

Received March 11, 2021, accepted April 1, 2021, date of publication April 12, 2021, date of current version April 19, 2021. Digital Object Identifier 10.1109/ACCESS.2021.3072640

DeepCEDNet: An Efficient Deep Convolutional Encoder-Decoder Networks for ECG Signal Enhancement

PINGPING BING¹⁰, **WEI LIU¹⁰**, **AND ZHIHUA ZHANG¹⁰**

²College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029, China
 Corresponding author: Pingping Bing (bpping@163.com)

This work was supported by the project of Scientific Research Fund of Hunan Provincial Education Department under Grant 19C0185.

ABSTRACT Electrocardiogram (ECG) signal can be thought of as an effective indicator for detection of various arrhythmias. However, the acquired ECG data is always corrupted by amounts of noise, which have a great influence on the diagnosis of cardiovascular diseases. In this paper, an efficient deep convolutional encoder-decoder network framework is proposed to remove the noise from ECG signal, which is termed as 'DeepCEDNet'. This network is able to learn a sparse representation of data in the time-frequency domain via the high-order synchrosqueezing transform (FSSTH) and a nonlinear function that maps the noisy data into the clean one based on the distribution difference between signal and noise from the training set. Extensive experiments are conducted on ECG signals from the MIT-BIH Arrhythmia database and MIT-BIH Long-Term ECG database, and the added noise is obtained from the MIT-BIH Noise Stress Test database. The denoising performance is evaluated by means of signal to noise ratio (SNR), root mean squared error (RMSE) and percent root mean square difference (PRD). The results indicate that the proposed DeepCEDNet can obtain superior performance in both noise reduction and details preservation with higher SNR and lower RMSE and PRD compared to the traditional convolutional neural network (CNN) and the fully convolutional network-based denoising auto-encoder (FCN). We believe that the DeepCEDNet has a wide application prospect in the biomedical field.

INDEX TERMS Electrocardiogram signal, noise reduction, deep neural network, sparse representation, time-frequency domain.

I. INTRODUCTION

Recorded electrocardiogram (ECG) signal is inevitably contaminated by various types of noise (coined artifacts) [1], [2], such as baseline wander (BW), muscle artifact (MA), and electrode motion (EM), and so on. All of the noise has a severe impact on ECG waveform and covers the weak characteristics of ECG signal, which poses a challenge for the following cardiovascular diseases diagnosis [3]–[6]. Therefore, the noise removal from ECG signal is becoming urgent [7], [8].

Over the past decades, numerous efforts have been made to develop different methods for denoising ECG signal, for example adaptive filter [9]–[11], wavelet transform [12],

The associate editor coordinating the review of this manuscript and approving it for publication was Chih-Yu Hsu¹⁰.

principle component analysis (PCA) [13], independent component analysis (ICA) [14], and empirical mode decomposition (EMD) [15], [16]. Adaptive filter can effectively remove noise outside ECG signal frequency band, however, it will fail when the signal and noise share a common frequency range. Wavelet transform is capable of suppressing noise well by shrinking the wavelet coefficients in the transformed domain. Unfortunately, a suitable wavelet basis function and the threshold strategy need to be selected with prior knowledge, which is usually a troublesome process in practice. In addition, the threshold algorithm is likely to affect the ECG waveform. The key idea of the methods based on PCA and ICA is to eliminate the dimensions corresponding to noise, however, the obtained mapping model is more sensitive to small disturbances in the signal or noise. EMD-based approaches decompose the noisy signal into a series of intrinsic mode

functions (IMFs), and then the noise-dominant IMFs are removed while the remaining IMFs are used to reconstruct the signal. As some useful signals are embedded in the IMFs containing noise, such methods may not give satisfactory result. Meanwhile, the mode-mixing is also the main drawback of EMD. Although the variations of EMD (e.g. ensemble empirical mode decomposition (EEMD) [17], [18] and complete ensemble empirical mode decomposition (CEEMD) [19], [20]) and variational mode decomposition (VMD) [21], [22] have greatly alleviated this issue, the nature of redundancy is still not been fundamentally solved, which limits the potential application of these methods in a lot of fields.

Recently, deep learning techniques have been gaining attention due to their powerful capability in learning the characteristics of ECG signal [23]–[27]. The noise suppression techniques based on denoising autoencoder (DAE) have shown the excellent performance than conventional denoising methods. In [28], an improved DAE reformed by a wavelet transform was created, in which a scale-adaptive thresholding algorithm is employed to attenuate most of the noise. In [29], a stacked contractive DEA through multi-level feature extraction was developed for noise reduction. In [30], a fully convolutional network based DEA is proposed for ECG signal denoising. However, these methods are usually carried out in the time domain, and cannot fully exploit the capability of auto-encoders in learning a sparse representation of data.

In this paper, we present DeepCEDNet, a novel sparsely promoted deep convolutional encoder-decoder network in the time-frequency domain. This network is able to simultaneously learn a sparse representation of input data and a nonlinear function that maps the noisy data into the clean one according to the learned features from signal and noise based on the training set. We utilize real ECG signals contaminated with various types of noise such as BW, MA and EM to train the network and demonstrate its performance, and compare with state-of-the-art methods, CNN and FCN (fully convolutional network). Our contributions are as follows: (1) we propose a sparsely promoted deep neural network to extract ECG signal features, (2) DeepCEDNet has powerful advantage in learning a sparse representation of data (3) our method can achieve impressive denoising of ECG signal and preserve the details well.

The rest of this paper is organized as follows. Section II reviews the basic concepts of the autoencoder (AE) and DAE, the theory of the FSSTH, and then the proposed DeepCEDNet is depicted in detail. In Section III, the experimental results from two available MIT-BIH databases are exhibited in order to verify the effectiveness of the proposed architecture. The discussion with respect to experimental results is given in Section IV. Section V concludes this paper.

II. METHODOLOGY

A. REVIEW OF AE AND DAE

As a deep learning model, the aim of AE is to reconstruct the input as accurately as possible through the constraint of a loss function. A basic AE architecture is comprised of encoder and decoder. The former maps an input vector x to a hidden representation y by a deterministic mapping expression. In the second stage, the latent representation y is mapped back to a reconstructed vector z. The two parts can be formulated as:

$$y = \varphi \left(Wx + b \right), \tag{1}$$

$$z = \varphi' \left(W' y + b' \right). \tag{2}$$

where W and b are a weight matrix and a bias vector of the encoder, respectively. Similarly, W' and b' are the corresponding weight matrix and bias vector of the decoder, respectively. φ and φ' are the non-linear activation functions.

The parameters in the aforementioned model are optimized by minimizing the reconstruction error:

$$L = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} \|x_i - z_i\|_2^2.$$
 (3)

where θ is a parameter set $\{W, b, W', b'\}$, N is the number of data samples, and *i* is the sample index.

DAE, originally developed by Vincent *et al.* [31], is a variant of classic AE. In the DAE, the input \tilde{x} is a corrupted version of data, which is created by means of a stochastic mapping $\tilde{x} \sim q(\tilde{x}|x)$. The corrupted input \tilde{x} is first mapped to a hidden representation using Eq. (1), and then one can obtain its reconstruction using Eq. (2). Throughout the mapping process, the parameters are trained to minimize the reconstruction error (Eq. (3)) over a training set in order to make z as close as possible to the uncorrupted input x.

B. HIGH-ORDER SYNCHROSQUEEZING TRANSFORM

The key idea of FSSTH is to sharpen the short-time Fourier transform (STFT) representation of a signal by computing a new local instantaneous frequency estimate, using higher order approximations both for the amplitude and phase [32].

Considering an AM-FM signal:

$$f(t) = A(t) e^{i2\pi\phi(t)}$$
. (4)

where A(t) denotes the instantaneous amplitude and $\phi(t)$ is the instantaneous phase.

The Taylor expansion of the signal f in Eq. (4) for τ close to t is described as:

$$f(\tau) = \exp\left(\sum_{k=0}^{N} \frac{\left[\log\left(A\right)\right]^{(k)}(t) + i2\pi\phi^{(k)}(t)}{k!}(\tau-t)^{k}\right).$$
(5)

where $T^{(k)}(t)$ is the *k*th derivative of *T* with respect to *t*.

Therefore, the local instantaneous frequency estimate $\omega_f(t, \eta)$ can be written as:

$$\omega_{f}(t,\eta) = \frac{\left[\log\left(A\right)\right]'(t)}{i2\pi} + \phi'(t) + \sum_{k=2}^{N} \frac{\left[\log\left(A\right)\right]^{(k)}(t) + i2\pi\phi^{(k)}(t)}{i2\pi(k-1)!} \frac{V_{f}^{t^{k-1}g}(t,\eta)}{V_{f}^{g}(t,\eta)}.$$
 (6)

where $V_f^g(t, \eta)$ is the STFT of f(t).



FIGURE 1. Architecture of the proposed DeepCEDNet.

Based on the above-mentioned high order Taylor expansions of the amplitude and phase of a signal, a frequency modulation operator $q_{\eta,f}^{[k,N]}$ is defined as:

$$q_{\eta,f}^{[k,N]} = \frac{\left[\log\left(A\right)\right]^{(k)}(t) + i2\pi\phi^{(k)}(t)}{i2\pi(k-1)!}.$$
(7)

Then, the Nth-order local complex instantaneous frequency, $\omega_{n,f}^{[N]}$ at time t and frequency η , can be expressed by:

$$\omega_{\eta,f}^{[N]}(t,\eta) = \begin{cases} \omega_{f}(t,\eta) + \sum_{k=2}^{N} q_{\eta,f}^{[k,N]}(\eta,t) \left(-x_{k,1}(t,\eta)\right), \\ V_{f}^{g}(t,\eta) \neq 0, \, \partial_{\eta}x_{j,j-1}(t,\eta) \neq 0 \, (j \geq 2) \, . \end{cases}$$

$$\omega_{f}(t,\eta), \\ otherwise. \end{cases}$$
(8)

Finally, the FSSTH is defined as follows:

$$\begin{aligned} T_{N,f}^{g,\gamma}(t,\omega) \\ &= \frac{1}{g^*(0)} \int_{\left\{\eta, \left|V_f^g(t,\eta)\right| > \gamma\right\}} V_f^g(t,\eta) \,\delta\left(\omega - \omega_{\eta,f}^{[N]}(t,\eta)\right) d\eta. \end{aligned}$$
(9)

where γ denotes some threshold.

The *i*th mode can be approximately reconstructed by:

$$f_i(t) \approx \int_{\{\omega, |\omega - \varphi(t)| < d\}} T_{N, f}^{g, \gamma}(t, \omega) d\omega.$$
(10)

where $\varphi(t)$ is an estimate for $\phi'_i(t)$, and d is a compensation factor.

C. PROPOSED DeepCEDNet

frequency domain via the FSSTH [32], a variant of STFT-based SST (FSST) [35], which can achieve a highly concentrated time-frequency representation [33], [34]. In the time-frequency domain, the noisy signal X(t, f) can be expressed as:

The noisy signal x(t) is firstly transformed into time-

$$X(t,f) = S(t,f) + N(t,f).$$
 (11)

where S(t, f) and N(t, f) denote the useful signal and noise, respectively.

Our aim is to estimate the underlying signal $\hat{s}(t)$ from its noise corrupted version x(t) by minimizing the following square error:

$$E = \left\| \hat{s}(t) - s(t) \right\|_{2}^{2}.$$
 (12)

where s(t) is the true signal in the time domain. $\hat{s}(t)$ is the estimated signal in the time domain, which is obtained by implementing an inverse transform of FSSTH on $\hat{S}(t, f)$:

$$\widehat{S}(t,f) = M(t,f)X(t,f).$$
(13)

Herein, M(t, f) is a nonlinear function that maps X(t, f) to a time-frequency representation of the estimated signal.

Now, the above-mentioned problem is regarded as a supervised learning problem, a deep neural network is designed to learn a sparse representation of data in the time-frequency domain and create an nonlinear mapping function through a training set, in which the distribution difference of signal and noise is described well.

Inspired by the advantage of auto-encoders in learning a sparse representation of data, we construct a sparsely promoted deep convolutional encoder-decoder network



FIGURE 2. Flow diagram of the proposed denoising method. Inputs are the real and imaginary parts of the time-frequency representation of noisy signal.

framework in the time-frequency domain (Figure 1), which is mainly composed of two modules: an encoder denoted by the convolutional layers and a decoder denoted by the deconvolutional layers. Meanwhile, the skip connections between two corresponding convolutional and deconvolutional layers are used, with which the training converges much faster and attains a higher-quality local optimum [36].

(1) Encoder module

DeepCEDNet projects the input data X into a highdimensional feature space in order to obtain the vector y via a nonlinear mapping function En(X):

$$y = En\left(X\right). \tag{14}$$

It is worth noting that the inputs to the first layer are the real and imaginary parts of the time-frequency coefficients of the noisy data X. En (X) is carried out by the encoder network that consists of a series of 2D convolutional layers, rectified linear unit (Relu) and batch normalization (BN). The size of convolution filter is a constant (3×3), and the feature space is gradually reduced using strides of 2×2 . These layers make up a feature extractor that can achieve learning with respect to a sparse representation of time-frequency coefficients and capture the abstract content of signal without the noise.

(2) Decoder module

The decoder module is almost symmetric to the encoder part, it can be thought of as the inverse process of the encoder. The aim of decoder is to remap *y* back into the time-frequency space in order to recover the input details:

$$\stackrel{\wedge}{S} = De(y). \tag{15}$$

Here, the deconvolution is utilized to map this learned highdimensional feature into the desired sparse representation. Then, the denoised signal \hat{s} can be reconstructed by the inverse FSSTH with the constraint of Eq. (12). Through the whole training process, the network learns the sparse representation of noisy data and the optimal map that recovers the desired signal by minimizing a loss function.

Figure 2 shows the flow diagram of the proposed denoising method. First, the noisy ECG signal is transformed into the time-frequency domain via the FSSTH. Then, DeepCEDNet extracts the real and imaginary parts of the obtained timefrequency coefficients as the input and produces the mapping for signal as the output. The estimated time-frequency coeffi-

cients \hat{S} associated with ECG signal can be obtained by applying the mapping to the real and imaginary of time-frequency coefficients of noisy ECG signal X. Finally, the denoised ECG signal \hat{s} is obtained by transforming \hat{S} back into the time domain using an inverse FSSTH, and followed by the least square constraint. Compared with the conventional denoising algorithms, DeepCEDNet can automatically learn more abstract features from the noisy ECG data, and remove the noise in the time-frequency domain.

III. EXPERIMENTS

A. EVALUATION CRITERIA

In the paper, three criteria will be used for quantitative evaluation in denoising performance, which are signal to noise ratio (SNR), root mean squared error (RMSE) and percent root mean square difference (PRD), respectively.

$$SNR = 10\log_{10} \frac{\sum_{t=1}^{N} [s(t)]^2}{\sum_{t=1}^{N} \left[s(t) - \hat{s}(t) \right]^2}.$$
 (16)

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left[s(t) - \hat{s}(t) \right]^{2}}.$$
 (17)

$$PRD = \sqrt{\frac{\sum_{t=1}^{N} \left[s(t) - \hat{s}(t) \right]^{2}}{\sum_{t=1}^{N} \left[s(t) \right]^{2}}} \times 100.$$
(18)

where s(t) is the noise-free ECG signal, $\hat{s}(t)$ is the denoised ECG signal, and N is the length of ECG signal.

It should be noteworthy that the SNR describes the level of noise suppression, thus the higher the SNR, the better the denoising performance. The RMSE depicts the difference between the desired output and the actual one, a lower RMSE indicates the smaller difference between both outputs. The PRD delineates the signal recovery capability. A lower PRD means a better reconstruction.

B. DATASET DESCRIPTION

The proposed DeepCEDNet is tested using two standard ECG datasets, namely, MIT-BIH Arrhythmia Database and MIT-BIH Long-Term ECG Database, because both of them have the long duration, which is helpful to train the deep neural network. The first dataset is composed of 48 ECG records with length of 30 minutes, and are sampled at 360 Hz and quantized with 11-bit resolution. The second one contains 7 long-term ECG records and each record lasts about 14h to 22h at the sampling rate of 128 Hz. For each record, we partition 400 fragments, each with a length of 1024 samples. Real noise is comprised of BW, MA and EW, which are from the MIT-BIH Noise Stress Test Database.

We randomly divide the ECG datasets into three parts. More specifically, 90% of the samples are split into the training set and 10% into the test set. Meanwhile, in the training set, 90% of the samples are split into the training set, and 10% into the validation set. Similarly, the noise dataset is also divided into three sections, the training set, validation set and test set. In the three segments, the noise contents are similar but not the same. To increase the randomness, we randomly choose the noise samples from the training set with respect to each noise source (BW, MA and EW), and randomly scale its amplitude. Subsequently, the three kinds of noise are mixed with equal weight to form the complex noise. And, the noisy ECG data is generated by adding the above-mentioned complex noise to the original ECG data. The FSSTH is applied to generate the time-frequency representation of the noisy ECG signal, a detailed description of FSSTH algorithm, please refer to [37]. Then, the obtained time-frequency matrix is normalized by removing the mean and dividing by the standard deviation. During prediction, such information is temporarily saved in order to transform the processed time-frequency matrix into the original scale after denoising. The real and imaginary parts of timefrequency coefficients are respectively fed to the deep neural network as two channels. The same procedure is applicative to both validation and test sets. We utilize the test set to analyze the final performance and demonstrate the denoising results.



FIGURE 3. Training and validation losses during training process. (a) MIT-BIH Arrhythmia Database, (b) MIT-BIH Long-Term ECG Database.

 TABLE 1. Parameters in the proposed DeepCEDNet.

Parameter	Value
Batch size	128
Epoches	300
Optimizer	Adam
Loss	Mean square error (Mse)

C. EXPERIMENTAL RESULTS

In this section, we first evaluate the performance of the proposed DeepCEDNet. Figure 3 shows the losses of training and validation of the presented model for two datasets. As can be obviously seen that the loss function of the training set decreases rapidly and then converges to a smaller value. Meanwhile, the validation set also shows a similar trend. Besides, when the training reaches a certain epoch, the loss difference between the training set and the validation set is relatively small, which means that the proposed DeepCEDNet has almost no overfitting, in other words, the DeepCEDNet has strong learning ability.

CNN and FCN are two classic neural networks, and have been widely applied in ECG signal analysis [27], thus, they are employed for comparison. In the paper, the CNN consists of 22 convolutional layers, while the FCN is composed of 11 layers of convolution and deconvolution operators. Now, the four examples from the test set are selected, and the denoising performance is compared with the CNN and FCN. We select the number of decomposed modes for the FSSTH algorithm as 5. The parameters used by DeepCEDNet are summarized in Table 1. In addition, we also apply the dropout technique with a rate of 0.5 for all layers since it obtain the



FIGURE 4. Denoised results regarding record 101m.dat. (a) Original ECG signal, (b) Noisy ECG signal, (c) CNN, (d) FCN, (e) CEDN.

best performance in the validation set. Figures 4 and 5 show the denoised results regarding record 101m.dat and 223m.dat from MIT-BIH Arrhythmia Database. In these figures, (a) is the original ECG signal, (b) is the noisy ECG signal with a SNR of 0 dB. The denoised results using CNN, FCN and CEDN are shown in (c), (d) and (e), respectively. As reported in Figure 4(b), the noise has a certain influence on P wave, T wave and QRS complex wave. However, in Figure 5(b), the noise corrupts the ECG signal severely so that these characteristic waveforms, P, T and QRS, are completely distorted and the useful signal features cannot be effectively extracted. The denoised results indicate that the CEDN can successfully recover the signal with a less distortion of the waveform (see magenta rectangles in Figures 4(c), (d) and (e) and Figures 5(c), (d) and (e)), meanwhile, the signal leakage is minimal, and the ECG waveform and amplitude characteristics are well preserved after denoising. In contrast, both of CNN and FCN can also suppress most of the noise but the residual noise is still found, which maybe lead to an inappropriate diagnosis of the disease. In addition, from the perspective of local details, it seems that FCN is superior to CNN because of a less loss of amplitudes of ECG signal (see black rectangles in Figures 4(c), (d) and Figures 5(c), (d)).



FIGURE 5. Denoised results regarding record 223m.dat. (a) Original ECG signal, (b) Noisy ECG signal, (c) CNN, (d) FCN, (e) CEDN.

Similarly, our method is also applied to 14046m.dat and 14149m.dat from MIT-BIH Long-Term ECG Database. In this data, it is difficult to accurately capture the features of P, T and QRS waves in the presence of strong noise, which directly results in the inability to make correct diagnosis about the diseases. Figures 6 and 7 show the denoised results based on CNN, FCN and CEDN, respectively. It can be clearly seen that CNN seems to produce a severe loss of amplitudes, especially in Figure 6(c). This phenomenon can also be found in FCN (Figure 6(d)), but is not as serious as in CNN. However, CEDN achieves a better denoising performance that maintains the waveform characteristics of ECG signal well compared to the other two methods (see magenta rectangles in Figures 6(c), (d) and (e)). In the example 14149m.dat, CEDN also yields a satisfactory denoised result, which is obviously superior to the CNN and FCN. The difference is that the overall denoising performance of the mentioned methods is better than that in 14046m.dat. Besides, it is worth noting that FCN has more advantage in denoising compared to CNN (see black rectangles in Figures 6(c), (d) and Figures 7(c), (d)).



FIGURE 6. Denoised results regarding record 14046m.dat. (a) Original ECG signal, (b) Noisy ECG signal, (c) CNN, (d) FCN, (e) CEDN.

To further demonstrate the effectiveness of the proposed DeepCEDNet, we calculate the average of the SNR, RMSE and PRD for all ECG records from the test set about two datasets, which are plotted in Figures 8 and 9, respectively. The test set consists of 3 different levels of input SNR of 0, 4, and 8 dB. As it can be seen, CEDN obtains the higher SNR than CNN and FCN for all records (Figures 8(a) and 9(a)), which means that CEDN performs better in noise removal. In Figures 8 and 9, CEDN yields the lower RMSE and PRD, which indicates that the denoised signal after running DeepCEDNet is closer to the original signal, in other words, CEDN does a better job in preserving the details from ECG signal.

IV. DISCUSSION

As an indicator, ECG signal can provide a large amount of valuable information that can be used to diagnose early-stage cardiovascular disorders. However, the acquisition and transmission enable the collected ECG signal to be contaminated with various types of noise, thus, it is necessary to remove noise from ECG signal.



FIGURE 7. Denoised results regarding record 14149m.dat. (a) Original ECG signal, (b) Noisy ECG signal, (c) CNN, (d) FCN, (e) CEDN.

In this paper, we proposed a deep convolutional encoderdecoder network framework, DeepCEDNet. It can learn the sparse representation of data in the time-frequency space in order to predict a mask that maps the noise corrupted signal into the clean one by optimizing a loss function. The mask determined by the deep neural network effectively decomposes the input data into signal and noise. Experimental results show that DeepCEDNet performs clearly better in noise removal and details preservation compared with the traditional CNN and FCN, which can be attributed to three aspects: (1) the difference between signal and noise is more obvious in the time-frequency domain; (2) Deep-CEDNet makes full use of the capability of auto-encoders in learning a sparse representation of data; (3) the skip connections between two corresponding convolutional and deconvolutional layers also help to handle the problem of gradient vanishing, improve the signal reconstruction performance, and enhance the robustness of deep neural network. In addition, our tests also indicate that DeepCEDNet significantly improves the SNR with minimal distortion to the underlying signal, which is extremely beneficial to clinical applications. As a deep learning model, DeepCEDNet also



FIGURE 8. Comparative performance in terms of the average of SNR, RMSE and PRD, with CNN, FCN and CEDN for the test set from MIT-BIH Arrhythmia Database.

provides an end-to-end mapping, which is a complicated process. We visualize the reactions of the partial convolutional layers from record 101m.dat in order to exhibit the learned abstract features in CEDN, which is shown in Figure 10. In addition, our results indicate that FCN has the slight advantage over CNN in denoising performance. This may be due to the framework of DAE, which is helpful to recover the



characteristics of a signal of interest during the denoising process.

Although the denoising performance of DeepCEDNet is impressive, it does not achieve a perfect separation of signal and noise. Several issues should be discussed in depth, for example, (1) the number of data segments is not sufficient, (2) high-quality signal labels need to be taken into account,



FIGURE 10. Visualization of the partial convolutional layers from record 101m.dat in CEDN.

(3) the network leads to a large computational burden,(4) the parameter setting of CEDN needs further optimization. Solving all above-mentioned problems will help to improve the current DeepCEDNet so that it can be applied to more complex noise suppression in the future.

V. CONCLUSION

We have developed a novel deep learning based denoising framework for ECG signal, DeepCEDNet. This network can learn a sparse representation of data in the time-frequency domain and a nonlinear mapping function that aims at signal and noise separation. Experimental results indicate that DeepCEDNet significantly outperforms the traditional CNN and FCN in both noise removal and details preservation, meanwhile, our network shows the higher SNR and lower RMSE and PRD. Thus, DeepCEDNet has the potential to provide a more effective denoising tool for ECG signal processing. Future work will focus on data segments, signal labels, parameter setting optimization and computational efficiency so that DeepCEDNet can be better applied in practice.

ACKNOWLEDGMENT

The authors would like to thank associate editor Prof. Chih-Yu Hsu and four anonymous reviewers for their constructive comments that improve the article greatly.

REFERENCES

- X. Wang, Y. Zhou, M. Shu, Y. Wang, and A. Dong, "ECG baseline wander correction and denoising based on sparsity," *IEEE Access*, vol. 7, pp. 31573–31585, 2019.
- [2] B. Chen, Y. Li, X. Cao, W. Sun, and W. He, "Removal of power line interference from ECG signals using adaptive notch filters of sharp resolution," *IEEE Access*, vol. 7, pp. 150667–150676, 2019.
- [3] R. Hoekema, G. J. H. Uijen, and A. van Oosterom, "Geometrical aspects of the interindividual variability of multilead ECG recordings," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 5, pp. 551–559, May 2001.

- [4] S. Sahoo, P. Biswal, T. Das, and S. Sabut, "De-noising of ECG signal and QRS detection using Hilbert transform and adaptive thresholding," *Procedia Technol.*, vol. 25, pp. 68–75, Jan. 2016.
- [5] B. R. D. Oliveira, M. A. Q. Duarte, C. C. E. D. Abreu, and J. V. Filho, "A wavelet-based method for power-line interference removal in ECG signals," *Res. Biomed. Eng.*, vol. 34, no. 1, pp. 73–86, Jan. 2018.
- [6] X. Tian, Y. Li, H. Zhou, X. Li, L. Chen, and X. Zhang, "Electrocardiogram signal denoising using extreme-point symmetric mode decomposition and nonlocal means," *Sensors*, vol. 16, pp. 1584–1602, Oct. 2016.
- [7] W. Liu and W. Chen, "Recent advancements in empirical wavelet transform and its applications," *IEEE Access*, vol. 7, pp. 103770–103780, 2019.
- [8] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, A. K. Abasi, and S. N. Makhadmeh, "EEG signals denoising using optimal wavelet transform hybridized with efficient Metaheuristic methods," *IEEE Access*, vol. 8, pp. 10584–10605, 2020.
- [9] N. V. Thakor and Y.-S. Zhu, "Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 8, pp. 785–794, Aug. 1991.
- [10] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A nonlinear Bayesian filtering framework for ECG denoising," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2172–2185, Dec. 2007.
- [11] S. Poungponsri and X.-H. Yu, "An adaptive filtering approach for electrocardiogram (ECG) signal noise reduction using neural networks," *Neurocomputing*, vol. 117, pp. 206–213, Oct. 2013.
- [12] D. Zhang, "Wavelet approach for ECG baseline wander correction and noise reduction," in *Proc. 27th Annu. Conf. IEEE Eng. Med. Biol.*, Jan. 2006, pp. 1212–1215.
- [13] I. Romero, "PCA and ICA applied to noise reduction in multi-lead ECG," *Comput. Cardiol.*, vol. 23, no. 1, pp. 613–616, 2011.
- [14] A. K. Barros, A. Mansour, and N. Ohnishi, "Removing artifacts from electrocardiographic signals using independent components analysis," *Neurocomputing*, vol. 22, nos. 1–3, pp. 173–186, Nov. 1998.
- [15] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
- [16] A. Chacko and S. Ari, "Denoising of ECG signals using empirical mode decomposition based technique," in *Proc. Int. Conf. Adv. Eng., Sci. Manage.*, 2012, pp. 6–9.
- [17] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.
- [18] K.-M. Chang and S.-H. Liu, "Gaussian noise filtering from ECG by Wiener filter and ensemble empirical mode decomposition," *J. Signal Process. Syst.*, vol. 64, no. 2, pp. 249–264, Aug. 2011.

- [19] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.
- [20] Y. Xu, M. Luo, T. Li, and G. Song, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition," *Sensors*, vol. 17, pp. 2754–2769, Jan. 2017.
- [21] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
- [22] W. Liu and Z. Duan, "Seismic signal denoising using f x variational mode decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1313–1317, Oct. 2020.
- [23] M. Wasimuddin, K. Elleithy, A.-S. Abuzneid, M. Faezipour, and O. Abuzaghleh, "Stages-based ECG signal analysis from traditional signal processing to machine learning approaches: A survey," *IEEE Access*, vol. 8, pp. 177782–177803, 2020.
- [24] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, Oct. 2017.
- [25] Z. Zhao, C. Liu, Y. Li, Y. Li, J. Wang, B.-S. Lin, and J. Li, "Noise rejection for wearable ECGs using modified frequency slice wavelet transform and convolutional neural networks," *IEEE Access*, vol. 7, pp. 34060–34067, 2019.
- [26] E. Ihsanto, K. Ramli, D. Sudiana, and T. Gunawan, "An efficient algorithm for cardiac arrhythmia classification using ensemble of depthwise separable convolutional neural networks," *Appl. Sci.*, vol. 10, no. 2, pp. 483–498, 2020.
- [27] M. Jia, F. Li, J. Wu, Z. Chen, and Y. Pu, "Robust QRS detection using highresolution wavelet packet decomposition and time-attention convolutional neural network," *IEEE Access*, vol. 8, pp. 16979–16988, 2020.
- [28] P. Xiong, H. Wang, M. Liu, S. Zhou, Z. Hou, and X. Liu, "ECG signal enhancement based on improved denoising auto-encoder," *Eng. Appl. Artif. Intell.*, vol. 52, pp. 194–202, Jun. 2016.
- [29] P. Xiong, H. Wang, M. Liu, F. Lin, Z. Hou, and X. Liu, "A stacked contractive denoising auto-encoder for ECG signal denoising," *Physiol. Meas.*, vol. 37, no. 12, pp. 2214–2230, Dec. 2016.
- [30] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, and S.-Y. Chien, "Noise reduction in ECG signals using fully convolutional denoising autoencoders," *IEEE Access*, vol. 7, pp. 60806–60813, 2019.
- [31] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 1096–1103.
- [32] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis—With an application to gravitationalwave signal," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3168–3178, Jun. 2017.
- [33] W. Liu, S. Cao, Z. Wang, K. Jiang, Q. Zhang, and Y. Chen, "A novel approach for seismic time-frequency analysis based on high-order synchrosqueezing transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 8, pp. 1159–1163, Aug. 2018.
- [34] W. Liu, W. Chen, and Z. Zhang, "A novel fault diagnosis approach for rolling bearing based on high-order synchrosqueezing transform and detrended fluctuation analysis," *IEEE Access*, vol. 8, pp. 12533–12541, 2020.

- [35] G. Thakur and H.-T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM J. Math. Anal.*, vol. 43, no. 5, pp. 2078–2095, Jan. 2011.
- [36] X. Mao, C. Shen, and Y. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Proc. 30th Conf. Neural Inf. Process. Syst.*, Barcelona, Spain, 2016, pp. 1–9.
- [37] P. Bing, W. Liu, Z. Wang, and Z. Zhang, "Noise reduction in ECG signal using an effective hybrid scheme," *IEEE Access*, vol. 8, pp. 160790–160801, 2020.



PINGPING BING received the Ph.D. degree in geophysics from the China University of Petroleum, Beijing, China, in 2012. She is currently a Lecturer with Changsha Medical University. Her research interests include optimization algorithm, signal analysis and model prediction, and deep learning.



WEI LIU received the B.E. degree in exploration geophysics from the China University of Petroleum (East China), Qingdao, China, in 2006, and the M.S. degree in exploration geophysics and the Ph.D. degree in geological resources and geological engineering from the China University of Petroleum, Beijing, China, in 2009 and 2016, respectively. He is currently working as a Lecturer with the Beijing University of Chemical Technology. His research interests include signal analysis

and processing, seismic data processing and interpretation, and mechanical fault diagnosis.



ZHIHUA ZHANG received the B.E. degree in safety science and engineering from the North University of China, in 2018. He is currently pursuing the M.E. degree with the Beijing University of Chemical Technology. His research interests include signal analysis, deep learning, and intelligent fault diagnosis.

. . .