals are marked in red, blue, and green wave lines, respectively. Figure 6 can reflect that the DMAM-ECG achieves effective noise removal for all noise segments. Even for extremely noisy signals (the blue) that have undergone significant deviation, the denoised signals (the green) are restored and match the clean signals (the red). In specific, From the Figure 6(a) to (d), for the lower noise segments, including 0.2 to 0.6 and 0.6 to 1.0, the denoised results of the DMAM-ECG with 10-shot reconstructions are matching better with the original clean signal overall. And for the higher noise segments, including 1.0 to 1.5 and 1.5 to 2.0 noise segments, there are certain degrees of wave line mismatch. For example, in those signal regions with steeper frequency changes, it can clearly see the red wave lines of the original clean signal. This indicates that our DMAM-ECG processes better for low level noisy ECG signals.

Additionally, to validate that our DMAM-ECG can be more effective for high noise removal, we compared it with the best comparing algorithm DeSocD-ECG (10-shot) on the extreme noise segment (2.0-2.5). As shown in Figure 7, it can be observed that DMAM-ECG surpasses DesocD-ECG (10-shot) with only 3-shot, and the results after 5-shot and 10-shot can be further improved, manifested as being close to the clean signal wave lines.

TABLE 2. Quantitative comparison results of different methods on processing 0.2 to 0.6 noise segments. The best results are marked in bold.

Methods	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	0.2 to 0.6			
DRNN [9]	4.012 ± 7.751	40.415 ± 19.997	0.402 ± 0.296	0.919 ± 0.084
FCN-DAE [11]	7.421 ± 8.052	76.752 ± 41.890	0.493 ± 0.286	0.839 ± 0.121
CGAN [13]	3.461 ± 6.054	39.090 ± 19.696	0.297 ± 0.175	0.928 ± 0.060
DeepFilter [12]	2.421 ± 3.442	34.341 ± 14.564	0.296 ± 0.244	0.945 ± 0.068
DeSocD-ECG(10-shot) [27]	2.421 ± 3.297	28.794 ± 15.534	0.218 ± 0.178	0.959 ± 0.040
Ours(3-shot)	1.804 ± 4.035	27.126 ± 14.006	0.214 ± 0.165	0.965 ± 0.038
Ours(5-shot)	1.744 ± 4.307	26.490 ± 13.359	0.210 ± 0.167	0.967 ± 0.035
Ours(10-shot)	1.630 ± 3.853	25.902 ± 13.313	0.206 ± 0.161	0.969 ± 0.033

TABLE 3. Quantitative comparison results of different methods on processing 0.6 to 1.0 noise segments. The best results are marked in bold.

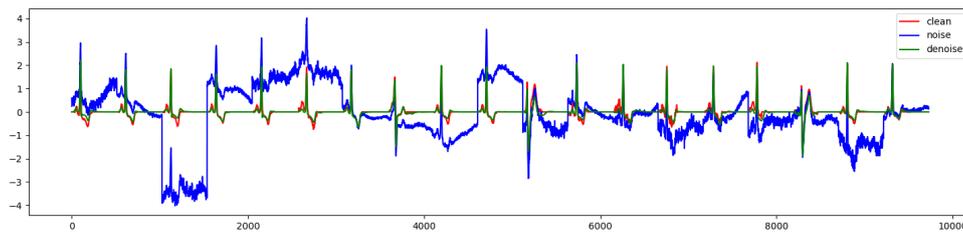
Methods	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	0.6 to 1.0			
DRNN [9]	4.988 ± 6.782	46.382 ± 22.103	0.452 ± 0.301	0.903 ± 0.087
FCN-DAE [11]	8.742 ± 8.325	87.145 ± 57.057	0.551 ± 0.361	0.863 ± 0.151
CGAN [13]	5.518 ± 7.994	48.972 ± 24.694	0.370 ± 0.215	0.891 ± 0.085
DeepFilter [12]	3.945 ± 4.841	43.834 ± 20.093	0.341 ± 0.222	0.913 ± 0.084
DeSocD-ECG(10-shot) [27]	3.249 ± 4.424	38.470 ± 19.219	0.321 ± 0.246	0.936 ± 0.056
Ours(3-shot)	2.841 ± 4.561	35.162 ± 16.110	0.298 ± 0.194	0.945 ± 0.053
Ours(5-shot)	2.716 ± 4.613	34.236 ± 15.755	0.291 ± 0.189	0.948 ± 0.049
Ours(10-shot)	2.607 ± 4.238	33.769 ± 15.543	0.288 ± 0.187	0.950 ± 0.046

TABLE 4. Quantitative comparison results of different methods on processing 1.0 to 1.5 noise segments. The best results are marked in bold.

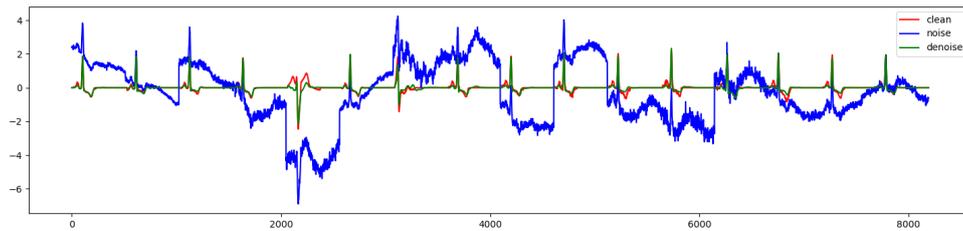
Methods	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	1.0 to 1.5			
DRNN [9]	7.324 ± 11.034	53.527 ± 25.436	0.502 ± 0.419	0.863 ± 0.151
FCN-DAE [11]	10.987 ± 12.993	99.101 ± 68.672	0.625 ± 0.430	0.751 ± 0.177
CGAN [13]	8.579 ± 12.565	60.514 ± 36.254	0.468 ± 0.306	0.838 ± 0.131
DeepFilter [12]	5.881 ± 7.923	54.103 ± 27.361	0.415 ± 0.283	0.874 ± 0.109
DeSocD-ECG(10-shot) [27]	4.436 ± 5.756	45.782 ± 22.999	0.378 ± 0.257	0.916 ± 0.072
Ours(3-shot)	4.139 ± 6.543	42.474 ± 20.829	0.368 ± 0.253	0.923 ± 0.071
Ours(5-shot)	3.913 ± 6.338	41.506 ± 20.474	0.362 ± 0.2489	0.927 ± 0.068
Ours(10-shot)	3.752 ± 6.148	40.890 ± 20.296	0.354 ± 0.242	0.931 ± 0.063

TABLE 5. Quantitative comparison results of different methods on processing 1.5 to 2.0 noise segments. The best results are marked in bold.

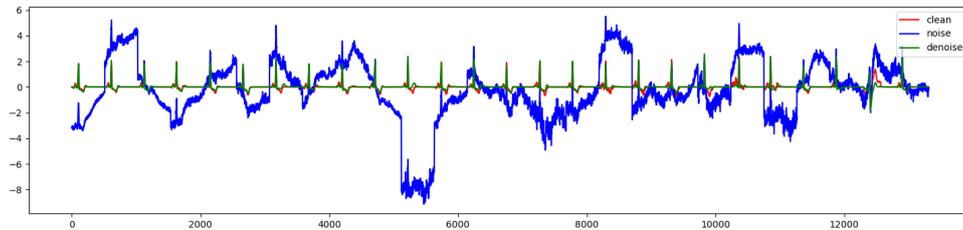
Methods	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	1.5 to 2.0			
DRNN [9]	10.151 ± 15.893	59.983 ± 27.833	0.596 ± 0.451	0.863 ± 0.151
FCN-DAE [11]	14.843 ± 21.831	112.325 ± 95.021	0.741 ± 0.485	0.751 ± 0.177
CGAN [13]	12.382 ± 18.598	74.830 ± 59.722	0.587 ± 0.419	0.838 ± 0.131
DeepFilter [12]	8.152 ± 11.513	66.712 ± 38.841	0.490 ± 0.371	0.773 ± 0.184
DeSocD-ECG(10-shot) [27]	6.161 ± 8.533	56.061 ± 41.694	0.468 ± 0.331	0.883 ± 0.112
Ours(3-shot)	5.828 ± 8.445	50.359 ± 26.899	0.451 ± 0.294	0.894 ± 0.101
Ours(5-shot)	5.525 ± 8.015	49.301 ± 26.663	0.441 ± 0.290	0.901 ± 0.094
Ours(10-shot)	5.205 ± 7.370	48.671 ± 27.008	0.433 ± 0.284	0.906 ± 0.088



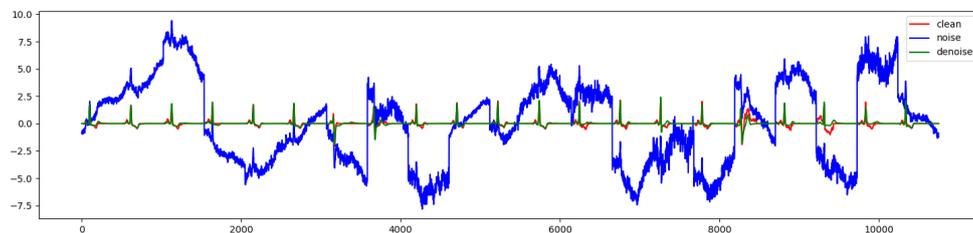
(a) Ours (10-shot) on 0.2 to 0.6



(b) Ours (10-shot) on 0.6 to 1.0



(c) Ours (10-shot) on 1.0 to 1.5

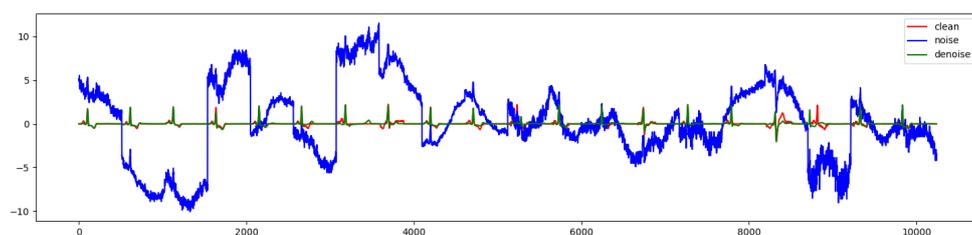


(d) Ours (10-shot) on 1.5 to 2.0

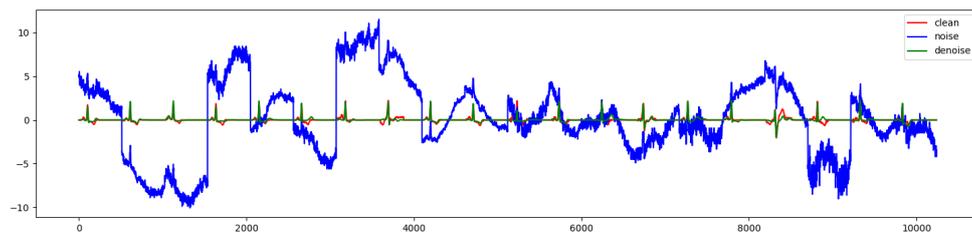
FIGURE 6. Visual denoised results of ours (10-shot) on processing different noise segments. The clean, noisy, and denoised ECG signals are marked in red, blue, and green wave lines, respectively.

TABLE 6. Quantitative comparison results of DeSocD-ECG [27] with 10-shot reconstruction and ours with 3-, 5-, and 10-shot reconstruction. The best results are marked in bold.

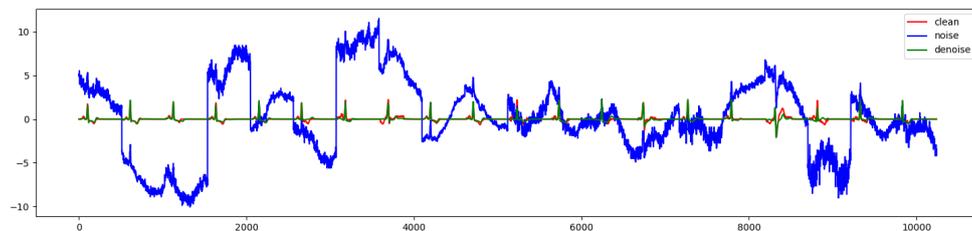
Methods	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	2.0 to 2.5			
DeSocD-ECG(10-shot)	7.936 ± 11.358	68.518 ± 53.907	0.558 ± 0.427	0.835 ± 0.183
Ours(3-shot)	7.031 ± 9.016	58.187 ± 35.147	0.520 ± 0.347	0.863 ± 0.134
Ours(5-shot)	6.474 ± 7.795	56.792 ± 32.396	0.507 ± 0.337	0.872 ± 0.127
Ours(10-shot)	6.094 ± 7.433	55.463 ± 32.011	0.491 ± 0.326	0.881 ± 0.118



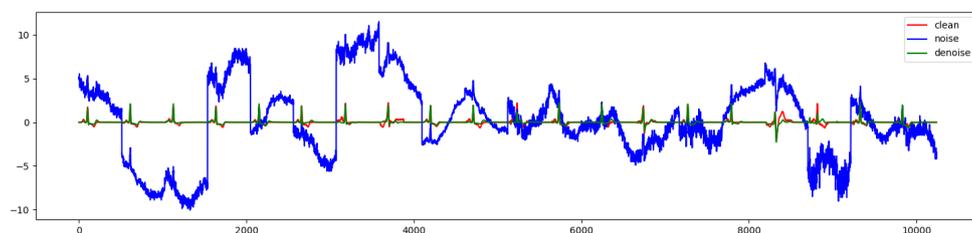
(a) DeScoD-ECG (10-shot)



(b) Ours (3-shot)



(c) Ours (5-shot)



(d) Ours (10-shot)

FIGURE 7. Qualitative comparison results of DeScoD-ECG (10-shot) and ours on processing extreme noise segment (2.0 to 2.5). The clean, noisy, and denoised ECG signals are marked in red, blue, and green wave lines, respectively.

4.2.2. *Quantitative Results.* Quantitative compared results of different methods for processing different noise segments are listed in Tables 2-6, the data marked in bold of all Tables indicate the best denoising results. We demonstrate the compared results of the DMAM-ECG (3-, 5-, and 10-shot reconstruction) with the five state-of-the-art baselines. Among them, we take the best results of DeScoD-ECG (10-shot) [27] for comparison. It can be observed that our DMAM-ECG (10-shot) achieves the best denoising effects on all the noise segments (including 0.2 to 0.6, 0.6 to 1.0, 1.0 to 1.5, 1.5 to 2.0, and the extreme noise 2.0 to 2.5) in all metrics. Specifically, as shown in Tables 2 and 3, the SSD, PRD, and MAD values of ours (10-shot) are less (better) than the best baseline (10-shot) of $1.0 \pm$, $4.0 \pm$, and $0.02 \pm$ in average, respectively. And the Cos_Sim value are higher than it

of $0.03 \pm$ in average. Especially, compared to the best baseline (Descod-ECG with 10-shot reconstruction), the DMAM-ECG has surpassed it with only 3-shot reconstruction for all noise segments, which reflects the superiority of the DMAM-ECG in extreme ECG signal noise removal. Besides, as the number of reconstructions increases, its denoising effect can continue to improve.

TABLE 7. Quantitative comparison results of the DMAM-ECG without and with SAM, as well as different MHAM (-2head, -4head, and -8head) used for different noise segments.

Metrics	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	0.2 to 0.6			
Without SAM	1.817 ± 3.317	27.680 ± 12.951	0.224 ± 0.166	0.965 ± 0.032
With SAM	1.630 ± 3.853	25.901 ± 13.313	0.206 ± 0.161	0.969 ± 0.033
MHAM-2head	2.251 ± 4.193	29.749 ± 14.726	0.229 ± 0.162	0.958 ± 0.040
MHAM-4head	8.484 ± 9.749	50.367 ± 20.094	0.298 ± 0.172	0.842 ± 0.157
MHAM-8head	2.423 ± 4.913	30.203 ± 14.359	0.233 ± 0.154	0.960 ± 0.035
Noise Segment	0.6 to 1.0			
Without SAM	2.919 ± 3.674	36.661 ± 16.803	0.314 ± 0.214	0.942 ± 0.050
With SAM	2.604 ± 4.238	33.769 ± 15.543	0.288 ± 0.187	0.950 ± 0.046
MHAM-2head	3.190 ± 4.021	36.648 ± 17.656	0.310 ± 0.220	0.940 ± 0.053
MHAM-4head	9.700 ± 8.921	55.079 ± 19.672	0.371 ± 0.223	0.822 ± 0.161
MHAM-8head	3.411 ± 4.572	37.413 ± 18.937	0.310 ± 0.215	0.941 ± 0.053
Noise Segment	1.0 to 1.5			
Without SAM	4.136 ± 5.515	44.054 ± 21.544	0.379 ± 0.254	0.919 ± 0.071
With SAM	3.752 ± 6.148	40.89 ± 20.296	0.354 ± 0.242	0.931 ± 0.063
MHAM-2head	4.380 ± 6.381	43.459 ± 21.452	0.379 ± 0.279	0.921 ± 0.069
MHAM-4head	11.099 ± 12.171	60.173 ± 20.709	0.438 ± 0.285	0.799 ± 0.167
MHAM-8head	4.797 ± 7.536	44.965 ± 23.548	0.377 ± 0.274	0.919 ± 0.076
Noise Segment	1.5 to 2.0			
Without SAM	5.939 ± 8.689	52.814 ± 31.434	0.460 ± 0.314	0.885 ± 0.118
With SAM	5.205 ± 7.370	48.671 ± 27.008	0.433 ± 0.284	0.906 ± 0.088
MHAM-2head	5.686 ± 8.011	50.584 ± 26.772	0.448 ± 0.294	0.898 ± 0.088
MHAM-4head	12.171 ± 11.341	65.296 ± 24.255	0.502 ± 0.303	0.779 ± 0.170
MHAM-8head	6.250 ± 9.533	52.630 ± 29.821	0.449 ± 0.284	0.892 ± 0.101
Noise Segment	2.0 to 2.5			
Without SAM	8.016 ± 10.861	64.636 ± 45.136	0.545 ± 0.413	0.827 ± 0.190
With SAM	6.094 ± 7.433	55.463 ± 32.011	0.491 ± 0.326	0.881 ± 0.118
MHAM-2head	6.696 ± 8.238	57.875 ± 33.635	0.505 ± 0.347	0.869 ± 0.124
MHAM-4head	13.582 ± 12.650	71.115 ± 28.153	0.557 ± 0.347	0.748 ± 0.189
MHAM-8head	7.160 ± 9.145	59.851 ± 36.226	0.501 ± 0.329	0.863 ± 0.134

4.3. **Ablation Experiments.** Besides, to validate the significance of the introduced Self-Attention Module (SAM) and the smooth-L1 loss function for our DMAM-ECG, we further conducted related ablation experiments on them.

4.3.1. *Effect of the SAM.* We compared the denoising results of the DMAM-ECG without and with the SAM, as well as different Multi-Head Attention Module (MHAM, including -2head, -4head, and -8head) used for different noise segments. As shown in Figure 8 and Table 7, it can be observed that the denoising ECG results with SAM used to get

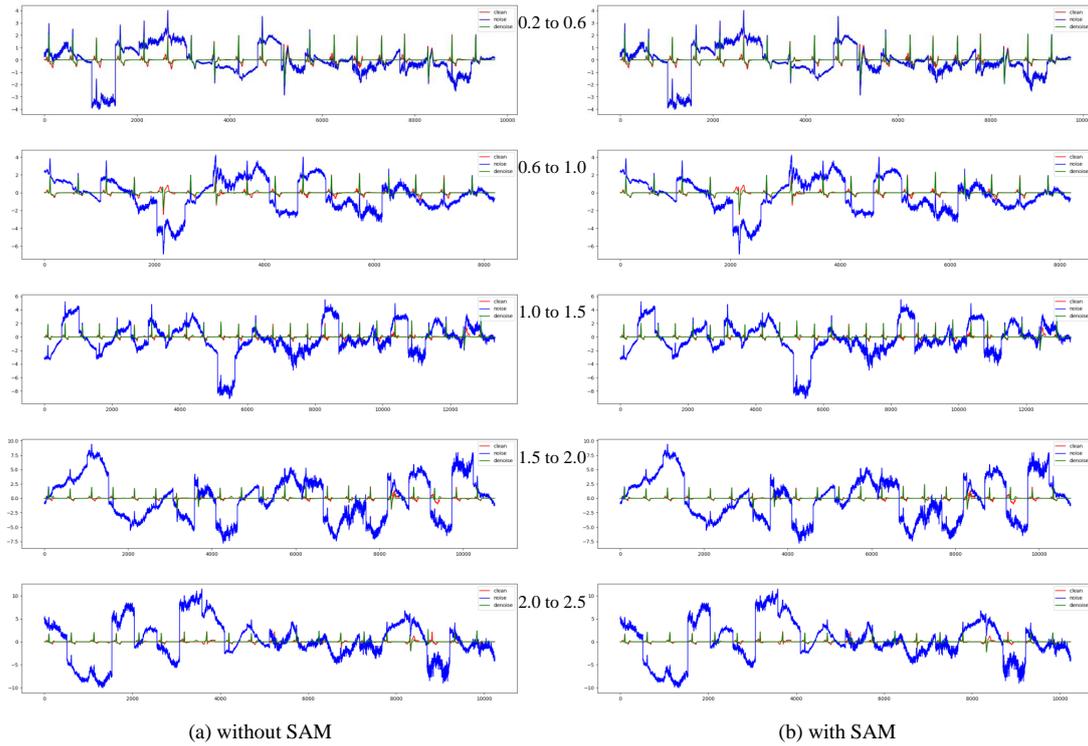


FIGURE 8. Quantitative comparison results of the DMAM-ECG with and without SAM for different noise segments.

higher similarity-based metric values, while MHAM-4head gets the worth metric values for all the noise segments. It is worth noting that the greater the noise intensity, the more significant the improvement in effect, it is thus reflecting the important significance of the introduced SAM for extreme noise conditions.

TABLE 8. Quantitative comparison results of the DMAM-ECG using L1 loss or Smooth L1 loss for different noise segments.

Metrics	$SSD \downarrow$	$PRD \downarrow$	$MAD \downarrow$	$Cos_Sim \uparrow$
Noise Segment	0.2 to 0.6			
L1 loss	2.039 ± 4.935	27.194 ± 15.138	0.199 ± 0.190	0.965 ± 0.039
Smooth L1 loss	1.630 ± 3.853	25.901 ± 13.313	0.206 ± 0.161	0.969 ± 0.033
Noise Segment	0.6 to 1.0			
L1 loss	2.811 ± 4.077	33.899 ± 15.447	0.264 ± 0.174	0.948 ± 0.046
Smooth L1 loss	2.604 ± 4.238	33.769 ± 15.543	0.288 ± 0.187	0.950 ± 0.046
Noise Segment	1.0 to 1.5			
L1 loss	4.242 ± 6.436	42.271 ± 22.230	0.348 ± 0.234	0.925 ± 0.0709
Smooth L1 loss	3.752 ± 6.148	40.89 ± 20.296	0.354 ± 0.242	0.931 ± 0.063
Noise Segment	1.5 to 2.0			
L1 loss	5.516 ± 7.856	51.177 ± 31.071	0.489 ± 0.329	0.900 ± 0.094
Smooth L1 loss	5.205 ± 7.370	48.671 ± 27.008	0.433 ± 0.284	0.906 ± 0.088
Noise Segment	2.0 to 2.5			
L1 loss	6.672 ± 7.899	58.734 ± 36.390	0.489 ± 0.329	0.873 ± 0.115
Smooth L1 loss	6.094 ± 7.433	55.463 ± 32.011	0.491 ± 0.326	0.881 ± 0.118

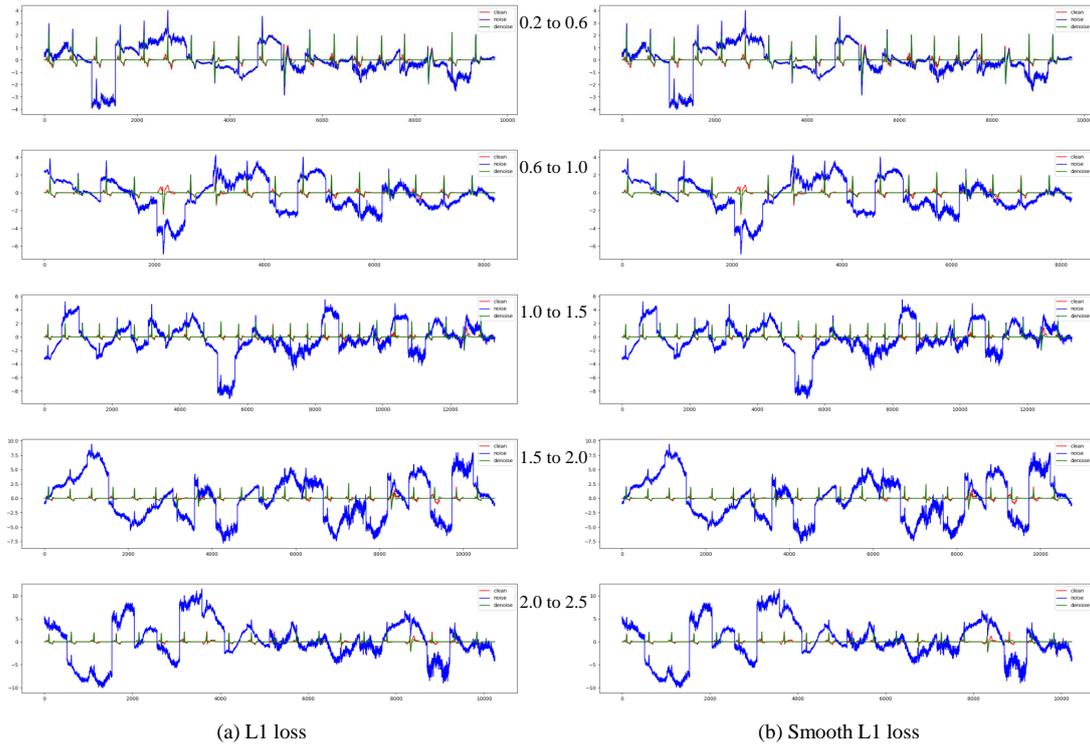


FIGURE 9. Qualitative comparison results of the DMAM-ECG using L1 loss or Smooth L1 loss for different noise segments.

4.3.2. *Effect of Smooth L1 loss.* We compared the effect of L1 loss used in DeScoD-ECG [27] and Smooth L1 loss on ECG signal denoising task. As shown in Figure 9 and Table 8, it can be found that the DMAM-ECCG using Smooth L1 loss gets better denoising effects with higher similarity metric values.

4.3.3. *Time complexity of the DAMA-ECG.* Since the time complexity is a very important performance metric for signal denoising tasks, we conducted the related experiments to validate the efficiency of the DMAM-ECG to achieve real-time performance. We test the average time consuming of processing one cycle ECG signal. The specific time complexities of the DMAM-ECG (including 3-, 5-, and 10-shot) for different noise segments are shown in Table 9. It indicates that the time results show that the average processing time of the DMAM-ECG for one cycle of ECG signals increases overall with the increase of noise level. But it will not exceed 1 second, which can better meet the denoising time requirements of EEG signals.

5. Conclusions. In this paper, we proposed a diffusion model with self-attention module for ECG signal denoising (DMAM-ECG), which can reconstruct the clean ECG signal from random Gaussian noise by learning the conditional distribution in the reverse diffusion process conditioned by noisy ECG observations. Specifically, the DAMA-ECG contains two main branches, each consisting of six designed CIE-Filter blocks, for feature extractions of noisy observations and latent variables. Meanwhile, to reduce the dependence on external information and better capture internal correlations of ECG features, the DMAM-ECCG introduces corresponding six SAM-Fusion blocks to fuse the two branches. A related ablation study was performed to validate the significance of the self-attention module for extreme noise removal. Experimental results demonstrate that

TABLE 9. Time efficiency analysis of the DMAM-ECG (including 3-, 5-, and 10-shot).

DMAM-ECG	3-shot	5-shot	10-shot
Noise Segment	0.2 to 0.6		
Time(seconds)	0.08885	0.14555	0.29117
Noise Segment	0.6 to 1.0		
Time(seconds)	0.08852	0.14796	0.29611
Noise Segment	1.0 to 1.5		
Time(seconds)	0.09170	0.15270	0.30543
Noise Segment	1.5 to 2.0		
Time(seconds)	0.09521	0.15858	0.31766
Noise Segment	2.0 to 2.5		
Time(seconds)	0.09475	0.15803	0.31687

the proposed DMAM-ECG outperforms the existing five state-of-the-art baselines on different degrees of noise interference removal. This indicates that the DMAM-ECG has certain potential application significance in monitoring cardiac activity in various clinical settings.

Acknowledgment. This work was jointly supported by the Young and Middle-aged Teacher Education Research Project of Fujian Province (Science and Technology, Project No. JAT220829), the project of Digital Media Art, Key Laboratory of Sichuan Province (Sichuan Conservatory of Music, Project No. 21DMAKL01), the first batch of industry-university cooperation collaborative education project funded by the Ministry of Education of the People's Republic of China (Minjiang University, Project No. 202101071001), Minjiang University 2021 school-level scientific research project (Minjiang University, Project No. MYK21011), Open Fund Project of Fuzhou Technology Innovation Center of Intelligent Manufacturing Information System (Minjiang University, Grant No. MJUKF-FTICIMIS2022), Open Fund Project of Engineering Research Center for ICH Digitalization and Multi-source Information Fusion (Fujian Polytechnic Normal University, Grant No. G3-KF2204), Guiding Project of Fujian Province (Minjiang University, Project No. 2020H0046). Key Technology Research and Industrialization Project for Software Industry Innovation in Fujian Province (Minjiang University and Fujian Guotong Information Technology Co., Ltd., Project No. 36). Supported by Digital Media Art, Key Laboratory of Sichuan Province, Sichuan Conservatory of Music, Project No.: 23DMAKL02; Supported by The Ministry of Education's first batch of industry-university cooperation collaborative education projects in 2024 (project number: 231104497282106), and The Higher Education Scientific Research Planning Project of Fujian province. No. ZD202309.

REFERENCES

- [1] P. J. Zimetbaum and M. E. Josephson, "Use of the electrocardiogram in acute myocardial infarction," *New England Journal of Medicine*, vol. 348, no. 10, pp. 933–940, 2003.
- [2] E. K. Wang, C.-M. Chen, M. M. Hassan, and A. Almogren, "A deep learning based medical image segmentation technique in internet-of-medical-things domain," *Future Generation Computer Systems*, vol. 108, pp. 135–144, 2020.
- [3] K.-K. Tseng, J. Lin, C.-M. Chen, and M. M. Hassan, "A fast instance segmentation with one-stage multi-task deep neural network for autonomous driving," *Computers & Electrical Engineering*, vol. 93, p. 107194, 2021.

- [4] G. U. Reddy, M. Muralidhar, and S. Varadarajan, "Ecg de-noising using improved thresholding based on wavelet transforms," *International Journal of Computer Science and Network Security*, vol. 9, no. 9, pp. 221–225, 2009.
- [5] F. Zhang, T.-Y. Wu, and G. Zheng, "Video salient region detection model based on wavelet transform and feature comparison," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 58, 2019.
- [6] S. Ari, M. K. Das, and A. Chacko, "Ecg signal enhancement using s-transform," *Computers in Biology and Medicine*, vol. 43, no. 6, pp. 649–660, 2013.
- [7] J. Huang, C. Wang, J. Huang, H. Fan, A. Grau, and F. Zhang, "Ldgcn: An edge-end lightweight dual gcn based on single-channel eeg for driver drowsiness monitoring," *arXiv preprint arXiv:2407.05749*, 2024.
- [8] J. Huang, C. Wang, W. Zhao, A. Grau, X. Xue, and F. Zhang, "Ltdnet-eeg: A lightweight network of portable/wearable devices for real-time eeg signal denoising," *IEEE Transactions on Consumer Electronics*, 2024.
- [9] K. Antczak, "Deep recurrent neural networks for eeg signal denoising," *arXiv preprint arXiv:1807.11551*, 2018.
- [10] P. Xiong, H. Wang, M. Liu, S. Zhou, Z. Hou, and X. Liu, "Ecg signal enhancement based on improved denoising auto-encoder," *Engineering Applications of Artificial Intelligence*, vol. 52, pp. 194–202, 2016.
- [11] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, and S.-Y. Chien, "Noise reduction in eeg signals using fully convolutional denoising autoencoders," *IEEE Access*, vol. 7, pp. 60806–60813, 2019.
- [12] F. P. Romero, D. C. Piñol, and C. R. Vázquez-Seisdedos, "Deepfilter: An eeg baseline wander removal filter using deep learning techniques," *Biomedical Signal Processing and Control*, vol. 70, p. 102992, 2021.
- [13] B. Xu, R. Liu, M. Shu, X. Shang, and Y. Wang, "An eeg denoising method based on the generative adversarial residual network," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1–23, 2021.
- [14] Z. Barati and A. Ayatollahi, "Baseline wandering removal by using independent component analysis to single-channel eeg data," in *International Conference on Biomedical and Pharmaceutical Engineering*, pp. 152–156, IEEE, 2006.
- [15] K. S. Kumar, B. Yazdanpanah, and P. R. Kumar, "Removal of noise from electrocardiogram using digital fir and iir filters with various methods," in *International Conference on Communications and Signal Processing (ICCSP)*, pp. 0157–0162, IEEE, 2015.
- [16] Taouli, A. S., Bereksi-Reguig, and F., "Noise and baseline wandering suppression of eeg signals by morphological filter," *Journal of Medical Engineering & Technology*, vol. 34, no. 2, pp. 87–96, 2010.
- [17] X. Meng, J. Huang, Z. Li, C. Wang, S. Teng, and A. Grau, "Dedustgan: Unpaired learning for image dedusting based on retinex with gans," *Expert Systems with Applications*, vol. 243, p. 122844, 2024.
- [18] J. Yan, Y. Wang, H. Fan, J. Huang, A. Grau, and C. Wang, "Lepf-net: Light enhancement pixel fusion network for underwater image enhancement," *Journal of Marine Science and Engineering*, vol. 11, no. 6, p. 1195, 2023.
- [19] Z. Chen, C. Wang, F. Zhang, L. Zhang, A. Grau, and E. Guerra, "All-in-one aerial image enhancement network for forest scenes," *Frontiers in plant science*, vol. 14, p. 1154176, 2023.
- [20] C. Wang, A. Grau, E. Guerra, Z. Shen, J. Hu, and H. Fan, "Semi-supervised wildfire smoke detection based on smoke-aware consistency," *Frontiers in plant science*, vol. 13, p. 980425, 2022.
- [21] A. Fu, J. Ma, C. Wang, C. Zhou, Z. Li, and S. Teng, "Traditional chinese medicine health status identification with graph attention network," in *International Conference on Machine Learning for Cyber Security*, pp. 1–14, Springer, 2022.
- [22] J. Huang, H. Xu, G. Liu, C. Wang, Z. Hu, and Z. Li, "Sidnet: a single image dedusting network with color cast correction," *Signal Processing*, vol. 199, p. 108612, 2022.
- [23] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," *arXiv preprint arXiv:2010.02502*, 2020.
- [24] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.
- [25] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [26] P. Vincent, "A connection between score matching and denoising autoencoders," *Neural Computation*, vol. 23, no. 7, pp. 1661–1674, 2011.

- [27] H. Li, G. Ditzler, J. Roveda, and A. Li, “Descod-ecg: Deep score-based diffusion model for ecg baseline wander and noise removal,” *IEEE Journal of Biomedical and Health Informatics*, pp. 1–11, 2023.
- [28] A. Q. Nichol and P. Dhariwal, “Improved denoising diffusion probabilistic models,” in *International Conference on Machine Learning*, pp. 8162–8171, PMLR, 2021.
- [29] V. Dumoulin, E. Perez, N. Schucher, F. Strub, H. d. Vries, A. Courville, and Y. Bengio, “Feature-wise transformations,” *Distill*, vol. 3, no. 7, p. e11, 2018.
- [30] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in Neural Information Processing Systems*, vol. 30, p. 5998–6008, 2017.
- [31] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., “Pytorch: An imperative style, high-performance deep learning library,” *Advances in Neural Information Processing Systems*, vol. 32, pp. 8026–8037, 2019.
- [32] P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody, “A database for evaluation of algorithms for measurement of qt and other waveform intervals in the ecg,” in *Computers in Cardiology*, pp. 673–676, IEEE, 1997.
- [33] G. B. Moody and R. G. Mark, “The mit-bih arrhythmia database on cd-rom and software for use with it,” in *Computers in Cardiology*, pp. 185–188, IEEE, 1990.
- [34] R. Nygaard, G. Melnikov, and A. K. Katsaggelos, “Rate distortion optimal ecg signal compression,” in *International Conference on Image Processing (Cat. 99CH36348)*, vol. 2, pp. 348–351, IEEE, 1999.
- [35] R. Nygaard, G. Melnikov, and A. K. Katsaggelos, “A rate distortion optimal ecg coding algorithm,” in *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 28–40, IEEE, 2001.