IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 August 2023, accepted 12 September 2023, date of publication 18 September 2023,
date of current version 27 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3316487

RESEARCH ARTICLE

ECG Signals Deep Compressive Sensing Framework Based on Multiscale Feature Fusion and SE Block

JING HUA[1], JIAWEN ZOU[2], JUE RAO[2], HUA YIN[1], AND JIE CHEN[2]
[1]School of Software, Jiangxi Agricultural University, Nanchang 330045, China
[2]School of Computer and Information Engineering, Jiangxi Agricultural University, Nanchang 330045, China

Corresponding author: Hua Yin (2454877607@qq.com)

This work was supported in part by the Natural Science Foundation of Jiangxi Province under Grant 20224BAB202038, and in part by the National Natural Science Foundation of China under Grant 61861021.

ABSTRACT Electrocardiogram (ECG) is nowadays an important technology to be applied in the clinical diagnosis for the detection of the heart disease. But the large storage and high-burden transmission of the ECG data is a challenge. Therefore, the compressive sensing (CS) is appropriate to deal with those signals for it can compress and sample the signal at the same time. In order to get rid of the constraints in the traditional CS methods, we propose a compressive sensing framework based on multiscale feature fusion and SE block. In the compression process we use sequential convolutional layers instead of the traditional compressive sensing using measurement matrix projection for ECG signals. In the reconstruction process, the multi-scale feature fusion method is first used to fuse multiple feature maps output from the convolution layer to better extract signal features. Subsequently, Squeeze-and-Excitation (SE) block is used to further enhance the feature representation. Finally, sequence modeling of the ECG signal is performed using LSTM to obtain the reconstructed signal. The results show that the proposed method performs well on various datasets and evaluation metrics, in the case of SR = 0.4, the PRD and SNR of the experiments on the MIT-BIH Arrhythmia database are 1.55% and 37.66dB, respectively. The PRD and SNR of the experiments on the Non-Invasive Fetal ECG Arrhythmia Database were 2.48% and 34.57dB, respectively, which were the lowest among all the comparison methods, indicating that the proposed method has good ECG signal processing capability.

INDEX TERMS Deep learning, compressive sensing, multi-scale feature, SE block, LSTM.

I. INTRODUCTION

Electrocardiogram (ECG) plays a more and more important role in clinical diagnosis and healthcare application because it can be used in the detection of the latent cardiac diseases of the patients in early time to avoid further exacerbation of the situation. The ECG monitoring devices usually work hours to record the beats of heart, leading to large amounts of data to be stored, or even be transmitted when it comes to telemedicine. Thus, the need to reduce the amount of data is generated. In Nyquist-Shannon sampling theory, recovering the analog signal without distortion means the sampling frequency is no less than two times the highest frequency in the analog signal spectrum. Compressive sensing [1] (CS) is a technology which can sample the signal and recover the data from fewer measurement suggested by Nyquist-Shannon sampling theory. Thus, it is regarded as a proper approach to deal with the large number of ECG data. In the clinical situation, CS compresses the original ECG signal to measurement with lower dimension which need less storage room, and measurement can be recovered to reconstructed signal which approaches the original signal, which is suitable for clinical diagnosis.

Traditional CS methods always exploit known measuring matrix to compress the original signal, and reconstruction involves finding a solution in the underdetermined

The associate editor coordinating the review of this manuscript and approving it for publication was Fang Yang.

system. Besides, traditional CS methods have some strict constraints. Firstly, the original signal needs to be sparse in some domain. However, many signals in nature are not sparse, so they need to go through some transformation, such as wavelet transformation [2] or Fourier transformation [3]. Secondly, we should find a proper measuring matrix to compress the original signal efficiently. The matrix is always a known one satisfying the Restricted Isometry Property (RIP) principle. RIP principle can be defined as follow:

$$(1-\delta)\|x\|_2^2 \leq \|\mathbf{\Phi}x\|_2^2 \leq (1+\delta)\|x\|_2^2 \tag{1}$$

where δ is a small constant, x is original signal and $\mathbf{\Phi}$ is the measuring matrix. For a given $\mathbf{\Phi}$, if there exists a δ satisfying Equation (1), $\mathbf{\Phi}$ has the RIP principle, which means the matrix can be used as measuring matrix in CS application to compress the original signal. Finally, in the process of finding the solution in the underdetermined system, researchers face the non-convex optimization problem which is difficult to solve accurately in a reasonable amount of time. Thus, they often convert the problem into a convex one by altering some conditions.

Based on the above three points, many researchers proposed relative methods. Tropp et al. [4] propose the OMP algorithm to implement fast reconstruction from random linear measurements, significantly narrow the gap between the theoretical performance of greedy acid sum and linear programming approach. Zhang and Rao [5] establish a sparse representation model without parameter coupling, then propose BSBL-EM algorithm based on sparse representation to recover block sparse signals, and the signal is reconstructed with high quality by deriving DOA and polarization parameters using intra-block correlations. Zhang et al. [6] exploited the spatial ad temporal structure of the signals and the relationship between them to compress and recover the fetal ECG signal. Vito et al. [7] proposed an algorithm based on the optimization of the dictionary. The method used overcomplete wavelet dictionary which is reduced by the training phase and took advantage of the features of ECG signals of different scale to conduct the reconstruction. Aiming at the problem in the CS domain. Kumar et al. [8] proposed a low-ranking approach and used Kroneker's sparse bases to exploit the spatial-temporal correlations in the ECG signals. The optimization problem in the recovering process was solved by an algorithm called ''Emperor Penguin Colony'', which has higher accuracy of reconstruction and lower complexity. Liu et al. [9] exploited a method using basis pursuit to denoise and compress sensing the ECG signal. In the process, the method adopted low-pass filtering and alternating direction method of multipliers (ADMM) algorithms, achieving good denoising quality and low recovering error. Liu et al. [10] aimed at getting an approximated norm constraint method for sparse expression using accelerated gradient descent method, and getting a self-training dictionary. This scheme performs well in terms of signal-to-noise ratio, percent norm difference and running time. Shinde et al. [11] put forward a scheme using an enhanced bi-orthogonal wavelet filter and finetuned the scaling coefficients, revealing that the influence of the method on the variation in mutation rate and awareness probability for ECG signal corresponding to different bi-orthogonal wavelets. Erkoc and Karaboga [12] designed a sparse reconstruction method based on non-dominated sorting genetic algorithm and verified the effectiveness of the optimization ability of the method.

Although traditional CS methods reduce the amount of the data, there are still some limits in the practical situation. Methods recover the single signal using optimization algorithm, and the measuring matrix is fixed when compressing the signal, so the overall structure information of the original signals is not exploited during the process. Therefore, the performance of the reconstruction of original signals is not so good as expected when the sensing rate is rather low.

Thus, as the deep learning theory developed, it has been gradually put into use in the CS field, because reconstruction methods using deep learning are almost data-driven and self-adaptive, achieving better performance. Usually, the labels in a Convolutional Neural Network (CNN) framework are fitting values or the categories. But in deep learning CS methods, the labels are the original signals.

In recent years, there are more and more researches about deep learning CS approaches. They can be roughly divided into three types. The first category includes the networks using unrolling iterative algorithm. Each stage of one network corresponds to one of the iterative steps it exploits. The examples of them are ADMM-Net proposed by Yang et al. [13] and ISTA-Net proposed by Zhang and Ghanem [14] The two methods used Alternating Direction Method of Multipliers (ADMM) algorithm and Iterative Shrinkage-Thresholding Algorithm (ISTA) to optimize the model to get better quality, respectively. The second category contains those based on Deep Neural Network (DNN) framework and similar structure. Many researchers have proposed relative approaches. Wei et al. [15] utilized a privacy-aware sensing and transmission scheme, exploiting a model-based sparsity-aware DNN to extract the features of the original signal and reconstruct them accurately. Zhang et al. [16] combined CNN and long short-term memory (LSTM) to learn the mapping from the measurement after rising dimension to the original signal. The methods can recover the signal fast and gain good reconstruction quality. Muduli et al. [17] adopted Stacked Denoised Autoencoder to conduct the CS process of Fetal ECG. In the framework, the compression is done in the encoder and the reconstruction in decoder with a model similar to CS-Based Multiple Measurement Vector model. Sheykhivand et al. [18] utilized seven convolutional layers and three LSTMs to construct an framework for EEG compressive sensing, enabling the detection of driver fatigue with a minimal EEG electrodes. There are also some

DNN frameworks without LSTM for ECG compressive sensing, e.g., Al-Marridi et al. [19] used a convolutional autoencoder (CAE) to implement efficient compression and reconstruction of ECG signals. The third category is comprised of methods using Generative Adversarial Network (GAN). Li et al. [20] defined a new structure regularization called Patch Correlation Regularization to extract the information of signal efficiently. Besides, they adopted a generator using SU-Net to enhance the ability of revealing the information. Yaqub et al. [21] constructed a GAN model based on transfer learning and examined the capability of the GAN model for under sampled multi-channel MR images in terms of difference between training and test data, which produced superior results with fewer data.

Deep learning based compressive sensing can obtain better reconstruction performance than conventional methods, once the network is trained, it can reconstruct signals at a very fast speed, but most of the existing deep learning based compressive sensing methods tend to be over complicated in the signal compression module, but if a simple compression model is used may lead to insufficient feature extraction that affects the reconstruction quality.

In order to address the issues mentioned above, in our paper, we propose a deep compressed sensing framework based on multi-scale feature fusion and SE block to achieve highquality reconstruction of ECG signals while reducing the complexity of the compression module, and our contributions are listed below:

- The compression module consists of three sequential convolutional layers, and the channels of the convolutional layers are adjusted according to the sensing rate to implement adaptive compression. In addition the space complexity of the compression module is reduced due to the parameter sharing of the convolutional layers.
- We used a multi-scale convolutional feature fusion structure, where two feature maps are convolved with three different kernel sizes respectively, and the feature maps obtained from the same kernel sizes are summed to extract the different scales of ECG signals.
- We introduce SE block in the reconstruction process to improve the performance of the ECG signal reconstruction model and thus reconstruct the ECG signal more accurately.
- We conduct our experiment on two different ECG data-bases, validating the performance of our method. Compared with other methods, the performance of the proposed method shows the best performance on different ECG databases.

The rest of the paper is organized as follow. Section II illustrates the relative technology in the CS domain. Section III describes the method we propose and the structure of our network. Section IV lists the details in our experiment and the comparison results between all the five methods. Section V is the discussion about the method. Section VI concludes this paper.

II. BACKGROUND

A. COMPRESSED SENSING

In Compressed sensing, the whole process can be comprised of compression and reconstruction parts. The compression module can be expressed as follow:

$$y = \Phi x \tag{2}$$

where x denotes the original signal, y denotes the measurement obtained in the compression module and Φ is a fixed measuring matrix with dimension of $M \times N$ ($M \ll N$). In order to get accurate reconstruction signal, Φ must satisfy RIP principle and x must be sparse in some domain.

In practice, measuring matrixes can be divided into two categories. One is random measuring matrix, such as Gaussian random matrix [22] and Bernoulli random matrix [23]. The other is deterministic measuring matrix, such as matrix generating from chaotic sequence [24].

Original signal x can be expressed as below:

$$x = \Psi\alpha \tag{3}$$

where $\Psi = [\psi_1, \psi_2, \cdots, \psi_N]$ denotes a set of orthogonal bases, α denotes the representation of x under Ψ. If there are k ($k \ll N$) non-zero elements in α, we call α is k-sparse. Therefore, a reconstruction of the signal accurately needs x or α to be sparse.

B. DEEP LEARNING CS METHODS

The deep learning CS method can be divided into three categories. Thus, we briefly introduce the three types of models in the rest of the subsection.

1) CS NETWORK BASED ON UNROLLING ITERATIVE ALGORITHM

In the unrolling iterative CS network, If the iterative process is mapped into a data flow diagram, we can observe that easily the dataflow is comprised of the repetition of the alternating optimization of different operators in iterative algorithm. There are nodes mapped from the operators in the iterative algorithm in each stage, and the ith iterative in the algorithm corresponds to the ith stage in the dataflow. That is why we call this type of models unrolling iterative CS networks. The parameters in the framework are updated using back propagation, and the performance of the networks is better than the traditional methods. The basic framework is depicted in Figure. 1.

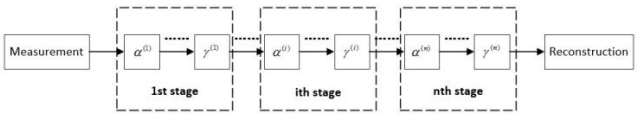


FIGURE 1. Basic framework of unrolling iterative CS network.

2) CS NETWORK BASED ON DNN

The CS network based on DNN usually include input layer, hidden layers and output layer. If the input is original signal,

the network includes compressing and reconstruction parts. In the compressing module, multiple convolutional layers are adopted to work as compression function. And in the reconstruction module, we often use convolutional layers, pooling layers, upsample layers or else to recover the signal from the compression. If the input is the measurement of the original signal, the network only includes reconstruction module and the measurements are often obtained through the multiplication between the measuring matrix and the original signal. The output of the network is the recovered signal. Unlike the DNN model in classification tasks, the label of the network is the original signal rather than a single value. Through the comparison between the output and the label, we can get the evaluation of the performance of the methods. Figure. 2. shows an example of DNN CS model.

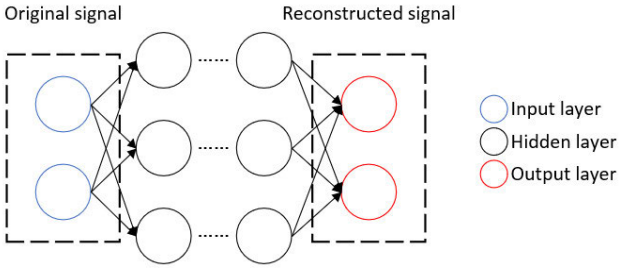


FIGURE 2. An example of DNN CS model.

3) CS NETWORK BASED ON GAN

In general GAN, the generator always generates synthetic data from a random noise z, and the discriminator plays the role of a binary classifier to categorize the real data and the synthetic data. Through the back propagation of the loss function which contains adversarial loss (the cross-entropy loss), the generator and discriminator get updated alternatively. In this case, the generated synthetic data get closer and closer to the real data. In CS field, the generator doesn't generate data from a random z, instead, it is a whole process of general CS process. That is to say, the input and output of the generator is the original signal and reconstructed signal, respectively, which has the similar structure as that of the autoencoder, firstly the original signal is compressed by several layers of neural network to get the compressed value, secondly the reconstructed signal is obtained by several layers of neural network. And then the discriminator distinguishes the original signal from recovered one.

C. SQUEEZE AND EXCITATION (SE) BLOCK

SE block is a method for enhancing feature channels in CNNs, proposed by Hu et al. [25] The main characteristic is the introduction of a block called "Squeeze-and-Excitation" in the traditional convolutional neural network structure, which is used to adaptively adjust the importance of features between different channels to improve the classification performance of the network. The core idea of SE block is to model global features on each channel to learn the correlation

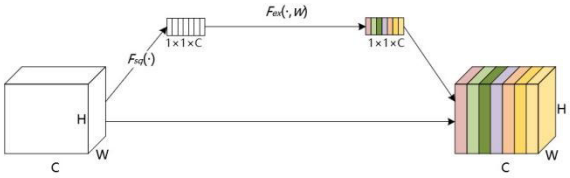


FIGURE 3. The structure of SE block.

between feature channels and adaptively adjust the importance of features between channels. As shown in Figure. 3., this process consists of two main phases: Squeeze and Excitation. in the Squeeze phase (Fsq(·) in Figure. 3.), the SE block first performs a global average pooling of the input features in order to obtain a feature value on each channel. Then, a fully connected layer is used to map these feature values to a smaller dimension in order to reduce the computational cost and spatial complexity. In the Excitation phase (Fex(·, W) in Figure. 3.), the SE block adaptively adjusts the feature importance of each channel using a sigmoid function. Specifically, SE block inputs the globally averaged pooled features into a fully connected layer and outputs a vector with a dimension equal to the number of channels of the input features. Then, each element in this vector is mapped to a value between 0 and 1 by a sigmoid function, indicating the importance of the feature corresponding to the channel. Finally, these values are multiplied by the input features, thus adaptively adjusting the importance of the features of each channel, allowing the network to distinguish more accurately between the different features. By adaptively adjusting the importance of features between different channels, the SE block can significantly improve the classification performance of the network and has achieved excellent performance on several image classification datasets. In addition, the design ideas of SE block can be applied to other fields, such as target detection and semantic segmentation tasks.

III. MATERIALS AND METHODS

In this paper, we propose a deep compressive sensing framework for ECG signals based on multiscale feature fusion and SE block. The framework consists of four main modules: preprocessing, compression, initial reconstruction and secondary reconstruction. The data preprocessing module normalizes the input ECG signal, the compression module utilizes sequential convolutional layers to compact the features of the ECG signal, and the initial reconstruction upscales the compressed signal and recovers the same size as the original ECG signal. The secondary reconstruction module extracts features of different scales using different kernel sizes, subsequently, and fuses the different scale features, then the interdependencies of features between individual channels are learned using SE block to enhance the expressiveness of the network. Finally, the temporal features of the ECG signal are extracted using LSTM and the ECG signal is reconstructed by fully connected layers. The overall framework is

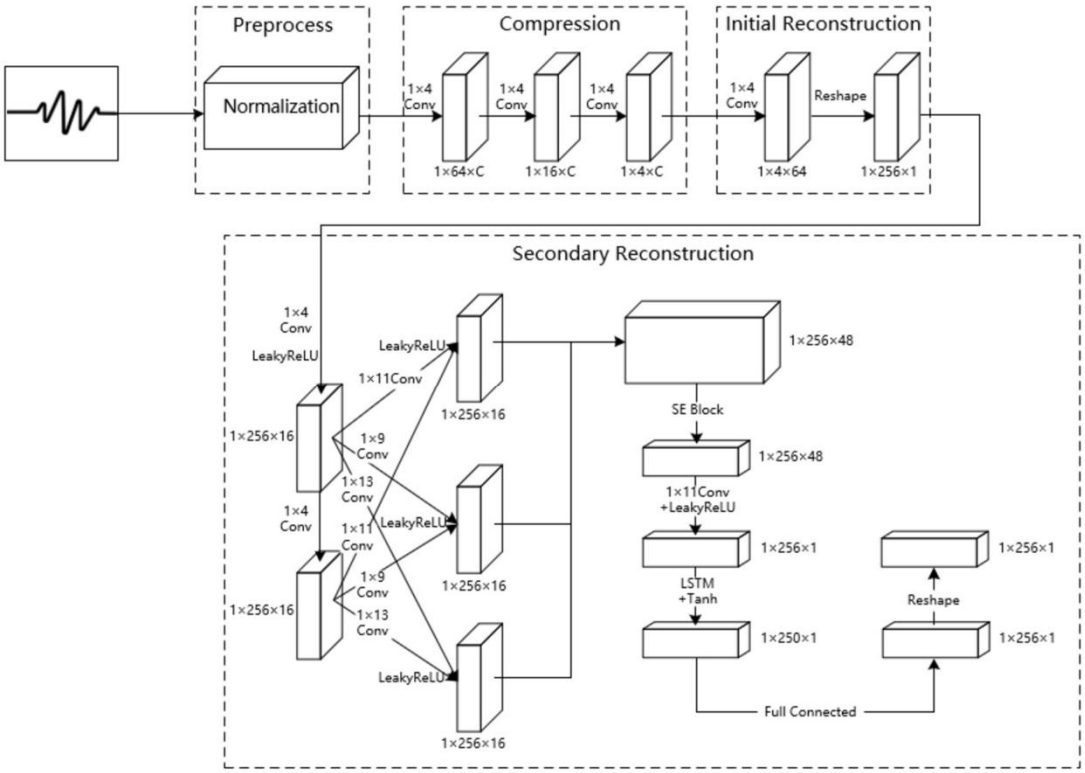


FIGURE 4. The overall framework of our network.

shown in Figure. 4. This section focuses on introducing our network architecture and calculating the number of parameters to analyze its complexity.

A. DATASET

In this paper, we validated the performance of the proposed method on the MIT-BIH Arrhythmia Database [26], which was established with the support of Boston's Beth Israel Hospital and MIT, and it contains 48 ECG signal records from 47 subjects, each record includes two channels of ECG signals with a sampling frequency of 360 Hz. We utilized the ECG signals from the lead MLII in the experiments, we excluded records 102 and 104 because these two only recorded lead V2 and lead V5. In addition, we validated the robustness of the proposed model using the Non-Invasive Fetal ECG Arrhythmia Database, which contains 12 arrhythmia records and 14 normal rhythm records, each of which contains four or five abdominal channels and one chest maternal channel, with a sampling frequency of 500 Hz or 1000 Hz.

B. DATA PREPROCESSING

The range of values of ECG signal sampling points is not fixed. In order to avoid the possible gradient exploding problem, signal should be normalized before being input into the network. Let $x = \{x_{(1)}, x_{(2)}, \cdots, x_{(i)}\}$ be the original ECG signals, and $x_{(i)}$ represents the i-th signal with 256 sampling points. The normalization is expressed as follow:

$$\hat{x}_{(i)} = \frac{x_{(i)} - min\left(x_{(i)}\right)}{max\left(x_{(i)}\right) - min\left(x_{(i)}\right)} \tag{4}$$

where $min\left(x_{(i)}\right)$ is the minimum value in $x_{(i)}$, $max\left(x_{(i)}\right)$ is the maximum value in $x_{(i)}$ and $\hat{x}_{(i)}$ is the transformation of the signal $x_{(i)}$ after normalization. After the normalization operation, the values of the input signal is between zero and one, which contributes to a proper gradient and the stable convergence of the network.

C. COMPRESSION

In the compression module, we use three sequential convolutional layers to compress the ECG data. The kernel size of the convolutional layers is 1×4 and the stride is 4. After three convolutional layers, the size of the feature map is reduced to 1×4, and we need to calculate the channels according to the sensing rate. For example, the ECG signal we have chosen contains 256 sampling points, in the case of the sensing rate is 0.05, we need to compress the signal into $floor(256 \times 0.05) = 12$ measurements, then we need to set the output channels of the convolution layer to $floor(12 \div 4) = 3$, where $floor(\cdot)$ means that the decimal part is discarded and only the integer part is retained. The output of the convolution layer can be considered as the response of the signal under different filters, which helps to understand the frequency

domain characteristics of the signal and the activity in specific frequency bands.

D. RECONSTRUCTION

The reconstruction module includes initial and secondary reconstruction. The main function of the initial reconstruction is to recover the dimension of the ECG signal. Since the dimensionality of the measurements is much smaller than that of the original signal, we need to extend the dimensionality of the measurements to ensure that the size of the reconstructed signal is always the same as that of the original signal before the secondary reconstruction module. First, we use a convolutional layer with a kernel size of 1×3, an input channel equal to the output channels in the compression module described above, and the output channels is 64, and the stride is 1. Then, using the LeakyReLU activation function, we convert the feature map to a size of $1 \times 4 \times 64$, matching the total number of sample points of the original signal. Finally, the reshaping layer is then applied to the feature map to keep it at the same size as the original signal.

The input of the secondary reconstruction module is the $1 \times 256 \times 1$ signal obtained from the initial reconstruction. Firstly, the features of the signal are extracted utilizing two sequential convolutional layers and LeakyReLU, where the size of the convolutional kernel is 1×11 and the number of channels is 16, and these two convolutional layers output feature maps having the same size and the same number of channels. Secondly, we use 1×9, 1×11 and 1×13 convolutional layers with stride of 1 and LeakyReLU activation function to extract features at different scales for the feature maps obtained from the above two convolutional layers, respectively, and get six groups of feature maps, and sum the feature maps obtained from the same kernel size to obtain three groups of feature maps at different scales. and concatenate the three features maps. Third, we use the SE block (will be described below) to compute the dependencies between channels and reassign the weights to obtain a signal of length 1×256 with 48 channels. The signal is then converted to a signal of length 1×256 with 1 channel using a convolutional layer with a kernel size of 1×11 and a Leaky-ReLU activation function. Finally the temporal features of the signal are extracted employing an LSTM with 250 hidden units, and the signal is restored to the same dimension of the original signal using a fully concatenated layer and a dimensional reshaping layer. This method sums the multi-scale features generated from two different signals and finally splices all the features to improve the characterization ability of the network and the signal reconstruction performance.

E. SE-BLOCK

The input of the SE module is a signal of length 1×256 with 48 channels, firstly, a global pooling average layer is utilized to obtain a feature map of length 1×1 with 48 channels. In order to reduce the complexity of the network, a fully-connected layer is used before the ReLU activation function, the output of this fully-connected layer is a signal of length 1×1 with 16 channels, and a fully-connected layer is used again before the Sigmoid activation function to recover the number of channels of the signal, which results in a signal of length 1×1 with 48 channels. After the ReLU and Sigmoid activation functions, the weights of the channels are reassigned, and then the weights are multiplied with the input of the SE module by channel to get the output of the SE module, so that the output of the SE module is the reassigned channel relationship, and the specific process is shown in Figure. 5. The use of this module can emphasize the more informative features and suppress the less informative ones, which makes it easier to generate the representation of the information and improve the reconfiguration performance of the network.

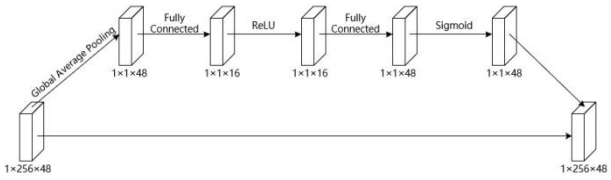


FIGURE 5. The SE block structure in our approach.

F. COMPLEXITY ANALYSIS

The proposed model consists of a total of $593{,}209 + 36 \times C$ parameters, including $20{,}609 + 36 \times C$ parameters in convolutional layers, 507,000 parameters in LSTM layers, and 65,600 parameters in fully connected layers. The detailed breakdown is shown in TABLE 1, where C represents the number of channels in the compression module's convolutional layers, which needs to be calculated based on the sensing rate. As mentioned earlier, in the case of sensing rate is 0.05, C equals 3.

IV. RESULTS

In this section, we compare our CS method with two traditional CS methods and three deep learning CS methods, showing that our method gains the best performance among all the five methods in terms of the two evaluating metrics. In addition, we choose another database to repeat the experiments and find that our approach also performs better than other comparative methods. The experiment details and results are listed below.

A. EXPERIMENT SETTING

1) DATASET

The proposed method is validated on the database MIT-BIH Arrhythmia database [26]. The database contains two-lead 30 min heartbeats records from 48 patients. In our experiments, we adopt MLII signals. Nos. 101, 107, 109, 112, 116, 119, 121, 200, 205, 210, 215, 219, 221, 228, 230, 234 are chosen as our testing datasets, and the rest records are chosen as our training datasets. We select 18000 ECG signal segments evenly from training datasets and 4000 from testing datasets, each with 256 sampling points.

TABLE 1. The model parameter.

Module	Layer	Input Size	Kernnel Size	Biases	Filters	Parameters
Compressiion	Conv1d	1×256×1	1×4	No	C	1×1×4×C = 4×C
	Conv1d	1×64×1	1×4	No	C	4×1×4×C = 16×C
	Conv1d	1×16×1	1×4	No	C	4×1×4×C = 16×C
Initial Reconstruction	Conv1d	1×4×1	1×1	No	64	1×1×1×64 = 64
	Reshape	1×64×1	-	-	-	0
Secondary Reconstruction	Conv1d	1×256×1	1×11	Yes	16	1×1×11×16+16 = 192
	Conv1d	1×256×16	$\begin{Bmatrix}1\times11\\1\times9\\1\times11\\1\times13\end{Bmatrix}$	Yes	16	$16\times\begin{Bmatrix}1\times11\\1\times9\\1\times11\\1\times13\end{Bmatrix}\times16+16\times4 = 11328$
	Conv1d	1×256×16	$\begin{Bmatrix}1\times9\\1\times11\\1\times13\end{Bmatrix}$	Yes	16	$16\times\begin{Bmatrix}1\times9\\1\times11\\1\times13\end{Bmatrix}\times16+16\times3 = 8496$
	Concatnate	$3\times\{1\times256\times16\}$	-	-	-	0
	SE Block	1×256×48				(48×16)+16+(16×48)+48 = 1600
	Conv1d	1×256×48	1×11	Yes	1	48×1×11×1+1 = 529
	LSTM	1×256×1	-	-	-	((256+250)×250+250)×4 = 507000
	Full Connected	1×250×1	-	-	-	250×256 = 64000

In CS domain, the sensing rate (SR) can be defined as follow:

$$\mathrm{SR} = \frac{M}{N} \tag{5}$$

where N is the number of sampling points in one original signal, and M is the number of sampling points in one measurement. As for deep learning CS, the sampling points refer to the total numbers (length × width × channel number) in a feature map. In our experiments, N is 256 and the we choose 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5 as our sensing rates.

2) METRICS

To evaluate the performance of the CS methods quantitatively, we use Percentage Root-mean-square Difference [28] (PRD) and Signal-to-Noise [29] (SNR) to measure the difference between the original signal and reconstructed signal.

PRD can be expressed as follow:

$$\mathrm{PRD}\,(\%) = \frac{\|x-\bar{x}\|_2}{\|x\|_2}\times 100 \tag{6}$$

where x and $\bar{x}$ are the original signal and the reconstructed signal generated by the generator, respectively. PRD describes the ratio of the l_2 norm of the distance between the original signal and reconstructed signal to l_2 norm of the original signal, indicating the recovering error. The smaller the PRD, the better the reconstruction quality. According to the value of the PRD, we can divide the reconstruction quality to four levels, as depicted in TABLE 2.

SNR can be expressed as follow:

$$\mathrm{SNR}\,(\mathrm{dB}) = 10\log_{10}\frac{\|x\|_2^2}{\|x-\bar{x}\|_2^2} \tag{7}$$

TABLE 2. PRD and reconstruction quality level.

PRD(%)	Signals Reconstruction Quality
0–2	Very good
2–9	Good
9–19	No good
19–60	Bad

SNR is relative to the reciprocal of the PRD, showing the recovering quality. The bigger the SNR, the better the recovering quality.

3) LOSS FUNCTION

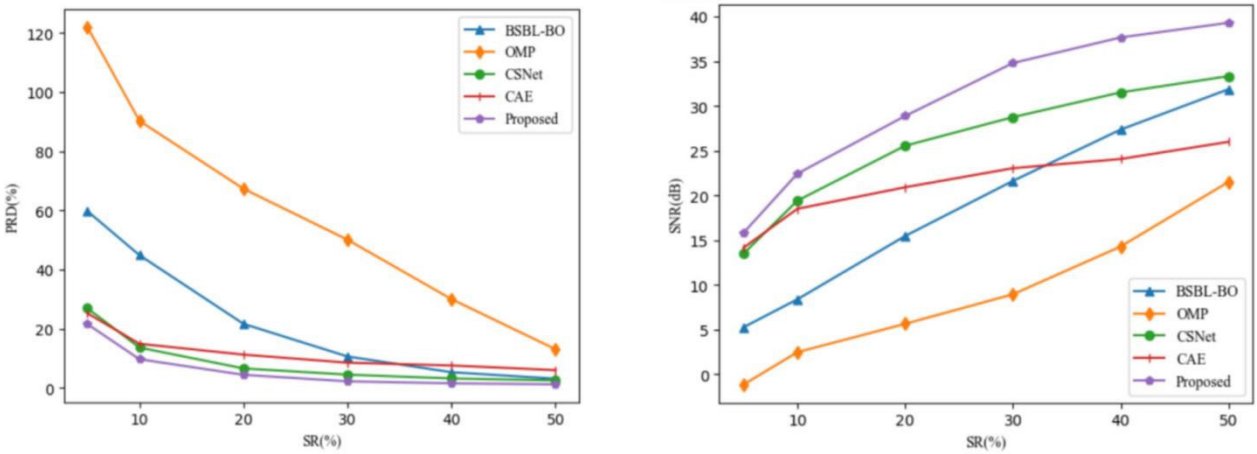
In this paper, we need to compress the original signal using continuous convolution to map the high-dimensional original signal into the low-dimensional compressed domain, and measure the difference between the reconstructed signal and the original signal by the loss function during decompression. Since the mean square error has good mathematical properties and intuitive physical meaning, such as derivability and convexity, making it easy to solve and optimize during the optimization process, in this paper we use MSE as the loss function used to train the proposed framework to improve the quality of the reconstructed signal, MSE can be expressed as follow:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{x}_i)^2 \tag{8}$$

where x_i denotes the input of our methods, namely, the original signal, and $\hat{x}_i$ denotes the reconstructed signal.

TABLE 3. The PRD and SNR of diff8rent compressive sensing methods at different sensing rates on MIT-BIH database.

SR	BSBL-BO		OMP		CSNet		CAE		Proposed	
	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)
0. 05	59. 72	5. 25	121. 88	−1. 16	26. 90	13. 47	25. 22	14. 14	21. 55	15. 80
0. 1	44. 85	8. 38	90. 23	2. 47	13. 63	19. 41	14. 91	18. 50	9. 70	22. 43
0. 2	21. 65	15. 46	67. 30	5. 64	6. 59	25. 55	11. 25	20. 91	4. 43	28. 88
0. 3	10. 62	21. 62	50. 07	8. 94	4. 49	28. 75	8. 53	23. 04	2. 23	34. 80
0. 4	5. 30	27. 36	30. 00	14. 29	3. 20	31. 52	7. 58	24. 07	1. 55	37. 66
0. 5	3. 12	31. 87	13. 08	21. 51	2. 56	33. 33	6. 02	25. 99	1. 24	39. 30



(a) PRD curve **(b)** SNR curve

FIGURE 6. The curve of PRD and SNR of different compressive sensing methods at different sensing rates on MIT-BIH database.

4) TRAINING PARAMETERS

In the process of training, we choose Adam as our optimizer. We set the initial learning rate of the generator and the discriminator to 0.0005, with the first and second moment decay rate as 0.9 and 0.999, respectively. The total training epoch is 200 and the batch size is 32. In our experiment, we conduct all the experiments at sensing rates 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5. Our experiments are all conducted on Windows 10, Pytorch 1.9.0, GPU NVIDIA Tesla K80 and CPU Intel Xeon E5-2678 v3.

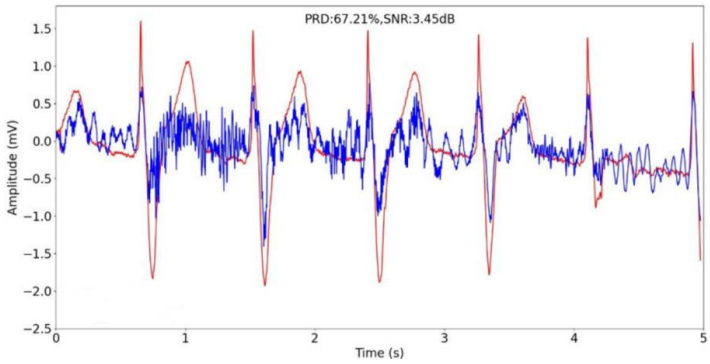
B. COMPARISON WITH OTHER CS METHODS

We compare our methods with two traditional CS methods (BSBL-BO [12] and OMP [4]) and two deep learning CS methods (CSNet [16] and CAE [19] in terms of PRD and SNR. In our methods, we use signal segments of 256 sampling points as the input, however, the CAE methods converted the one-dimension signal into a square one using zigzag method. Thus, when we conduct the contrast experiment using CAE methods, we reshape the original signal into size of 16 × 16. As for the BSBL-BO and OMP methods, we exploit sparse random matrix and random Gaussian matrix for compression.

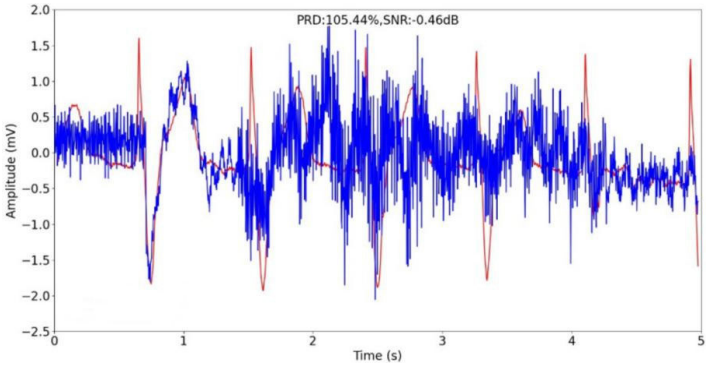
The average PRDs and SNRs of different CS methods at different sensing rates are listed in TABLE 3. As can be seen from the table, the reconstruction performance of the proposed method is optimal at all sensing rates compared to other benchmark methods. When the sensing rate is 0.1, the PRD of BSBL-BO, OMP, CSNet, CAE and the proposed method are 44.85%, 90.23%, 13.63%, 14.91% and 9.70%, respectively, and the corresponding SNRs are 8.38dB, 2.47dB, 19.41dB, 18.50dB and 22.43dB, and our proposed method has the lowest PRD and the highest SNR values. When the sensing rate is 0.4, the PRDs of BSBL-BO, OMP, CSNet, CAE and our proposed method are 5.30%, 30.00%, 3.20%, 7.58% and 1.55%, respectively, and the corresponding SNRs are 27.36dB, 14.29dB, 31.52dB, 24.07dB and 37.66dB, our proposed method also has the lowest PRD and the highest SNR values.

Figure. 6. shows the trend of the PRDs and SNRs of different CS methods at different sensing rates. We can observe that the curve of PRD of our methods is the flattest, that is to say, our methods gains the lowest construction error at all the sensing rates among all the comparative methods.

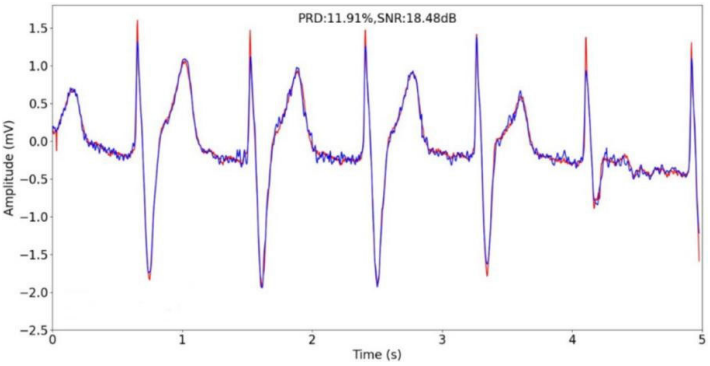
We choose a signal segment about 5 s from No.219 in the testing datasets, and recover the signal at sensing rate



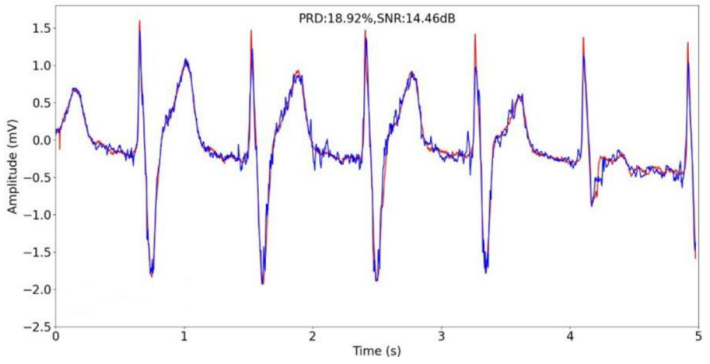
(a) BSBL-BO



(b) OMP



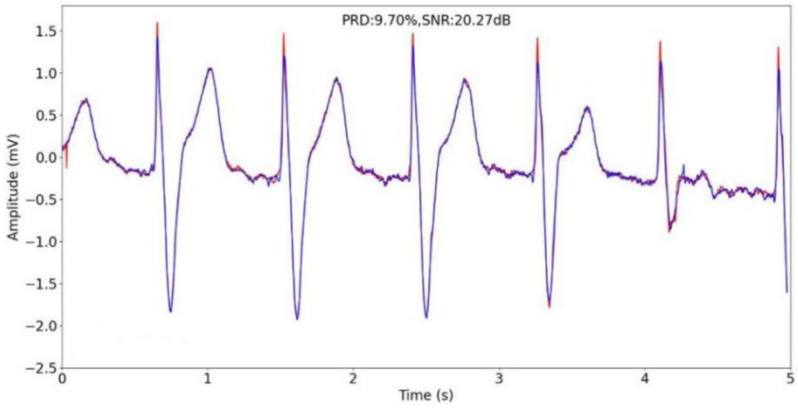
(c) CSNet



(d) CAE

FIGURE 7. The comparison of original signal and reconstructed signal using different compressive sensing methods at sensing rates of 0.1 when reconstructing signal about 5s in the 219 record on MIT-BIH database.

of 0.1 for the visual comparison. Figure. 7. shows the comparison of the reconstructed signal obtained using different compressed perception methods to the original signal. As can be seen from the figure, our proposed method obtains the

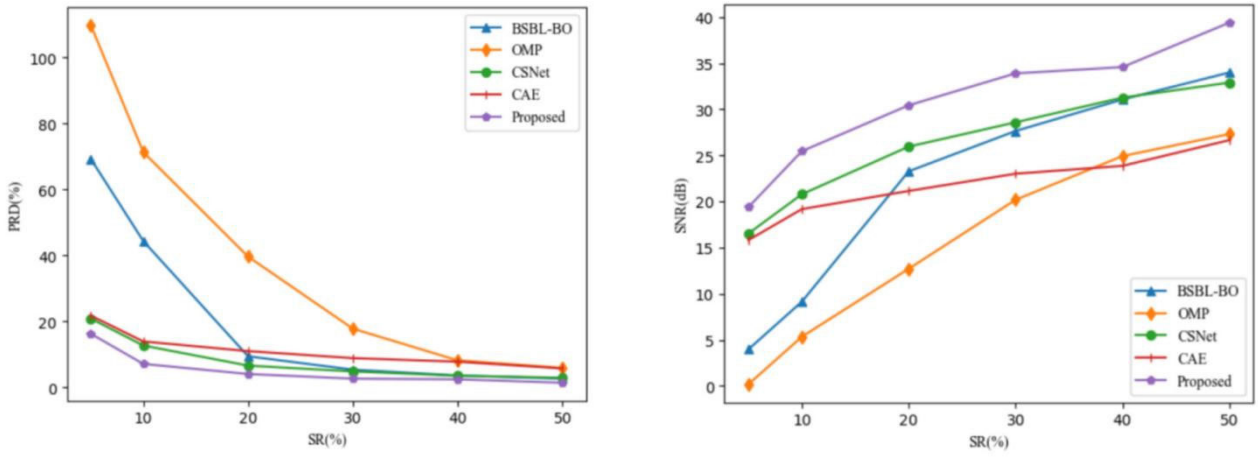


(e) Proposed

FIGURE 7. *(Continued.)* **The comparison of original signal and reconstructed signal using different compressive sensing methods at sensing rates of 0.1 when reconstructing signal about 5s in the 219 record on MIT-BIH database.**

TABLE 4. **The PRD and SNR of different compressive sensing methods at different sensing rates on MIT-BIH database.**

SR	BSBL-BO		OMP		CSNet		CAE		Proposed	
	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)	PRD (%)	SNR (dB)
0.05	69.23	3.96	109.77	0.17	20.87	16.52	21.77	15.81	16.27	19.38
0.1	44.37	9.13	71.31	5.33	12.71	20.78	13.94	19.17	7.14	25.45
0.2	9.44	23.27	39.64	12.70	6.62	25.95	11.01	21.14	4.08	30.43
0.3	5.39	27.64	17.80	20.18	4.86	28.58	8.88	23.01	2.67	33.89
0.4	3.63	31.06	8.21	24.92	3.55	31.26	7.81	23.86	2.48	34.57
0.5	2.66	33.98	5.93	27.31	2.94	32.89	5.76	26.65	1.43	39.39



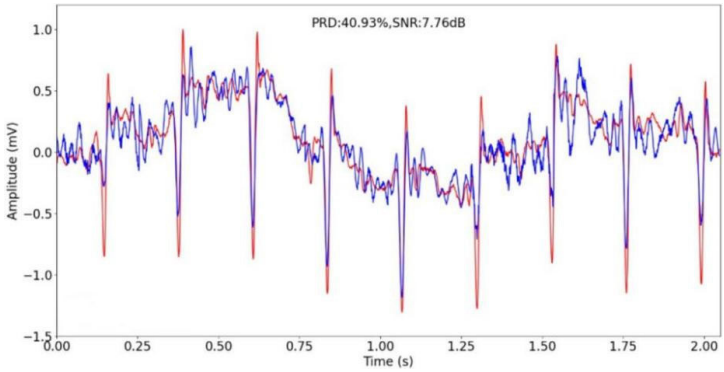
(a) PRD curve **(b)** SNR curve

FIGURE 8. **The curve of PRD and SNR of different compressive sensing methods at different sensing rates on NIFEA database.**

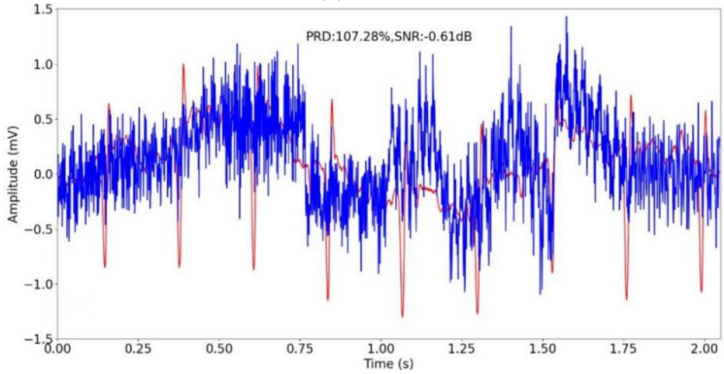
lowest PRD (9.72%) and the highest SNR (20.27 dB), produces the reconstructed signal closest to the original signal, and the peak of the recovered signal matches the original signal the best, and obtains the best reconstructed performance.

C. EXPERIMENT ON ANOTHER DATABASE

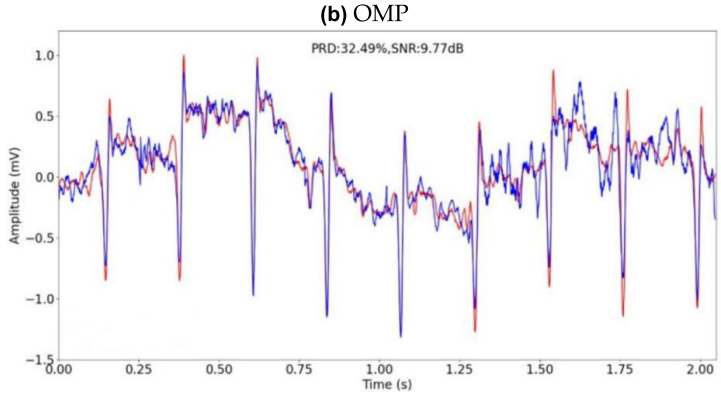
We have obtained satisfactory results on the MIT-BIH Arrhythmia Database and in order to validate the robustness of our methods, we choose another database—Non-Invasive



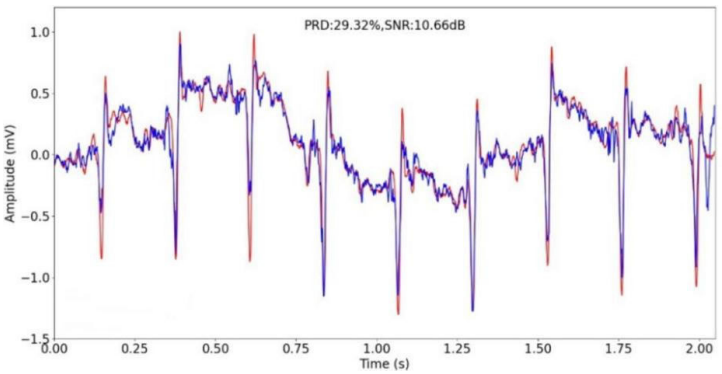
(a) BSBL-BO



(b) OMP



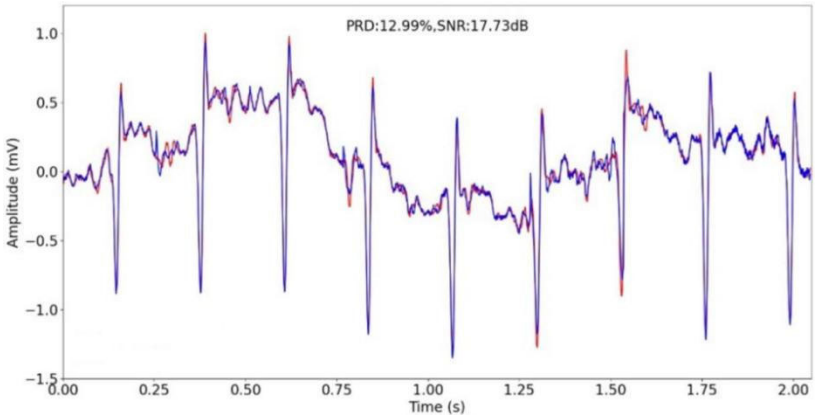
(c) CSNet



(d) CAE

FIGURE 9. The comparison of the original signal and reconstructed signal at sensing rate of 0.1 when reconstructing signal about 2s in the ARR_09 record on NIFEA database.

Fetal ECG Arrhythmia Database [27] to conduct experiments using all the five methods at different sensing rates. This database consists of the chest signals and four to five leads abdomen signals from 26 objects. In our experiment,



(e) Proposed

FIGURE 9. ***(Continued.)*** **The comparison of the original signal and reconstructed signal at sensing rate of 0.1 when reconstructing signal about 2s in the ARR_09 record on NIFEA database.**

we choose the signal in ECG channel. Records ARR_09, ARR_10, ARR_11, ARR12, NR_11, NR_12, NR_13 and NR_14 are adopted as testing datasets, and the rest as the training datasets. We select 16000 ECG signal segment consisting of 256 sampling points evenly from the training datasets and 4000 from the testing datasets. Consistent with the experiments on the MIT-BIH arrhythmia database, the sensing rates were also 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5 in the experiments, and the learning rates and other hyperparameters remained consistent.

The average PRDs and SNRs of different CS methods at different sensing rates on the Non-Invasive Fetal ECG Arrhythmia Database are listed in TABLE 4. As can be seen from the table, proposed method has the highest SNR and the lowest PRD at all sensing rates, and presents the best reconstruction performance overall. When the sensing rate is 0.1, the PRDs of BSBL-BO, OMP, CSNet, CAE and the proposed method are 44.37%, 71.31%, 12.71%, 13.94% and 7.14%, respectively, and the corresponding SNRs are 9.13 dB, 5.33 dB, 20.78 dB, 19.17 dB and 25.45 dB. Proposed method has the lowest PRD and the highest SNR value. When the sensing rate is 0.4, the PRD of BSBL-BO, OMP, CSNet, CAE and the proposed method are 3.63%, 8.21%, 3.55%, 7.81% and 2.48%, respectively, and the corresponding SNRs are 31.06dB, 24.92dB, 31.26dB, 23.86dB34.57dB, and the proposed method also has the lowest PRD and the highest SNR values.

Figure. 8. shows the trend of the PRDs and SNRs of different CS methods at different sensing rates. We can also conclude that our method has the best performance in Non-Invasive Fetal ECG Arrhythmia Database as well.

We choose a signal segment about 2 s from record ARR_09 in the testing datasets, and reconstruct the signal at sensing rate of 0.1 for visual comparison. The original (Indicated by the red line) and recovered(Indicated by the blue line) signal are depicted in Figure. 9. From the figure, we can see that the reconstructed signal produced by our method is the closed to the original one and has the lowest PRD among all the comparative methods.

The experiment results on the Non-Invasive Fetal ECG Arrhythmia Database show that our methods gain good performance as well on the new database, proving the robustness of our scheme.

V. DISCUSSION

From the experiment results, we can see that deep learning CS methods outperform the traditional CS methods generally on the reconstruction of ECG signals in terms of PRD and SNR. The traditional CS methods usually use iterative algorithm to reconstruct signals, therefore they can't exploit the structure information in all the data. However, the deep learning CS methods learn the data distribution from the training datasets, and can apply it in the recovering process. That is why the deep learning CS methods have a higher construction quality.

Our approach is mainly based on multiscale feature fusion and SE block. using SE block can help the model to better capture important information when learning features by learning correlations between feature channels to adjust the importance of each channel, thus improving the performance of the model on a specific task and improving the performance of the model. In addition, the SE block is relatively simple in design and can be easily integrated with existing CNN models. In contrast, LSTM is able to automatically learn and capture long-term dependencies in ECG signal sequences and encode them into compact feature representations, solving the problem of modeling time-series data and long sequences for efficient reconstruction. However, the proposed method performs not very well in reconstruction at low sensing rates and may lead to misdiagnosis in some particular cases.

VI. CONCLUSION

In this paper, we propose a compressive sensing framework based on multiscale feature fusion and SE block. The framework consists of four main components. We first normalize the original signal to facilitate the training of the model, and then use successive convolutions to compress the ECG signal for measurement. In the reconstruction process we first extract local features at different scales using different scales of convolution and fuse different local features with each other, then use SE block to capture important feature information, and finally use LSTM to extract the time dependence of the ECG signal to improve the reconstruction performance. Through the comparison experiments of MIT-BIH arrhythmia database and non-invasive fetal ECG arrhythmia in this paper, our method obtains better reconstruction quality than other comparison methods. At different sensing rates, our method obtains the lowest PRD and the highest SNR. When the sensing rate is 0.4, the proposed method obtains a PRD of 1.55%, which shows a very good signal reconstruction quality according to the metrics in Table 1, and the SNR is 37.66 dB, which is significantly better than the traditional compressed sensing methods and other deep compressed sensing methods. In future work, we will investigate how to reconstruct ECG signals with high quality at low sensing rates, and explore the application of deep compressive sensing approaches for multi-lead ECG signals.

AUTHOR CONTRIBUTIONS

Conceptualization: Jing Hua and Jue Rao; Methodology: Jiawen Zou; Software: Jiawen Zou and Jue Rao; Validation: Hua Yin and Jie Chen; Formal Analysis: Jue Rao; Investigation: Hua Yin; Resources: Jiawen Zou; Data Curation: ,Jie Chen; Writing—Original Draft Preparation: Jue Rao; Writing—Review and Editing: Jing Hua; Visualization: Jue Rao; Supervision: Jing Hua; Project Administration: Hua Yin; Funding Acquisition: Jing Hua and Hua Yin. All authors have read and agreed to the published version of the manuscript.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] A. Mishra, F. Thakkar, C. Modi, and R. Kher, "ECG signal compression using compressive sensing and wavelet transform," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 3404–3407.

[3] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic MRI," *Magn. Reson. Med.*, vol. 59, no. 2, pp. 365–373, Feb. 2008.

[4] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[5] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2009–2015, Apr. 2013.

[6] Z. Zhang, T.-P. Jung, S. Makeig, and B. D. Rao, "Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ECG via block sparse Bayesian learning," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 300–309, Feb. 2013.

[7] L. De Vito, E. Picariello, F. Picariello, S. Rapuano, and I. Tudosa, "A dictionary optimization method for reconstruction of ECG signals after compressed sensing," *Sensors*, vol. 21, no. 16, p. 5282, Aug. 2021.

[8] P. V. Kumar, C. N. Marimuthu, S. Sivaranjani, and S. Punitha, "Reconstruction of multi-channel ECG using compressive sensing based emperor penguin colony in WBSN," *Int. J. Mech. Eng.*, vol. 6, no. 3, pp. 958–967, 2021.

[9] R. Liu, M. Shu, and C. Chen, "ECG signal denoising and reconstruction based on basis pursuit," *Appl. Sci.*, vol. 11, no. 4, p. 1591, Feb. 2021.

[10] S. Liu and F.-Y. Wu, "Self-training dictionary based approximated l_0 norm constraint reconstruction for compressed ECG," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102768.

[11] A. N. Shinde, S. L. Nalbalwar, and A. B. Nandgaonkar, "Algorithmic analysis of optimized bi-orthogonal wavelet filters on compressed sensing," in *Proc. 5th Int. Conf. Trends Electron. Informat. (ICOEI)*, Tirunelveli, India, Jun. 2021, pp. 199–206.

[12] M. E. Erkoc and N. Karaboga, "Multi-objective sparse signal reconstruction in compressed sensing," in *Nature-Inspired Metaheuristic Algorithms for Engineering Optimization Applications*, 1st ed., S. Carbas, A. Toctas, and D. Ustun, Eds. Singapore: Springer, 2021, pp. 373–396.

[13] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: A deep learning approach for image compressive sensing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 3, pp. 521–538, Mar. 2020, doi: 10.1109/TPAMI.2018.2883941.

[14] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1828–1837.

[15] T. Wei, S. Liu, and X. Du, "Learning-based efficient sparse sensing and recovery for privacy-aware IoMT," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9948–9959, Jun. 2022, doi: 10.1109/JIOT.2022.3163593.

[16] H. Zhang, Z. Dong, Z. Wang, L. Guo, and Z. Wang, "CSNet: A deep learning approach for ECG compressed sensing," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 103065.

[17] P. R. Muduli, R. R. Gunukula, and A. Mukherjee, "A deep learning approach to fetal-ECG signal reconstruction," in *Proc. 22nd Nat. Conf. Commun. (NCC)*, Mar. 2016, pp. 1–6.

[18] S. Sheykhivand, T. Y. Rezaii, S. Meshgini, S. Makoui, and A. Farzamnia, "Developing a deep neural network for driver fatigue detection using EEG signals based on compressed sensing," *Sustainability*, vol. 14, no. 5, p. 2941, Mar. 2022.

[19] A. Z. Al-Marridi, A. Mohamed, and A. Erbad, "Convolutional autoencoder approach for EEG compression and reconstruction in m-Health systems," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2018, pp. 370–375.

[20] Z. Li, T. Zhang, and D. Zhang, "SEGAN: Structure-enhanced generative adversarial network for compressed sensing MRI reconstruction," in *Proc. 33rd AAAI Conf. Artif. Intell. (AAAI)*, Honolulu, HI, USA, Jan./Feb. 2019, pp. 1012–1019, doi: 10.1609/aaai.v33i01.33011012.

[21] M. Yaqub, F. Jinchao, S. Ahmed, K. Arshid, M. A. Bilal, M. P. Akhter, and M. S. Zia, "GAN-TL: Generative adversarial networks with transfer learning for MRI reconstruction," *Appl. Sci.*, vol. 12, no. 17, p. 8841, Sep. 2022.

[22] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006, doi: 10.1109/TIT.2005.862083.

[23] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.

[24] L. Yu, J. P. Barbot, G. Zheng, and H. Sun, "Compressive sensing with chaotic sequence," *IEEE Signal Process. Lett.*, vol. 17, no. 8, pp. 731–734, Aug. 2010.

[25] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, Aug. 2020.

[26] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.
[27] J. A. Behar, L. Bonnemains, V. Shulgin, J. Oster, O. Ostras, and I. Lakhno, "Noninvasive fetal electrocardiography for the detection of fetal arrhythmias," *Prenatal Diagnosis*, vol. 39, no. 3, pp. 178–187, Feb. 2019.
[28] F. M. Dias, M. Khosravy, T. W. Cabral, H. L. M. Monteiro, L. M. A. Filho, L. M. Honório, R. Naji, and C. A. Duque, "Compressive sensing of electrocardiogram," in *Compressive Sensing in Healthcare: Advances in Ubiquitous Sensing Applications for Healthcare*, M. Khosravy, N. Dey, and C. A. Duque, Eds. Amsterdam, The Netherlands: Elsevier, 2020, ch. 9, pp. 165–184.
[29] G. Box, "Signal-to-noise ratios, performance criteria, and transformations," *Technometrics*, vol. 30, no. 1, pp. 1–17, Feb. 1988.

JING HUA received the B.S. degree from the Shanghai University of Electric Power, in 2005, the M.S. degree from East China Normal University, in 2007, and the Ph.D. degree in mechanical engineering from Nanchang University, in 2018. She was a Lecturer with the School of Software, Jiangxi Agricultural University, where she is currently an Associate Professor. Her current research interests include compressive sensing, biosignal processing, and embedded systems.

JIAWEN ZOU received the bachelor's degree in bioengineering from Beijing Union University. He is currently pursuing the master's degree in computer science and technology with Jiangxi Agricultural University, China. His current research interests include deep learning, biosignal processing, and compressive sensing.

JUE RAO received the bachelor's degree in information management and information systems from the Dalian University of Technology, and the master's degree in computer science and technology from Jiangxi Agricultural University. Her current interests include machine learning, deep learning, and compressive sensing.

HUA YIN received the B.S. degree in computer science and technology from Nanchang University, Nanchang, China, in 2003, the M.S. degree in computer software and theory from Donghua University, Shanghai, China, in 2009, and the Ph.D. degree in mechanicaJIECHEN.pdfl engineering from Nanchang University, in 2017. His current research interests include smart agriculture and machine vision.

JIE CHEN received the bachelor's degree in software engineering from Jiangxi Agricultural University, where she is currently pursuing the master's degree in computer science and technology. Her current research interests include 3D modeling and blockchain.