Singular Value Decomposition-Based Adaptive Sampling Approximate Message Passing Net for Sparse-View CT Reconstruction

Zhenhua Wu , Jiafei Xu , and Lixia Yang , *Member, IEEE*

Abstract—Sparse-viewcomputed tomography (CT) imaging is a promising technique for reducing radiation dose and accelerating data acquisition in medical imaging. However, the challenges of handling a reduced number of projection views persist for both iterative estimation and deep neural reconstruction methods. In this paper, to address these challenges, we present a singular value decomposition-based adaptive sampling approximate message passing network (ASAMP-Net) sparse-view CT imaging method. To achieve multiple sparse views projection within a single scene imaging and alleviate the computational burden, our proposed ASAMP-Net method incorporates an adaptive sampling module into the AMP deep unrolling network. This module dynamically adjusts the data samples used during the learning process, making our method highly adaptable to various projection matrices. Moreover, by decomposing the projection matrix into its principal components, our approach identifies the respective contributions of independent structures. We then select the most significant principal components to construct a projection matrix model with increased orthogonality, thereby enhancing reconstruction performance. Extensive experiments on public datasets demonstrate the superiority of our method. Notably, ASAMP-Net handles various sparse projection views with just a single training process, achieving prominent imaging results compared to other methods in the literature.

Index Terms—Adaptive sampling, ASAMP-Net, deep unfolding, image reconstruction, sparse-view CT.

I. INTRODUCTION

COMPUTED tomography (CT) is a non-destructive testing method extensively employed in medical, industrial, and materials applications [1], [2], [3], [4], [5]. Despite its versatility in clinical medicine, the use of X-ray has raised concerns about radiation exposure. Consequently, sparse-view CT has emerged as a promising strategy to mitigate radiation exposure, while there is still the streak artifacts problem in the image formation process due to insufficient projection data.

To date, a brunch of methods have been developed for sparse-view CT reconstruction ranging from iterative estimation techniques to supervised deep neural reconstruction networks. Iterative estimation techniques including FBP [6], ART [7], SART [8], SIRT [9], EM [10] leverage analytical mathematical models and iterative techniques to iteratively refine the image, and image reconstruction performance heavily relies on a large quantities of projection views. Needless to say, the streak artifacts can still be troublesome when comes to the limited projections. By imposing regularization penalty term, Total Variation (TV) [11], [12] exploits the total variation of the image and able to promote smoothness while preserving edges from highly undersampling data. Besides, sparsity penalty term and optimization techniques are leveraged in compressed sensing (CS) [13], [14] algorithms to achieve considerable performance in sparse-view CT imaging. Note that the reconstruction performance of these regularization based methods are highly dependent on the regularization term and the orthogonality of projection matrix, the memory and storage costs of matrix operation of large dimensionality can still be troublesome in the iterative optimization process.

In more recent years, with the ability to achieve underlying hierarchical feature learning and complex nonlinear mapping capability from quantities of training data, deep neural reconstruction networks have received considerable attention in sparse-view literature [15], [16], [17], [18]. Boublil et al. [19] proposed a supervised machine learning method to enhance image restoration, using convolutional neural networks (CNNs) to integrate the reconstruction results with different bias or variance in the image restoration process to improve the reconstruction quality. Chen et al. [20] mapped the low-dose CT image to the corresponding normal dose image block by block by using the depth Convolutional neural network, thereby reducing artifacts and preserving the image structure. Chen et al. [21] combined autoencoders, deconvolutional networks, and shortcut connections into the residual encoder-decoder convolutional neural network (RED-CNN) to achieve low-dose CT reconstruction. In order to reconstruct good images from highly sparse views, Wu [22] proposed a deep embedding-attention-refinement (DEAR) network to achieve finer image features and structure.In

general, data-driven deep neural sparse-view reconstruction networks leverage large quantities of collected training projection data and sample images to learn optimal non-linear mapping relations specific to the projection matrix, which has poor generalization ability and requires retraining when the projection view changes.

Typically, aiming at avoiding iterations optimization and sophisticated regularization parameters turning, model-driven methods [23], [24], [25] are built based on deep unfolding techniques that stem from the standard linear optimization algorithms, including IHT/IST [23], ADMM networks [24] and AMP networks [25]. Sun et al. [26] expanded the ADMM algorithm for MRI image reconstruction. J. Adler and O. Öktem [27] proposed a primal-dual algorithm for CT reconstruction. J. Zhang and B. Ghanem [28] expanded the ISTA algorithm for visual image reconstruction. W. Dong et al. [29] developed the Half Quadratic Splitting (HQS) algorithm to solve the image inverse problem.

Inspired by the recent advancements in *beyond deep unfolding* reconstruction techniques, in this paper, we propose the adaptive sampling approximate message passing deep unfolding network (ASAMP-Net). The network consists of three modules, including a sampling module, an initialization module and a reconstruction module. By embedding a controllable sampling module into the AMP backbone unrolling network, multiple sensing matrices with different projection views are trained in parallel. Besides, for mode orthogonality and redundancy minimization, we propose to truncate insignificant principal components of the projection matrix via singular value decomposition (SVD). Measurements are acquired dynamically via the adaptive sampling module. Then, the initialization module selects the most important principal components of the projection matrix and adjusts the initial input of the network. Finally, the structure of the image is reconstructed using the deep network layer of the reconstruction module. Unlike recent work using matrix-inversion-based and data-driven deep reconstruction networks, our generic approach is directly adapted to multi-sparse views sampling ratios and multi-scene image reconstruction, and requires no need for heavy matrix operations and massive amount of training scene targets and measurements datasets.

The following subsection of this paper is organized as follows. Section II details the proposed reconstruction network. Section III presents the experiments and results. Section IV summarizes the paper.

II. METHOD

A. Imaging Principle

In an ideal CT imaging system under noise-free conditions, the CT measurement can be converted into the form of the line integral of the absorption function of the object at a certain position on the cross section:

$$g\left(\vec{r}(\lambda),\widehat{\theta}\right) = \int_0^\infty dt f\left(\vec{r}(\lambda) + t\widehat{\theta}\right) \quad (1)$$

where f($\vec{r}$) represents the absorption function; $g(\vec{r},\widehat{\theta})$ is the sinogram, where $\vec{r}$ and $\widehat{\theta}$ denote the radial coordinate and

the angle of the projection point. If the trajectory of the X-ray source is a circle, then

$$\vec{r}(\lambda) = R(\cos\lambda, \sin\lambda, 0) \quad (2)$$

where $\lambda \in [0, 2\pi)$. R is the radius of the source trajectory.

The imaging model (1), can be approximated by the following discrete linear system:

$$\mathbf{g} = \mathbf{Mf} + \mathbf{e} \quad (3)$$

where $\mathbf{f} \in \mathfrak{R}^N$ represents the original image, N represents the number of pixels in the original image. $\mathbf{g} \in \mathfrak{R}^M$ is projection data through the $M \times N$ measurement matrix $\mathbf{M} \in \mathfrak{R}^{M \times N}$, ($M < N$), and $\mathbf{e} \in \mathfrak{R}^M$ is the noise. The traditional CS-CT method reconstructs the original image $\mathbf{f}$ by solving the following optimization problems:

$$\min_{\mathbf{f}} \frac{1}{2} \|\mathbf{g} - \mathbf{Mf}\|_2^2 + \lambda \|D\mathbf{f}\|_1 \quad (4)$$

where $D\mathbf{f}$ denotes the transform coefficients of $\mathbf{f}$ with re-spect to some transform D and the sparsity of the vector $D\mathbf{f}$ is encouraged by the ℓ_1 norm. λ is a regularization parameter for sparsity, and D is the optimal transform.

B. AMP Algorithm

Approximate Message Passing (AMP) algorithm is an efficient iterative algorithm, which is mainly used in high-dimensional signal processing, and can be used to solve compressed sensing, signal recovery, matrix decomposition, signal classification, channel estimation and other problems. The basic idea is to gradually approximate the true value of the signal through iterative estimation. The AMP algorithm iteratively estimates signals through a series of linear mappings, soft threshold functions, and other operations, and updates parameters through a series of exchange information transfers. This iterative process can be expressed as:

$$\mathbf{z}^{(N-1)} = \mathbf{g} - \mathbf{Mf}^{(N-1)} \quad (5)$$

$$\mathbf{f}^{(N)} = \mathfrak{F}_N(\mathbf{M}^T\mathbf{z}^{(N-1)} + \mathbf{f}^{(N-1)}) \quad (6)$$

where $\mathbf{M}^T$ is the transpose of measurement matrix M, $\mathfrak{F}_N(\cdot)$ is the non-linear function, and N is the number of the iteration. If the initial data is $\mathbf{f}_0$ and the original data is $\bar{\mathbf{f}}$, then

$$\mathbf{M}^T\mathbf{z}^0 + \mathbf{f}^0 = \bar{\mathbf{f}} + (\mathbf{M}^T\mathbf{M} - \mathbf{I})(\bar{\mathbf{f}} - \mathbf{f}^0) \quad (7)$$

where $\mathbf{I}$ is the identity matrix in the size of $N \times N$. The N-th iteration of (7) is

$$\mathbf{M}^T\mathbf{z}^{(N-1)} + \mathbf{f}^{(N-1)} = \bar{\mathbf{f}} + (\mathbf{M}^T\mathbf{M} - \mathbf{I})(\bar{\mathbf{f}} - \mathbf{f}^{(N-1)}) \quad (8)$$

where the entries M_{ij} of the measurement matrix are independent and identically distributed as $M_{ij} \sim \mathcal{N}(0, 1/a)$, and $N(\mu, \sigma^2)$ denotes the Gaussian distribution with the mean value μ and the variance σ^2. Under this assumption, $(\mathbf{M}^T\mathbf{M} - \mathbf{I})(\bar{\mathbf{f}} - \mathbf{f}^{(N-1)})$ is also a Gaussian distributed vector with the variance $a^{-1}\|\bar{\mathbf{f}} - \mathbf{f}^{(N-1)}\|_2^2$. Then, (9) can be reformulated into the sum of the original signal and a noise term as follows:

$$\mathbf{M}^T\mathbf{z}^{(N-1)} + \mathbf{f}^{(N-1)} = \bar{\mathbf{f}} + \mathbf{e} \quad (9)$$

where $\mathbf{e} = (\mathbf{M}^T\mathbf{M} - \mathbf{I})(\bar{\mathbf{f}} - \mathbf{f}^{(N-1)})$ denotes the noise term.
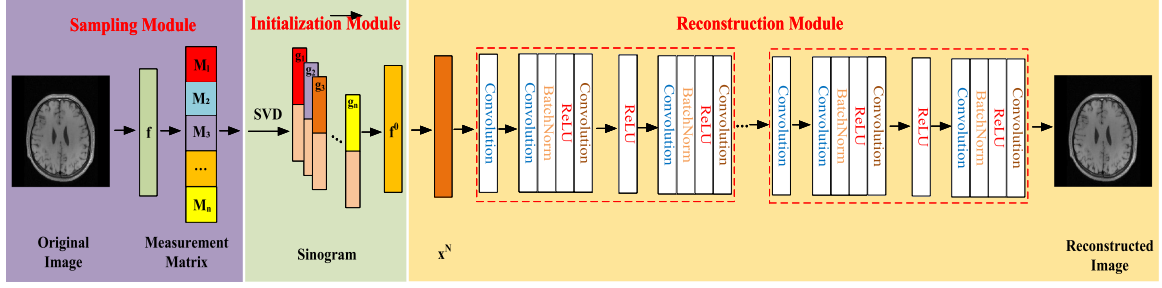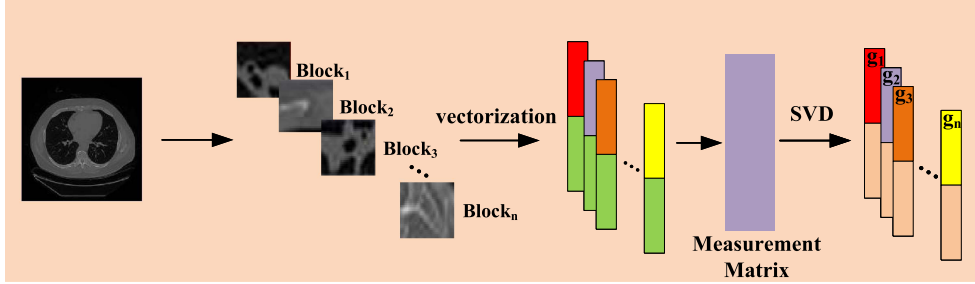
Fig. 1. Proposed ASAMP-Net network architecture.

Fig. 2. Schematic diagram of measurement value acquisition.

C. Imaging Network Model

By deeply unfolding (6), we convert or map each iteration of the traditional AMP algorithm into a single-level module of the network, and stack multiple modules to form a complete deep unfolded network. The details are as follows:

$$\mathbf{f}^{(N)} = D^{(N)^T} \eta D^{(N)} (\mathbf{M}^T \mathbf{z}^{(N-1)} + \mathbf{f}^{(N-1)}) \qquad (10)$$

where $D^{(N)}$ satisfies $D^{(N)^T} D^{(N)} = \mathbf{I}$, D can be different in each iteration. η is threshold function for ℓ_1 regulation.

The overview structure of ASAMP-Net is illustrated in Fig. 1. The our network consists of a sampling module, an initialization module and a reconstruction module.

1) Sampling Module: In traditional block-based compressed sensing, the patterns of the measurement matrices for training and testing are fixed, resulting in inefficient training of the network model and testing phases. For our network, as shown in Fig. 1, we improve the computational efficiency by introducing additional measurement matrices under different projection views. During the training process, we can generate network models under different projection views by changing the mode of the measurement matrices. The sampling operation adaptive to different sparse views is obtained by training the sampling sub-network. Our network can handle network models under any projected view with only one training.

2) Initialization Module: The original image **f** is cropped into an image patch $\vec{f}_i$ of size 33×33. The corresponding measurement values can be calculated through the measurement matrix $\vec{g}_i$ ($i = 1, 2, 3...N$). The process of obtaining the image block $\vec{f}_i$ and the corresponding measurement value $\vec{g}_i$ is shown in Fig. 2. Before network training, singular value decomposition is performed on the measurement matrix to obtain the optimized measurement value $\mathbf{g}_{SVD}$ and semi-orthogonal measurement

matrix $\mathbf{M}_{SVD}$. The original image blocks and SVD optimized measurement values are used as inputs to the network.

Assuming that the non-singular matrix is $\mathbf{M} \in \mathfrak{R}^{M \times N}$, then $\mathbf{M}$ can be decomposed into three sub-matrix multiplication forms using the SVD technique [30]:

$$\mathbf{M} = U \Sigma V^T \qquad (11)$$

In (12), U and V are orthonormal matrices and $\Sigma (\Sigma \geq 0)$ is arranged in descending order.

Due to the fact that the number of projections M is smaller than the number of pixels N in the original image, there are M non zero singular values σ_n, where $K < M$. Substitute (12) into (3):

$$\mathbf{g} = U \Sigma V^T \mathbf{f} + \mathbf{e} \qquad (12)$$

$$= U [\overbrace{\Sigma_1}^{M \times M} \overbrace{\mathbf{0}}^{M \times (N-M)}] \begin{bmatrix} \overbrace{V_1^T}^{M \times N} \\ \overbrace{V_2^T}^{(N-M) \times N} \end{bmatrix} \mathbf{f} + \mathbf{e} \qquad (13)$$

$$= \sum_{n=1}^{N} \sigma_n \mathbf{u}_n \mathbf{v}_n^T \mathbf{f} + \mathbf{e} \qquad (14)$$

Then,

$$\mathbf{g} \approx \sum_{n=1}^{K} \sigma_n \mathbf{u}_n \mathbf{v}_n^T \mathbf{f} = U \Sigma_1 V_1^T \mathbf{f} \qquad (15)$$

Finally, the optimized discrete linear system is expressed as:

$$\mathbf{g}_{SVD} = \mathbf{M}_{SVD} \mathbf{f} \qquad (16)$$

where $\mathbf{g}_{SVD} = \Sigma_1^{-1} U^T \mathbf{g}$, $\mathbf{M}_{SVD} = V_1^T$.

Algorithm 1: ASAMP-Net Training Algorithm.

Input: Sampling matrix: $M = \{M_1, M_2, \ldots, M_n\}$; Image patch: $f = \{f_1, f_2, \ldots, f_k\}$; Number of iterations: N; T: maximum training epochs.

Output: $\tilde{\mathbf{f}}^{(N)}$

Sampling process:
$(g_i)_{i=1}^n = \{(M_i f_i)\}_{i=1}^n$

Initialization process:
$(g'_i)_{i=1}^n = \{(M_i)_{SVD} f_i\}_{i=1}^n$
$\mathbf{f}^{(0)} = \mathbf{M}_{SVD}^T \mathbf{g}'$

Reconstruction process:
For k < N Do:
  k = k + 1
  $\mathbf{x}^k = \mathbf{M}_{SVD}^T \mathbf{z}^{(k-1)} + \mathbf{f}^{(k-1)}$
  $\mathbf{f}^{(k)} = D^{(k)^T} \eta D^{(k)} (\mathbf{M}_{SVD}^T \mathbf{z}^{(k-1)} + \mathbf{f}^{(k-1)})$
  Adam optimization algorithm updates loss function:
  $\mathcal{L}_{\text{total}} = \frac{1}{N_a N_b} \sum_{i=1}^{N_b} \|\mathbf{f}_i - \mathbf{f}_i^k\|_2^2$
  $\mathbf{z}^{(k-1)} = \mathbf{g}_{SVD} - \mathbf{M}_{SVD} \mathbf{f}^{(k-1)}$
Return $\tilde{\mathbf{f}}^{(k)}$

From (5) and (10), it can be seen that:

$$\mathbf{z}^{(N-1)} = \mathbf{g}_{SVD} - \mathbf{M}_{SVD} \mathbf{f}^{(N-1)} \tag{17}$$

$$\mathbf{f}^{(N)} = D^{(N)^T} \eta D^{(N)} (\mathbf{M}_{SVD}^T \mathbf{z}^{(N-1)} + \mathbf{f}^{(N-1)}) \tag{18}$$

3) Reconstruction Module: Influenced by the idea of deep unfolding network, the iterative denoising process of AMP algorithm is mapped to the deep network to obtain the reconstruction module. Each module represents one iteration. According to (8), if $\bar{\mathbf{f}} - \mathbf{f}^{N-1}$ is known, it can be directly solved through linear reconstruction. In addition, the result of $\bar{\mathbf{f}} - \mathbf{f}^{N-1}$ is not necessarily the same after each iteration. In order to obtain more features of the input data and improve the reconstruction accuracy, it is solved by introducing a deep neural network. The nonlinear function $\mathfrak{F}_N(\cdot)$ can be replaced by a multi-layer CNN. $\mathfrak{F}_N(\cdot)$ consists of five convolutional layers. The first four convolutional layers are all set with bias items, and the numbers of output channels are 32, 32, 32, and 32. Except for the first and last layers, each layer is followed by a rectified linear unit (ReLU) [31]. The last convolutional layer has no bias term and the number of output channels is 1. The filter size of each convolutional layer is set to 3×3. The padding size of each convolutional layer is set to 1 to ensure that the output and input have the same size. Batch normalization (BN) [32] is introduced to accelerate the training time of the neural network, improve the robustness and generalization ability of the network and reduce the probability of overfitting.

If $\mathbf{x}^N = \mathbf{M}_{SVD}^T \mathbf{z}^{(N-1)} + \mathbf{f}^{(N-1)}$, then there is

$$\mathbf{f}^{(N)} = D^{(N)^T} \eta D^{(N)} \mathbf{x}^N \tag{19}$$

4) Loss Function: Given the training data pairs $\{(\mathbf{g}_i, \mathbf{f}_i)\}_{i=1}^{N_b}$, AMP-Net first takes the preprocessed CS measurements $\mathbf{g}_{\mathbf{SVD}}$ as input and generates the reconstruction results, denoted as $\mathbf{f_i}^N$, as output. We seek to reduce the difference between $\mathbf{f}_i$ and $\mathbf{f_i}^N$. Therefore, we use mean square error (MSE) to describe the differences between the original image and the restored image. As follows:

$$\mathcal{L}_{\text{total}}(\Theta) = \frac{1}{N_a N_b} \sum_{i=1}^{N_b} \|\mathbf{f}_i - \mathbf{f}_i^N\|_2^2 \tag{20}$$

where $\mathbf{f}_i$ is the i-th original image in the training set, N_a denotes the size of $\mathbf{f}_i$ and N_b denotes the size of the training set.

III. EXPERIMENTS AND RESULTS

A. Experiment Settings

We adopt LIDC-IDRI [33] as our dataset, which includes 1012 cases and about 240 thousand CT images with a size of 512×512. We use Microdicom software to convert the picture size to 256×256. The slice thickness of CT images ranges from 0.6 mm to 5.0 mm. Case 1–50 is used as the test set, case 50–150 is used as the validation set, and the rest is the training set. The original CT image is converted into parallel beam CT projection data of 180 projection views, as the ground truth value of Radon transform.

We optimized the loss function using the Adam algorithm [34]. The learning rate is set as 0.0001 and the epoch is 100. All methods are performed on a workstation with the model of 11th Gen Intel(R) Core(TM) i7-11700 @ 2.50 GHz CPU, NVIDIA GeForce RTX 4070 GPU and 48 GB memory. We use root mean square error (RMSE), peak signal-to-noise ratio (PSNR) [35] and structural similarity (SSIM) [36] to quantitatively evaluate the performance of reconstructed images.

To test the performance of the reconstruction algorithms in various sparse sampling conditions, we set different numbers of projections to reconstruct CT images and analyzed qualitatively as well as quantitatively.

B. Comparison

In this experiment, we choose the LIDC-IDRI dataset as the training set of the network and randomly cut the images into 500000 small patches with a size of 33×33, and use the image patches and their corresponding CS measurement vectors to construct training data pairs. The dataset is divided into 70% training set, 20% validation set and 10% test set. At the same time, different sparse views are selected to train the network model. In order to verify the effectiveness of the proposed singular value optimization on the network reconstruction results, we preprocess the dataset to generate the SVDLIDC-IDRI dataset to train the ASAMP network. In addition, four different extremely low sampling methods are set to simulate the adaptive sparse sampling process of the proposed algorithm. It is worth noting that the experiments were all simulated without noise. In order to conduct numerical experiments and performance evaluation more reasonably, we selected different test pictures in the above datasets for comparison of results. The experimental results are shown in Fig. 3.
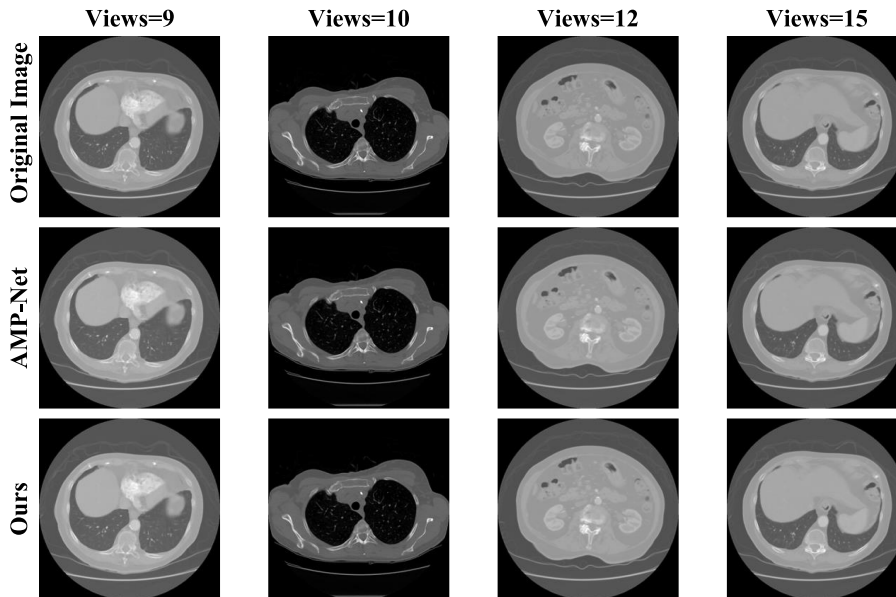
Fig. 3. Comparison of results between AMP-Net and the proposed method under different sparse views.

TABLE I
QUANTITATIVE RESULTS OBTAINED WITH DIFFERENT SPARSE VIEWS

View	9	10	12	15
RMSE	0.0076/**0.0070**	0.0075/**0.0069**	0.0049/**0.0043**	0.0043/**0.0038**
PSNR	40.066/**41.601**	41.323/**42.926**	45.276/**46.823**	46.321/**47.899**
SSIM	0.9775/**0.9816**	0.9791/**0.9826**	0.9855/**0.9883**	0.9897/**0.9938**

The bold entities represent the quantitative indicators of the proposed algorithm.

Fig. 3 shows that training the AMP network model using a dataset without SVD preprocessing can reconstruct clear images. As the viewing angle increases, the quality of the reconstructed image becomes better. However, the our network can achieve more accurate reconstruction of the original image. As the viewing angle continues to increase, the reconstructed image becomes clearer. Comparing AMP-Net with our network under different observation perspectives, the results show that our network has better performance. The first item of the quantitative indicators shown in Table I represents the result of AMP-Net reconstruction, and the latter item represents the result of our network reconstruction. Table I further verifies that the reconstruction performance of our network is better than traditional AMP-Net. At the same time, the reconstruction performance of our network continues to improve with the increase of observation angles.

C. Comparison of Reconstruction Methods

We compare our network with six reconstruction algorithms, including FBP [6], ART [7], TVAL3 [37], ReconNet [38], CSNet$^+$ [39] and DEAR [22]. The FBP algorithm is the most widely used analytical reconstruction algorithm, and the ART algorithm is the most commonly used classical iterative algorithm for sparse view reconstruction. The TVAL3 algorithm uses the ADMM algorithm to solve the TV-constrained minimization CT model, which has better convergence speed and reconstruction performance than other methods. When rebuilding, the parameter μ is set to 2^8, and the parameter β is set to 2^5. ReconNet is a classic deep network method based on CNN. CSNet$^+$ is a method based on convolutional neural network. Its sampling network can adaptively learn the sampling matrix from the training image and retain more image structure information. DEAR is a deep Embedding- Attention-Refinement deep learning network, which consists of three modules: deep embedding, deep attention, and deep refinement. Among these reconstruction algorithms, its reconstruction quality is the best. For a fair comparison, the deep learning based methods are all trained on the same dataset with epoch set to 100. For the ReconNet method, the initial input of the network is set to 0, and the learning rate is set to $1e^{-4}$. For the CSNet$^+$ method, the learning rates for the first 50 epochs, 51 to 80 epochs and the other 20 epochs are $1e^{-3}$, $1e^{-4}$, and $1e^{-5}$, respectively. For the DEAR method, the loss functions of the deep embedding and attention networks are 0.9 and 0.999, respectively. The learning rate decreases with the number of epochs.

To verify the performance of the proposed network, we evaluate the network on the test dataset. Considering that the reconstruction performance of the model-driven method is very poor under the condition of extremely sparse projection views, we set the lowest projection view to 12. We use different reconstruction methods to reconstruct images from 12
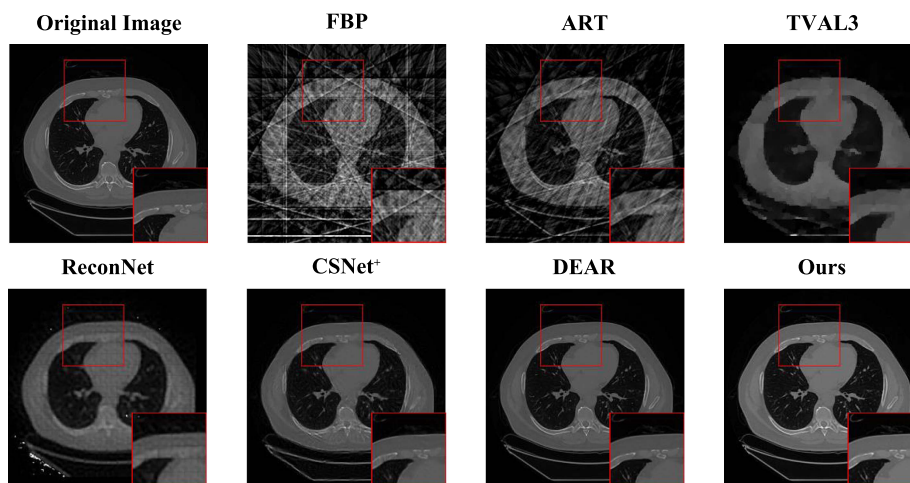
Fig. 4.   Reconstruction results using different methods when Views = 12.

projection views, as shown in Fig. 4. Due to the extremely sparse projection, it is difficult for traditional reconstruction methods to obtain detailed information, resulting in poor reconstructed image quality. As can be seen from Fig. 4, the reconstruction result of the FBP algorithm has very serious streak artifacts, and the details of the image cannot be restored. The reconstruction results of the ART algorithm also have serious streak artifacts, and the quality of the reconstructed image is so blurred that it is impossible to see the specific details of the image. The reconstruction result of the TVAL3 algorithm suppresses the streak artifact very well and retains the edge information of the original image, but the reconstruction result is also very blurred and cannot display specific details. The reconstruction results based on deep learning methods have achieved better reconstruction results. The ReconNet algorithm effectively suppresses the streak artifacts and preserves the edge information of most organs, but the quality of the reconstructed image is still relatively blurred. During the reconstruction process, most of the details between the reconstructed image blocks are lost due to block artifacts. The structure of the reconstruction result of the CSNet$^+$ algorithm is relatively complete and clear, and most of the image details are preserved. Although the reconstruction results of this algorithm have good visual effects, some details are still very blurred. The structure of the reconstruction results of the DEAR algorithm is complete and clear, and a large number of image details are preserved. However, the effectiveness of this algorithm in reconstructing some edge details is still relatively blurred. The reconstruction effect of our proposed algorithm is more delicate and clearer than the reconstruction effect of the DEAR algorithm in restoring the edge of the image. The reconstruction results present the most details and the smallest differences compared to the reference image. At the same time, it can be observed from the red rectangular box in Fig. 4 that the results of different reconstruction algorithms are locally magnified, and the local regions of the proposed reconstruction algorithm retain more detailed information. In order to further demonstrate the ability of the proposed algorithm to preserve image details, we also presented the difference images between the reconstruction results of different algorithms and the original

image, as shown in Fig. 5. From the difference images, it can be seen that the proposed reconstruction algorithm obtained the smallest difference from these difference images. This result once again verifies the superiority of the proposed reconstruction algorithm.

The quality of the reconstructed images improves significantly with increasing number of projected views. Although the reconstruction results of FBP and ART algorithms have been further improved, the reconstruction results obtained by these algorithms still have obvious streak artifacts. The results of the FBP algorithm have the most serious streak artifacts due to the severe sparse projection. The results of the ART algorithm have relatively few streak artifacts. The TVAL3 algorithm effectively suppresses most of the streak artifacts and provides more accurate image detail information. The ReconNet algorithm still cannot reconstruct a good image, and the reconstructed image has obvious image blocks. The reconstruction results of CSNet$^+$ algorithm have been effectively improved, but some areas are relatively blurred. The reconstruction results of the DEAR algorithm and the proposed method both have good image quality. Most of the details of the reconstructed image obtained by the two algorithms can be accurately presented, and the areas with relatively low contrast can also be well reconstructed. The red rectangular box in Fig. 6 shows the slight differences in the reconstruction results between the two algorithms. To further highlight the differences between these two algorithms, we use difference images for evaluation. The visualization results in Fig. 7 demonstrate that the proposed algorithm can more accurately reconstruct the pixel values of the original image and has a better ability to preserve structure.

When the projected views reach 18, the results of each reconstruction algorithm are significantly improved. Due to the severe sparsity of the projected view, the reconstruction results of the FBP algorithm still exhibit significant streak artifacts. The reconstruction results of the ART algorithm contain less streak artifacts. The TVAL3 algorithm effectively suppresses a large number of streak artifacts and preserves the edge information of the image well. However, information in areas of low contrast cannot be reconstructed. Although the
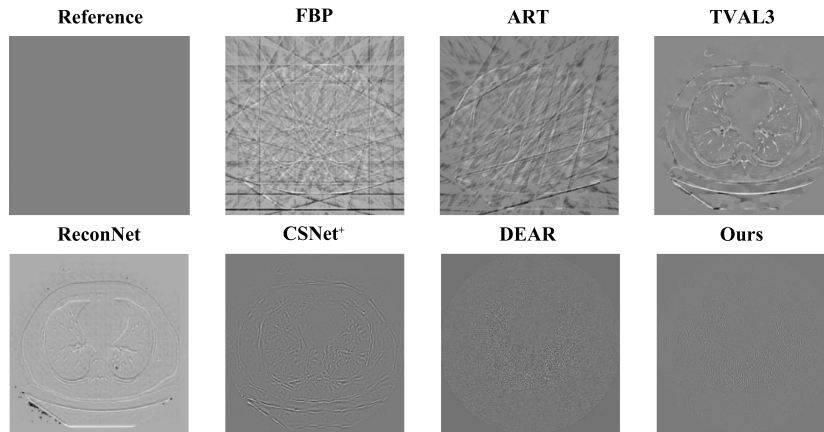
Fig. 5. Difference images between the reconstruction results of different algorithms and the original image when Views = 12.
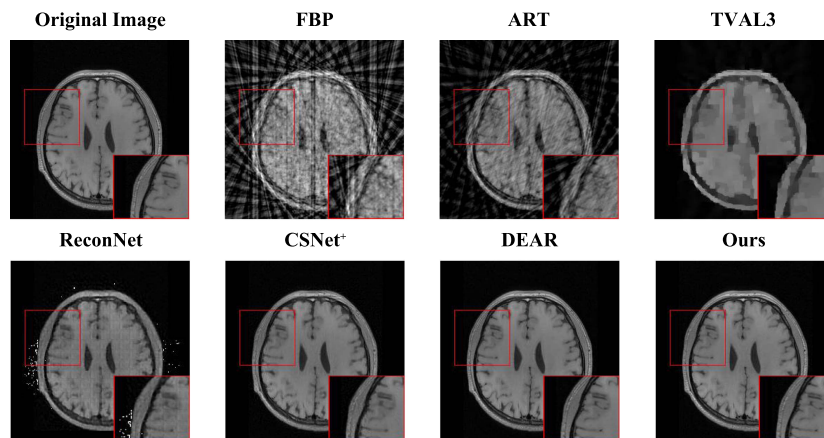

Fig. 6. Reconstruction results using different methods when Views = 15.
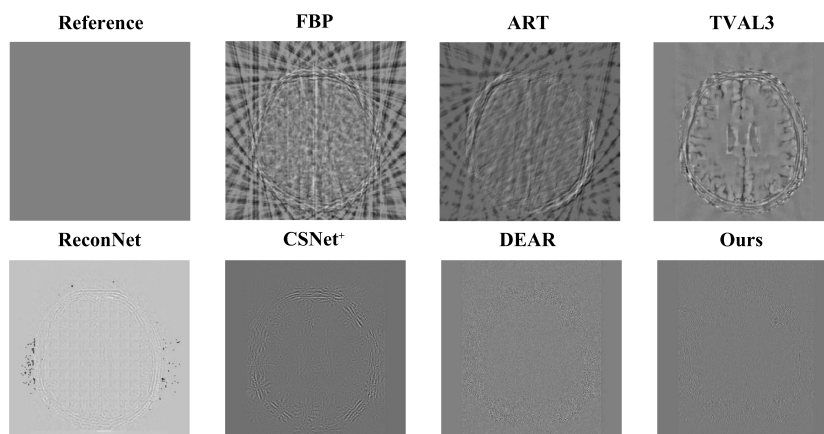

Fig. 7. Difference images between the reconstruction results of different algorithms and the original image when Views = 15.

reconstruction results of the ReconNet algorithm have been greatly improved, the restoration of image details is still not ideal. Moreover, during the reconstruction process, a small amount of noise appears in the reconstruction result of the RconNet algorithm and the edge information of the reconstructed

image is missing. The reconstruction result of the CSNet$^+$ algorithm has a good visual effect and reconstructs a large amount of detailed information, but it still cannot reliably obtain fine details. As shown in the red rectangular box of Fig. 8, the difference between the reconstruction results of the DEAR algorithm and
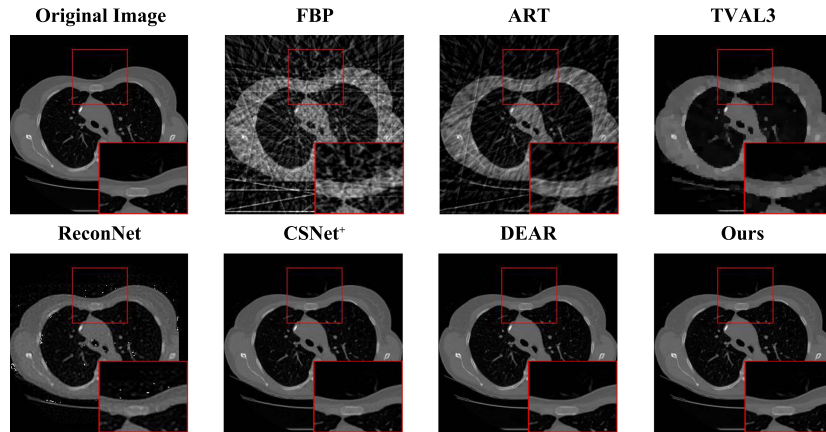
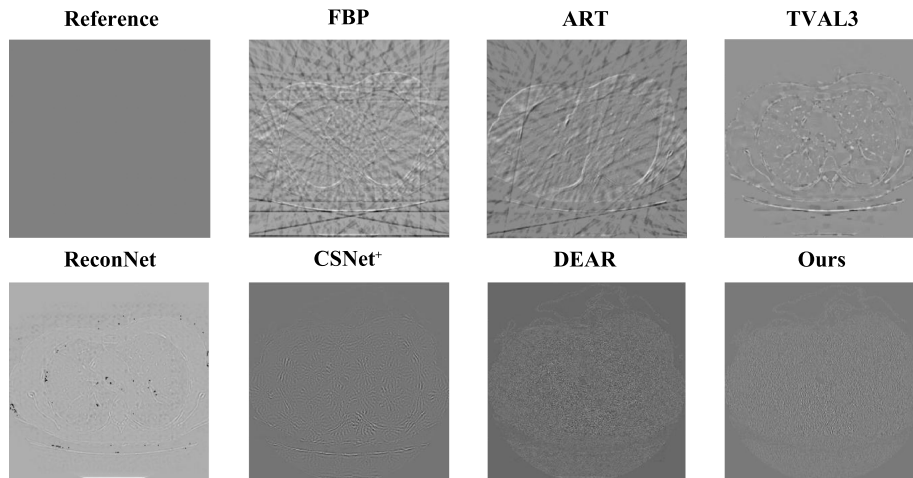Fig. 8. Reconstruction results using different methods when Views = 18.

Fig. 9. Difference images between the reconstruction results of different algorithms and the original image when Views = 18.

TABLE II
COMPARISON OF QUANTITATIVE RESULTS OBTAINED BY DIFFERENT RECONSTRUCTION METHODS

View	Index	FBP	ART	TVAL3	ReconNet	CSNet$^+$	DEAR	Ours
	RMSE	0.1719	0.0980	0.0479	0.0568	0.0248	0.0086	**0.0080**
12	PSNR	15.297	20.179	26.386	24.909	32.423	41.442	**41.896**
	SSIM	0.1260	0.2170	0.7478	0.6952	0.8942	0.9743	**0.9779**
	RMSE	0.1716	0.1011	0.0478	0.0564	0.0202	0.0099	**0.0095**
15	PSNR	15.308	19.908	26.418	24.973	34.168	40.270	**40.343**
	SSIM	0.1961	0.2698	0.7733	0.7861	0.9379	0.9730	**0.9742**
	RMSE	0.1538	0.0954	0.0420	0.0631	0.0141	0.0059	**0.0045**
18	PSNR	16.260	20.409	27.536	24.005	37.206	45.182	**46.916**
	SSIM	0.1545	0.2254	0.7894	0.7503	0.9565	0.9881	**0.9912**

The bold entities represent the quantitative indicators of the proposed algorithm.

our proposed algorithm cannot be clearly distinguished from the qualitative results. Both networks have reconstructed the image well. In order to more clearly demonstrate the difference in reconstruction performance between the two algorithms, Fig. 9 shows the difference images between the reconstruction results of the different algorithms and the original image. From the visualization results, it can be seen that the reconstruction result of the proposed algorithm has less loss and is closer to the original image. The result verifies the superiority of the proposed algorithm.

Table II shows the quantitative results of different reconstruction algorithms under different projection views. RMSE can evaluate the degree of change in data, and a smaller value of RMSE indicates higher reconstruction accuracy. PSNR and SSIM are two commonly used image quality evaluation indicators. PSNR can better reflect the difference of signal pixels.

SSIM is an indicator that measures the similarity between two images. The larger the value of PSNR and SSIM, the better the quality of the reconstructed image. As can be seen from Table II, the values of the three metrics used to evaluate our network are better than those of other algorithms. This results once again verify the effectiveness of the proposed algorithm.

IV. CONCLUSION

In this paper, a model and data co-driven ASAMP reconstruction method is proposed for sparse-view CT imaging. To address the challenges of achieving multiple sparse views projection in single scene imaging and enhancing computational efficiency, we introduce an adaptive sampling module within the AMP deep unfolding network. Furthermore, by incorporating a pre-processing step and employing singular value decomposition, we identify the contribution levels of independent structures within the projection matrix. Subsequently, the most significant principal components are selected to construct a more orthogonal projection matrix model, effectively enhancing the quality of reconstructed images. The superior imaging performance of the proposed method against other state-of-art method in the literature are both quantitatively and qualitatively verified on public datasets. In conclusion, our ASAMP method's adaptability to diverse projection matrices and prominent imaging performance make it a promising solution for enhancing medical imaging capabilities while reducing radiation dose and acquisition time.

REFERENCES

[1] X. Jing, D. Yu, and L. Wan, "Non-destructive testing of through silicon vias by high-resolution X-ray/CT techniques," in *Proc. IEEE 14th Electron. Packag. Technol. Conf.*, 2012, pp. 533–536.

[2] S. Han, L. Yu, and D. Qi, "Application of X-ray computed tomography to automatic wood testing," in *Proc. IEEE Int. Conf. Automat. Logistics*, 2007, pp. 1325–1330.

[3] B. Pan and X. Zhu, "18F-FDG PET/CT imaging of posttransplant lymphoproliferative disorder following renal transplantation: Three clinical cases," *Heliyon*, vol. 9, no. 4, 2023, Art. no. e14746.

[4] M. Hasni et al., "An efficient approach based on 3D reconstruction of CT scan to improve the management and monitoring of COVID-19 patients," *Heliyon*, vol. 6, no. 11, 2020, Art. no. e05453.

[5] M. Nishio et al., "Convolutional auto-encoder for image denoising of ultra-low-dose CT," *Heliyon*, vol. 3, no. 8, 2017, Art. no. e00393.

[6] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Philadelphia, PA, USA: SIAM, 2001.

[7] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theor. Biol.*, vol. 29, no. 3, pp. 471–481, 1970.

[8] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): A superior implementation of the art algorithm," *Ultrason. Imag.*, vol. 6, no. 1, pp. 81–94, 1984.

[9] J. Gregor and T. Benson, "Computational analysis and improvement of SIRT," *IEEE Trans. Med. Imag.*, vol. 27, no. 7, pp. 918–924, Jul. 2008.

[10] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[11] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-ray Sci. Technol.*, vol. 14, no. 2, pp. 119–139, 2009.

[12] Y. Zhang, Y. Wang, W. Zhang, F. Lin, Y. Pu, and J. Zhou, "Statistical iterative reconstruction using adaptive fractional order regularization," *Biomed. Opt. Exp.*, vol. 7, no. 3, pp. 1015–1029, 2016.

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

[14] E. J. Candès, 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.

[15] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A sparse-view CT reconstruction method based on combination of densenet and deconvolution," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1407–1417, Jun. 2018.

[16] H. Lee, J. Lee, H. Kim, B. Cho, and S. Cho, "Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 109–119, Mar. 2019.

[17] D. Hu et al., "Hybrid-domain neural network processing for sparse-view CT reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 1, pp. 88–98, Jan. 2021.

[18] C. Sun, K. Deng, Y. Liu, and H. Yang, "A lightweight dual-domain attention framework for sparse-view CT reconstruction," in *Proc. IEEE 8th Int. Conf. Comput. Commun.*, 2022, pp. 2102–2106.

[19] D. Boublil, M. Elad, J. Shtok, and M. Zibulevsky, "Spatially-adaptive reconstruction in computed tomography using neural networks," *IEEE Trans. Med. Imag.*, vol. 34, no. 7, pp. 1474–1485, Jul. 2015.

[20] H. Chen et al., "Low-dose CT via convolutional neural network," *Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 679–694, 2017.

[21] H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.

[22] W. Wu, X. Guo, Y. Chen, S. Wang, and J. Chen, "Deep embedding-attention-refinement for sparse-view CT reconstruction," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 4501111.

[23] M. Wang et al., "RMIST-Net: Joint range migration and sparse reconstruction network for 3-D mmW imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5205117.

[24] R. Li, S. Zhang, C. Zhang, Y. Liu, and X. Li, "Deep learning approach for sparse aperture ISAR imaging and autofocusing based on complex-valued ADMM-Net," *IEEE Sensors J.*, vol. 21, no. 3, pp. 3437–3451, Feb. 2021.

[25] Z. Zhang, Y. Liu, J. Liu, F. Wen, and C. Zhu, "AMP-Net: Denoising-based deep unfolding for compressive image sensing," *IEEE Trans. Image Process.*, vol. 30, pp. 1487–1500, 2021.

[26] J. Sun et al., "Deep ADMM-Net for compressive sensing MRI," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2016.

[27] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, Jun. 2018.

[28] 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.*, 2018, pp. 1828–1837.

[29] W. Dong, P. Wang, W. Yin, G. Shi, F. Wu, and X. Lu, "Denoising prior driven deep neural network for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 10, pp. 2305–2318, Oct. 2019.

[30] T. Fromenteze, O. Yurduseven, P. Del Hougne, and D. R. Smith, "Lowering latency and processing burden in computational imaging through dimensionality reduction of the sensing matrix," *Sci. Rep.*, vol. 11, no. 1, 2021, Art. no. 3545.

[31] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 807–814.

[32] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[33] S. G. Armato III et al., "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, 2011.

[34] D. Kingma and J. Ba, "Adam: A. method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, 2015.

[35] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. IEEE 20th Int. Conf. Pattern Recognit.*, 2010, pp. 2366–2369.

[36] 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.

[37] C. Li, W. Yin, H. Jiang, and Y. Zhang, "An efficient augmented Lagrangian method with applications to total variation minimization," *Comput. Optim. Appl.*, vol. 56, pp. 507–530, 2013.

[38] 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.*, 2016, pp. 449–458.

[39] W. Shi, F. Jiang, S. Liu, and D. Zhao, "Image compressed sensing using convolutional neural network," *IEEE Trans. Image Process.*, vol. 29, pp. 375–388, 2020.