DFuse: A Dual-Branch Adaptive Fusion Network for Multi-channel MCG Denoising

Shuai Huang[1], Hao Li[2], You Li[3], Jingbo Zhou[3], Jian Cui[3(✉)], Zhong Zheng[3(✉)], Xin Zhang[4], and Tianzi Jiang[4]

[1] Fudan University, Shanghai, China
[2] ShanghaiTech University, Shanghai, China
[3] Harbin Institute of Technology, Shenzhen, China
[4] Chinese Academy of Sciences, Beijing, China
cuijian@hit.edu.cn, zhengz@hit.edu.cn

Abstract. Magnetocardiography (MCG) offers superior spatial resolution and sensitivity over electrocardiography (ECG) for non-invasive cardiac assessment, yet clinical adoption is limited by costly, complex superconducting quantum interference devices (SQUIDs) and optically pumped magnetometers (OPMs). Tunneling magnetoresistance (TMR) sensors provide a room-temperature, cost-effective alternative, but their elevated intrinsic noise impedes clinical translation. We propose DFuse, a dual-branch adaptive fusion network for multichannel MCG denoising that jointly captures long-range temporal dependencies and preserves diagnostically critical morphological features under cross-channel heterogeneous signal-to-noise ratio (SNR). Evaluated on a self-developed 24-channel TMR system, DFuse achieves SNR improvements of 20.31 dB on simulated data (4,800 ten-second samples, 12 channels) and 15.79 dB on recordings from 22 healthy volunteers. The residual denoising error, converted to an equivalent noise level of $6 \pm 6\,\mathrm{fT}/\sqrt{\mathrm{Hz}}$ at 1 Hz, approaches the sensitivity range reported for clinical-grade SQUID and OPM systems. These results demonstrate that deep learning–enhanced TMR-based MCG enables accessible, high-fidelity cardiac diagnostics in resource-constrained settings. Code is publicly available at `https://github.com/HHarr1son/DFuse`.

Keywords: MCG · TMR · Multi-channel denoising.

1 Introduction

Magnetocardiography (MCG) has emerged as a promising non-invasive modality for cardiac assessment, offering superior spatial resolution and immunity to volume conduction effects compared to electrocardiography (ECG) [16,20]. Conventional MCG measurement relies on superconducting quantum interference devices (SQUIDs) or optically pumped magnetometers (OPMs). While SQUIDs

S. Huang and H. Li contributed equally to this work.
Corresponding author: J. Cui and Z. Zheng

achieve ultra-low noise floors (1–10 fT/$\sqrt{\text{Hz}}$)[21], their dependence on liquid helium cooling and high costs (often exceeding $1 million) severely limit clinical deployment. OPMs offer a cryogen-free alternative at comparable sensitivity (10–100 fT/$\sqrt{\text{Hz}}$)[26], yet require complex optical configurations, precise alkali vapor cell management, and mandatory magnetic shielding, imposing considerable operational burden. Tunneling magnetoresistance (TMR) sensors have emerged as a cost-effective room-temperature alternative, enabling closer skin proximity and stronger signal capture [8,15,23]. However, their relatively high noise floor (3–5 pT/$\sqrt{\text{Hz}}$)[18] causes subtle cardiac features such as P-waves, T-waves, and fine QRS morphology to submerge in noise, precluding their direct use in clinical diagnosis without effective denoising.

Traditional MCG denoising methods—including FIR/IIR filtering [11,14,10], wavelet decomposition [7], VMD [4], and EEMD [17]—are inherently limited in handling nonstationary noise while preserving subtle cardiac features. Although deep learning has achieved remarkable success in ECG denoising via CNNs [5,19,12] and attention-based autoencoders [6,3], these frameworks cannot be directly transferred to TMR-based MCG. Unlike ECG, where noise comprises baseline wander (0.05–2 Hz), electrode motion (0.1–10 Hz), and muscle artifacts (5–500 Hz), TMR-MCG is dominated by sensor-specific 1/f electrical noise (0.1–100 Hz) with non-uniform spectral attenuation, presenting fundamentally different denoising challenges. Regarding MCG-specific denoising, Xing et al. [25] proposed MGU-Net for single-channel MCG denoising, leveraging long-term temporal contexts and inter-beat similarities. However, its single-channel design neglects spatial correlations and inter-channel dependencies inherent in multi-channel MCG recordings. Wang et al. [24] introduced SkipDAEFormer and was validated it on optically pumped magnetometer (OPM)-derived MCG with simulated random mixed noise. However, it diverges significantly from real-world TMR-based MCG scenarios characterized by sensor-specific noise patterns.

To bridge this gap, we propose DFuse, a dual-branch adaptive fusion network tailored for multi-channel TMR-MCG denoising, explicitly addressing channel-dependent and spatially heterogeneous noise corruption. Our contributions can be summarized as:

- We construct two multi-channel MCG benchmarks, including a physics-driven 12-channel simulated dataset derived from PTB-XL and a real 24-channel dataset acquired with a TMR sensor array, filling a critical gap in multi-channel MCG evaluation resources.
- We introduce DFuse, a dual-branch adaptive fusion architecture that explicitly decouples long-range context modeling from morphology-aware representation and integrates them via time- and sample-varying fusion to handle the pronounced cross-channel signal heterogeneity inherent in multi-channel TMR-based MCG.
- Extensive leave-one-subject-out experiments demonstrate state-of-the-art denoising performance, achieving SNR gains of 20.31 dB and 15.79 dB on the simulated and real multi-channel datasets, respectively. In terms of equivalent noise level, the proposed method reduces the noise floor to $6 \pm 6\,\text{fT}/\sqrt{\text{Hz}}$

at 1Hz on the real dataset, thereby achieving a noise level comparable to clinical-grade SQUID and OPM systems (1–100 fT/$\sqrt{\text{Hz}}$) at 1Hz.

2 Method

2.1 Architecture Overview

Multi-channel MCG denoising faces a critical challenge of heterogeneous SNR across channels, where proximal sensors capture strong cardiac signals while distant ones are noise-dominated. Existing methods typically trade off between preserving local morphological details and capturing global structural patterns, as can be seen in Fig. 3(a). We find that a single encoder cannot optimally serve both objectives simultaneously, motivating our dual-branch design that separates these two tasks into dedicated encoders. The **context-aware encoder** extracts global cross-channel structural patterns, while the **morphology-aware encoder** focuses on local signal details.

Our architecture follows a U-Net with K levels, as illustrated in Fig. 1. At each level, the two encoders operate in parallel branches, followed by strided convolution (stride=2) for downsampling. We denote the multi-scale representations as $\{F_\ell^{\text{ctx}}\}_{\ell=0}^{K-1}$ and $\{F_\ell^{\text{mor}}\}_{\ell=0}^{K-1}$, with bottleneck representations Z^{ctx} and Z^{mor}, ensuring identical shapes at each level for subsequent fusion.

Representations from both branches are fused via an **enhanced fusion block** at each level, yielding $\{F_\ell\}_{\ell=0}^{K-1}$ to retain the most salient features from both regimes. The bottleneck features undergo the same fusion, augmented by a two-layer convolutional enhancement to boost representational capacity, producing the fused bottleneck Z. The symmetric decoder comprises K levels: starting from Z, the feature at each level is merged with the corresponding fused representation F_ℓ through a **cross-channel reweighting (CCR)** module, followed by a convolutional block for upsampling. A final 1×1 convolution maps the output back to the input space.

2.2 Details on Core Components

Context-Aware Encoder: Transformer-based architectures with self-attention naturally model global dependencies regardless of temporal distance, making them well-suited for capturing long-range cardiac rhythm structure. However, pure transformer encoders applied to raw 1-D signals lack the local inductive bias critical at early high-resolution stages. To address this, we implement a hybrid feature extraction with two parallel paths. The **convolutional path** comprises a convolutional layer followed by batch normalization, providing local inductive bias. The **transformation path** employs a layer normalization (LN) multi-head self-attention (MHSA) block, with sinusoidal positional encoding (PE) injected prior to attention to overcome permutation invariance. The two paths are merged via channel-wise concatenation and a 1×1 convolution that restores the target channel dimension. This hybrid block is applied to the

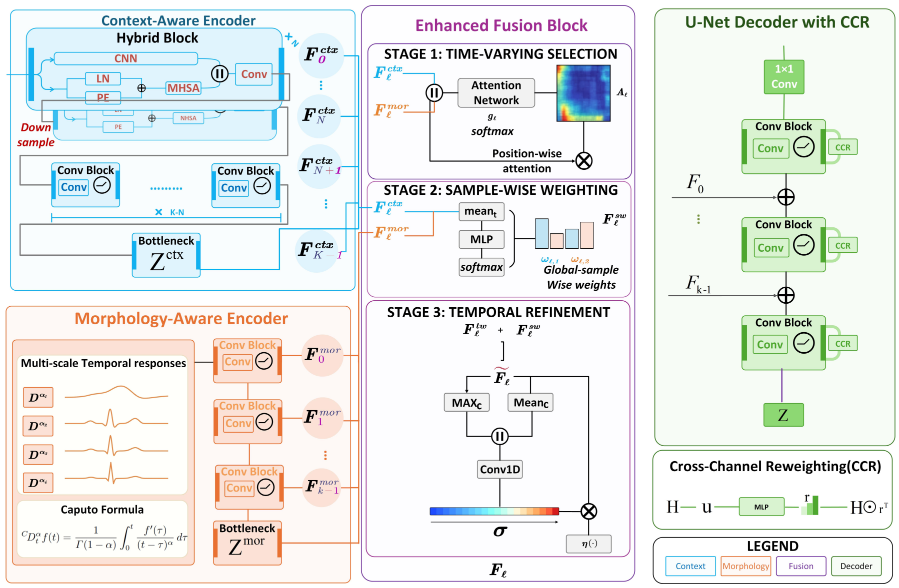

Fig. 1: Overview of DFuse. The network adopts a U-Net structure with five major parts, which are context-aware encorder, morphology-aware encoder, enhanced fusion block and U-Net decorder with CCR.

first N levels. For levels $N+1$ to $K-1$, only the convolutional path is retained, as successive downsampling has reduced spatial resolution sufficiently that convolutional kernels already cover a large temporal span.

Morphology-Aware Encoder: Standard first-order derivatives amplify high-frequency components uniformly, which makes QRS complexes conspicuous but simultaneously amplifies noise and obscures the smooth transitions of P- and T-waves. To address this limitation, we employ Caputo fractional derivatives [2] to decompose MCG signals into sharp high-amplitude features (QRS complexes) and smooth low-amplitude events (P- and T-waves). Specifically, evaluating the fractional operator at M distinct orders transforms the input signal $\mathbf{X} \in \mathbb{R}^{B\times C\times L}$ into $\tilde{\mathbf{X}} \in \mathbb{R}^{B\times C(M+1)\times L}$, where B, C, and L denote batch size, channel number, and sequence length, respectively. The expanded representation is subsequently processed by stacked convolutional blocks to yield multi-scale features $\{F_\ell^{\mathrm{mor}}\}_{\ell=0}^{K-1}$ and the bottleneck representation Z^{mor}.

Enhanced Fusion Block: The input representations F_ℓ^{mor} and F_ℓ^{ctx} are fused in three stages. First, a time-varying selection computes temporal attention vectors $A_{\ell,1}, A_{\ell,2} \in \mathbb{R}^{B\times 1\times L_\ell}$ via a lightweight convolutional gate g_ℓ as $A_\ell = \mathrm{softmax}(g_\ell([F_\ell^{\mathrm{ctx}}; F_\ell^{\mathrm{mor}}])) \in \mathbb{R}^{B\times 2\times L_\ell}$, yielding $F_\ell^{\mathrm{tw}} = A_{\ell,1} \odot F_\ell^{\mathrm{ctx}} + A_{\ell,2} \odot F_\ell^{\mathrm{mor}}$. Second, sample-wise weighting suppresses subject-level variations via recording-level scalar weights $w_{\ell,1}, w_{\ell,2} \in \mathbb{R}^{B\times 1\times 1}$, which are computed by $w_\ell = \text{softmax}(\text{MLP}(\text{mean}_t([F_\ell^{\text{ctx}}; F_\ell^{\text{mor}}])))$, giving $F_\ell^{\text{sw}} = w_{\ell,1} \cdot F_\ell^{\text{ctx}} + w_{\ell,2} \cdot F_\ell^{\text{mor}}$. Finally, the two fused features are then averaged: $\tilde{F}_\ell = \frac{1}{2}(F_\ell^{\text{tw}} + F_\ell^{\text{sw}})$. A spatial attention gate identifies temporally important positions by computing channel-wise max- and mean-pooling at each time step, and producing a scalar temporal mask $s_\ell = \sigma(\text{Conv1D}([\max_c(\tilde{F}_\ell); \text{mean}_c(\tilde{F}_\ell)]))$, and the masked feature is refined by a convolutional block η as $F_\ell = \eta(\tilde{F}_\ell \odot s_\ell)$.

Cross-Channel Reweighting : Suppose the input is $\mathbf{H}$, the module computes a channel-importance vector $\mathbf{u} = \text{mean}_t(\mathbf{H}) \in \mathbb{R}^{B\times C}$ from the temporal mean of the current feature map, treating each channel's average activation as a proxy for its signal reliability. A bottleneck excitation network with reduction ratio $\rho = 4$, formed by $W_1 \in \mathbb{R}^{(C/\rho)\times C}$ and $W_2 \in \mathbb{R}^{C\times(C/\rho)}$, produces channel weights $r = \sigma(W_2 \,\text{ReLU}(W_1\mathbf{u})) \in \mathbb{R}^{B\times C}$, which are broadcast over time and applied as $\mathbf{H} \leftarrow \mathbf{H} \odot r^\top$. Channels with consistently strong activations thus receive higher weights, focusing reconstruction fidelity on the most informative leads while attenuating heavily corrupted ones.

Implementation Details: DFuse is trained end-to-end by minimising the mean squared error $\mathcal{L} = \frac{1}{BCL}\|\hat{\mathbf{y}} - \mathbf{y}\|_F^2$ between the network output $\hat{\mathbf{y}}$ and the reference signal $\mathbf{y}$, where $\|\cdot\|_F$ denotes the Frobenius norm. The encoder has $K = 4$ scales with base channel width 64 doubling at each scale, and $N = 2$ hybrid blocks at shallow stages. Caputo orders of $\{0.3, 0.5, 0.7, 1.0\}$ were used in the morphology-aware encoder. The model is optimised with AdamW ($\beta_1 = 0.9$, $\beta_2 = 0.999$, weight decay 10^{-5}) at an initial learning rate of 2×10^{-4} with gradient clipping at norm 1.0. The learning rate is halved when validation loss plateaus for 15 epochs, with a lower bound of 10^{-6}. Training proceeds for at most 200 epochs with early stopping after 50 epochs without improvement; the checkpoint with the best validation SNR under LOSO is retained.

3 Experiments

3.1 Data acquisition

We design a multi-channel TMR-based MCG acquisition system to enable synchronized magnetic signal recording, as summarized in Fig. 2. Each TMR sensor integrates a Wheatstone bridge formed by four matched TMR elements, and each element comprises tens of thousands of magnetic tunnel junctions connected in hybrid series–parallel configurations, as illustrated in Fig. 2(a–c). A low-noise analog front-end (AFE) with 50 Hz notch filtering is used before digitization, as shown in Fig. 2(d,e). The amplified analog signals are then digitized by a National Instruments data acquisition card (DAQ), providing synchronous sampling across channels. The resulting system achieves a noise floor of 3–5 $\text{pT}/\sqrt{\text{Hz}}$, which supports high-sensitivity MCG measurements.

Simulated Multi-channel MCG dataset We construct a simulated MCG dataset using a Helmholtz coil-based magnetic signal generation platform, as

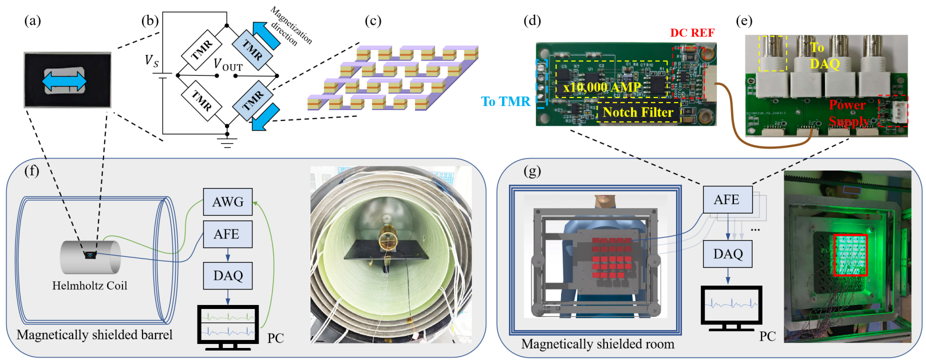

Fig. 2: Multi-channel TMR-based MCG system and experimental setups. (a) A TMR sensor. The blue arrow indicates the sensitive axis direction. (b) A Wheatstone bridge structure consists of four TMR elements. (c) A TMR element is formed by an array comprising tens of thousands of magnetic tunnel junctions (MTJs) arranged in series-parallel configurations. (d) A low-noise amplification circuit board. (e) A 4-channel adapter module. (d) and (e) together form the analog front end (AFE). (f) Experimental setup for simulated Multi-channel data acquisition. (g) Experimental setup for real Multi-channel data acquisition.

shown in Fig. 2(f). The excitation sources are taken from 4,800 normal subjects in the PTB-XL database [22]. The 12-lead ECG recordings are first bandpass filtered from 1 to 35 Hz, after which the lead with the largest peak-to-peak amplitude is normalized to 100–200 pT, and the remaining leads are proportionally scaled using the same factor. The resulting lead signals are sequentially injected into the simulated acquisition system to obtain noisy signals. The simulated dataset therefore contains true signal labels (original waveforms before noise injection) and authentic TMR noise, validating whether the models can learn global structured patterns rather than fitting arbitrary noise, and whether they can suppress TMR-specific noise while preserving morphology. Each sample spans 10 seconds, yielding 4,800 simulated 12-channel MCG samples in total.

Real Multi-channel MCG dataset Real MCG data are collected in a magnetically shielded room using a 24-channel tunneling magnetoresistance (TMR) system with dedicated amplification circuits. The 24 sensors are arranged in a 5×5 grid with one corner position vacant, while a height-adjustable support structure maintains stable placement above the subject's chest, as shown in Fig. 2(g). Twenty-two healthy volunteers aged 18–35 years participate, with recordings performed in the supine position. This study adheres to the ethical guidelines of the Harbin Institute of Technology (HIT), China and was approved by Human Subjects Research Ethics Review Committee of HIT (Ref: HIT-26YS-024). The raw signals are bandpass filtered from 1 to 35 Hz, and R-wave peaks are used to align cardiac cycles and average 101 cycles to obtain a reference

Table 1: Performance comparison on MCG denoising under LOSO (mean±std). ↓ means lower is better and ↑ means higher is better. Best , Second , Third .

Stimulated 12-channel MCG						
Method	**SSD ↓**	**MAD ↓**	**CosineSim ↑**	**SNR ↑**	**SNR$_{\text{imp}}$ ↑**	**NoisLvL ↓**
FCN_DAE [5]	9.787±2.826	0.412±0.121	0.000±0.000	-8.41±0.03	0.33±0.11	0.447±0.129
CBAM-DAE [6]	0.330±0.930	0.163±0.056	0.640±0.184	0.56±1.70	9.30±2.48	0.015±0.042
Ude_Net [19]	0.033±0.016	0.098±0.016	0.906±0.012	5.91±0.67	14.65±0.65	0.002±0.001
TCDAE [3]	0.125±0.150	0.148±0.053	0.652±0.168	0.65±0.58	9.39±1.81	0.006±0.007
Swin-UNet [1]	0.173±0.023	0.197±0.004	0.002±0.001	-1.80±0.57	6.94±0.49	0.008±0.002
SkipDAEformer [24]	0.060±0.039	0.133±0.003	0.807±0.012	2.64±0.49	11.38±0.41	0.003±0.001
LUnet [12]	0.728±0.353	0.220±0.027	0.001±0.004	-7.57±0.74	1.17±0.95	0.033±0.016
DWT [7]	94.708±94.564	0.254±0.031	0.222±0.022	-14.46±0.04	-5.72±0.17	4.326±4.319
FIR [13]	48.135±6.565	0.257±0.021	0.288±0.016	-12.10±0.06	-3.36±0.47	2.199±0.300
IIR [9]	51.335±1.893	0.261±0.018	0.286±0.018	-10.57±0.06	-1.83±0.44	2.345±0.086
Baseline	29.737±3.646	0.204±0.008	-0.223±0.004	-3.15±0.25	5.59±0.52	1.358±0.167
DFuse (Ours)	**0.008±0.004**	**0.055±0.002**	**0.945±0.003**	**11.57±0.19**	**20.31±0.19**	**0.000±0.000**
Real 24-channel MCG						
Method	**SSD ↓**	**MAD ↓**	**CosineSim ↑**	**SNR ↑**	**SNR$_{\text{imp}}$ ↑**	**NoisLvL ↓**
FCN_DAE [5]	1.416±0.434	0.116±0.018	0.000±0.001	-5.25±0.97	3.72±2.25	0.062±0.020
CBAM-DAE [6]	0.328±0.297	0.073±0.360	0.644±0.171	1.70±1.68	10.67±3.39	0.015±0.014
Ude_Net [19]	0.297±0.288	0.063±0.020	0.733±0.058	2.53±1.15	11.50±3.87	0.014±0.013
TCDAE [3]	0.310±0.289	0.067±0.018	0.696±0.093	1.87±1.27	10.84±3.52	0.014±0.013
Swin-UNet [1]	0.701±0.340	0.104±0.020	0.034±0.025	-1.45±0.01	7.52±3.00	0.032±0.015
SkipDAEformer [24]	0.366±0.246	0.087±0.015	0.653±0.074	0.88±0.89	9.85±3.00	0.017±0.011
LUnet [12]	0.698±0.324	0.103±0.026	-0.001±0.007	-1.43±0.08	7.54±3.24	0.032±0.015
DWT [7]	34.398±4.303	0.221±0.038	0.435±0.113	-11.36±0.80	-2.39±1.05	1.571±0.196
FIR [13]	57.074±8.154	0.251±0.039	0.449±0.106	-10.60±0.58	-1.64±1.78	2.607±0.372
IIR [9]	61.122±8.419	0.259±0.039	0.447±0.109	-11.96±0.68	-2.99±1.03	2.792±0.385
Baseline	31.246±17.416	0.284±0.037	-0.049±0.087	-3.70±0.95	5.27±1.50	1.427±0.796
DFuse (Ours)	**0.133±0.123**	**0.045±0.016**	**0.834±0.050**	**6.82±0.79**	**15.79±1.62**	**0.006±0.006**

waveform. Continuous recordings are segmented into 10-second clips containing multiple cycles, producing approximately 2,000 24-channel samples.

3.2 Results

Comparison with State-of-the-Art Methods We compare DFuse with ten state-of-the-art baselines under the LOSO protocol. Results are shown in Table 1. DFuse achieves the best overall performance across all metrics, yielding SNR improvements of 20.31 dB and 15.79 dB on the simulated and real datasets. The baseline is obtained by cycle-averaging within each sample and rearranging to match the original length. We further estimate residual Noise Level (NoisLvL) by converting SSD to noise PSD at 1 Hz following [18]. Specifically, SSD was normalized by the sample count per segment to obtain root-mean-square (RMS) noise, from which equivalent PSD was estimated via the integral relationship between PSD and squared RMS over the bandwidth. DFuse achieves $0.363 \pm 0.178\,\mathrm{fT}/\sqrt{\mathrm{Hz}}$ and $6 \pm 6\,\mathrm{fT}/\sqrt{\mathrm{Hz}}$ on the two benchmarks, approaching the sensitivity range of SQUID and OPM systems. Qualitative comparisons in Fig. 3 further support these findings.

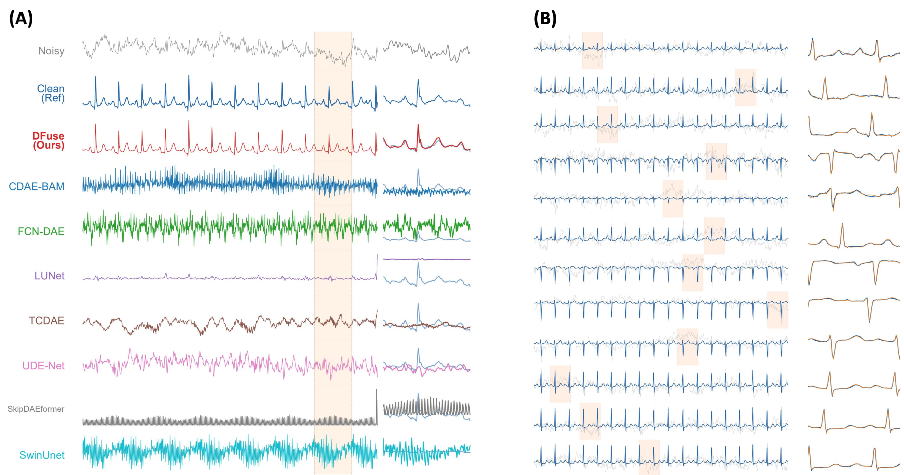

Fig. 3: Qualitative comparisons on the simulated 12-channel MCG benchmark under LOSO. (A) Representative-channel denoising results for DFuse and competing methods. It shows DFuse suppresses strong interference while preserving QRS, P-, and T-wave morphology, whereas competitors suffer from residual oscillations, ringing, or over-smoothing. (B) Lead-wise visualization across all 12 channels. DFuse demonstrates consistent recovery across all 12 channels with effective suppression of spatially varying artifacts.

Table 2: Ablation study on both simulated 12-channel and real 24-channel MCG under LOSO. ✓ indicates the component is included, ✗ indicates it is removed.

Model Configuration	**Dual-Branch**	**Enhanced Fusion**	**CCR**	**12-channel MCG**			**24-channel MCG**		
				SSD ↓	**CosineSim** ↑	$\mathbf{SNR_{imp}}$ ↑	**SSD** ↓	**CosineSim** ↑	$\mathbf{SNR_{imp}}$ ↑
(a) Progressive Component Addition									
Baseline (single-branch U-Net)	✗	✗	✗	17.36	0.0019	6.94	0.7008	0.0340	7.52
+ Dual-branch encoder	✓	✗	✗	5.91	0.8068	11.38	0.3662	0.6529	9.85
+ Enhanced Fusion	✓	✓	✗	3.31	0.9063	14.65	0.2974	0.7329	11.50
Full model (DFuse)	✓	✓	✓	**0.0794**	**0.9445**	**20.31**	**0.1333**	**0.8335**	**15.79**
(b) Context-Aware Encoder Variants									
Pure convolutional	✓	✓	✓	4.23	0.8712	12.87	0.3162	0.7012	10.45
Pure Transformer	✓	✓	✓	5.67	0.8234	11.56	0.4009	0.6387	9.23
Hybrid (Conv + Transformer)	✓	✓	✓	**0.0794**	**0.9445**	**20.31**	**0.1333**	**0.8335**	**15.79**
(c) Morphology-Aware Encoder Variants									
Standard integer-order derivative	✓	✓	✓	4.78	0.8589	12.34	0.3384	0.6892	10.12
Single fractional order ($\alpha = 0.5$)	✓	✓	✓	2.87	0.8912	13.67	0.2651	0.7456	11.89
Multi-order Caputo (ours)	✓	✓	✓	**0.0794**	**0.9445**	**20.31**	**0.1333**	**0.8335**	**15.79**
(d) Fusion Strategy Variants									
Simple addition	✓	✗	✓	4.56	0.8623	12.56	0.3291	0.6934	10.34
Simple concatenation	✓	✗	✓	4.12	0.8745	13.12	0.3112	0.7123	10.89
Time-varying selection only	✓	✗	✓	2.34	0.9123	15.87	0.2201	0.7689	12.67
Sample-wise weighting only	✓	✗	✓	2.56	0.9078	15.43	0.2326	0.7578	12.34
Time-varying + Sample-wise (no refine)	✓	✗	✓	1.89	0.9234	17.23	0.1925	0.7923	13.78
Enhanced Fusion (all three stages)	✓	✓	✓	**0.0794**	**0.9445**	**20.31**	**0.1333**	**0.8335**	**15.79**
(e) Cross-Channel Reweighting Variants									
Without CCR	✓	✓	✗	3.31	0.9063	14.65	0.2974	0.7329	11.50
Channel-wise average pooling	✓	✓	✗	1.67	0.9287	17.89	0.2040	0.7856	13.89
Attention-based CCR (ours)	✓	✓	✓	**0.0794**	**0.9445**	**20.31**	**0.1333**	**0.8335**	**15.79**

Ablation Study We conduct leave-one-subject-out ablations on both benchmarks, with results in Table 2. Starting from a single-branch U-Net baseline, we progressively add each component of DFuse. The dual-branch encoder, enhanced fusion block, and CCR each yield consistent gains, confirming that decoupling context and morphology modeling, adaptive cross-branch integration, and inter-lead dependency modeling are all critical.

4 Conclusion

We proposed DFuse for multi-channel MCG denoising and achieved significant denoising performance on TMR-based MCG—yielding SNR gains of 20.31 dB (simulated) and 15.79 dB (real data from 22 healthy volunteers), with equivalent noise levels approaching clinical SQUID/OPM systems. Future work will extend DFuse to patient cohorts with diverse cardiac conditions, and will explore multi-loss frameworks to improve model stability across populations. Nevertheless, this study demonstrates that deep learning–enhanced TMR-MCG holds strong potential for affordable, accessible cardiac diagnostics upon successful clinical translation.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (62327805, 62336007) and the STI2030-Major Projects (2021ZD0200201).

Disclosure of Interests. The authors have no competing interests to disclose that are relevant to the content of this article.

References

1. Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M.: Swin-unet: Unet-like pure transformer for medical image segmentation. In: Lecture Notes in Computer Science. vol. 13662, pp. 205–218 (2022)
2. Caputo, M.: Linear models of dissipation whose q is almost frequency independent—ii. Geophysical Journal International **13**(5), 529–539 (11 1967). `https://doi.org/10.1111/j.1365-246X.1967.tb02303.x`, `https://doi.org/10.1111/j.1365-246X.1967.tb02303.x`
3. Chen, M., Li, Y., Zhang, L., Liu, L., Han, B., Shi, W., Wei, S.: Elimination of random mixed noise in ecg using convolutional denoising autoencoder with transformer encoder. IEEE Journal of Biomedical and Health Informatics **28**(4), 1993–2004 (2024)
4. Chen, M., Cheng, Q., Feng, X., Zhao, K., Zhou, Y., Xing, B., Tang, S., Wang, R., Duan, J., Wang, J., et al.: Optimized variational mode decomposition algorithm based on adaptive thresholding method and improved whale optimization algorithm for denoising magnetocardiography signal. Biomedical Signal Processing and Control **88**, 105681 (2024)
5. Chiang, H.T., Hsieh, Y.Y., Fu, S.W., Hung, K.H., Tsao, Y., Chien, S.Y.: Noise reduction in ecg signals using fully convolutional denoising autoencoders. IEEE Access **7**, 60806–60813 (2019)
6. Chorney, W., Wang, H., He, L., Lee, S., Fan, L.W.: Convolutional block attention autoencoder for denoising electrocardiograms. Biomedical Signal Processing and Control **86**, 105242 (2023)
7. Dong, Y., Shi, H., Luo, J., Fan, G., Zhang, C.: Application of wavelet transform in mcg-signal denoising. Modern Applied Science **4**(6), 20–24 (2010)
8. Fujiwara, K., et al.: Magnetocardiography and magnetoencephalography measurements at room temperature using tunnel magneto-resistance sensors. Applied Physics Express **11**(2), 023001 (2018)
9. Han, X., Pang, J., Xu, D., Wang, R., Xie, F., Yang, Y., Sun, J., Li, Y., Li, R., Yin, X., Xu, Y., Fan, J., Dong, Y., Wu, X., Yang, X., Yu, D., Wang, D., Gao, Y., Xiang, M., Xu, F., Sun, J., Chen, Y., Ning, X.: Magnetocardiography-based coronary artery disease severity assessment and localization using spatiotemporal features. Physiological Measurement **44**, 125002 (2023)
10. Han, X., Pang, J., Xu, D., Wang, R., Xie, F., Yang, Y., Sun, J., Li, Y., Li, R., Yin, X., et al.: Magnetocardiography-based coronary artery disease severity assessment and localization using spatiotemporal features. Physiological Measurement **44**(12), 125002 (2023)
11. Heinonen, P., Saramaki, T., Malmivuo, J., Neuvo, Y.: Periodic interference rejection using coherent sampling and waveform estimation. IEEE transactions on circuits and systems **31**(5), 438–446 (1984)
12. Hu, L., Cai, W., Chen, Z., Wang, M.: A lightweight u-net model for denoising and noise localization of ecg signals. Biomedical Signal Processing and Control **88**, 105504 (2024)
13. Lu, L., Yang, K., Kong, X., Yang, R., Wang, Y.: A real-time, automatic mcg signal quality evaluation method using the magnetocardiography and electrocardiography. IEEE Transactions on Applied Superconductivity **28**(7), 1602106 (2018)
14. Lu, L., Yang, K., Kong, X., Yang, R., Wang, Y.: A real-time, automatic mcg signal quality evaluation method using the magnetocardiography and electrocardiography. Ieee Transactions on Applied Superconductivity **28**(7), 1–6 (2018)

15. Lu, Z., Ji, S., Yang, J.: Measurement of T wave in magnetocardiography using tunnel magnetoresistance sensor. Chinese Physics B **32**(2), 020703 (2023)
16. Malmivuo, J.: Comparison of the properties of eeg and meg in detecting the electric activity of the brain. Brain Topography **25**(1), 1–19 (2012). https://doi.org/10.1007/s10548-011-0202-1
17. Mariyappa, N., Sengottuvel, S., Parasakthi, C., Gireesan, K., Janawadkar, M., Radhakrishnan, T., Sundar, C.: Baseline drift removal and denoising of mcg data using eemd: role of noise amplitude and the thresholding effect. Medical Engineering & Physics **36**(10), 1266–1276 (2014)
18. Oogane, M., et al.: Sub-pT magnetic field detection by tunnel magneto-resistive sensors. Applied Physics Express **14**(12), 123002 (2021)
19. Qiu, L., Cai, W., Zhang, M., Zhu, W., Wang, L.: Two-stage ecg signal denoising based on deep convolutional network. Physiological Measurement **42**, 115002 (2021)
20. Steinberg, B.A., Roguin, A., Watkins III, S.P., Hill, P., Fernando, D., Resar, J.R.: Magnetocardiogram recordings in a nonshielded environment—reproducibility and ischemia detection. Annals of Noninvasive Electrocardiology **10**(2), 152–160 (2005)
21. Sternickel, K., Braginski, A.I.: Biomagnetism using SQUIDs: status and perspectives. Superconductor Science and Technology **19**(3), S160 (2006)
22. Wagner, P., Strodthoff, N., Bousseljot, R.D., Kreiseler, D., Lunze, F.I., Samek, W., Schaeffter, T.: Ptb-xl, a large publicly available electrocardiography dataset. Scientific Data **7**(1), 1–15 (2020)
23. Wang, M., Peng, L., Ye, C.: Magnetocardiography measurements with a high sensitivity tunnel magnetoresistance sensor. In: Proceedings of the IEEE International Conference on Sensing and Nanotechnology. pp. 1–4 (2019)
24. Wang, R., Liu, Z., Pang, J., Sun, J., Xiang, M., Ning, X.: Skipdaeformer: A high-precision representation learning method for removing random mixed noise in mcg signals. IEEE Journal of Biomedical and Health Informatics **29**(11), 8086–8099 (2025)
25. Xing, Z., Li, H., Dou, H., Zheng, Z., Dai, J., Wang, C., Cui, J., Zhang, X., Jiang, T.: Multi-level gated u-net for denoising tmr sensor-based mcg signals. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 428–437. Springer (2025)
26. Zhao, B., Li, L., Zhang, Y., Tang, J., Liu, Y., Zhai, Y.: Optically pumped magnetometers recent advances and applications in biomagnetism: A review. IEEE Sensors Journal **23**(17), 18949–18962 (2023)