SOUL-Net: A Sparse and Low-Rank Unrolling Network for Spectral CT Image Reconstruction  

Xiang Chen, Wenjun Xia, Member, IEEE, Ziyuan Yang, Hu Chen, Member, IEEE, Jiliu Zhou, Senior Member, IEEE, Zhe Wang, Yang Chen, Senior Member, IEEE, Bihan Wen, Senior Member, IEEE, and Yi Zhang, Senior Member, IEEE  

Abstract—Spectral computed tomography (CT) is an emerging technology that generates a multi-energy attenuation map for the interior of an object and extends the traditional image volume into a 4-D form. Compared with traditional CT based on energy-integrating detectors, spectral CT can make full use of spectral information, resulting in high resolution and providing accurate material quantification. Numerous model-based iterative reconstruction methods have been proposed for spectral CT reconstruction. However, these methods usually suffer from difficulties such as laborious parameter selection and expensive computational costs. In addition, due to the image similarity of different energy bins, spectral CT usually implies a strong low-rank prior, which has been widely adopted in current iterative reconstruction models. Singular value thresholding (SVT) is an effective algorithm to solve the low-rank constrained model. However, the SVT method requires a manual selection of thresholds, which may lead to suboptimal results. To relieve these problems, in this article, we propose a sparse and low-rank unrolling network (SOUL-Net) for spectral CT image reconstruction, that learns the parameters and thresholds in a data-driven manner. Furthermore, a Taylor expansion-based neural network backpropagation method is introduced to improve the numerical stability. The qualitative and quantitative results demonstrate that the proposed method outperforms several representative state-of-the-art algorithms in terms of detail preservation and artifact reduction.

Index Terms—Deep learning, image reconstruction, low-rank prior, spectral computed tomography (CT).

I. INTRODUCTION

Due to the great potential for lesion detection and material decomposition, spectral computed tomography (CT) has attracted increasing attention [1], [2], [3] recently. By discriminating photon energies during data acquisition, photon-counting detector (PCD)-based spectral CT can collect multiple sets of measured data with different spectral information within a single exposure. Due to the limited numbers of photons in different energy bins and the impact of pile-up, spectral projections acquired by PCD usually contain strong noise [4]. Therefore, it would be of great significance to develop more efficient spectral CT reconstruction algorithms to aid in medical diagnosis.

To date, dose-reduced scanning configuration has been widely used to alleviate the radiation risk. In this situation, directly reconstructing the spectral CT image with traditional reconstruction algorithms such as filtered back projection (FBP) and algebraic reconstruction technique (ART) [5] will inevitably lead to obvious artifacts and noise, which negatively affect the subsequent diagnosis procedure. To this end, a great number of works have been developed to improve the imaging quality of spectral CT.

Recently, numerous sparsity regularization-based methods have been developed for CT reconstruction, including total variation (TV) [6], tight frames [7], and dictionary learning [8]. These methods can be extended to spectral CT directly by applying these sparsity constraints to each energy bin separately. However, such a strategy ignores the correlation among different energy bins, resulting in suboptimal results. As a result, many algorithms have been developed to fully exploit the data correlation in the spectral domain [9], [10], [11]. For instance, tensor dictionary learning (TDL) was introduced to explore the correlation among the spectral CT images form.
different energy bins [9]. Wu et al. [10] integrated the gradient \(\ell_0\) norm and TDL to eliminate artifacts effectively. In [12], an adaptive weight sparsity method, combining the \(l_0\) norm and spectral-spatial self-similarity, termed SISTER, was proposed.

On the other hand, low-rank is an important prior in spectral CT image reconstruction. This has been proven in numerous works [13], [14], [15]. The nuclear norm is an efficient constraint for low-rank regularization, which can be solved by the singular value thresholding (SVT) method [16]. Semerci et al. [17] proposed an iterative spectral CT reconstruction algorithm based on the tensor nuclear norm regularization. Wu et al. [20] represented a low-rank tensor factorization method based on Kronecker-basis-representation (KBR), which enforces the nonlocal self-similarity constraint into the reconstruction model.

Meanwhile, to enjoy the best of both worlds, preliminary attempts that integrate both sparsity and low-rank prior information for spectral CT have been made [18], [19]. Furthermore, Li et al. [21] applied Tucker decomposition to low-rank tensors and simultaneously employed TV to impose sparsity constraints on perfusion CT. Niu et al. [22] leveraged the nonlocal similarity of spectral images to construct the sparse matrix and then solved it with low-rank matrix decomposition. Xia et al. [23] utilized spatial and spectral nonlocal self-similarity to expand the matrix decomposition to tensor form and obtained more efficient results. Chen et al. [24] proposed an effective similarity measure to construct tensor units for low-rank sparse decomposition. These studies have demonstrated the merit that combining the sparse and low-rank priors is beneficial to the reconstruction performance.

However, although these iterative methods are effective in improving the reconstruction results, they are usually computationally expensive, even for single-energy reconstruction [12]. The processing of multispectral images inevitably brings extra time consumption that is unexpected in the clinic. Another disadvantage of iterative methods is that there are usually several parameters that need to be determined manually [12], [25]. Manual selection of these parameters is very labor-intensive and time-consuming and may lead to a suboptimal result.

As an emerging technology, deep learning has been widely used in many low-level vision tasks [26], [27], [28]. Inspired by these encouraging results, a large amount of CT image reconstruction methods based on deep learning have been developed [3], [29], [30], [31]. When these methods are separately applied to each monochromatic slice, the similarities among different energy bins are neglected. In the field of deep-learning-based spectral CT reconstruction, related research are quite limited. Wu et al. [32] proposed a skip-encode U-net network equipped with anisotropic TV and \(\ell_p^b\)-loss to improve the performance on detail preservation. Mustafa et al. [33] proposed a fast reconstruction model based on the U-net architecture with multichannel input and output for sparse-view spectral CT (DSIR). However, these methods are all postprocessing-based networks for spectral CT, ignoring data consistency. Nor do these methods fully explore the prior information inherent in spectral CT images, such as low rank and sparsity.

To conquer the above challenges, in this article, we propose a sparse and low-rank unrolling network (SOUL-Net) for spectral CT image reconstruction. On the one hand, to better exploit the similarities among different energy bins, we impose a low-rank constraint on the reconstruction model and SVT is adopted to solve the low-rank constraint problem. Since SVT usually requires the manual selection of thresholds, which may lead to suboptimal results, a learnable threshold is embedded into SVT. On the other hand, backpropagation in the learned SVT may lead to numerical instability, so a Taylor expansion-based neural network backpropagation method is introduced to relieve this problem. The qualitative and quantitative results demonstrate that the proposed method outperforms several representative state-of-the-art algorithms in terms of image detail preservation and artifact reduction.

The rest of this article is organized as follows. Section II elaborates the proposed method. Section III presents the experiment configuration and the results. Concluding remarks are provided in Section IV.

II. METHODOLOGY

During the CT scan, when the X-ray passes through an object, some of the photons are absorbed in the process following Lambert–Beer’s law. In spectral CT, according to the energy of photons, we can divide the photon spectrum into multiple energy bins. With a narrow energy bin, the detected photon flux in each energy bin can be expressed as

\[
I(E_k) \approx \tilde{I}_0(E_k) \exp \left( -\int_{I(x)} \bar{\mu}(x,E_k) dx \right), \quad k = 1, 2, \ldots, K
\]  

(1)

where \(E_k\) is the \(k\)th energy bin and there are \(K\) energy bins in total. \(\tilde{I}_0(E_k)\) is the number of incident photons in \(E_k\) energy bin. \(I(x)\) represents the X-ray path. \(\bar{\mu}(x,E_k)\) stands for average attenuation coefficient of the material at position \(x\) to the photons in \(E_k\). Applying the transformation to (1), we can obtain the following formula:

\[
\log \frac{\tilde{I}_0(E_k)}{I(E_k)} = \int_{I(x)} \bar{\mu}(x,E_k) dx, \quad k = 1, 2, \ldots, K.
\]  

(2)

After discretizing (2), we have the following linear form:

\[
y_k = Ax_k, \quad k = 1, 2, \ldots, K
\]  

(3)

where \(x_k \in \mathbb{R}^{N_h \times N_w}\) represents the attenuation coefficient distribution for the slice with a spatial resolution of \(N_h \times N_w\) in the \(k\)th energy bin, and \(y_k \in \mathbb{R}^{N_h \times N_v}\) is the corresponding projection data, detected with a scanning geometry of \(N_f\) detector elements and \(N_v\) projection views. \(A \in \mathbb{R}^{(N_h \times N_w) \times (N_h \times N_v)}\) is the measurement matrix generated according to the scanning geometry. For spectral CT imaging, our purpose is to solve the above inverse problem to obtain \(x_k\).

For (3), a tensor version for \(\ell_1\) norm-based regularized reconstruction model is obtained as follows [34], [35], [36], [37]:

\[
\min_x \frac{1}{2} \| A(x) - \zeta \|_F^2 + \lambda \| \Psi(x) \|_1
\]  

(4)
where $\mathbf{\chi} \in \mathbb{R}^{N_h \times N_w \times N_i}$ denotes the spectral CT image with a spatial resolution of $N_h \times N_w$ and $N_i$ energy bins. $\mathbf{\xi} \in \mathbb{R}^{N_h \times N_w \times N_i}$ is the projection data acquired with a scanning geometry of $N_d$ detector elements and $N_i$ projection views. $\Psi$ denotes the sparsifying transform for $\mathbf{\chi}$. $\|\cdot\|_F$ and $\|\cdot\|$ denote the Frobenius norm and $\ell_1$ norm, respectively. $\lambda$ is the regularization parameter. As mentioned above, many works [6], [9] have been proposed based on the low-rank regularization terms and have shown the effectiveness of the low-rank prior in spectral CT reconstruction. Due to the NP-hardness of the rank minimization problem, it is typically approximated using nuclear norm minimization. This convex optimization problem can be solved directly by applying a soft threshold to the singular values of the matrix, a technique called SVT. SVT is widely employed in low-rank problems, as evidenced by its use in papers [38], [39], and [40]. To exploit the merits of both low-rank prior and sparse regularization, (4) can be extended as

$$\min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{A}(\mathbf{\chi}) - \mathbf{\xi} \right\|_F^2 + \lambda_1 \|\mathbf{\phi} (\mathbf{\chi})\|_1 + \lambda_2 \|\Psi (\mathbf{\chi})\|_1 \tag{5}$$

where $\lambda_1$ and $\lambda_2$ are the weighting parameters for the low-rank and sparse regularization terms, respectively. $\mathbf{\phi}$ denotes the operator that reshapes the tensor $\mathbf{\chi}$ to a matrix with a size of $\mathbb{R}^{(N_i \times N_w) \times N_h}$. $\|\cdot\|_1$ represents the nuclear norm. In short, we used a low-rank prior for the spectral domain and a sparse prior for the spatial domain [17], [18], [41].

To simplify the optimization, $\mathbf{\phi} (\mathbf{\chi})$ can be replaced with an auxiliary variable $\mathbf{\Gamma}$. Then, (5) can be rewritten as

$$\min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{A}(\mathbf{\chi}) - \mathbf{\xi} \right\|_F^2 + \lambda_1 \|\mathbf{\Gamma}\|_1 + \lambda_2 \|\Psi (\mathbf{\chi})\|_1, \text{ s.t. } \mathbf{\Gamma} = \mathbf{\phi} (\mathbf{\chi}). \tag{6}$$

The augmented Lagrange form of (6) is formulated as

$$\min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{A}(\mathbf{\chi}) - \mathbf{\xi} \right\|_F^2 + \lambda_1 \|\mathbf{\Gamma}\|_1 + \lambda_2 \|\Psi (\mathbf{\chi})\|_1 + \alpha \|\mathbf{\Gamma} - \mathbf{\phi} (\mathbf{\chi})\|_F^2 \tag{7}$$

where $\alpha$ is the Lagrangian multiplier and $\rho$ is the penalty factor. Let $\mathbf{\Omega} = \alpha / \rho$. Then, a simplified version of (7) can be obtained as

$$\min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{A}(\mathbf{\chi}) - \mathbf{\xi} \right\|_F^2 + \lambda_1 \|\mathbf{\Gamma}\|_1 + \lambda_2 \|\Psi (\mathbf{\chi})\|_1 + \frac{\rho}{2} \|\mathbf{\Gamma} - \mathbf{\phi} (\mathbf{\chi})\|_F^2 \tag{8}$$

which can be iteratively solved using the alternating direction method of multipliers (ADMMs) algorithm with scaled dual variables [42] as follows:

$$\mathbf{\Gamma}^k = \arg \min_{\mathbf{\Gamma}} \frac{\rho}{2} \|\mathbf{\Gamma} - \mathbf{\phi} (\mathbf{\chi}^{k-1}) + \mathbf{\Omega}^{k-1}\|_F^2 + \lambda_1 \|\mathbf{\Gamma}\|_1 \tag{9a}$$

$$\mathbf{\chi}^k = \arg \min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{A}(\mathbf{\chi}) - \mathbf{\xi} \right\|_F^2 + \lambda_1 \|\mathbf{\Gamma}\|_1 + \frac{\rho}{2} \|\mathbf{\Gamma}^k - \mathbf{\phi} (\mathbf{\chi}) + \mathbf{\Omega}^{k-1}\|_F^2 \tag{9b}$$

$$\mathbf{\Omega}^k = \mathbf{\Omega}^{k-1} + \mathbf{\Gamma}^k - \mathbf{\phi} (\mathbf{\chi}^k). \tag{9c}$$

The subproblem in (9a) is a nuclear norm minimization problem that can be solved with SVT [16]. We perform singular value decomposition (SVD) as $\mathbf{\phi} (\mathbf{\chi}^{k-1}) - \mathbf{\Omega}^{k-1} = \mathbf{U} \Sigma \mathbf{V}^T$. $\mathbf{U}$ and $\mathbf{V}$ denote the left singular matrix and right singular matrix of $\mathbf{\phi} (\mathbf{\chi}^{k-1}) - \mathbf{\Omega}^{k-1}$, respectively. $\Sigma$ is the diagonal matrix of singular values of $\mathbf{\phi} (\mathbf{\chi}^{k-1}) - \mathbf{\Omega}^{k-1}$. The result is obtained as

$$\mathbf{\Gamma}^k = \mathbf{\Sigma} \max (\mathbf{\sigma}_i - \lambda_1 / \rho, 0) \mathbf{u}_i \mathbf{v}_i^T \tag{10}$$

where $\sigma_i$ is the $i$th entry of the diagonal matrix $\mathbf{\Sigma}$. $\mathbf{u}_i$ and $\mathbf{v}_i$ are the $i$th column vectors of $\mathbf{U}$ and $\mathbf{V}$, respectively.

Equation (9b) is a combination of two terms based on L2 norm (referred to as the “smooth” term) and one term based on L1-norm (referred to as the “nonsmooth” term). Typically, an iterative soft thresholding algorithm (ISTA) is used to solve (9b) [43] as

$$\mathbf{\chi}^{k+1} = \mathbf{\chi}^{k-1} - t \cdot \left( \mathbf{A}^T \left( \mathbf{A} (\mathbf{\chi}^{k-1}) - \mathbf{\xi} \right) + \mathbf{\rho} \mathbf{\phi} (\mathbf{\chi}^{k-1}) - \mathbf{\Gamma}^k - \mathbf{\Omega}^{k-1} \right) \tag{11a}$$

$$\mathbf{\chi}^k = \arg \min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{\chi} - \mathbf{\chi}^{k-1} \right\|_F^2 + \lambda_2 \|\Psi (\mathbf{\chi})\|_1 \tag{11b}$$

where $t$ denotes the step size, $\phi'$ denotes the conjugate operator of $\mathbf{\phi}$, and $\mathbf{A}^T$ is the transpose of $\mathbf{A}$.

The corresponding optimization procedure is summarized as follows:

$$\mathbf{L}^k = \mathbf{\Gamma}^k = \mathbf{\Sigma} \max (\mathbf{\sigma}_i - \lambda_1 / \rho, 0) \mathbf{u}_i \mathbf{v}_i^T, \tag{12a}$$

$$\mathbf{R}^k = \mathbf{\chi}^{k+1} = \mathbf{\chi}^{k-1} - t \cdot \left( \mathbf{A}^T \left( \mathbf{A} (\mathbf{\chi}^{k-1}) - \mathbf{\xi} \right) + \mathbf{\rho} \mathbf{\phi} (\mathbf{\chi}^{k-1}) - \mathbf{\Gamma}^k - \mathbf{\Omega}^{k-1} \right) \tag{12b}$$

$$\mathbf{S}^k = \mathbf{\chi}^k = \arg \min_{\mathbf{\chi}} \frac{1}{2} \left\| \mathbf{\chi} - \mathbf{\chi}^{k-1} \right\|_F^2 + \lambda_2 \|\Psi (\mathbf{\chi})\|_1 \tag{12c}$$

$$\mathbf{B}^k = \mathbf{\Omega}^k = \mathbf{\Omega}^{k-1} + \mathbf{\Gamma}^k - \mathbf{\phi} (\mathbf{\chi}^k). \tag{12d}$$

In our proposed SOUL-Net, the above iterative procedure is unrolled into a neural network, whose general architecture and components are shown in Fig. 1. In the unrolled 4th iteration block shown in Fig. 1(a), there are four main modules: low-rank module $\mathbf{L}^k$, reconstruction module $\mathbf{R}^k$, sparse module $\mathbf{S}^k$ and Lagrangian multiplier module $\mathbf{B}^k$. Each module corresponds to the subproblem in (12). The structures of each module are shown in Fig. 1(b)–(e), respectively. Each module is elaborated in the following parts.

The L2 norm loss function usually leads to smoother results, while the L1 norm loss function is able to preserve image detail better. Therefore, in our work, we chose to use the L1 norm [44], [45], [46]. The loss function of the proposed SOUL-Net is defined as

$$\mathcal{L} = \frac{1}{N_s} \sum_{n=1}^{N_s} \left\| \mathbf{\bar{X}}_n - \mathbf{X}_n \right\|_1 \tag{13}$$

where $\mathbf{\bar{X}}_n$ denotes the label image and $N_s$ denotes total sample number.
CHEN et al.: SOUL-Net FOR SPECTRAL CT IMAGE RECONSTRUCTION

Fig. 1. Schematic of the proposed SOUL-Net. SOUL-Net is obtained based on the iteration process defined in (12). (a) Optimization procedure for \( K \) iterative modules of SOUL-Net method. Four modules \( L^k, R^k, S^k, \) and \( B^k \) in black dashed box correspond to four equations of (12). (b) \( L^k \) module of each iteration block. (c) \( R^k \) module of each iteration block. (d) Corresponds to \( S^k \) module of each iteration block. (e) Corresponds to \( B^k \) module of each iteration block. \( N_c \) is the number of energy bins.

A. Low-Rank Module

The update of the low-rank module in (12a) is shown in Fig. 1(b). In the traditional SVT method, the selection of the threshold \( \lambda_1/\rho \), which has a great impact on the final performance, is empirically set and time-consuming. To circumvent this obstacle, we introduce a Taylor expansion-based learnable threshold method [47] to solve the low-rank problem. This is shown as

\[
0 = 6_i \text{Relu} (\sigma_i - \text{Sigmoid} (d) \times \sigma_1) u_i v_i^T
\]

(14)

where \( \sigma_1 \) is the largest value in the singular value matrix and \( d \) represents the learnable parameter initialized to 0. The Relu activation function is essentially the max function in (12a) and is used to guarantee the nonnegativity. \( \text{Sigmoid} (d) \times \sigma_1 \) is the learned threshold and has a range of \((0, \sigma_1)\). Equation (14) guarantees that the output is meaningful.

However, the backpropagation of SVD may lead to numerical instability in the PyTorch framework [48]. In an SVD-based network, we assume \( X \in \mathbb{R}^{m \times n} = U \Sigma V^T, m > n \) and \( \mathcal{L} \) is the loss function of the network. According to [48], only one of \( \partial \mathcal{L} / \partial V \) and \( \partial \mathcal{L} / \partial U^T \) is needed to complete the backpropagation, and the partial derivatives can be computed as

\[
\frac{\partial \mathcal{L}}{\partial X} = U \left( 2 \Sigma \left( K^T \cdot \left( V^T \frac{\partial \mathcal{L}}{\partial V} \right) \right) \right)_{\text{sym}} + \left( \frac{\partial \mathcal{L}}{\partial \Sigma} \right)_{\text{diag}} V^T
\]

(15)

and

\[
K_{i,j} = \begin{cases} 
\frac{1}{\sigma_i^2 - \sigma_j^2}, & i \neq j \\
0, & i = j 
\end{cases}
\]

(16)

where \( \sigma_i \) and \( \sigma_j \) are the \( i \)th and \( j \)th largest singular values. For a square matrix \( M \), \( M_{\text{sym}} = (M^T + M)/2 \) and \( M_{\text{diag}} \) is the resultant matrix by setting all off-diagonal elements from \( M \) to 0.

However, if the distance between any two singular values is close to zero, then the partial derivative with respect to \( X \) will be numerically unstable. Wang et al. [47] proposed a Taylor expansion-based method to compute the eigenvector gradients for covariance matrices. This achieves fast, accurate, and stable performance and can be easily embedded into a deep-learning framework. As a result, to relieve the problem of SVT in gradient backpropagation, in this article, we use a similar approach to calculate the gradient of SVD for a matrix. With the \( Z \)th order Taylor expansion of \( f(x) = (1/1 - x), x \in (0, 1), f(x) \) can be rewritten as

\[
f(x) = 1 + x + \cdots + x^Z + R_{Z+1}
\]

(17)

where \( R_{Z+1} = (x^{Z+1})/(1 - x) \) is the remainder of the expansion. Equation (16) can be split into two cases: \( \sigma_i > \sigma_j \) and

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
\[ \sigma_i < \sigma_j. \] For \( \sigma_i > \sigma_j \), (16) is converted into
\[ K_{i,j} = \frac{1}{\sigma_i + \sigma_j} \frac{1}{\sigma_i} \left( 1 + \frac{\sigma_j}{\sigma_i} \right)^Z + R_{Z+1}, \quad i < j. \]

The Zth order Taylor expansion of (18) is formulated as
\[ K_{i,j} = \frac{1}{\sigma_i + \sigma_j} \frac{1}{\sigma_i} \left( 1 + \frac{\sigma_j}{\sigma_i} + \ldots + \left( \frac{\sigma_j}{\sigma_i} \right)^Z \right) + R_{Z+1}, \quad i < j. \] (19)

where
\[ R_{Z+1} = \frac{1}{\sigma_i^Z - \sigma_j^Z} \left( \frac{\sigma_j}{\sigma_i} \right)^{(Z+1)}, \quad i < j. \] (20)

The solution for the case \( \sigma_i < \sigma_j \) can be derived in the same manner. To avoid the numerical instability caused by the case \( \sigma_i = 0 \), a small positive number \( \epsilon \) is added to the diagonal of the matrix \( X \), that is \( X = X + \epsilon I \). This operation guarantees that \( \sigma_i \geq \sqrt{\epsilon} \) in the backpropagation of the neural network. More details can be found in [47]. In the backpropagation of (12a), we use this approach for threshold updating. \( \epsilon \) is set to \( 1 \times 10^{-8} \). \( Z \) is set to 9.

### B. Reconstruction Module

Fig. 1(c) shows the reconstruction module in (12b). Since the projection \( A \) and backprojection \( A^T \) operations are usually time-consuming, compute unified device architecture (CUDA) is adopted to accelerate the computation of (12b) [49]. In our experiments, we used a C++ extension of PyTorch. Specifically, we implemented the CT projection operator \( A \) and the backprojection operator \( A^T \) in CUDA and compiled them to generate a library. This library is then used to encapsulate forward and backward propagation in C++ into a function, which we used in PyTorch. More details can be found at https://github.com/xwj01/CTLIB. The parameters we use to generate \( A \) and \( A^T \) are shown in Section III-A. To avoid the suboptimal results caused by the manual selection of parameters, in this section, parameters \( t \) and \( \rho \) are learned from the training data.

### C. Sparse Module

The update of sparse module in (12c) is shown in Fig. 1(d). Equation (12c) is a standard \( \ell_1 \) norm problem that can be solved by a nonlinear method, and the solution is related to the specific form of the sparse transform. For instance, the solution of the widely used TV norm problem [6] can be obtained using the Chambolle-Pock projection algorithm [50] as follows:
\[ \chi^k = \chi^{k-\frac{1}{2}} - \pi_{t,\lambda_2} \left( \chi^{k-\frac{1}{2}} \right) \] (21)
where \( \pi_{t,\lambda_2}(\cdot) \) is a nonlinear projection operator. According to [50], the nonlinear projection \( \pi_{t,\lambda_2}(\chi^{k-\frac{1}{2}}) \) can be solved by the \( \ell_2 \) div \( q \), and \( q \) can be solved iteratively as
\[ q_{i,j,r}^{k+1} = q_{i,j,r}^k + \tau \left( \nabla \left( \text{div} \left( q^k - \chi^{m'/\mu} \right) \right) \right)_{i,j,r} \]
\[ + \frac{1}{1 + \tau \nabla \left( \text{div} \left( q^k - \chi^{m'/\mu} \right) \right)_{i,j,r}} \] (22)
where \( \nabla \) denotes the gradient operator, the \( \text{div} \) denotes the divergence operator, and \( \tau \) is the step size.

To improve the performance, in our proposed SOUL-Net, an attention-based network is utilized to learn the nonlinear projection operator instead of the handcrafted sparse transform. As shown in the sparse module in Fig. 1(d), \( \chi \) is fed into a two-layer convolutional neural network (CNN). Then, the obtained feature maps are scaled with the learned channel attention [51] and fed into another two-layer network. Finally, the output adds \( \chi^{k-(1/2)} \) to complete the update of the sparse module.

The attention mechanism is widely used in medical imaging [52], [53], which identifies the parts of features that most contribute to the results. For spectral CT imaging, we introduce the channel attention mechanism [51] to better fuse the information across the energy bins and learn the nonlinear projection operator. The attention block is shown in Fig. 1(d). The kernel sizes are \( 3 \times 3 \).

### D. Lagrangian Multiplier Module

The update of the Lagrangian multiplier module in (12d) is shown in Fig. 1(e). We initialized \( B_0 \) to 0 based on our experience with traditional iterative reconstruction methods, but it can also be initialized with the FBP result, which, as we analyzed in our previous work [54], does not make much difference.

### III. EXPERIMENTS AND RESULTS

#### A. Data Preparation

In order to verify the performance of the proposed method, we use simulation clinical data. The images were initially collected from the GE Discovery dual-energy CT 750 HD scanner of Ruijin Hospital, Shanghai, China. Virtual multichannel monochromatic images are generated using GE software. The spectrum ranges from 60 to 100 keV in increments of 10 keV [58]. We directly use the images to simulate the photon counting projection data. In our experiments, we randomly selected 400 images from four patients as the training set and 100 images from another patient as the test set. The image size is \( 512 \times 512 \). In our experiments, the projection data are produced using the distance-driven algorithm with fan beam geometry [59]. The distance from the X-ray source to the scanner rotation center is 35 cm. The distance from the scanner rotation center to the radiation detector is 30 cm. The physical width and height of a pixel are both set to 0.25 mm. The number of detector elements is set to 1024 and each element is 0.30 mm in length. We use sparse-view data to validate the performance of SOUL-Net. We sampled projection data from 64 views uniformly distributed over \( 360^\circ \) for the experiments. It’s worth noting that sparse-view data represents a type of undersampled data, and leveraging the sparse prior image data can be an effective way to address undersampling issues. Furthermore, Poisson and electronic noise are inserted into the generated projection data to assess the robustness of the proposed SOUL-Net in clinical situations [60] as follows:
\[ I = P(I_0 \exp(-\xi)) + N(0, \sigma^2) \] (23)
where \( I_0 \) denotes the incident flux of the X-ray, the background electronic noise is represented by \( \sigma^2_e \), and \( \zeta \) stands for the noise-free projection data. In the experiment, \( I_0 \) and \( \sigma^2_e \) are set to \( 1 \times 10^6 \) and 6, respectively.

The real data are from the Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China. The real data are scanned with a mouse. The distance between the scanner rotation center and the X-ray source is set to 200 mm. The distance between the radiation detector and scanner rotation center is set to 250 mm. There are a total of 1027 detector elements with a length of 0.23 mm. The X-ray source is divided into a total of 4 energy bins, 12, 26, 34, and 42 keV. The image size is \( 512 \times 512 \). The projection data consists of a total of 1080 view data. Meanwhile, only 60 of 1080 views of data are involved to verify the effectiveness of the method. The dataset size of real data is 13, we use 12 of them for training and the remaining one for testing.

**B. Experiment Configuration**

The experiments are conducted in Python 3.8 with the PyTorch library on a server computer (Intel Xeon Gold 5220 CPU, 128 GB RAM, and RTX 3090 GPU*8). The software code for SOUL-Net is released at https://github.com/scuchenxiang/SOUL-Net. In the experiment, the learning rate is \( 1 \times 10^{-4} \). The Adam solver is adopted to optimize our models with \((\beta_1, \beta_2) = (0.9, 0.999) \) [61]. The batch size is 1. The number of iterations blocks is set to 10. \( t \) is initialized to 0. \( \rho \) is initialized to 1. \( Z \) is set to 9.

The peak signal-to-noise ratio (PSNR) and structural similarity index metric (SSIM) are adopted to quantitatively evaluate our method [62], [63]. In addition, the noise power spectrum (NPS) is adopted to assess the noise texture [64], [65], which is defined as

\[
\text{NPS}(f_i, f_j) = \frac{\Delta_i \Delta_j}{N_i N_j} \langle |\text{DFT}(\Delta I(i, j))|^2 \rangle
\]

where \( \Delta_i \) and \( \Delta_j \) are the physical sizes of the image pixels and \( N_i \) and \( N_j \) denote the sizes of the selected region of interest (ROI) for the calculation of NPS. In this article, \( \Delta_i = \Delta_j = 0.6641 \) mm and \( N_i = N_j = 254 \). \( \Delta I(i, j) \) represents the error component obtained by subtracting the label image from the reconstructed image in the ROI. In [64], the ROI is obtained from images with equal extraction intervals. In this article, a total of 44 \( \times \) 44 ROIs are chosen from the image of \( 512 \times 512 \). The operator \( \langle \cdot \rangle \) denotes the average of all ROIs. DFT denotes the discrete Fourier transform. In NPS
maps, the content located toward the center represents the structural information of the image, while the content toward the periphery represents the edge and detail information. The protection of the edge and detail information improves as the periphery of the NPS image gets closer. Conversely, the protection of the structural information improves as the center of the NPS image gets closer. If the peak of the NPS is concentrated in the center low frequency, it can result in significant errors in the structural information. For the same signal-to-noise ratio, images that peak at high frequencies tend to look better than those that peak at low frequencies.

In the experiments, four state-of-the-art methods NLSMD [22], LEARN [54], DSIR [33], DenseNet [55], ResNet [56], Unet [57], and ULTRA [32] are used for comparison. NLSMD is a recently proposed iterative reconstruction method based on nonlocal low-rank and sparse priors. LEARN is a classic unrolling network method for CT reconstruction. We reconstruct the images separately for each energy bin using the LEARN network in our experiments. DSIR and ULTRA are two postprocessing methods for spectral CT image reconstruction. On the other hand, three popular CNN architectures for generic image processing, including DenseNet, ResNet, and Unet, were also included for comparison. We removed the pooling operation from these networks and adjusted the output size of the networks with minor architectural adjustments to make these three networks applicable to the denoising problem. Six hundred images were used as the training data for these three networks. We first extracted patches from noisy and corresponding clean images with a fixed size. In addition, data augmentation, including up and down flipping, left and right flipping, and rotating was performed to further enlarge the training dataset. Finally, 38 400 patches were used for training and the remaining 169 images were used for testing. The configuration and initialization of these competing methods are implemented according to the recommendations of the original papers.

C. Results of Simulated Noisy Data

1) Results of Simulated Noisy Abdominal Image: Fig. 2 shows one representative abdominal slice reconstructed with simulated noisy data using different methods. Three commonly used energy bins including 60, 80, and 100 keV are used in this article. It can be seen that some details in the NLSMD result are blurred, although most noise and artifacts are suppressed. Similarly, some important structural details are smoothed by LEARN, DSIR, DenseNet, ResNet, Unet, and ULTRA. In order to better visualize the effectiveness of detail preservation, we zoomed-in view the ROI indicated by the red box in Fig. 2. We can observe that compared with other methods, SOUL-Net not only eliminates the noise and artifacts effectively but also maintains the structures effectively, especially in the parts indicated by red arrows.

To better verify the detail recovery ability, Fig. 3 shows the absolute difference images associated with the label image. In Fig. 3, it is clear that SOUL-Net has the smallest residual compared with other methods. This indicates that the proposed method can maintain the structural details effectively.

The NPS of all results are shown in Fig. 4. We can see that in both high- and low-frequency bands, the error in the results of SOUL-Net is lower than that of the other methods. This
Fig. 4. NPS maps of abdominal images from different reconstruction methods in 60, 80, and 100 KeV energy bins.

Fig. 5. Reconstruction results for different methods with simulated noisy thoracic data. Display window is [−1200,800]HU.

can be considered as additional evidence of the effectiveness of SOUL-Net.

2) Results of Simulated Noisy Thoracic Image: A typical thoracic case reconstructed by different methods is shown in Fig. 5. The lung window is used to show the details of the blood vessels. In the results of NLSMD, although noise and artifacts are well removed, most structural details in the lungs are blurred. LEARN, DSIR, DenseNet, and ResNet achieve better performance than NLSMD but still suffer from visible detail loss. Unet, ULTRA, and SOUL-Net obtain promising results in terms of both artifact suppression and detail preservation. The ROI marked by the red boxes in the first row of Fig. 5 is enlarged. It can be seen that the results of SOUL-Net maintain sharper edges and clearer structures, which can be better identified in the areas indicated by the red arrows and demonstrate the impressive potential in detail preservation of SOUL-Net.

To better show the merits of SOUL-Net, Fig. 6 shows the absolute difference images associated with the label image. It is obvious that our proposed SOUL-Net has minimal
errors compared to other methods, which further confirms the effectiveness of our proposed method in noise and artifact suppression.

To further analyze the noise in the results of different methods in different frequency bands, the NPS maps are shown in Fig. 7. It is easy to notice that SOUL-Net outperforms all other methods in noise suppression.

3) Quantitative Results: Tables I and II show the average PSNR and SSIM and their standard deviations across the whole test set in different energy bins. It is obvious that SOUL-Net achieves the best results in terms of both PSNR and SSIM in different energy bins, which demonstrates the effectiveness of our approach. To further demonstrate the effectiveness of our method, we conducted the student t-test with $p < 0.05$ on the mean SSIM/PSNR values across different energy bins of the test set, comparing SOUL-Net to other methods. The $p$-values for NLSMD, LEARN, DSIR, and ULTRA are significantly less than 0.05, which demonstrates the statistical significance of our method.

D. Ablation Experiments

In this section, an experiment was conducted to verify the effectiveness of the attention mechanism. The quantitative results of SOUL-Net with and without attention (called SOUL-Net-NA) mechanism are shown in Table III. Compared with SOUL-Net-NA, SOUL-Net is slightly higher in the PSNR index, but slightly lower in the SSIM index.

To verify the robustness of the experiment, we have showcased the loss curves for both training and test datasets in Fig. 8. The blue curve represents the loss curve of the training set, while the red curve represents the loss curve of the test set. The steady decrease in the loss demonstrates the numerical stability of our method. On the other side, we selected 200 abdominal images and 200 thoracic images from four patients. The energy bins are 40, 70, 120, and 200 keV, respectively. The geometrical parameters are consistent with the simulated clinical data in Section III-A. We trained SOUL-Net in two different ways. In the first approach, we trained with abdominal data and then tested with thoracic data, called SOUL-Net-A. In the second method, we trained with thoracic data and then tested with abdominal data, called SOUL-Net-T. Quantitative results are shown in Table IV. The minimal difference in the results serves as additional evidence for the robustness of our method.

To examine the primary role of each part in SOUL-Net, an ablation study was carried out by activating each part individually while setting the other parts as identity functions. Specifically, we examined the effects of updating only part $L$, only part $B$, only part $R$, and only part $S$. For instance, when only part $L$ was updated, other parts were fixed to identity, and only $\Gamma$ was updated while $\chi^k$ in $S$ was kept constant; this resulted in the input and output being the same. Similarly, when part $B$ was the only part being updated, $\chi^k$ remained unchanged, thereby preventing SOUL-Net from learning. Consequently, $\chi^k$ cannot be learned when SOUL-Net is processed separately for $L$ and $B$. We labeled the SOUL-Net-R and SOUL-Net-S for updating only the $R$ part and $S$ part, respectively. Table V displays the experiment outcomes for each part. The results showed that all four components, $L$, $R$, $S$, and $B$, played a significant role in the reconstruction process. Direct use of the $R$ part did not result in good reconstruction. While using either $L$ or $B$ alone did not produce significant results, they supported the reconstruction

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Fig. 7. NPS maps of thoracic images from different reconstruction methods in 60, 80, and 100 KeV energy bins.

TABLE I
AVERAGE PSNR VALUES OF RECONSTRUCTIONS WITH DIFFERENT ALGORITHMS ON TEST SET

| Methods  | FBP   | NLSMD | LEARN | DSIR  | DenseNet | ResNet | Unet  | ULTRA | SOUL-Net |
|----------|-------|-------|-------|-------|----------|--------|-------|-------|----------|
| 60keV    | 38.26±0.26  | 38.62±0.56  | 38.84±0.26  | 37.79±1.76  | 37.29±1.58  | 37.79±1.85  | 37.95±1.44  | 38.47±1.33  | 39.42±1.33  |
| 70keV    | 39.33±1.72  | 39.53±1.69  | 40.09±0.32  | 39.48±1.72  | 39.41±1.63  | 39.78±1.67  | 39.48±1.53  | 40.18±1.68  | 41.22±1.63  |
| 80keV    | 39.92±1.67  | 39.60±1.64  | 40.19±0.53  | 39.92±1.67  | 39.48±1.63  | 39.92±1.64  | 39.70±1.45  | 40.28±1.85  | 41.28±1.74  |
| 90keV    | 40.08±1.56  | 39.79±1.67  | 40.31±1.99  | 39.89±1.64  | 40.28±1.67  | 40.07±1.48  | 40.76±1.97  | 41.76±1.77  | 41.76±1.77  |
| 100keV   | 40.35±1.70  | 40.10±0.36  | 40.38±1.51  | 40.05±1.77  | 40.63±1.71  | 40.45±1.50  | 41.20±2.11  | 42.24±1.91  | 42.24±1.91  |

Fig. 8. Loss curves of SOUL-Net for training and test datasets during the training stage.

process. In comparison, part S played the most direct role in the reconstruction process. Using only part S generated higher SSIM and PSNR values than using part R alone. Therefore, part S was identified as the main contributor to the reconstruction process.

E. Results of True Data

1) Visual Inspection: To assess the effectiveness of our proposed method, we have conducted experiments on real mouse data. Fig. 9 shows the results of the different methods on four energy bins. The NLSMD method does not remove noise and artifacts well. The LEARN method result is smooth, resulting in a loss of detail in the reconstructed images. The result of the DSIR method still contains some noise and artifacts. DenseNet, ResNet, Unet, ULTRA, and SOUL-Net

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
TABLE V
QUANTITATIVE RESULTS OF SOUL-NET WITH DIFFERENT PARTS

| Methods | SOUL-Net | SOUL-Net-R | SOUL-Net-S |
|---------|----------|------------|------------|
| 60keV   | 39.42±0.33/0.9365±0.007 | 30.33±6.38/0.5773±0.031 | 38.21±1.87/0.9273±0.011 |
| 70keV   | 41.22±1.63/0.9561±0.008 | 30.52±6.39/0.5804±0.031 | 39.86±1.95/0.9454±0.009 |
| 80keV   | 41.28±1.74/0.9542±0.008 | 30.47±6.38/0.5773±0.031 | 39.94±1.91/0.9419±0.009 |
| 90keV   | 41.76±1.77/0.9591±0.008 | 30.48±6.53/0.5805±0.031 | 40.33±1.93/0.9475±0.008 |
| 100keV  | 42.24±1.91/0.9641±0.008 | 30.64±6.40/0.5823±0.031 | 40.73±1.96/0.9497±0.008 |

Fig. 9. Results of SOUL-Net in the true data. The first to fourth rows represent the results of the different methods in the first to fourth channels. The display window of images is $[0,40]$ cm$^{-1}$.

are able to remove the artifacts and retain the image details well. In the bottom-left corner of Fig. 9, we have a zoomed-in view of the ROI region selected by the red box. The red arrows indicate where SOUL-Net shows up well and other methods have lost some image detail. This shows the effectiveness of the SOUL-Net method.
TABLE VI

| Methods | FBP   | NLSMD | LEARN | DSIR | DenseNet | ResNet | Unet  | ULTRA | SOUL-Net |
|---------|-------|-------|-------|------|----------|--------|-------|-------|---------|
| 12keV   | 28.51 | 39.50 | 39.46 | 39.10| 38.12    | 37.43  | 39.35 | 39.23 | 40.10   |
| 26keV   | 27.70 | 38.51 | 38.51 | 38.25| 37.50    | 36.82  | 38.45 | 37.96 | 39.33   |
| 34keV   | 28.66 | 39.80 | 39.48 | 39.30| 38.35    | 37.63  | 39.51 | 39.38 | 40.21   |
| 42keV   | 31.26 | 42.11 | 41.51 | 41.81| 40.72    | 40.59  | 42.17 | 42.07 | 42.84   |

TABLE VII

| Methods | FBP   | NLSMD | LEARN | DSIR | DenseNet | ResNet | Unet  | ULTRA | SOUL-Net |
|---------|-------|-------|-------|------|----------|--------|-------|-------|---------|
| 12keV   | 0.4253| 0.9039| 0.9098| 0.9004| 0.8873   | 0.8675 | 0.9128| 0.9043| 0.9177  |
| 26keV   | 0.3863| 0.8929| 0.8914| 0.8957| 0.8761   | 0.8636 | 0.9011| 0.8885| 0.9058  |
| 34keV   | 0.4233| 0.9082| 0.9113| 0.9016| 0.8883   | 0.8772 | 0.9131| 0.9085| 0.9147  |
| 42keV   | 0.5388| 0.9407| 0.9371| 0.9379| 0.9245   | 0.9186 | 0.9415| 0.9409| 0.9445  |

2) Quantitative Result: We calculated the SSIM and PSNR values between the reconstructed results of different methods and the results of full view FBP result, which are shown in Tables VI and VII. We can find that consistent with the results of visual inspection, our proposed SOUL-Net achieves the best results on both PSNR and SSIM compared with other methods. The results in true data demonstrate the validity of our proposed method.

IV. DISCUSSION AND CONCLUSION

In this article, we proposed a sparse and low-rank unrolling network for spectral CT image reconstruction. To make full use of the prior information of the spectral CT images, we introduced a learnable SVT method to learn low-rank regularization and used the Taylor expansion-based method to calculate the backpropagation gradient. To better utilize the sparsity of spectral CT images, we employed a CNN aided by an attention mechanism to implement sparse regularization. The experiment results demonstrated the effectiveness of our proposed SOUL-Net qualitatively and quantitatively in both noise suppression and detail preservation.

It is noticed that there are two possible improvements to our model. First, the parameter d in the learned SVT is independent of image features. A possible improvement is to learn from the image features. Second, the convolution layers in this network are 2-D convolutions. Three-dimensional convolution has the potential to improve performance in spectral CT data. We will evaluate these modifications in our future work.

Although the proposed SOUL-Net achieved promising results in spectral CT image reconstruction, it still has some drawbacks. One of the limitations is that the SVD operation in the SOUL-Net model requires additional time. Another limitation is the GPU memory consumption in SOUL-Net. If the data have more energy bins, then we need more GPU memory. The SVD operation also requires extra GPU memory. In addition, SOUL-Net requires paired data for training, which is usually difficult to obtain in clinical practice. Therefore, in the future, we will explore how to combine SOUL-Net with other unsupervised models, such as generative or flow models [66], [67], [68], [69], [70], [71].
Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Xiang Chen received the B.A. degree in computer science and technology from the College of Computer Science, Sichuan University, Chengdu, China, in 2013. He is currently working as an Algorithm Engineer at Alibaba, Hangzhou, China. He is also an Assistant Researcher with the School of Cyber Science and Engineering and the Key Laboratory of Data Protection and Intelligent Management, Ministry of Education, Sichuan University. His research interests include medical imaging and deep learning.

Wenjun Xia (Member, IEEE) received the B.A. and Ph.D. degrees from Sichuan University, Chengdu, China, in 2012 and 2021, respectively. He is currently a Lecturer with the School of Cyber Science and Engineering and the Key Laboratory of Data Protection and Intelligent Management, Ministry of Education, Sichuan University.

Ziyuan Yang received the M.S. degree in computer science from the School of Information Engineering, Nanchang University, Nanchang, China, in 2021. He is currently pursuing the Ph.D. degree with the College of Computer Science, Sichuan University, Chengdu, China.

His research interests include biometrics, distributed learning, and security analysis.

Hu Chen received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2011. He was a Visiting Ph.D. Student at the University of California at San Diego, San Diego, CA, USA, from 2008 to 2009. He is currently an Associate Professor with the School of Computer Science, Sichuan University, Chengdu. His current research interests include image and video processing, computer vision, and artificial intelligence. He has authored or coauthored over 40 articles in these areas.

Yan Liu received the Ph.D. degree in engineering from the School of Electronic Information, Sichuan University, Chengdu, China, in December 2011. She did her post-doctoral work at the University of Texas Southwestern Medical Center, Dallas, TX, USA, from July 2015 to July 2016. Her research interests include pattern recognition and machine learning.

Jiliu Zhou (Senior Member, IEEE) received the Ph.D. degree from Sichuan University, Chengdu, China, in 1999. He is a Full Professor and a Doctoral Supervisor with the School of Computer Science and Technology, Sichuan University. He is also an Academic and Technical Leader and an Outstanding Expert in Sichuan Province and has held 17 state or provincial scientific projects, including key projects supported by the National Science Foundation. He has authored more than 300 articles, of which more than 200 articles are indexed by SCI, EI, and ISTP. His current research interests include X-ray CT imaging technology and image processing.

Zhe Wang received the B.S. degree in computer science and technology from the College of Computer Science and Technology, Jilin University, Changchun, China, in 2006, and the Ph.D. degree in computer application technology from the School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China, in 2011.

Now, he is a Professor at the Nuclear Technology and Application Division, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing. His research interests include X-ray CT imaging technology and image processing.

Yang Chen (Senior Member, IEEE) received the M.S. and Ph.D. degrees in biomedical engineering from Southern Medical University, Guangzhou, China, in 2004 and 2007, respectively. Since 2008, he has been a Faculty Member with the Department of Computer Science and Engineering, Southeast University, Nanjing, China. His research interests include medical image reconstruction, image analysis, pattern recognition, and computerized aid diagnosis.
Bihan Wen (Senior Member, IEEE) received the B.Eng. degree in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2012, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 2015 and 2018, respectively. He is currently a Nanyang Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University. His research interests include machine learning, computational imaging, computer vision, image and video processing, and big data applications.

Dr. Wen was a recipient of the 2016 UIUC Yee Fellowship, the 2012 Singapore PEB Gold Medal, the Best Paper Runner-Up Award from IEEE ICME 2020, the Best Paper Award from IEEE ICIAE 2023, and the Best Paper Award from IEEE MIPR 2023. He was a Guest Editor of IEEE Signal Processing Magazine from 2021 to 2023 and IEEE Journal of Selected Topics in Signal Processing from 2023 to 2025. He is now an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.

Yi Zhang (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in computer science and technology from the College of Computer Science, Sichuan University, Chengdu, China, in 2005, 2008, and 2012, respectively. From 2014 to 2015, he was with the Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA, as a Postdoctoral Researcher. He is currently a Full Professor with the School of Cyber Science and Engineering, Sichuan University, and the Director of the Deep Imaging Group (DIG). He has authored more than 80 articles in the field of image processing. These articles were published in several leading journals, including IEEE TRANSACTIONS ON MEDICAL IMAGING, IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING, Medical Image Analysis, European Radiology, and Optics Express, and reported by the Institute of Physics (IOP) and during the Lindau Nobel Laureate Meeting. He received major funding from the National Key Research and Development Program of China, the National Natural Science Foundation of China, and the Science and Technology Support Project of Sichuan Province, China. His research interests include medical imaging, compressive sensing, and deep learning.

Dr. Zhang is a Guest Editor of the International Journal of Biomedical Imaging and Sensing and Imaging and an Associate Editor of IEEE TRANSACTIONS ON MEDICAL IMAGING and IEEE TRANSACTIONS ON RADIATION AND PLASMA MEDICAL SCIENCES.