IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 18, 2021, accepted March 23, 2021, date of publication March 26, 2021, date of current version April 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3069086

Image Reconstruction Based on Fused Features and Perceptual Loss Encoder-Decoder Residual Network for Space Optical Remote Sensing Images Compressive Sensing

SHUMING XIAO, SHAOJU WANG, AND LIN CHANG
Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
Daheng College, University of Chinese Academy of Sciences, Beijing 100039, China
Key Laboratory of Space-Based Dynamic and Rapid Optical Imaging Technology, Chinese Academy of Sciences, Changchun 130033, China

Corresponding author: Shaoju Wang (wangshaoju@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61805244, in part by the Key Science and Technology Research and Development Projects of Jilin Province under Grant 20190303094SF, and in part by the Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology (CAST) under Grant GZZKFJJ2020003.

ABSTRACT Compressive sensing (CS) technology is introduced into space optical remote sensing image acquisition stage, which could make wireless image sensor network node quickly and accurately obtain images in the case of two constraints of limited battery power and expensive sensor costs. On this basis, in order to further improve the quality of CS image reconstruction, we propose fused features and perceptual loss encoder-decoder residual network (FFPL-EDRNet) for image reconstruction. FFPL-EDRNet consists of a convolution layer and a reconstruction network. We train FFPL-EDRNet end-to-end, thus greatly simplifying the pre-processing and post-processing process and eliminating the block effect of reconstructed images. The reconstruction network is based on residual network, which introduces multi-scale feature extraction, multi-scale feature combination and multi-level feature combination. Feature fusion integrates low-level information with high-level information to reduce reconstruction error. The perceptual loss function based on pretrained InceptionV3 uses the weighted mean square error to define the loss value between the reconstructed image feature and the label image feature, which makes the reconstructed image more semantically similar to label image. In the measurement procedure, we use convolution to achieve block compression measurement, so as to obtain full image measurements. For image reconstruction, we firstly use a deconvolution layer to initially reconstruct the image and then use the residual network to refine the initial reconstructed image. The experimental results show that: in the case of measurement rates (MRs) of 0.25, 0.10, 0.04 and 0.01, the peak signal-to-noise ratio (PSNR) = 27.502, 26.804, 24.593, 21.359 and structural similarity (SSIM) = 0.842, 0.816, 0.720, 0.568 of the reconstructed images obtained by FFPL-EDRNet. Therefore, Our FFPL-EDRNet could enhance the quality of image reconstruction.

INDEX TERMS Image reconstruction, compressive sensing, encoder-decoder network, fused features, perceptual loss, residual block.

I. INTRODUCTION

Space optical remote sensing (SORS) technology has become an indispensable part of obtaining intelligence information. The results of SORS image processing and analysis could be applied to many fields, such as environmental monitoring, disaster prevention and mitigation, and surveying and mapping [1]–[4]. However, the prerequisite for image processing and analysis is the accuracy of image acquisition. The work of acquiring SORS images is completed by wireless image sensor network node, whose main function depends on battery power supply and sensor image sampling. At present, the two main constraints that restrict image acquisition work are limited battery power and expensive sensor costs. Therefore, under the condition of limited battery power and expensive sensor costs, it is an urgent problem to ensure the quality

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

and quantity of image acquisition. The algorithm designed by compressive sensing (CS) is an effective method to satisfy these constraints to a large extent.

CS is an emerging technology. It firstly completes the sample value compression while the signal is being sampled, and then carries on the effective recovery of the signal through the reconstruction algorithm [5], [6]. Compared with the traditional sampling and compression techniques, the image acquisition method based on CS has the advantages of simple coding and good compression performance [7]. On the one hand, since the sample value compression is achieved at the same time as the SORS image is sampled, the number of sensors used to obtain SORS images of the same width is greatly reduced, thereby reducing sensor costs. On the other hand, since compressed sampling values are used for data transmission, the energy consumption of data transmission is reduced, thereby reducing battery energy supply.

Research on high performance CS image reconstruction methods is a hot topic in CS field. Traditional methods [8]–[13] to solve CS reconstruction problem are mostly based on physical-driven approaches, such as convex optimization [8], [11], greedy algorithm [12], [13], and non-convex algorithm [9], [10]. However, these methods commonly adopt iterative optimization strategies to solve the problem of image signal reconstruction. This iterative optimization algorithm is computationally expensive and time-consuming.

Deep learning (DL) has been widely used in the image field, such as image classification [14], target detection [15], semantic segmentation [16], image denoising [17] and image super-resolution [18]. Recently, there have been a lot of researches on CS image reconstruction network based on DL. Thanks to the powerful self-learning capacity, deep neural networks effectively avoid a large amount of calculation in traditional iterative methods and achieves excellent reconstruction performance. Mousavi *et al.* [19] used SDA network to reconstruct compressed sampling measurements, which is the first time that DL was used to solve the problem of compressive reconstruction. Kulkarni *et al.* [20] proposed ReconNet based on CNN, and realized the reconstruction of compressed sensor images with a non-iterative strategy. Lohit *et al.* [21] added an adversarial network on the basis of ReconNet, and used the generative confrontation network to further improve the quality of reconstruction. Yao *et al.* [22] proposed deep residual reconstruction network (DR^2-Net) for image CS. DR^2-Net replaced ReconNet with a residual network and improved the reconstruction performance by adding a residual network. The common disadvantage of the above methods is that the independent reconstruction of image blocks leads to block effect. In order to solve this problem, Du *et al.* [23] used a fully convolutional measurement network (FCMN) to reconstruct the image. The convolutional network reconstructed the entire image measurements, thereby eliminating the block effect. FCMN is firstly introduced to remove block effect. Mousavi and Baraniuk [24] used the network to reconstruct CS measurements of the entire image, and at the same time forced the weight of fully connected layer to be set to Φ^T. It eliminated the block effect of the reconstructed image; Du *et al.* [25] proposed using deconvolution layer (DCL) to initially reconstruct the CS of the entire image, which also eliminated the block effect of the reconstructed image; Mousavi *et al.* [26] and Zhao *et al.* [27] both trained the measurement layer and reconstruction network together, describe a novel encoder-decoder network (EDNet) in order to eliminate blocking effect. Since the above reconstruction networks simply extract features without feature fusion, and their loss function only focuses on the difference between the corresponding pixels of the reconstructed image and label image, the quality of the reconstructed image could be further improved.

In summary, we investigate the SORS image acquisition method based on CS. This paper proposes FFPL-EDRNet for SORS image CS, which further enhances the reconstruction quality. Therefore, it will lay a reliable foundation for the following image processing and analysis. An intuitive description of FFPL-EDRNet is illustrated in Figure 1. The FFPL-EDRNet consists of two components: measurement layer and reconstruction network. Our main contributions are as follows:

1) In order to further improve the quality of CS image reconstruction, we propose a novel encoder-decoder residual network for image reconstruction, called FFPL-EDRNet. For CS image reconstruction of SORS images, this model is better than existing algorithms.

2) In order to simplify the image pre-processing and post-processing process and eliminate the block effect of reconstructed images, FFPL-EDRNet connects the measurement layer and the reconstruction network for end-to-end training. We also set the step size to be equal to the size of the convolution kernel in the convolution operation in measurement part, which makes the convolution operation simulate the block compression sampling process in CS.

3) In order to accurately extract the feature information of different scales in the initial reconstructed image, and make full use of the detailed feature information of the low-level and the overall feature information of the high-level, the CanNet unit adopts multi-scale feature extraction, multi-scale feature combination and multi-level feature combination. Feature fusion integrates low-level information with high-level information to reduce reconstruction errors, thereby improving the quality of image reconstruction.

4) In order to make the reconstructed image more semantically similar to label image, FFPL-EDRNet uses the perceptual loss function to improve reconstructed image quality. The loss function based on pretrained InceptionV3 compares the feature of reconstructed image with the feature of label image, so that the reconstructed image is more similar to the label image in semantics.

II. FFPL-EDRNet

The purpose of FFPL-EDRNet is to improve the quality of SORS image reconstruction in CS. The overall structure of

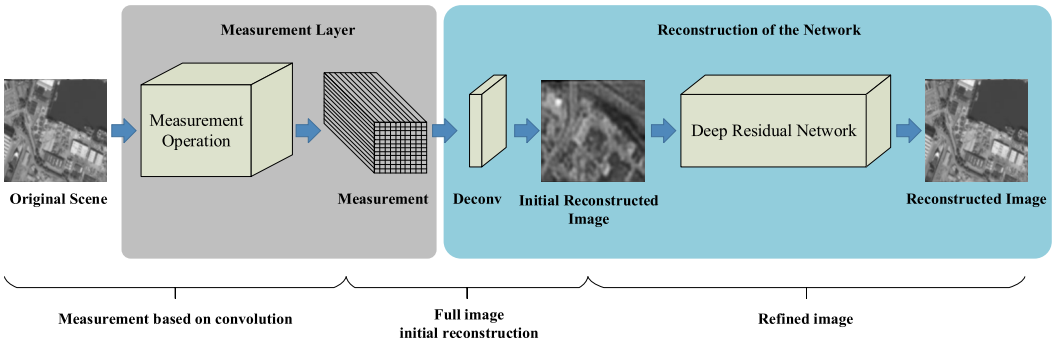

FIGURE 1. An intuitive description of FFPL-EDRNet.

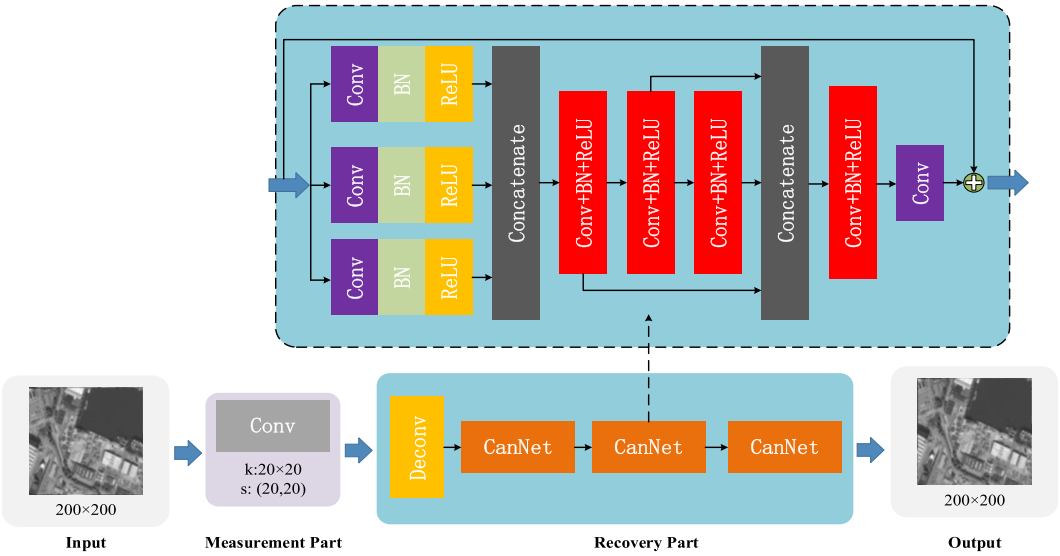

FIGURE 2. The overall structure of FFPL-EDRNet.

FFPL-EDRNet is shown in Figure 2. In the measurement part, the measurement layer simulates the compressive sampling process of the SORS camera (regardless of noise). Compared to the predefined measurement matrix, this method is just as easy to implement on the hardware equipment of optical system, and the image information will not also be missed during the measurement process. In the recovery part, the deconvolution layer and residual network could accurately reconstruct the compression measurements.

A. BLOCK MEASUREMENT BASED ON CONVOLUTION

The measurement layer $CN(\cdot)$ uses the non-overlapping sliding window convolution method to measure the image. We set the step size to be equal to the size of the convolution kernel, which is equivalent to block compression measurement. And the trained convolution weight is the measurement matrix. Particularly, MRs determines the number of convolution kernels. The number of convolution kernels could be determined with.

$$\boldsymbol{num}_{kernal} = \boldsymbol{MRs} \times \boldsymbol{n} \tag{1}$$

where n is the number of elements in the image block, which is also the number of elements in the convolution kernel.

The measurement process is shown in Figure 3. The measurement layer $CN(\cdot)$ is a convolution layer with a 20×20 convolution kernel, a step size of 20, and no bias value. Therefore, the value of n in Equation (1) is 400.

It takes the 200×200 full image as input and outputs $10 \times 10\times$ $\text{num}_{\text{kernal}}$ full image measurements, where $\text{num}_{\text{kernal}}$ is related to MRs, e.g., $\text{num}_{\text{kernal}} = 100, 40, 16,$ and 4 corresponding to MRs $= 0.25, 0.10, 0.04,$ and 0.01 [21], respectively. We denote the procedure of measurement as,

$$Y = CN(X, W) = W \times X \tag{2}$$

where Y is the measurements of full image. X is the original image. W is the weight matrix of the convolutional layer, which is also the measurement matrix in CS.

We connect the measurement layer $CN(\cdot)$ with the reconstruction network for end-to-end training, which makes the whole process of CS realized through the encoder-decoder network, thus greatly simplifying the process of image pre-processing and post-processing. Moreover, the reconstruction network reconstructs the complete image through the full image measurements obtained from the measurement layer, thereby inhibiting the block effect.

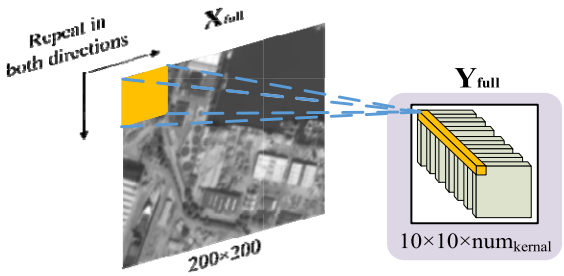

FIGURE 3. Block measurement process based on convolution.

B. FULL IMAGE INITIAL RECONSTRUCTION BASED ON DECONSTRUCTION

CS image reconstruction is to recover the image X from CS measurements Y. The reconstruction process from Y to X could be regarded as a mapping relationship, given the model:

$$X = QY \tag{3}$$

where Q is the reconstruction mapping matrix. Since Equation (3) is an over-determined equation, it has no exact solution. However, we could estimate a reconstruction mapping matrix Q^z, which minimizes the error between X and Q^zY.

We employ the deconvolution layer $F^f(\cdot)$ to infer the optimal mapping matrix Q^z. The process of initial image reconstruction is shown in Figure 4. The deconvolution layer $F^f(\cdot)$ is a deconvolution layer with a 20 × 20 convolution kernel, a step size of 20, and no bias value. It takes the 10 × 10 × num$_{kernal}$ full image measurements as input and outputs 200 × 200 initial reconstructed image.

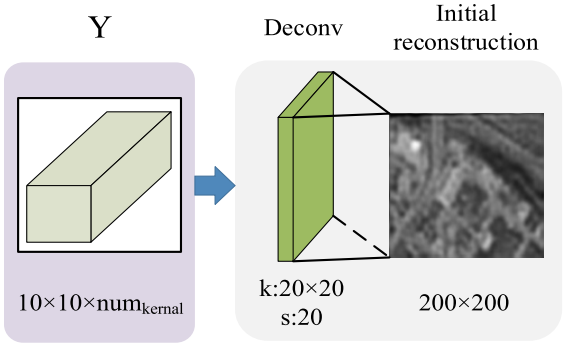

FIGURE 4. Full image initial reconstruction process based on deconvolution.

The trained deconvolution layer $F^f(\cdot)$ maps the full image Y into an initial reconstructed image $\hat{X}$. The process of generating the initial reconstructed image could be expressed as:

$$\hat{X} = F^f(Y, Q^f) = Q^f \times Y \tag{4}$$

where Q^f is the parameters for the trained deconvolution layer.

C. REFINEMENT BASED ON DEEP RESIDUAL NETWORK

The trained deconvolution layer could only obtain an approximate solution to the image X. In order to further narrow the difference between $\hat{X}$ and X, we introduce a deep residual network to predict the residual between two images.

Compared with the traditional deep neural network, such as AlexNet [28], VGG [29], and GoogleNet [30], the deep residual network [31] solves the shattering gradient problem. Therefore, for the refinement of the initial image, we use it as the main network to extract features.

In order to accurately extract the feature information of different scales in the initial reconstruction image $\hat{X}$, the residual network uses the convolution kernel of different scales to extract the multi-scale feature information of the $\hat{X}$ and fuse these features. In order to make full use of the detailed feature information of the low-level and the overall feature information of the high-level, the residual network integrates the features of different levels of low, medium and high.

As shown in Figure 2, deep residual network consists of three CanNet units. Each CanNet unit adopts multi-scale feature extraction, multi-scale feature combination and multi-level feature combination. In order to ensure that the generated feature map size is 200 × 200, we add corresponding padding to each layer of the residual network. We further add batch normalization [32] and ReLU layer to each convolution layer of each unit except the last convolution layer. The structure and paraments of CanNet unit in Table 1, where CONV stands for convolution layer, BN stands for batch standardization, and ReLU stands for activation function. The C1 layer is three convolutional layers connected in parallel. The sizes of the convolution kernels of the three convolution layers are 3 × 3, 5 × 5, and 7 × 7, respectively, and the numbers of the convolution kernels of the three convolution layers are 64, 64, and 64, respectively. The C2-C5 layers have the same structure, and they are all convolutional layers with a convolution kernel size of 3 × 3 and a convolution kernel number of 64. The C6 layer generates 1 feature map with 3×3 kernel. The output of C6 layer in the third CanNet unit is the output of the reconstruction network.

TABLE 1. Structure and paraments of CanNet unit.

Layer	Layer Structure	Kernel Number	Kernel Size
C1	CONV+BN+ReLU	64	3×3
C1	CONV+BN+ReLU	64	5×5
C1	CONV+BN+ReLU	64	7×7
C2-C5	CONV+BN+ReLU	64	3×3
C6	CONV	1	3×3

Specifically, the deep residual network $F^r(\cdot)$ uses $\hat{X}$ as input to generate a predicted residual $\hat{D}$ The process of generating the initial reconstructed image could be expressed as,

$$\hat{D} = F^r(\hat{X}, Q^r) = Q^r \times \hat{X} \tag{5}$$

where Q^r is the parameters of deep residual network.

The reconstruction part takes CS measurement Y as input, firstly obtains initial reconstructed image $\hat{X}$, and then adds it to the predicted residual $\hat{D}$ to get the final reconstructed

image $\hat{\hat{X}}$, *i.e.*,

$$\hat{\hat{X}} = \hat{X} + \hat{D} \tag{6}$$

We use Equation (2), Equation (4) and Equation (5) to replace the Y, $\hat{X}$ and $\hat{D}$. Therefore, the final reconstructed image $\hat{\hat{X}}$ of FFPL-EDRNet could be obtained, which could be expressed as:

$$\hat{\hat{X}} = \boldsymbol{F}^{f}(\boldsymbol{CN}(\boldsymbol{X}, \boldsymbol{W}), \boldsymbol{Q}^{f}) + \boldsymbol{F}^{r}(\boldsymbol{F}^{f}(\boldsymbol{CN}(\boldsymbol{X}, \boldsymbol{W}), \boldsymbol{Q}^{f}), \boldsymbol{Q}^{r}) \tag{7}$$

D. FFPL-EDRNet LOSS FUNCTION

The loss function of the traditional CS image reconstruction methods is Pixel-wise loss function, such as mean square error (MSE) and mean absolute error (MAE). Since Pixel-wise loss function only pays attention to the difference between the corresponding pixels of the reconstructed image and label image, it leads to blurred details of the reconstructed image. However, the perceptual loss function compares the feature of reconstructed image passing through CNN with the feature of label image passing through CNN, making the reconstructed image more semantically similar to label image.

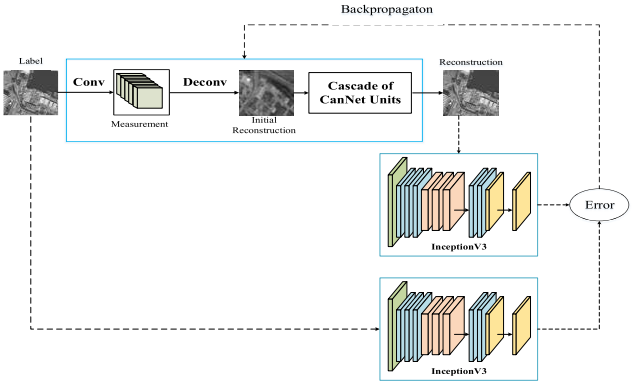

FIGURE 5. The error back propagation process of FFPL-EDRNet using the pretrained InceptionV3.

Since the pretrained InceptionV3 has excellent performance in feature extraction, we use it to define the loss function [25]. Figure 5 shows the error back propagation process of FFPL-EDRNet using the perceptual loss function based on pretrained InceptionV3. Firstly, the pretrained InceptionV3 extracts the features of the reconstructed image and label image. Secondly, we use the weighted mean square error (WMSE) to define the error between features. Finally, the best reconstructed image could be obtained by adjusting the network parameters according to the error.

The loss function could be determined with.

$$\boldsymbol{loss}^{\boldsymbol{InceptionV3}}(\hat{\hat{X}}, \boldsymbol{X}) = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{W}_j \times (\boldsymbol{Inception}_j(\hat{\hat{X}} - \boldsymbol{Inception}_j(\boldsymbol{X}))^2 \tag{8}$$

where $Inception_j(X)$ is the feature map of the j-th layer with the label image X, and $Inception_j(\hat{X})$ is the feature map of the j-th layer with the reconstructed image $\hat{X}$. Here j refers to conv2d_18, conv2d_25, conv2d_39, conv2d_49, conv2d_59, conv2d_69, conv2d_73, conv2d_83 and conv2d_93 layer. W_j is the weighted value of the above convolution layers. N is the number of convolution layers, here N refers to 9.

E. CS PROCEDURE

The flowchart of FFPL-EDRNet is shown in Figure 6. Given an image, We firstly use the measurement layer to directly perform non-overlapping sliding window convolution measurements on the image to obtain the full image measurements. Then, the deconvolution layer takes the CS measurements as input and outputs the initial reconstructed image, which is finally processed with residual network to refine the input image. We take the CS of SORS image as an example.

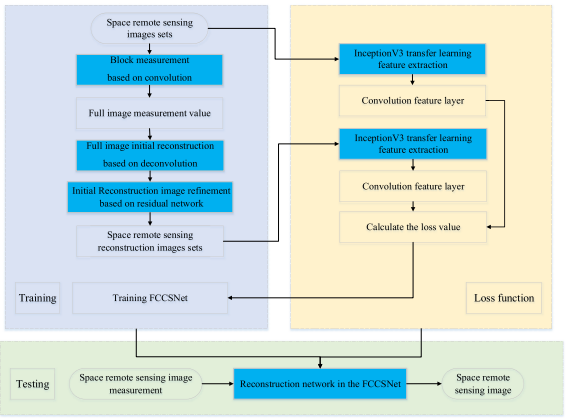

FIGURE 6. Flow chart of FFPL-EDRNet.

III. FFPL-EDRNet TRAINING

A. DATASET

The SpaceNet dataset [33] is a collection of SORS images provided by DigitalGlobe commercial satellite company. It contains some label information that could be used for machine learning research. Table 2 shows the number of samples in the training set, verification set and test set in this paper. We use 1500 images from SpaceNet dataset for training, 200 images from SpaceNet dataset for verifying and 400 images from SpaceNet dataset for testing. Specially, the validation set is run after each epoch to calculate the loss value of validation data during the training process. The number of iterations corresponding to the minimum loss value of the validation set is the optimal number of training of the network.

Figure 7 lists some of the dataset images used in this paper.

B. TRAINING ENVIRONMENT

DL requires a relatively high experimental environment, so the environment in Table 3 is adopted for FFPL-EDRNet

TABLE 2. Dataset partitioning.

Dataset	Images
Training set	1500
Validation set	200
Test set	400

FIGURE 7. Part of the SpaceNet SORS dataset.

TABLE 3. Experimental environment.

System	Windows10
RAM	32.0 GB
CPU	4.10GHz Intel processor
GPU	GeForce RTX 3070, memory 8 G
DL framework	Keras (backed as TensorFlow)

training. Specially, we use GeForce RTX 3070, which could greatly speed up training.

C. TRAINING PARAMETERS

In this paper, FFPL-EDRNet adopts the end-to-end training method, while the original image is used as both the input image and label image, *i.e.*, (X, X) as input-output pair for FFPL-EDRNet training. The training procedure of FFPL-EDRNet consists of measurement layer $CN(\cdot)$, deconvolution layer $F^f(\cdot)$ and deep residual network $F^r(\cdot)$, respectively. Table 4 shows the parameter setting for FFPL-EDRNet training. Specially, considering the constraints of GPU memory 8G, we set batch size to 5. Note that, the training process could obtain the optimal parameters W, Q^f and Q^r, respectively.

TABLE 4. Training parameters.

Convolution initialization	Gaussian distribution with a standard deviation of 0.001 [22]
Optimizer	Adadelta
Learn rate	0.0001
Batch size	5

IV. FFPL-EDRNet EXPERIMENTS

In this section, we conduct a series of experiments to test the performance of the proposed method.

A. EVALUATION METRICS

We utilize two objective quantitative indicators: PSNR [34] and SSIM [35] for evaluation. The calculation methods of PSNR and SSIM are shown in Equation (9) and (10) respectively.

$$\boldsymbol{PSNR}(\boldsymbol{X}, \hat{\hat{\boldsymbol{X}}}) = \mathbf{20} \times \boldsymbol{lg}\left(\frac{\mathbf{255}}{\sqrt{\boldsymbol{MSE}(\boldsymbol{X}, \hat{\hat{\boldsymbol{X}}})}}\right) \tag{9}$$

where X is the original image, $\hat{\hat{\boldsymbol{X}}}$ is the reconstruction image, $\boldsymbol{MSE}(\boldsymbol{X}, \hat{\hat{\boldsymbol{X}}})$ is the mean square error between the original image and the reconstruction image. PSNR could effectively reflect the similarity between the corresponding pixels of two images. The larger the value, the higher the similarity between the corresponding pixels of the image.

$$\boldsymbol{SSIM}(\boldsymbol{X}, \hat{\hat{\boldsymbol{X}}}) = \frac{(2\boldsymbol{\mu}_{\boldsymbol{X}}\boldsymbol{\mu}_{\hat{\hat{\boldsymbol{X}}}} + \boldsymbol{C}_1)(2\boldsymbol{\sigma}_{\boldsymbol{X}\hat{\hat{\boldsymbol{X}}}} + \boldsymbol{C}_2)}{(\boldsymbol{\mu}_{\boldsymbol{X}}^2 + \boldsymbol{\mu}_{\hat{\hat{\boldsymbol{X}}}}^2 + \boldsymbol{C}_1)(\boldsymbol{\sigma}_{\boldsymbol{X}}^2 + \boldsymbol{\sigma}_{\hat{\hat{\boldsymbol{X}}}}^2 + \boldsymbol{C}_2)} \tag{10}$$

where X is the original image, $\hat{\hat{\boldsymbol{X}}}$ is the reconstruction image, $\boldsymbol{\mu}_{\boldsymbol{X}}$ and $\boldsymbol{\mu}_{\hat{\hat{\boldsymbol{X}}}}$ are the mean values of the original image X and the reconstruction image $\hat{\hat{\boldsymbol{X}}}$, $\boldsymbol{\sigma}_{\boldsymbol{X}}$ and $\boldsymbol{\sigma}_{\hat{\hat{\boldsymbol{X}}}}$ are the variances of the original image X and the reconstruction image $\hat{\hat{\boldsymbol{X}}}$, $\boldsymbol{\sigma}_{\boldsymbol{X}\hat{\hat{\boldsymbol{X}}}}$ is the covariance of the original image X and the reconstruction image $\hat{\hat{\boldsymbol{X}}}$. SSIM could effectively reflect the structural similarity between two images. The value range is -1 to 1. The larger the value, the higher the structural similarity between the image.

B. COMPARATIVE EXPERIMENTS WITH OTHER STATE-OF-THE-ART IMAGE RECONSTRUCTUIN NETWORKS

We compare FFPL-EDRNet with other state-of-the-art image reconstruction networks on the dataset and experimental environment described in this paper. We compare FFPL-EDRNet with DR^2-Net [22], FCMN [23], EDNet [26], DCL+CanNet [25]. Table 5-8 shows the PSNR and SSIM value of six SORS reconstructed images in the test set from them. It could be seen from the mean in Table 5-8, Our FFPL-EDRNet outperforms them at high MRs. For instance, at MRs $= 0.25$, the FFPL-EDRNet outperforms DR^2-Net, FCMN, EDNet, and DCL+CanNet by 1.211 dB, 1.166 dB, 1.081dB and 0.918 dB respectively on PSNR, outperforms DR^2-Net, FCMN, EDNet, and DCL+CanNet by 0.021, 0.059, 0.02 and 0.052 respectively on SSIM. FFPL-EDRNet is also better than them at low MRs. For example, for the cases that MRs $= 0.04$ and 0.01, the PSNR and SSIM values of the FFPL-EDRNet are also higher than DR^2-Net, FCMN, EDNet, and DCL+CanNet. The appendix shows the reconstructed image diagrams of different CS image reconstruction algorithms.

V. FFPL-EDRNET DISCUSSION

A. LEARN OR PREDEFINED MEASUREMENT MATRIX

In this subsection, we compare the characteristics of the learned measurement matrix (LMM) and predefined measurement matrix (PMM) in the time domain and frequency domain. As shown in Figure 8, (a) is the time domain and frequency domain of LMM, (b) is the time domain and frequency domain of PMM. They were compared at four MRs.

Comparing the time domain of LMM and PMM, we observe that LMM has obvious structural information while PMM is irregular. Compared with the frequency

TABLE 5. PSNR (dB) and SSIM at MR 25%.

Test Images	ImageA	ImageB	ImageC	ImageD	ImageE	ImageF	Mean
DR^2-Net	24.299	27.437	25.864	24.038	25.292	22.769	24.938
	0.715	0.846	0.790	0.747	0.738	0.625	0.744
FCMN	24.427	27.388	25.768	23.891	25.266	23.160	24.983
	0.684	0.808	0.742	0.703	0.695	0.616	0.706
EDNet	24.509	27.607	25.923	23.887	25.401	23.081	25.068
	0.719	0.847	0.793	0.742	0.738	0.628	0.745
DCL + CanNet	24.542	27.507	25.962	24.349	25.564	23.465	25.231
	0.692	0.809	0.744	0.707	0.706	0.620	0.713
FFPL-EDRNet	26.940	29.510	28.422	26.952	27.848	25.341	**27.502**
	0.830	0.893	0.869	0.851	0.830	0.777	**0.842**

TABLE 6. PSNR (dB) and SSIM at MR 10%.

Test Images	ImageA	ImageB	ImageC	ImageD	ImageE	ImageF	Mean
DR^2-Net	23.099	26.262	24.375	22.326	23.884	21.531	23.580
	0.665	0.820	0.748	0.686	0.693	0.554	0.694
FCMN	23.545	26.396	24.744	22.731	24.359	22.314	24.045
	0.584	0.725	0.625	0.587	0.589	0.496	0.601
EDNet	23.408	26.095	24.313	22.310	24.012	21.815	23.659
	0.653	0.802	0.738	0.667	0.677	0.525	0.677
DCL + CanNet	23.643	26.507	24.980	23.030	24.575	22.680	24.236
	0.663	0.791	0.738	0.673	0.665	0.577	0.684
FFPL-EDRNet	26.099	29.029	27.660	26.252	27.099	24.682	**26.804**
	0.794	0.883	0.847	0.827	0.805	0.738	**0.816**

TABLE 7. PSNR (dB) and SSIM at MR 4%.

Test Images	ImageA	ImageB	ImageC	ImageD	ImageE	ImageF	Mean
DR^2-Net	21.125	23.726	22.008	19.324	21.321	19.686	21.198
	0.579	0.751	0.667	0.575	0.580	0.423	0.596
FCMN	22.519	25.078	23.438	21.282	22.981	21.111	22.735
	0.452	0.614	0.461	0.433	0.468	0.358	0.464
EDNet	23.036	25.793	24.003	21.911	23.525	21.580	23.308
	0.641	0.799	0.731	0.655	0.662	0.511	0.659
DCL + CanNet	22.867	25.588	23.972	21.958	23.717	21.754	23.309
	0.617	0.756	0.693	0.616	0.619	0.483	0.631
FFPL-EDRNet	24.124	27.168	25.426	23.489	24.874	22.476	**24.593**
	0.692	0.832	0.772	0.723	0.717	0.583	**0.720**

TABLE 8. PSNR (dB) and SSIM at MR 1%.

Test Images	ImageA	ImageB	ImageC	ImageD	ImageE	ImageF	Mean
DR^2-Net	18.302	20.442	19.106	17.036	19.218	17.848	18.659
	0.454	0.667	0.544	0.456	0.488	0.328	0.489
FCMN	20.278	22.768	20.909	18.039	20.457	19.297	20.336
	0.469	0.685	0.565	0.458	0.504	0.315	0.499
EDNet	21.234	23.553	21.599	18.965	21.372	19.768	21.082
	0.544	0.739	0.637	0.525	0.567	0.372	0.564
DCL + CanNet	20.886	23.070	21.241	18.835	21.277	19.559	20.811
	0.498	0.685	0.572	0.476	0.521	0.318	0.512
FFPL-EDRNet	21.522	23.934	21.821	19.278	21.701	19.893	**21.359**
	0.547	0.741	0.639	0.527	0.579	0.374	**0.568**

domain, the information captured by the LMM is more concentrated in the low frequency, which means that the LMM could capture more energy from the original image. According to the time-domain and frequency-domain diagrams, LMM could improve the quality of the later reconstructed image.

B. ABLATION STUDIES BETWEEN DEEP RESIDUAL NETWORK AND LOSS FUNCTION

In this subsection, we conduct ablation studies to confirm the performance of deep residual network and loss function. We define the network connecting the measurement layer $CN(\cdot)$ and the deconvolution layer $F^f(\cdot)$ as baseline

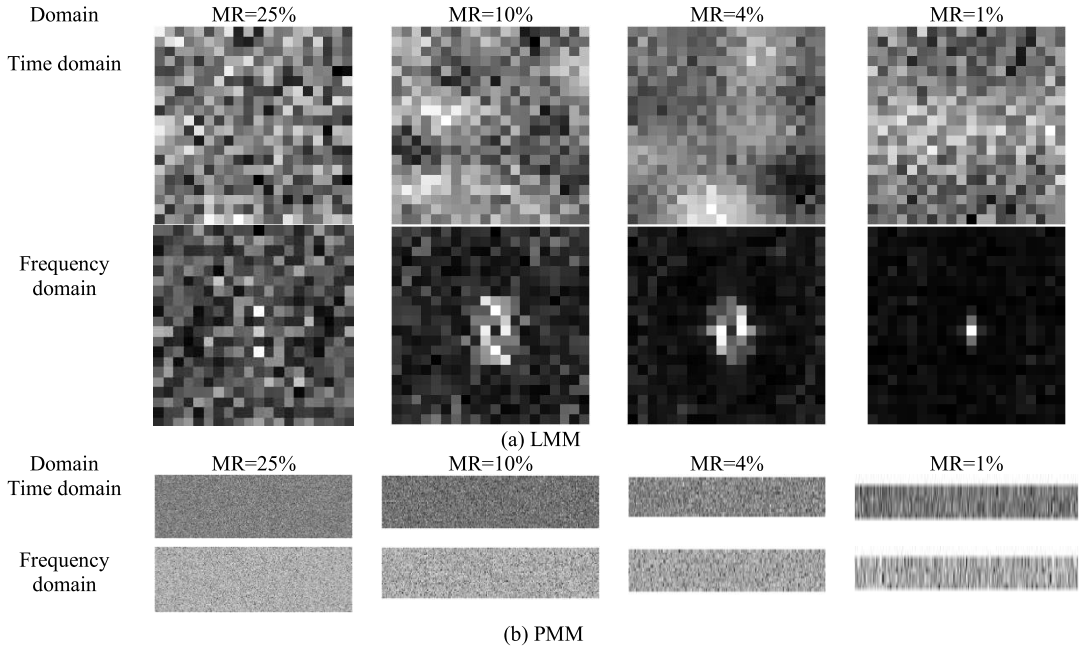

FIGURE 8. The time domain and frequency domain of the measurement matrix.

TABLE 9. Average PSNR (dB) and average SSIM OF THE reconstructed images in the test set.

network models	MRs= 25%	MRs= 10%	MRs= 4%	MRs= 1%
baseline network	19.725	19.436	18.139	15.845
+MSE loss function	0.493	0.460	0.403	0.296
baseline network	20.878	20.514	19.248	17.039
+pretrained InceptionV3 loss function	0.569	0.532	0.465	0.358
baseline-deep residual network	26.149	25.494	23.814	20.752
+MSE loss function	0.765	0.744	0.695	0.542
baseline-deep residual network	**27.502**	**26.804**	**24.593**	**21.359**
+pretrained InceptionV3 loss function	**0.842**	**0.816**	**0.720**	**0.568**

network. We define the network connecting baseline network and deep residual network as baseline-deep residual network.

We build four types of CS image reconstruction networks, which are: baseline network that uses MSE as loss function, baseline network that uses pretrained InceptionV3 as loss function, baseline-deep residual network that uses MSE as loss function and baseline-deep residual network that uses pretrained InceptionV3 as loss function.

Designed to ensure the accuracy of results, we take the average PSNR and average SSIM of the reconstructed image in the test set. Table 9 shows the average PSNR and average SSIM for the four network models at four MRs.

We firstly analyze the impact of pretrained InceptionV3 loss function on the performance of CS image reconstruction network. It could be seen from the results in Table 9, in the case of MRs of 0.25, 0.10, 0.04 and 0.01, the baseline network using pretrained InceptionV3 as the loss function outperforms baseline network using MSE as the loss function by 1.153 dB, 1.078 dB, 1.109 and 1.194 dB respectively on PSNR, outperforms baseline network using MSE as the loss function by 0.076, 0.072, 0.060 and 0.062 respectively on SSIM. This shows that the pretrained InceptionV3 loss function could improve the performance of the CS image reconstruction network.

Then, we analyze the impact of deep residual network on the performance of CS image reconstruction network. It could be seen from the results in Table 9, in the case of MRs of 0.25, 0.10, 0.04 and 0.01, the baseline-deep residual network using MSE as loss function outperforms baseline using MSE as loss function by 6.424 dB, 6.058 dB, 5.675 and 4.907 dB respectively on PSNR, outperforms baseline using MSE as loss function by 0.272, 0.284, 0.292 and 0.246 respectively on SSIM. This shows that the deep residual network could also improve the performance of CS image reconstruction network. And compared with pretrained InceptionV3 loss function, deep residual network improves CS image reconstruction network even more. However, it is not necessary to add too many CanNet units into the FFPL-EDRNet. Therefore, we add three CanNet units in FFPL-EDRNet. *i.e.*, the deep residual network consists of three CanNet units.

Finally, we analyze the impact of the combination of deep residual network and pretrained InceptionV3 loss function on

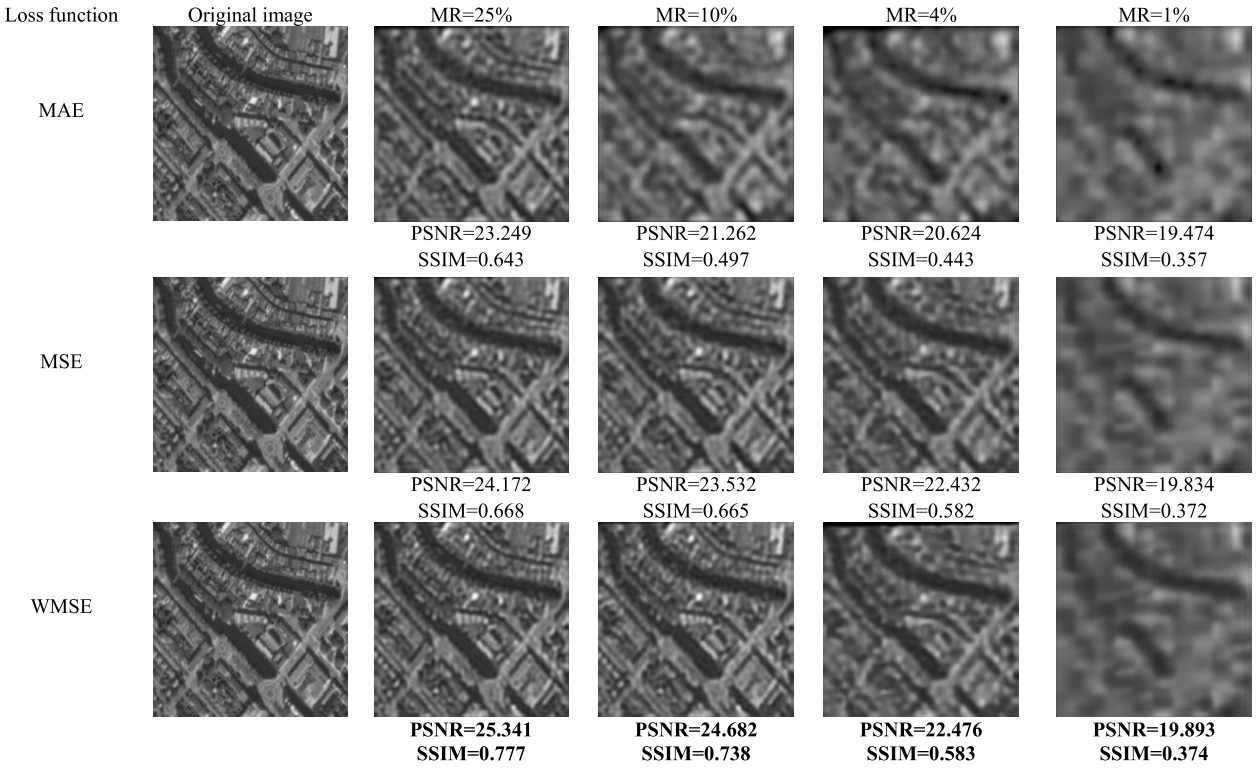

FIGURE 9. Comparison of loss function of MSE, MAE and WMSE to FFPL-EDRNet.

TABLE 10. Time complexity (in seconds) for reconstructing a single 200 × 200 image on GPU (GTX3070).

Algorithm	MR=25%	MR=10%	MR=4%	MR=1%
DR^2-Net	0.4277	0.4177	0.4314	0.4207
FCMN	0.5741	0.5520	0.5392	0.5319
EDNet	0.4744	0.4191	0.4095	0.3999
DCL + CanNet	0.5748	0.5110	0.5232	0.4896
baseline network	0.0575	0.0490	0.0504	0.0468
baseline network- 1 CanNet	0.4061	0.3854	0.3928	0.3773
baseline network- 2 CanNet	0.5045	0.4647	0.4765	0.4685
baseline network- 3 CanNet (FFPL-EDRNet)	**0.6019**	**0.5446**	**0.5565**	**0.5433**

the performance of CS image reconstruction network. It could be seen from the results in Table 9, in the case of MRs of 0.25, 0.10, 0.04 and 0.01, the baseline-deep residual network using pretrained InceptionV3 as loss function outperforms the other three CS image reconstruction networks. This shows that the combination of deep residual network and pretrained InceptionV3 loss function could greatly improve the performance of CS image reconstruction network.

C. PERCEPTUAL LOSS FUNCTION OR PIXEL-WISE LOSS FUNCTION

In this subsection, we compare the effects of Perceptual loss function and Pixel-wise loss function on the quality of reconstruction. It could be seen from the reconstructed images in Figure 9. The reconstructed images of FFPL-EDRNet have no block effect.

As shown in Figure 9, MSE and MAE using the Pixel-wise loss function will cause the reconstructed image to be smoother, *i.e.*, the details are lost. However, the reconstructed image with perceptual loss function solves this problem, and makes the reconstructed image more semantically similar to label image.

D. TIME COMPLEXITY

As the time complexity is an important factor in image reconstruction, we analyze the time complexity of FFPL-EDRNet. We compare it with DR^2-Net, FCMN, EDNet, DCL+CanNet, baseline network, baseline network-1CanNet and baseline network-2CanNet. The experimental results are summarized in Table 10 and Table 11.

Firstly, we show the time complexity of FFPL-EDRNet with different structures. From Table 10, we could observe that the time complexity has a linear relationship with the depth of network, *i.e.*, deeper FFPL-EDRNet requires more running time. Among the four networks, baseline network contains only two convolutional layers, thus shows the fastest speed. The baseline network-3CanNet (FFPL-EDRNet) contains 20 convolutional layers, and has the longest running

TABLE 11. Average PSNR (dB) and average SSIM of the reconstructed images in the test set.

Algorithm	MR=25%	MR=10%	MR=4%	MR=1%
DR^2-Net	24.938	23.580	21.198	18.659
	0.744	0.694	0.596	0.489
FCMN	24.983	24.045	22.735	20.336
	0.706	0.601	0.464	0.499
EDNet	25.068	23.659	23.308	21.082
	0.745	0.677	0.659	0.564
DCL + CanNet	25.231	24.236	23.309	20.811
	0.713	0.684	0.631	0.512
baseline network	20.878	20.514	19.248	17.039
	0.569	0.532	0.465	0.358
baseline network- 1 CanNet	27.235	24.545	24.335	21.136
	0.833	0.804	0.703	0.550
baseline network- 2 CanNet	27.464	24.735	24.528	21.316
	0.839	0.811	0.711	0.559
baseline network- 3 CanNet (FFPL-EDRNet)	**27.502**	**26.804**	**24.593**	**21.359**
	0.842	**0.816**	**0.720**	**0.568**

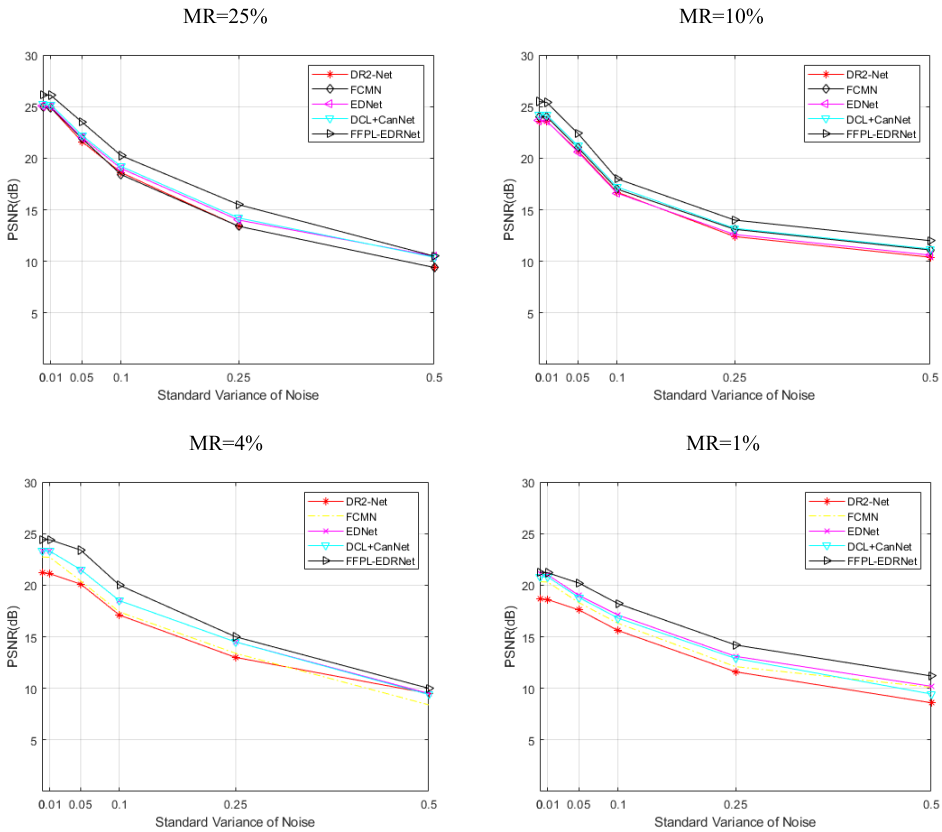

FIGURE 10. Comparison of robustness of DR^2-Net, FCMN, EDNet, DCL+CanNet and FFPL-EDRNet to Gaussian noise.

time. However, the increase in reconstruct time is at the millisecond level, which does not affect the real-time performance of image reconstruction.

Then, we compare them with DR^2-Net, FCMN, EDNet and DCL+CanNet. Compared with DR^2-Net, FCMN, EDNet and DCL+CanNet, the baseline network spends the least time to complete a basic image reconstruction. Therefore, when running time is a very important consideration, Fast image reconstruction using the baseline network is a good choice. From Tables 10-11, we could observe that FFPL-EDRNet with structure baseline network-1CanNet not only runs the fastest but also has the highest image reconstruction quality than DR^2-Net, FCMN, EDNet and DCL+CanNet.

E. ROBUST TO NOISE

In order to show the robustness of FFPL-EDRNet to noise, we study the effect of image reconstruction in the presence of measurement noise. We firstly add standard Gaussian noise to the CS measurement of the test set, $\sigma = 0.01$, 0.05, 0.1, 0.25, and 0.5 [22], which σ is the standard deviation of

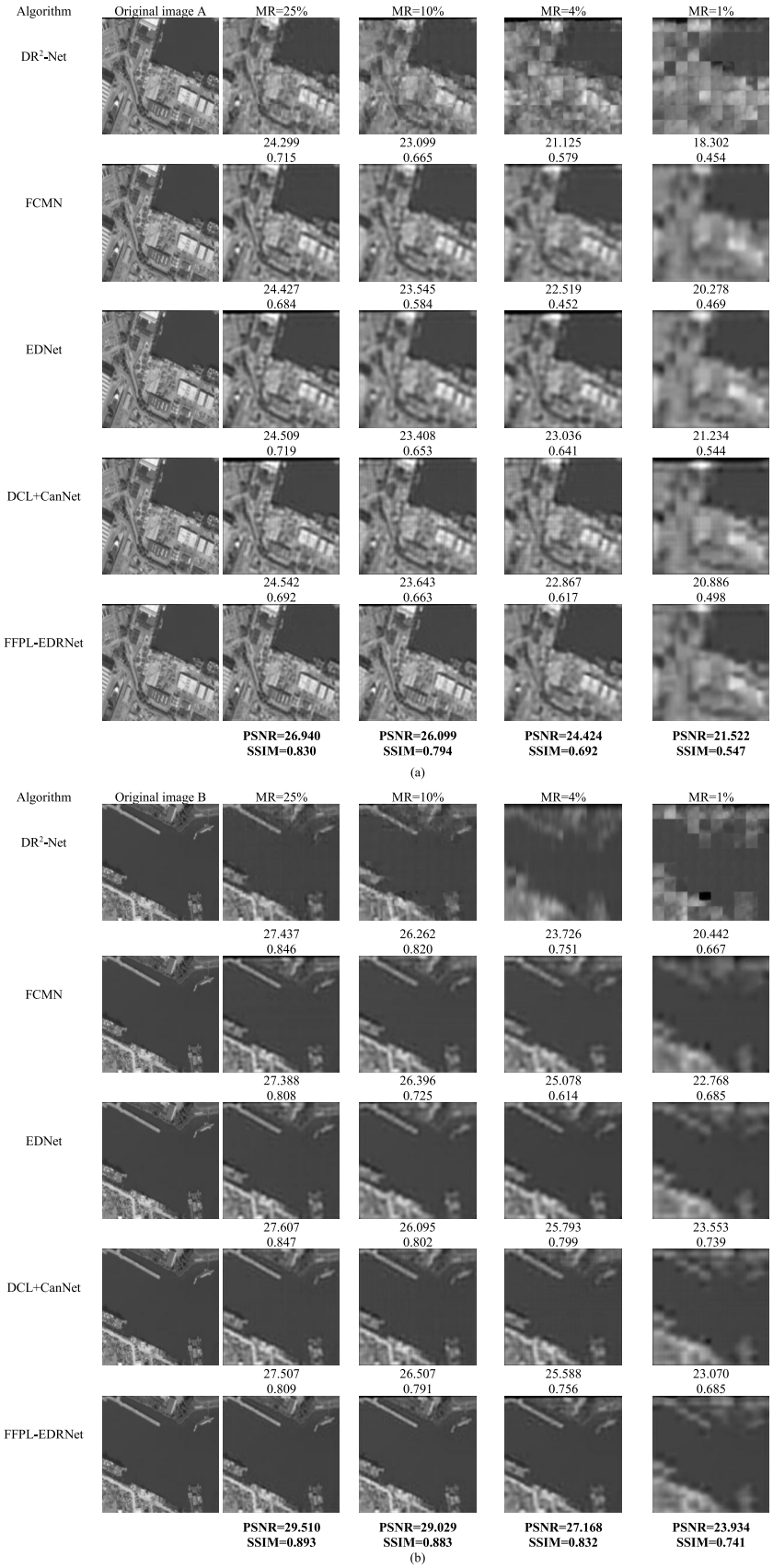

FIGURE 11. SORS images reconstruction results from DR^2 -Net, FCMN, EDNet, DCL+CanNet and FFPL-EDRNet: (a) Test image A, (b) Test image B.

Gaussian noise. Secondly, DR^2-Net, FCMN, EDNet, DCL+CanNet and FFPL-EDRNet trained under noiseless CS measurement takes the noisy CS measurement as the input and finally outputs the reconstructed image.

Designed to ensure the accuracy of results, we take the average PSNR of the reconstructed image in the test set. The results are summarized in Figure 10.

According to the comparative results in Figure 10, the quality of the reconstructed image decreases with the introduction of noise. However, the FFPL-EDRNet outperforms DR^2-Net, FCMN, EDNet, and DCL+CanNet for the $\sigma = 0.01$, 0.05, 0.1,0.25, 0.5 at four MRs.

From section IV, we get that the image reconstruction performance of FFPL-EDRNet is better than other image reconstruction algorithms. In this section, we have discussed five aspects of FFPL-EDRNet, these five aspects are learn or predefined measurement matrix, ablation studies between deep residual network and loss function, perceptual loss function or pixel-wise loss function, time complexity, and robust to noise. Section VA proves that LMM could improve the quality of the later reconstructed image. Section VB proves that both deep residual network and pretrained InceptionV3 loss function could improve the performance of CS image reconstruction network, it is worth noting that the deep residual network improves the reconstruction performance more. Section VC proves that the pretrained InceptionV3 loss function is better than the Pixel-wise loss function in the improvement of reconstruction performance. Section VD proves FFPL-EDRNet with structure baseline network-1CanNet not only runs the fastest but also has the highest image reconstruction quality than DR^2-Net, FCMN, EDNet and DCL+CanNet. Section VE proves the quality of the reconstructed image decreases with the introduction of noise, but the FFPL-EDRNet outperforms DR^2-Net, FCMN, EDNet, and DCL+CanNet under the influence of noise.

VI. CONCLUSION

After the introduction of CS technology in the data acquisition phase of SORS images, in order to further improve the quality of SORS image reconstruction, we propose FFPL-EDRNet. The network simulates the image acquisition process of the SORS camera based on CS, and further enhances the reconstruction quality. The structure of FFPL-EDRNet is composed of a convolution measurement layer and reconstruction network. In the measurement procedure, the learned measurement layer directly performs non-overlapping sliding window convolution measurement on the image to obtain the full image measurements. For the measurement reconstruction, the reconstruction network firstly uses the trained deconvolution layer to obtain the initial reconstructed image and then trained deep residual networks refine the initial result by predicting the residual between the initial reconstruction and label image. Experiment results shown that in the case of MRs of 0.25, 0.10, 0.04 and 0.01, our FFPL-EDRNet has excellent performance compared with state-of-the-art reconstruction network, with PSNR = 27.502dB, 26.804 dB, 24.593 dB, 21.359 dB and SSIM = 0.842, 0.816, 0.720, 0.568. Moreover, there is no block effect in the reconstructed images. Therefore, Our FFPL-EDRNet could enhance the quality of image reconstruction.

APPENDIX A
RECONSTRUCTED IMAGES DISPLAY OF FIVE CS IMAGE RECONSTRUCTION ALGORITHMS

In this section, we show the reconstruction results of two images randomly selected from the test set under five different image reconstruction algorithms. Figure 11(a) shows the reconstruction effect of test image A under five image reconstruction algorithms. Figure 11(b) shows the reconstruction effect of test image B under five image reconstruction algorithms.

It could be seen from the reconstructed images in Figure 11. The reconstructed images of FFPL-EDRNet have no block effect. Moreover, in the case of MRs of 0.25, 0.10, 0.04 and 0.01, our FFPL-EDRNet has the best performance compared with DR^2-Net, FCMN, EDNet and DCL+CanNet.

ACKNOWLEDGMENT

The authors thank the Key Laboratory of Space-Based Dynamic &Rapid Optical Imaging Technology, Chinese Academy of Sciences, the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences.

REFERENCES

[1] D. Li, M. Wang, and J. Jiang, "China's high-resolution optical remote sensing satellites and their mapping applications," *Geo-Spatial Inf. Sci.*, vol. 24, no. 1, pp. 85–94, 2020.
[2] E. Lutz, H. Bischl, and H. Ernst, "Development and future applications of satellite communications," in *Proc. 15th IEEE Int. Symp. Pers., Indoor, Mobile, Radio Commun. (PIMRC)*, Barcelona, Spain, 2004.
[3] Y. Xue, Y. Li, J. Guang, X. Zhang, and J. Guo, "Small satellite remote sensing and applications–history, current and future," *Int. J. Remote Sens.*, vol. 29, no. 15, pp. 4339–4372, 2008.
[4] P. Zhang, "The chinese meteorological satellite and applications," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 5516–5517.
[5] 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.
[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[7] R. Nan, G. Sun, Z. Wang, and X. Ren, "Research on image reconstruction of compressed sensing based on a multi-feature residual network," *Sensors*, vol. 20, no. 15, p. 4202, Jul. 2020.
[8] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[9] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.
[10] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.
[11] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[12] T. Jiang, X. Zhang, and Y. Li, "Bayesian compressive sensing using reweighted laplace priors," *AEU-Int. J. Electron. Commun.*, vol. 97, pp. 178–184, Dec. 2018.
[13] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
[14] K. Joshi, V. Tripathi, C. Bose, and C. Bhardwaj, "Robust sports image classification using InceptionV3 and neural networks," in *Proc. Int. Conf. Comput. Intell. Data Sci.* (Procedia Computer Science), V. Singh, V. K. Asari, and K. C. Li, Eds. Amsterdam, The Netherlands: Elsevier, 2020, pp. 2374–2381.
[15] Q. Wang, F. Shen, L. Cheng, J. Jiang, G. He, W. Sheng, N. Jing, and Z. Mao, "Ship detection based on fused features and rebuilt YOLOv3 networks in optical remote-sensing images," *Int. J. Remote Sens.*, vol. 42, no. 2, pp. 520–536, Jan. 2021.
[16] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New York, NY, USA, Jun. 2015, pp. 3431–3440.
[17] Y. Quan, Y. Chen, Y. Shao, H. Teng, Y. Xu, and H. Ji, "Image denoising using complex-valued deep CNN," *Pattern Recognit.*, vol. 111, Mar. 2021, Art. no. 107639.
[18] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
[19] A. Mousavi, A. B. Patel, and R. G. Baraniuk, "A deep learning approach to structured signal recovery," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, New York, NY, USA, Sep. 2015, pp. 1336–1343.
[20] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "ReconNet: Non-iterative reconstruction of images from compressively sensed measurements," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New York, NY, USA, Jun. 2016, pp. 449–458.
[21] S. Lohit, K. Kulkarni, R. Kerviche, P. Turaga, and A. Ashok, "Convolutional neural networks for noniterative reconstruction of compressively sensed images," *IEEE Trans. Comput. Imag.*, vol. 4, no. 3, pp. 326–340, Sep. 2018.
[22] H. Yao, F. Dai, S. Zhang, Y. Zhang, Q. Tian, and C. Xu, "DR2-net: Deep residual reconstruction network for image compressive sensing," *Neurocomputing*, vol. 359, pp. 483–493, Sep. 2019.
[23] J. Du, X. Xie, C. Wang, G. Shi, X. Xu, and Y. Wang, "Fully convolutional measurement network for compressive sensing image reconstruction," *Neurocomputing*, vol. 328, pp. 105–112, Feb. 2019.
[24] A. Mousavi and R. G. Baraniuk, "Learning to invert: Signal recovery via deep convolutional networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 2272–2276.
[25] J. Du, X. Xie, C. Wang, and G. Shi, "Color image reconstruction with perceptual compressive sensing," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, New York, NY, USA, Aug. 2018, pp. 1512–1517.
[26] A. Mousavi, G. Dasarathy, and R. G. Baraniuk, "DeepCodec: Adaptive sensing and recovery via deep convolutional neural networks," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, p. 744.
[27] Z. Zhao, X. Xie, C. Wang, S. Mao, W. Liu, and G. Shi, "ROI-CSNet: Compressive sensing network for ROI-aware image recovery," *Signal Process., Image Commun.*, vol. 78, pp. 113–124, Oct. 2019.
[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[29] Y. Zhang, W. Chan, and N. Jaitly, "Very deep convolutional networks for end-to-end speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, New York, NY, USA, Mar. 2017, pp. 4845–4849.
[30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New York, NY, USA, Jun. 2015, pp. 1–9.
[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New York, NY, USA, Jun. 2016, pp. 770–778.
[32] Z. Fu, S. Li, X. Li, B. Dan, and X. Wang, "Influence of batch normalization on convolutional neural networks in HRRP target recognition," in *Proc. Int. Appl. Comput. Electromagn. Soc. Symp.-China (ACES)*, Aug. 2019, pp. 1–2.
[33] A. Van Etten, D. Lindenbaum, and T. M. Bacastow, "SpaceNet: A remote sensing dataset and challenge series," 2018, *arXiv:1807.01232*. [Online]. Available: https://arxiv.org/abs/1807.01232
[34] J. Korhonen and J. You, "Peak signal-to-noise ratio revisited: Is simple beautiful?" in *Proc. 4th Int. Workshop Qual. Multimedia Exper.*, New York, NY, USA, Jul. 2012, pp. 37–38.
[35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

SHUMING XIAO was born in Weifang, Shandong, in 1997. He received the B.Eng. degree from Northeast Agricultural University, Harbin, China, in 2019. He is currently pursuing the Ph.D. degree in mechatronics engineering with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include machine vision, compressive sensing, deep learning, and image processing.

SHAOJU WANG was born in Liaocheng, Shandong, in 1983. He received the B.Eng. degree from Northeast Normal University, Changchun, China, in 2005, and the Ph.D. degree in mechatronics engineering from the University of Chinese Academy of Sciences, Changchun, in 2010. His research interests include micro-nano satellite overall design, space optical camera control, and image processing technology.

LIN CHANG was born in Changchun, Jilin, in 1985. She received the B.S. degree in physics from Nankai University, Tianjin, in 2009, and the Ph.D. degree in optical engineering from the University of Chinese Academy of Sciences, Changchun, in 2014. Her research interests include space satellite attitude control design and remote sensing camera imaging technology.