Contrastive Disentanglement Learning Framework for Multi-lead Wearable ECG Denoising

Yue Zhang[1,2,3], Chenyu Zhao[1,2,3], Wen Zhang[1,2,3], Jinliang Wang[4], Jun Guo[5], Wei Yang[1,2,3] ⋆, and Qianjin Feng[1,2,3] ⋆⋆

[1] School of Biomedical Engineering, Southern Medical University, Guangzhou, China
[2] Guangdong Provincial Key Laboratory of Medical Image Processing, Guangzhou, China
[3] Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Guangzhou, China
[4] CardioCloud Medical Technology (Beijing) Co., Ltd., Beijing, China
[5] Department of Cardiology, Chinese PLA General Hospital, Beijing, China
fengqj99@smu.edu.cn; weiyanggm@gmail.com

Abstract. Electrocardiogram (ECG) denoising enhances the clarity of noisy signals while preserving or even improving diagnostic performance. Most existing single-lead denoising algorithms require a preliminary noise assessment across all 12 leads, discarding clean leads and denoising only the noisy leads. In this paper, a novel disentanglement learning denoising network is proposed for 12-lead wearable ECG that directly processes 12-lead ECG, denoising noisy leads while preserving clean leads. Specifically, the proposed network takes both raw ECG and its corresponding simulated noisy ECG as inputs, disentangling them into noise codes and signal content codes. To ensure consistency between the content codes from two inputs, a discriminator is introduced. Additionally, considering that clean leads within the same ECG can provide valuable information for denoising noisy leads, a lead encoder is designed to extract lead specific features from the raw ECG. A contrastive loss is then applied between the features of noisy and clean leads to optimize the model. The results demonstrate that our method achieves superior denoising performance across two different lead system test datasets. Furthermore, evaluations on an ST-segment change multi-label classification task indicate that the denoised ECG improve diagnostic AUC and AUPRC. Furthermore, our model can be used into remote wearable ECG diagnostic workflows, providing preliminary noise reduction to assist cardiologists in subsequent clinical assessments.

Keywords: Multi-lead wearable ECG · Contrastive disentanglement learning · Denoising.

⋆ Corresponding author
⋆⋆ Corresponding author

1 Introduction

Electrocardiogram (ECG) is a cost-effective and non-invasive cardiac examination. It reflects the health status of the cardiovascular system and serves as an essential tool for diagnosing cardiovascular diseases. However, ECG is frequently affected by noises, especially during home self-monitoring due to improper operation or patient movement. Noisy ECG pose challenges for interpretation and hinders accurate diagnosis by cardiologists. To avoid the inconvenience of repeated data collection and to enhance the visual clarity of noisy ECG while preserving diagnostic performance, an effective denoising algorithm for wearable 12-lead ECG is required.

Previous research has made many attempts to enhance noisy ECG. Classical methods such as digital filtering remove noise through frequency domain transformations[10, 1]. However, these methods perform poorly when ECG is severely noised. With the increasing availability of datasets and computational resources, machine learning and deep learning have gained popular in ECG denoising[8, 2, 18]. Various generative models have been investigated, employing supervised denoising by introducing simulated noise to clean ECG. Notable examples include generative adversarial networks (GANs)[20], and autoencoders[4], all of which have demonstrated promising results in enhancing ECG signal quality[16, 12, 6].

Despite the remarkable performance of deep learning ECG denoising methods, certain limitations hinder their practical application. Most methods are designed for single-lead ECGs[17], but clinical and wearable ECG devices commonly employ 12 leads configurations[15, 7, 9]. Notably, noise does not affect all leads uniformly, some leads may be noisy while others remain clean, as shown in Fig. 1(a). Consequently, single-lead methods require prior labeling of noisy leads, increasing the workload for cardiologists[3]. Moreover, it is worth exploring whether the inherent contrastive relationships between clean and noisy leads within the same ECG can improve denoising performance.

In this study, we propose a contrastive representation disentanglement framework for 12-lead wearable ECG denoising, termed CD-ECGNet. This framework takes 12-lead ECG as input and removes noise from noised leads while preserving clean leads. It reduces the workload for cardiologists by avoiding prior identification of noise leads. We assume a noisy ECG consists of a noise code (e.g., baseline wander) and a signal content code (e.g., cardiac information). To achieve effective denoising, these two codes are disentangled in the latent space, and the clean ECG is reconstructed using only the signal content while discarding the noise code. Since paired clean and noisy ECG are unavailable, simulated noise is introduced into clean ECG leads. Both the original ECG and the simulated ECG take the same disentanglement process in the latent space. To ensure that the signal content code remains noise-invariant between the original and simulated ECG, a domain discriminator is introduced, and adversarial learning is employed[21].

Moreover, clean leads can provide auxiliary supervision to facilitate the denoising of noisy leads during training. To leverage the contrastive relationships between clean and noisy leads, we also introduce a lead encoder to capture inter-lead features. A contrastive loss is incorporated to encourage the model to learn the differences between noisy and clean leads, further enhancing denoising performance.

2 Methods

An overview of our proposed CD-ECGNet learning framewor is shown in Fig. 1(a). Let X and X_S denote the original ECG and simulated ECG, respectively. The X may contain a mixture of clean and noisy leads, or consist entirely of clean or noisy leads. The X_S is generated by artificially adding noise into clean leads. $X^{'}$ and $X_S^{'}$ represent the denoised version of X and X_S. The final objective of this study is to learn a mapping from X to $X^{'}$.

The proposed CD-ECGNet consists of three encoders, a decoder, and a discriminator. Encoders E_{noise} and E_{ECG} are designed to disentangle the noise code and the signal content code. The E_{lead} extracts lead-specific features. Decoder D_{ECG} reconstructs the clean ECG using the signal content code. Discriminator Dis ensure the disentangled signal content remains consistent across two inputs.

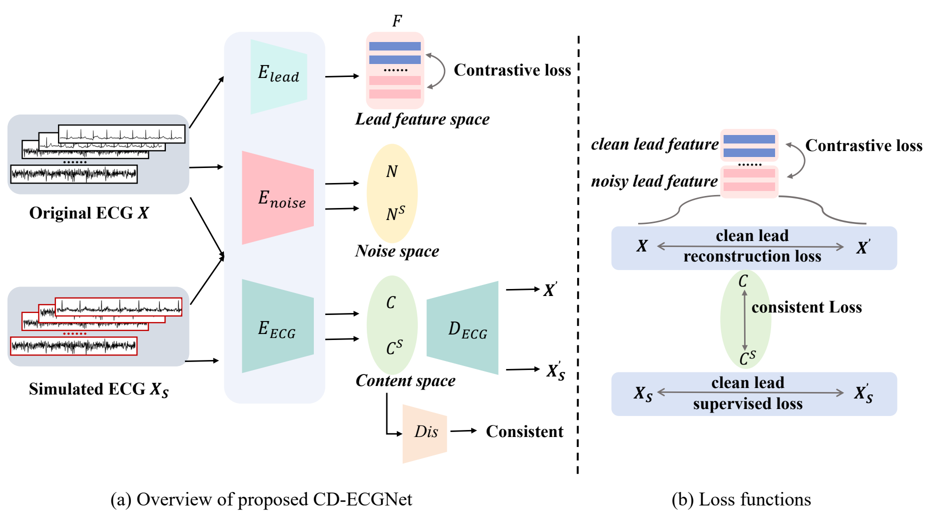


Fig. 1. Overview of proposed (a) CD-ECGNet and (b) loss functions for 12-lead ECG denoising.

2.1 Representation Distanglement and ECG Denoising

A possible approach for 12-lead ECG denoising is to introduce a clean-noisy lead classifier and only process noise leads. In this study, we aim to leverage clean leads to assist in denoising noisy leads during model training. Since representation disentanglement learning seeks to model latent space representations such that they can be disentangled into multiple independent latent factors, we naturally extend this idea to ECG denoising[14].

The wearable ECG dataset lacks paired ECG. Instead, cardiologists have only annotated which leads contain noise. Since unsupervised ECG reconstruction performs poorly, we introduce simulated noise augmentation. During training, the model processes two types of inputs:X (e.g., in Fig. 1(a), leads I and II are clean), used to compute contrastive loss; X_S where clean leads are artificially noised, used to compute supervised loss.

As shown in Fig. 1(a), X is input into E_{noise}, E_{ECG} and E_{lead}, X_S is input into E_{noise} and E_{ECG}:

$$c = E_{ECG}(X), c_S = E_{ECG}(X_S) \tag{1}$$

$$n = E_{noise}(X), n_S = E_{noise}(X_S) \tag{2}$$

$$F = E_{lead}(X) \tag{3}$$

where c and c_S represent the signal content codes of the X andX_S, n and n_S denote corresponding noise codes. The E_{lead} performs convolutional along the signal length dimension . The extracted feature F has 12 channels, corresponding to the 12 ECG leads.

Given c and c_S,we can reconstruct a noise-free ECG:

$$X^{'} = D_{ECG}(c), X_S^{'} = D_{ECG}(c_S) \tag{4}$$

Since simulated noise may differ from wearable noise, we introduce a discriminator to ensure that the E_{ECG} can extract the same signal content code from the ECG with real wearable noise X and the ECG with simulated noise X_S. The Dis enforce similarity between n and n_S, helps model learn to separate noise from cardiac information, making the representations invariant to variations in noise types.

2.2 Loss Function

The loss function in CD-ECGNet as shown in Fig. 1(b).
1) Supervised loss: the original ECG X^c contains 12 channel, with clean lead X^{clean} and noise X^{noise}. The denoised $X_S^{'}$ corresponds to the $X_S^{'clean}$ and $X_S^{'noisy}$. The clean lead supervised loss is:

$$\mathcal{L}_{sup} = \left\| X^{clean}, X_s^{'clean} \right\|_1 \tag{5}$$

2) Reconstruction loss: this loss ensures that clean leads are not affected by the denoising. The denoised $X^{'}$ corresponds to the $X^{'clean}$ and $X^{'noisy}$. The clean lead reconstruction loss is:

$$\mathcal{L}_{rec} = \left\| X^{clean}, X^{'clean} \right\|_1 \tag{6}$$

3) Contrastive Loss: ECG consists of clean and noisy leads, it is clear that the clean feature F^{clean} and noisy feature F^{noisy} should be different. It is benefit from recognizing this difference when optimizing the model. Therefore, we introduce a contrastive loss to measure the similarity between clean and noisy features.

$$\mathcal{L}_{cont} = 1 - log(\frac{exp(comsim(\frac{F^{clean}}{\sigma}))}{exp(cossim(\frac{F^{clean}}{\sigma}, \frac{F^{noisy}}{\sigma}))}) \tag{7}$$

where $comsim(\frac{F^{clean}}{\sigma})$ represents the average cosine similarity for clean leads, $cossim(\frac{F^{clean}}{\sigma}, \frac{F^{noisy}}{\sigma})$ represents the average cosine similarity between clean leads and noisy leads, σ is the temperature coefficient in the contrastive loss.
4) Consistent loss: similar to GANs, we use adversarial learning to measure the distance between c and c_S, encourages consistency between c and c_S. We treat c and c_S as two classes with labels l and l_S, and adopt the Dis that discriminates c from c_S and outputs a value between l and l_S. The min-max game between the generator and the discriminator ensures that the learned features are consistent across X and X_s.

$$\mathcal{L}_{cons} = (Dis(c) - l)^2 + (Dis(c_S) - l_S)^2 \tag{8}$$

The overall objective is:

$$E, D = \min_{E,D} \max_{Dis} \{\mathcal{L}_{sup} + \mathcal{L}_{rec} + \mathcal{L}_{cont} + \lambda \mathcal{L}_{cons}\} \tag{9}$$

where E and D represent encoders and decoder, λ is a hyperparameter that controls the importance of the consistent loss, in this study, λ is 0.5.

2.3 Network Architecture

1) Encoder: The E_{ECG} and E_{noise} adopt a multi-scale 1D convolutional network, incorporating squeeze-and-excitation layers and residual connections[11]. Given an input tensor $x \in \mathbb{R}^{12 \times l}$, where l represents signal length and 12 denotes the number of leads, a permutation transformation is first applied to rearrange the input as $x \in \mathbb{R}^{l \times 12}$. Then 1D convolutional kernels of size 1 in lead-wise feature extractor E_{lead} progressively reducing the temporal dimension while preserving the number of channels. This design maintains lead-specific characteristics, preventing noise and clean leads from interfering with each other.
2) Decoder: The D_{ECG} is an adaptation of the 1-D U-Net up-sampling process. Each upsampling operation contains two convolutional layers with a kernel size of 3 and a 1D linear interpolation step[22].

3 Dataset

The data in this study contains a private wearable 12-lead ECG dataset and the publicly PTB-XL clinical 12-lead ECG dataset[19].

The wearable ECG used for developing our model was collected by the CardioCloud Medical Technology(Beijing) Company using a CONX CC1612 ECG device in a Mason-Likar system. The devices are widely used for daily health monitoring, self-checkups, and by rural doctors for assisted diagnosis. The dataset comprises 6,224 15s ECG with a sampling frequency of 500 Hz. The experiments follow a five-fold hold-out, with the dataset partitioned into training, validation, and test sets in an 8:1:1 ratio.

The PTB-XL was collected in hospital settings using the Wilson lead system, contains 21,837 10s ECG with a sampling frequency of 500 Hz. The dataset used as an external test set. To ensure consistency in input length, these signals are zero-padded to 15 seconds.

To simulate noise conditions, we corrupt the ECG recordings using noise profiles from the Massachusetts Institute of Technology-Beth Israel Hospital Noise Stress Test Database (MIT-BIH NST)[13], which contains 3 half-hour real-world noise recordings. Electrodes were placed on limbs to avoid capturing visible ECG. We scale the noise amplitude by multiplying it with a random factor between 0.2 and 1.5.

4 Results and Analysis

4.1 Comparison with Existing Methods

In this section, we compared our CD-ECGNet against five representative from two main categories, two classical digital filter, FIR[10] and IIR[10]; three deep learning method, cGAN[20], Deep-Filter[16] and FCN-DAE[4]. To adapt the single-lead methods for our task, we modified the input and output channels to accommodate 12-lead ECG. As shown in Table 1, our method achieves improvements across multiple metrics in wearable testset. Table 2 presents the results on the PTB-XL testset, demonstrating that despite the differences in ECG devices and lead systems (wearable ECG using Mason-Likar system and PTB-XL using Wilson lead system), our approach outperforms others across all metrics, thereby highlighting its superior denoising performance and robustness.

Table 1. Wearable ECG testset evaluation results for different methods for random noise amplitude between 0.2 to 1.5.

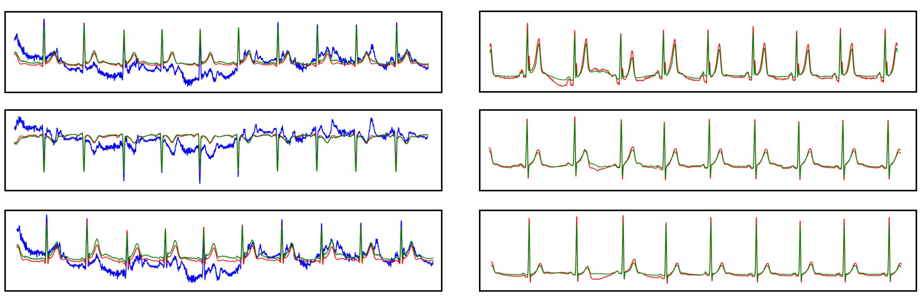
Method	SSD (au)	PRD (%)	Cos sim(%)	SNR	RMSE
FIR[10]	380.04±3.83	90.63±0.17	78.99±0.59	4.97±0.07	0.37±0.00
IIR[10]	378.92±3.58	90.74±0.29	79.10±0.54	4.98±0.07	0.37±0.00
cGAN[20]	37.39±1.62	41.11±0.95	99.02±0.07	18.26±0.13	0.22±0.00
Deep-Filter[16]	36.88±3.10	38.96±1.43	99.01±0.09	17.99±0.33	0.22±0.01
FCNDAE[4]	67.17±2.77	60.08±1.57	98.28±0.09	14.95±0.12	0.26±0.00
CD-ECGNet	**27.02±2.77**	**34.51±1.65**	**99.25±0.08**	**18.81±0.39**	**0.20±0.01**

Table 2. PTB-XL dataset evaluation results for different methods for random noise amplitude between 0.2 to 1.5.

Method	SSD (au)	PRD (%)	Cos sim(%)	SNR	RMSE
FIR[10]	305.76±5.62	90.48±0.08	88.36±0.19	4.73±0.08	0.43±0.00
IIR[10]	303.03±7.40	90.44±0.11	88.46±0.25	4.77±0.11	0.43±0.00
cGAN[20]	27.82±1.07	38.74±0.83	99.10±0.02	17.85±0.06	0.23±0.00
Deep-Filter[16]	24.83±2.24	36.09±2.04	99.16±0.06	17.11±0.32	0.23±0.00
FCNDAE[4]	50.08±1.89	56.68±2.63	98.27±0.05	14.16±0.26	0.27±0.00
CD-ECGNet	**23.52±1.33**	**35.61±1.83**	**99.21±0.03**	**17.95±0.30**	**0.22±0.00**

4.2 Interpretative Visualization

Some visualized cases are given to analysis the performance of wearable ECG denoising. From the Fig. 2(a), it can be observed that CD-ECGNet effectively removes noise from the simulated ECG(blue signal). In Fig. 2(b), it is notable that the original clean signal (red signal) contains slight wearable noise, the denoised output (red signal) from our model appears even smoother than the original clean signal, demonstrates that our model can effectively denoise real-world wearable ECG signals.



(a) (b)

Fig. 2. Performance of the proposed method on the wearable dataset. The noisy leads by blue signal, the original clean lead by red signal, the denoised lead by green signal.

4.3 Ablation Study

To evaluate the effectiveness of each module in our method, an ablation study is conducted using three variants: Baseline, Disentanglement Learning (DL), Disentanglement Learning with a Discriminator (DL+D). Our complete method is represented as CD-ECGNet, incorporates disentanglement learning and is optimized using both the consistent loss and lead contrastive loss. The results of the ablation study are summarized in Table 3. It can be observed that disentanglement leaning significantly enhances denoising compare to the baseline. Further improvements are observed by the discriminator, highlighting the importance of enforcing content code consistency. The full model CD-ECGNet achieves the best overall performance by leveraging lead contrastive loss to enhance feature learning.

Table 3. Performance of ablation experiments on wearable testset.

Method	SSD(au)	PRD(%)	Cos sim(%)	SNR	RMSE
Baseline	51.65±2.17	53.67±1.91	98.70±0.05	16.50±0.20	0.24±0.00
DL	40.88±5.42	42.23±2.98	98.92±0.12	17.56±0.59	0.22±0.01
DL+D	30.89±2.02	36.63±1.02	99.14±0.05	18.25±0.35	0.21±0.00
CD-ECGNet	**27.02±2.77**	**34.51±1.65**	**99.25±0.08**	**18.81±0.39**	**0.20±0.01**

4.4 Impact of Denoising on Diagnosis

To assess CD-ECGNet enhances diagnostic performance, we conduct an imbalanced multi-label classification for ST-segment changes (STC)[14]. STC includes ST-segment elevation (STE) and depression (STD). The dataset consists of 4,000 ECG, including 180 with STE, 180 with STD, 180 with both, and 3,460 normal samples, positive samples account for only 1:11.

A five-fold cross-validation strategy is employed. In the 'w/o denoising' setting, the model is trained and evaluated on normalized ECG without denoising. In the 'with denoising' setting, ECG are first processed using CD-ECGNet before classification. Both experiments are conducted under identical conditions. As shown in Table 4. After denoising, AUPRC improves for minority classes, with STD increasing by 0.449% and STE by 2.837%.

Table 4. AUC and AUPRC performance for STC.

Method	AUC		AUPRC	
	STE	STD	STE	STD
w/o denoising	82.44±1.62	82.89±3.62	55.07±4.54	50.28±7.33
with denoising	**84.32±2.58**	**83.75±3.86**	**55.52±3.89**	**53.12±4.83**

5 Conclusion

We propose a 12-lead ECG denoising framework based on disentanglement learning, which disentangle ECG into noisy and content code, using clean leads to assist in denoising noisy leads. Our model is trained on wearable ECG and tested on both wearable and hospital datasets, demonstrating its effectiveness. Additionally, we evaluated diagnostic performance after denoising using the STC classification task, where our method improved diagnostic AUC and AUPRC. Our approach can be extended to remote wearable ECG applications, enhancing visual clarity and diagnostic performance. Since the simulated noise is based on the Wilson lead system, a domain shift exists between it and wearable ECG[5]. In the future, we aim to address this domain shift to enhance model's robustness.

Acknowledgments. This work was supported by a grant from the National Key R&D Program of China (No. 2018YFC2001203).

Disclosure of Interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Barati, Z., Ayatollahi, A.: Baseline wandering removal by using independent component analysis to single-channel ecg data. In: 2006 International Conference on Biomedical and Pharmaceutical Engineering. pp. 152–156. IEEE (2006)
2. 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 (2024)
3. Chen, M., Li, Y., Zhang, L., Zhang, X., Gao, J., Sun, Y., Shi, W., Wei, S.: Multitask learning-based quality assessment and denoising of electrocardiogram signals. IEEE Transactions on Instrumentation and Measurement **73**, 1–13 (2024)
4. 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)
5. Du, M., Liang, K., Zhang, L., Gao, H., Liu, Y., Xing, Y.: Deep-learning-based metal artefact reduction with unsupervised domain adaptation regularization for practical ct images. IEEE Transactions on Medical Imaging **42**(8), 2133–2145 (2023)
6. Gao, N., Li, Y., Zheng, N., Shi, W., Cai, D., Huang, X., Chen, H.: Frequency information enhanced half instance normalization network for denoising electrocardiograms. Biomedical Signal Processing and Control **102**, 107225 (2025)
7. Hughes, J.W., Olgin, J.E., Avram, R., Abreau, S.A., Sittler, T., Radia, K., Hsia, H., Walters, T., Lee, B., Gonzalez, J.E., et al.: Performance of a convolutional neural network and explainability technique for 12-lead electrocardiogram interpretation. JAMA cardiology **6**(11), 1285–1295 (2021)
8. Jin, Y., Qin, C., Liu, J., Liu, Y., Li, Z., Liu, C.: A novel deep wavelet convolutional neural network for actual ecg signal denoising. Biomedical Signal Processing and Control **87**, 105480 (2024)
9. Joo, J., Joo, G., Kim, Y., Jin, M.N., Park, J., Im, H.: Twelve-lead ecg reconstruction from single-lead signals using generative adversarial networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 184–194. Springer (2023)

10. Kumar, K.S., Yazdanpanah, B., Kumar, P.R.: Removal of noise from electrocardiogram using digital fir and iir filters with various methods. In: 2015 International conference on communications and signal processing (ICCSP). pp. 0157–0162. IEEE (2015)
11. Lai, J., Zhang, Y., Zhao, C., Wang, J., Yan, Y., Chen, M., Ji, L., Guo, J., Han, B., Shi, Y., et al.: Multi-expert ensemble ecg diagnostic algorithm using mutually exclusive–symbiotic correlation between 254 hierarchical multiple labels. npj Cardiovascular Health **1**(1), 8 (2024)
12. Lomoio, U., Veltri, P., Guzzi, P.H., Liò, P.: Design and use of a denoising convolutional autoencoder for reconstructing electrocardiogram signals at super resolution. Artificial Intelligence in Medicine **160**, 103058 (2025)
13. Moody, G.B., Muldrow, W., Mark, R.G.: A noise stress test for arrhythmia detectors. Computers in cardiology **11**(3), 381–384 (1984)
14. Neshatavar, R., Yavartanoo, M., Son, S., Lee, K.M.: Cvf-sid: Cyclic multi-variate function for self-supervised image denoising by disentangling noise from image. In: Proceedings of the ieee/cvf Conference on Computer Vision and Pattern Recognition. pp. 17583–17591 (2022)
15. Ribeiro, A.H., Ribeiro, M.H., Paixão, G.M., Oliveira, D.M., Gomes, P.R., Canazart, J.A., Ferreira, M.P., Andersson, C.R., Macfarlane, P.W., Meira Jr, W., et al.: Automatic diagnosis of the 12-lead ecg using a deep neural network. Nature communications **11**(1), 1760 (2020)
16. Romero, F.P., Piñol, D.C., Vázquez-Seisdedos, C.R.: Deepfilter: An ecg baseline wander removal filter using deep learning techniques. Biomedical Signal Processing and Control **70**, 102992 (2021)
17. Singh, P., Sharma, A.: Attention-based convolutional denoising autoencoder for two-lead ecg denoising and arrhythmia classification. IEEE Transactions on Instrumentation and Measurement **71**, 1–10 (2022)
18. de Vries, I., de Jong, A., van Laar, J., Vullings, R., et al.: Deep learing for sparse domain kalman filtering with applications on ecg denoising and motility estimation. IEEE Transactions on Biomedical Engineering (2024)
19. 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)
20. Wang, X., Chen, B., Zeng, M., Wang, Y., Liu, H., Liu, R., Tian, L., Lu, X.: An ecg signal denoising method using conditional generative adversarial net. IEEE Journal of Biomedical and Health Informatics **26**(7), 2929–2940 (2022)
21. Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Prince, J.L., Xu, Z.: Unsupervised mr-to-ct synthesis using structure-constrained cyclegan. IEEE transactions on medical imaging **39**(12), 4249–4261 (2020)
22. Zhang, Y., Lai, J., Zhao, C., Wang, J., Yan, Y., Chen, M., Ji, L., Guo, J., Han, B., Shi, Y., et al.: Abnormal recognition-assisted and onset-offset aware network for pathological wearable ecg delineation. Artificial Intelligence in Medicine **157**, 102992 (2024)