Received April 19, 2021, accepted May 5, 2021, date of publication May 11, 2021, date of current version May 19, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3079323

A Dual-Domain CNN-Based Network for CT Reconstruction

FENGYUAN JIAO[1], ZHIGUO GUI[1], KUNPENG LI[1], HONG SHANGGUANG[2], YANLING WANG[3], YI LIU[1], AND PENGCHENG ZHANG[1]

[1]Shanxi Provincial Key Laboratory for Biomedical Imaging and Big Data, North University of China, Taiyuan 030051, China
[2]School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China
[3]School of Information Management, Shanxi University of Finance and Economics, Taiyuan 030006, China

Corresponding author: Pengcheng Zhang (zhangpc198456@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 11605160 and Grant 61801438, in part by the National Key Scientific Instrument and Equipment Development Project of China under Grant 2014YQ240445, in part by the Natural Science Foundation of Shanxi Province under Grant 201901D211246, in part by the Scientific and Technological Innovation Program of Higher Education of Shanxi Province under Grant 2020L0282, and in part by the Young Academic leader Project of the North University of China under Grant QX201801.

ABSTRACT Convolutional neural network (CNN)-based deep learning techniques have enjoyed many successful applications in the field of medical imaging. However, the complicated between-manifold projection from the projection domain to the spatial domain hinders the direct application of CNN techniques in computed tomography (CT) reconstruction. In this work, we proposed a novel CT reconstruction framework based on a CNN, i.e., an intelligent back-projection network (iBP-Net). The proposed iBP-Net method fused a pre-CNN, a back-projection layer, and a post-CNN into an end-to-end network. The pre-CNN adopted CNN techniques to model a filtering operation in the projection domain. In the back-projection layer, a back-projection operation was employed to perform between-manifold projection. Based on CNN techniques, the post-CNN worked together with the pre-CNN to recover reconstructed images from the outputs of the back-projection layer in the spatial domain while maintaining high visual sensitivity. The experimental results demonstrate the feasibility of the proposed iBP-Net framework for CT reconstruction.

INDEX TERMS Convolutional neural network, deep learning, between-manifold projection, CT reconstruction.

I. INTRODUCTION

X-ray computed tomography (CT) is one of the most valuable and widely used imaging techniques in clinical, industrial, and other applications [1]. The main problem in tomography is the process of reconstructing unknown images from their projections. In the past 30 years, with the wide application of tomography techniques in many fields, researchers have dedicated major efforts to the development of CT image reconstruction methods. Because CT image reconstruction is a typical example of an ill-posed inverse problem, it is difficult to find accurate reconstructions in practice. Current solutions for improving the quality of reconstructions can be roughly divided into three categories: analytical methods, iterative reconstruction (IR) methods and artificial neural network methods.

An analytical method is a direct reconstruction method based on the central slice theorem, which provides the relationship between the Radon transform of an object and its two-dimensional Fourier transform. The filtered back-projection (FBP) algorithm [2] is a classical analytical method and is the most widely used reconstruction method in actual clinical applications. The FBP algorithm is capable of computing accurate reconstructions with high computational efficiency. However, its limitations are also very obvious. The quality of the reconstructed images degrades significantly when only a small number of projections can be acquired or when the obtained projection data are heavily polluted [3].

Compared with analytical methods, IR methods can consider the statistical characteristics of both the obtained projection data and the corresponding reconstructed images. The IR methods dramatically improve CT reconstruction results by incorporating some elaborate prior knowledge/regularizations, such as total variation (TV) [4], [5], nonlocal patches [6], dictionary learning [7], and sparsifying transform learning [8], [9]. The existing regularizations indeed guide IR algorithms to yield better reconstruction results. Nevertheless, to obtain promising reconstructions, the corresponding hyperparameters of the IR methods need to be tuned carefully. More importantly, the wide application of IR methods is greatly hindered by the computational loads of the projection and back-projection operations during each

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

iteration. Though the reconstruction time can be greatly reduced by parallelizing the computations [10], [11], the IR methods still take longer to reconstruct large volumes than the FBP algorithm.

Recently, artificial neural network methods have achieved tremendous success in the fields of computer vision [12] and natural language processing [13], especially those that incorporate deep learning (DL) approaches. DL techniques can efficiently exploit high-level features from the pixel level of an image through a hierarchical multilayer framework [14]. Lately, convolutional neural networks (CNNs) have become the most popular DL frameworks. The fast development of CNN techniques has accelerated many successful applications in the field of medical imaging. Inspired by these successful applications [15], several network architectures have also been proposed for CT reconstruction, leading to promising experimental results.

The direct application of CNN-based DL techniques in CT reconstruction is greatly hindered by the complicated between-manifold projections from the sensor to the spatial domain. To address this problem, the stacked fully connected layer technique can be adopted to perform a domain transform from the feature space of the extracted data to image space [16], [17]. CNNs are utilized only to process the projection data or improve the reconstructed images. However, the application of the stacked fully connected layers in network architectures dramatically increases the number of utilized parameters, which results in a substantial computational cost for neural network training.

To avoid the application of stacked fully connected layers for between-manifold projections, researchers have tried to apply DL to the reconstruction problem by unrolling an iterative algorithm to a fixed N-step iterative algorithm. On the one hand, some regularizations in the objective functions of IR algorithms can be directly replaced by neural networks [18], [19]. On the other hand, an objective function can be divided into several subobjective functions according to operator splitting methods. The solutions of these subobjective functions are equivalent to those of the corresponding objective functions. Some of these subobjective functions can be solved with neural networks [20]-[23]. Such unrolling methods avoid the adoption of stacked fully connected layers in the network architectures and reduce the number of parameters used by these architectures. However, the complexity of such an algorithm is still much higher than that of an analytical method.

Different from the two kinds of DL methods above, a back-projection operation is applied to perform between-manifold projection in neural networks. Obtained by filtering the projection data with a special filter, the solution of back-projection is comparable to that of a certain IR algorithm. This special filter can be derived from a reformulation of an IR algorithm [24] or acquired by calculating the filter matrix with an IR algorithm [25]. Würfl *et al.* [26] attempted to learn the projection-domain weights of the FBP algorithm with a fully connected layer, and this algorithm can effectively balance the tradeoff between reconstruction performance and speed. He *et al.* [27] utilized two fully connected layers to model the filter of the FBP algorithm and further improved the overall algorithmic performance with a common network architecture. Because there are no ideal filtered sinograms, special filters cannot be directly learned through a common neural network. In the two above methods, via combination with a back-projection layer, the special filters of the FBP algorithm are obtained from big data with DL approaches. These two methods avoid the application of a fully connected layer in the between-manifold projection [16], [17]. However, the fully connected layer is also involved in the neural network to filter projection data, significantly increasing the difficulty of neural network training. Yin *et al.* [28] and Wang *et al.* [29] proposed the similar network architectures. These two neural networks are not end-to-end networks. Their special filters are learned from the ''ideal filtered sinograms'', which are the full-dose projections, or are generated from tomography images with a Radon transform. Thus, the latter two algorithms cannot yield optimal filters, thereby limiting the overall performance of CT reconstruction.

In this study, in order to reduce the difficulty of neural network training and build an end-to-end network, we proposed an intelligent back-projection network (iBP-Net) to reconstruct CT images. The proposed iBP-Net consists of three components, i.e., a pre-CNN, a back-projection layer, and a post-CNN. In the pre-CNN, a CNN was directly employed to filter the obtained projection data in the projection domain. The back-projection layer was utilized to perform a domain transform from the projection space to the image space. In the post-CNN, a CNN was applied to recover the desired images from the outputs of the back-projection layer. Our main contributions in this paper are summarized as follows:

1) Two CNNs were applied to build the proposed iBP-Net method for CT reconstruction. The proposed method avoids incorporating fully connected layers, thereby lightening the workload of neural network training.

2) A new neural network architecture was proposed for CT reconstruction. A pre-CNN and post-CNN work together to generate feasible reconstruction images. The post-CNN is more than a denoiser; it further improves the outputs of the back-projection layer. More importantly, the post-CNN was adopted to extract detailed information from the outputs of the back-projection layer, contributing to the reconstruction of feasible images.

3) The proposed iBP-Net model is an end-to-end network. All parameters in the iBP-Net model are directly learned from the big data. Given its excellent end-to-end learning capabilities, the proposed iBP-Net method can achieve promising performance in CT reconstruction.

Clinical patient data were used to evaluate the performance of iBP-Net. The experimental results demonstrate that iBP-Net possesses a promising reconstruction ability.

II. METHODS AND MATERIALS

A. DATA-DEPENDENT FBP METHOD

The present study employed a fan-beam geometry in a 2D CT scanning setting. The projection data y were acquired only for a finite set of N_θ projections and N_d detector elements, and the scanned object x was represented by a pixel grid of $N \times N$ pixels. Typically, the CT reconstruction problem can be formulated as solving a linear system:

$$Ax = y, \tag{1}$$

where A is the system matrix of $N_\theta N_\mathrm{d} \times NN$ elements corresponding to a specific configuration of the CT system. The purpose of CT reconstruction is to recover the unknown x from the system matrix A and the projection data y. Mathematically, a_{ij}, the element of A, stands for the intersection of the i-th X-ray path with the j-th pixel. In this context, the product of the system matrix A with an image x is called a forward projection, and the product of the back-projection matrix A^{T} with the projection data y is called a backward projection. Since the system matrix can be extremely large for typical problem sizes in practice, the matrix is usually not computed explicitly. Instead, each multiplication of a vector with A or A^{T} is computed on the fly [1].

The CT reconstruction problem in (1) can be analytically inverted with FBP. The FBP method starts by filtering the projection data y with a filter h. Afterwards, the filtered results are back projected to obtain the reconstructed image. The operation of the FBP algorithm can be modeled as a linear operator applied to the projection data y, and this can be written as:

$$x = FBP(y) = A^T F_h y, \tag{2}$$

where $F_{\mathrm{h}}y$ is the convolution of y by h, i.e., the filtering of the projection data y can also be modeled as a linear operation with the matrix F_{h}. The FBP method is the most widely used approach because it is computationally efficient. However, the images reconstructed with FBP suffer greatly from image artifacts and significantly degraded image quality when the number of projections is small compared to the size of the image or when noise is present in the projection data [25].

An IR method is a different type of approach for solving the CT reconstruction problem in (1). Based on iteratively solving the discrete linear system in (1), a popular class of IR algorithms tries to find images that minimize a data fidelity term and the suitable regularizers $R(x)$. The data fidelity term represents the difference between the acquired data and simulated projections of the reconstructed images, and the regularizers penalize images that do not fit the chosen prior knowledge about the scanned object. In this setting, the inverse problem is formally expressed as:

$$x^* = \underset{x}{\operatorname{argmin}} \|y - Ax\|_2^2 + \lambda R(x), \tag{3}$$

where λ controls the relative weighting between the data fidelity term and the prior knowledge penalty. A popular choice for determining $R(x)$ is TV minimization [4]. As they are capable of easily accommodating imaging physics and prior knowledge, IR algorithms can retrieve feasible images and outperform the FBP method. However, the main disadvantage of IR algorithms is their high computational costs, which hinder their extensive use in practice.

To balance the tradeoff between promising reconstruction performance and high reconstruction speed, major efforts have been dedicated to incorporating the advantages of both analytical methods and IR methods [24], [25]. For a given set of projection data y and a filter h, the reconstructed image x can be acquired by the FBP method in (2). By exchanging the positions of y and h, (2) can be rewritten as:

$$x = A^T F_y h, \tag{4}$$

where the matrix F_{y} is made up of the projected data y. Substituting (4) into (3) gives:

$$h^* = \underset{x}{\operatorname{argmin}} \left\| y - AA^{\mathrm{T}} F_y h \right\|_2^2 + \lambda R(A^{\mathrm{T}} F_y h). \tag{5}$$

After this substitution, the minimization problem in (3) turns into a problem of solving the optimal filter in (5). For a given system matrix A and projection data y, there is a linear relationship between the reconstructed images x and the filter h in (4). Thus, in theory, the minimization problem of solving the reconstruction images x in (3) is equivalent to the minimization problem of solving the filter h in (5). After filtering the projection data with the solved filter h^* in (5), the reconstructed image x is directly acquired by a back-projection operation. As a result, the reconstruction performance of such an FBP algorithm is similar to that of an IR model, while the corresponding reconstruction speed is much faster. However, although the time required to compute the filter h^* in (5) is close to that of a full iterative reconstruction process, the filter h^* can be precomputed and reused on problems with identical experimental setups.

B. OVERALL NETWORK ARCHITECTURE

As discussed above, after processing the projection data with a specific filter, the solution of the FBP method is comparable to those of IR methods. However, the wide applications of the data-dependent FBP method are severely hindered by the use of precomputed filters for all experimental setups. In this work, to cope with this problem, the iBP-Net method was proposed to obtain the optimal filter for a given problem by using a CNN. By taking the advantages of big data-driven approaches, the learned filter is suitable for general CT reconstruction setups. The overall architecture of the proposed iBP-Net is shown in Fig. 1(a). The proposed iBP-Net architecture consists of three components: a pre-CNN, a back-projection layer, and a post-CNN. In the proposed iBP-Net, the input includes the projection data, and the output is the reconstructed image. The projection data are filtered in the pre-CNN, as shown in Fig. 1(b). Using the back-projection layer, the projection domain is connected to the image domain. To generate realistic results, the post-CNN extracts the subtle structural features from the outputs of

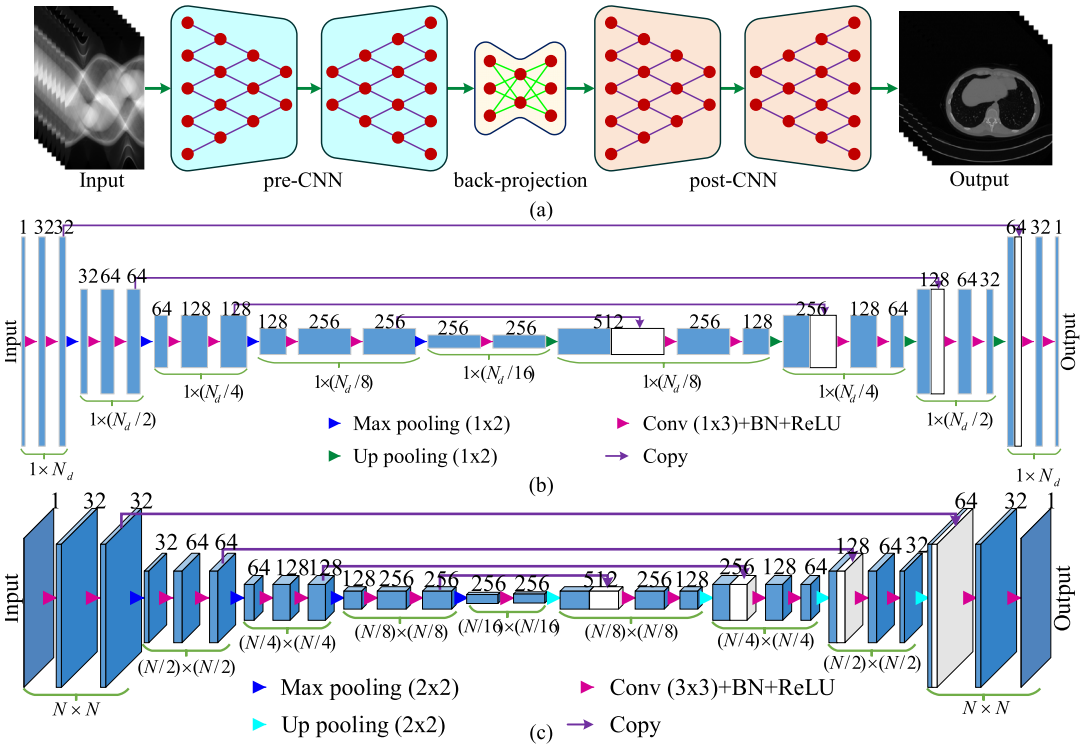


FIGURE 1. Network architecture of iBP-Net. (a) Overall architecture of iBP-Net, (b) the pre-CNN architecture, and (c) the post-CNN architecture.

the back-projection layer, as shown in Fig. 1(c). With the back-projection layer, the loss of the iBP-Net can be easily back propagated to update the weights in the pre-CNN. Thus, iBP-Net is an end-to-end neural network. All parameters in iBP-Net are directly determined by the neural network training process with the big data.

C. PRE-CNN ARCHITECTURE

In the FBP algorithm, a high-pass filter, such as the Ramp filter, is usually adopted to process the acquired projection data. The value of an element in a filtered dataset is calculated by all projection data in the corresponding view. To model this high-pass filter, the fully connected technique is employed for the filtering operation [26], [27] because connections exist from all inputs to each output. Adopting a fully connected structure dramatically increases the difficulty of neural network training. In the X-ray imaging process, an X-ray passes through the phantom and deposits its energy both in the human body and on the detector. Most of the energy deposited on the detector is deposited in the detector bins around the intersection point between the X-ray and the detector due to X-ray scattering. Therefore, in a filtering operation, it is unnecessary to connect all inputs to one output. In this work, a CNN-based network with a large receptive field was applied to model the filtering operation in the projection domain. For a large receptive field, an X-ray can be connected by the neural network to all detector bins in which this X-ray deposits energy. Theoretically, the CNN technique is capable of filtering projection data effectively. Moreover, a neural network based on a CNN architecture is much easier to train than one based on fully connected layers [14].

As the network depth increases, the complexity of the neural network is increased and the problem of vanishing gradients is increasingly likely to occur during neural network training, leading to the loss of network feature extraction ability. To address these problems, a pre-CNN based on U-Net [30] was proposed to learn the optimal filter for a given projection domain. One of the most important characteristics of U-Net is its exponentially large receptive field. As observed in Fig. 1(b), the pre-CNN architecture consists of an encoder subnetwork (left half) and a decoder subnetwork (right half), which are connected by a bridge section. The encoder and decoder subnetworks in the U-Net architecture consists of multiple stages. In our work, we set the number of stages, i.e., the number of encoder and decoder subnetworks used in U-Net, to 4. Specifically, each stage contained 4 sequential layers composed of convolutions with 1×3 kernels, batch normalization (BN), and (rectified linear unit) ReLU layers. After each stage of the encoder subnetwork, there is a max pooling operation (''Max pooling'') with a 1×2 window and 1×2 stride, which mainly realizes feature downsampling and increases the size of the receptive field. The number of channels is doubled after each pooling layer. After each stage of the decoder subnetwork, there is an upconvolution operation (''Up pooling'') with a 1×2 window and 1×2 stride. The upconvolution operation is mainly utilized to double the size of each feature map. To learn more features across different layers and effectively deal with the gradient vanishing problem in deeper network training, U-Net [30]

adopts skip connections [31] in some of its symmetric layers rather than in adjacent layers, as shown in Fig. 1(b).

D. BACK-PROJECTION LAYER

In (2), the reconstructed image x can be directly acquired by the multiplication of the filtered projection data y and the back-projection matrix A^{T}. Theoretically, A^{T} can be easily mapped to a neural network with the fully connected technique. However, for modest CT reconstruction problem sizes, an extremely massive number of parameters is involved in the fully connected layer, and the parameter size easily reaches several terabytes, making it infeasible to implement and train such a model [26], [27]. In our work, to address this problem, we computed A^{T} and A on the fly and constructed a new back-projection layer without adjustable parameters. In terms of a neural network, the forward pass operation of such a layer can be expressed as:

$$x_{l+1} = A^{\mathrm{T}} x_l, \tag{6}$$

where the vector x_{l+1} is the output of this layer in the image domain, while the input x_l is the output of the pre-CNN in the projection domain. To make the system capable of updating the parameters of the pre-CNN during neural network training, the error e_l from the post-CNN passes backward to the pre-CNN by back propagation:

$$e_{l-1} - = A e_l, \tag{7}$$

where e_{l-1} denotes the error back propagated to the pre-CNN. To enhance the performance of iBP-Net, A^{T} and A are computed with the unmatched projectors [32]. In this work, the forward projector was implemented in a ray-driven and back-projector pixel-driven manner.

E. POST-CNN ARCHITECTURE

As displayed in Fig. 1(c), the post-CNN, a neural network structure, was added after the back-projection layer to recover the reconstructed images from the outputs of the back-projection layer. Compared to that of the pre-CNN, a similar U-Net architecture was applied to the post-CNN in this work. As observed in Fig. 1(c), the post-CNN architecture consists of an encoder subnetwork (left half) and a decoder subnetwork (right half). The number of stages in the encoder and decoder subnetworks was set to 4. Each stage contained 4 sequential layers composed of convolutions with 3×3 kernels, BN, and ReLU layers. The convolution kernels were applied to obtain the features from the acquired feature maps. After each stage of the encoder subnetwork, there is a max pooling operation (''Max pooling'') with a 2×2 window and 2×2 stride. After each stage of the decoder subnetwork, there is an upconvolution operation (''Up pooling'') with a 2×2 window and 2×2 stride. In the selected U-Net, a skip connection (in Fig. 1(c)) was also employed to alleviate the vanishing gradient problem and improve the optimal convergence speed of the system. The input image of the post-CNN is the output of the back-projection layer, and the output image is the reconstructed image. Both the input and output images are in the image domain. Because of the high sensitivity of the network to features with different scales, the post-CNN is capable of capturing subtle structural features from the outputs of the back-projection layer. By cooperating with the pre-CNN, the post-CNN can generate feasible reconstructed images while maintaining high visual sensitivity. Due to its strong feature extraction performance, the post-CNN also contributes to shortening the computational time required to find the optimal hyperparameters for neural network training.

III. EXPERIMENTS

A. DATA

In this section, CT data provided by the AAPM Low Dose CT Grand Challenge [33] were used for evaluation purposes. The provided CT data included both raw projection data and reconstructed images from 20 different patients who were scanned from the chest under a normal dose of radiation. The raw projection data cannot be utilized directly for fan-beam CT reconstruction because it was acquired using a helix trajectory. In this work, the projection data were obtained by a forward projection of the provided reference volumes, which were reconstructed from the full-dose projections. We performed forward projections for 360, 180 and 90 degree views. In these three different scanning modes, the corresponding angular increments between every projection were 1°, 2° and 4°, respectively. The linear detector had 768 bins, and the space between two adjacent bins was 1 mm. The dimensionality of the reconstructed image was 512×512, and the voxel size was 0.5859 mm $\times$ 0.5859 mm. The distance from the X-ray source to the detector arrays was 1068.0 mm, and the distance from the X-ray source to the center of rotation was 595.0 mm. In total, we performed forward projection for 3,570 reconstructed images. The images of 19 patients were used for training iBP-Net, and the images of the last patient were used for evaluating the performance of the neural network.

B. IMPLEMENTATION DETAILS

We chose to implement the iBP-Net framework on the PyTorch DL toolbox [34]. The iBP-Net model is an end-to-end trainable neural network. The loss function of iBP-Net is the Euclidean loss function, which is the mean square error between the reconstructed image x and the reference x^*:

$$L - \left\| x - x^* \right\|_2^2. \tag{8}$$

Once the architecture of the proposed iBP-Net was configured, its parameters could be learned by optimizing the loss function. In our work, the loss function was optimized by the Adam algorithm [35], where the base learning rate was set to 10^{-4}, and it slowly decreased to 10^{-5}. The computer platform was configured as follows: the CPU was an Intel(R) Core (TM) i9-10900K at 3.70 GHz; the GPU was an NVIDIA RTX 2080 Ti with 11 GB of memory.

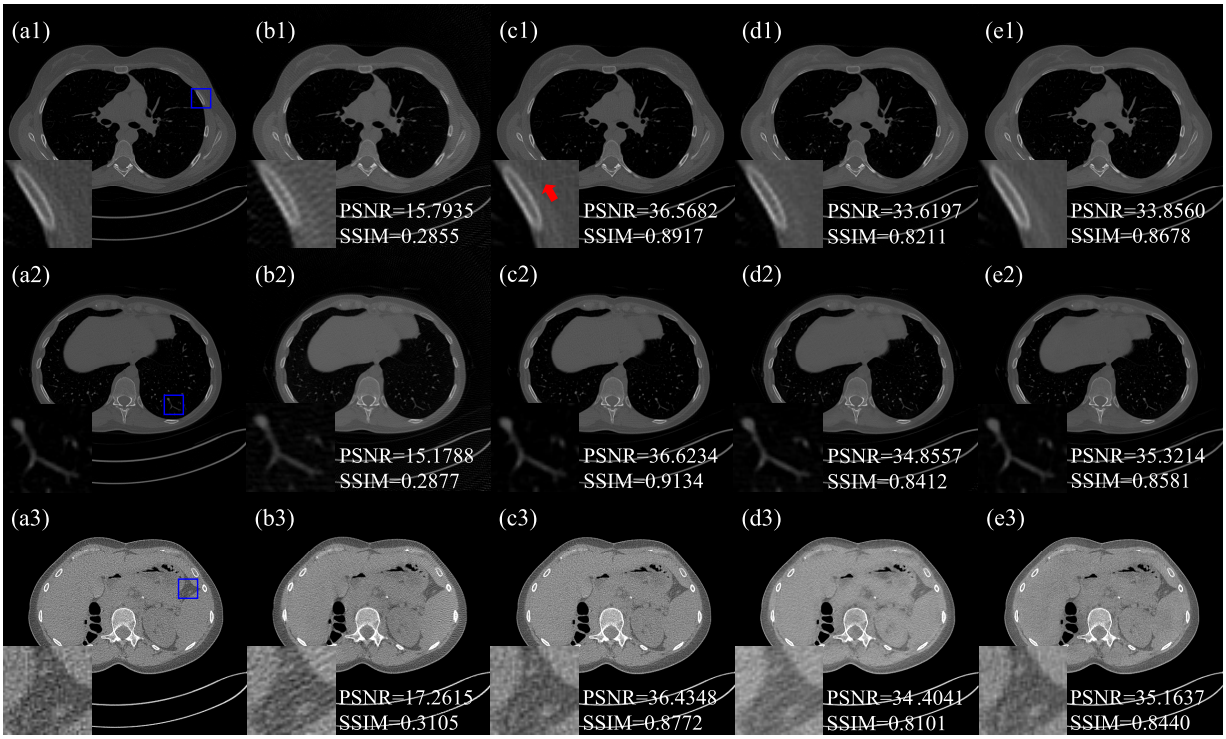


FIGURE 2. Reconstruction results obtained on the AAPM challenge dataset (chest subdata) under the 360-degree view scanning condition. (a1)-(a3) Ground truths; (b1)-(b3) images obtained with the FBP method; (c1)-(c3) images obtained with the TV method; (d1)-(d3) images obtained with the RED-CNN method; (e1)-(e3) images obtained with the iBP-Net method. The display window is [−900, 2100] HU for the images in the first two rows and [−360 440] HU for the images in the last row.

C. EVALUATION METRICS

We evaluated the quantitative performance of iBP-Net and several comparison models by adopting three popular indexes, i.e., the peak signal-to-noise ratio (PSNR), normalized mean square error (NMSE), and structural similarity index measure (SSIM) [36]. The PSNR and NMSE are the most widely adopted quantitative indexes for assessing CT reconstruction results, and they are written as follows:

$$PSNR(x,x^*) = 10 \times \log\left(\max\left(x \cdot x\right) / \left\|x - x^*\right\|_2^2\right), \quad (9)$$

$$NMSE(x,x^*) = \left\|x - x^*\right\|_2 / \left\|x^*\right\|_2, \quad (10)$$

where $\cdot$ denotes elementwise multiplication. Combining a luminance measure, a contrast measure, and a structure measure, the SSIM measures the structural similarity between two images, and it is capable of evaluating the degree of artifacts in a given image. The SSIM is defined as follows:

$$SSIM(x,x^*) = \frac{\left(2\mu_x\mu_{x^*} + \alpha_1\right)\left(2\sigma_x\sigma_{x^*} + \alpha_2\right)}{\left(\mu_x^2 + \mu_{x^*}^2 + \alpha_1\right)\left(\sigma_x^2 + \sigma_{x^*}^2 + \alpha_2\right)}, \quad (11)$$

where μ_x denotes the mean value of x, σ_x denotes the variance of x, and similar properties are defined for the reference image x^*. In (11), α_1 and α_2 are two constants that are used to stabilize a division operation with a weak denominator. The SSIM index lies between 0 and 1, and a higher value represents better image quality.

D. REFERENCE METHODS

To evaluate the effectiveness of the proposed algorithm, the iBP-Net model was compared with three different methods, i.e., FBP reconstruction (Ram-Lak filter), TV iterative reconstruction [4], and the RED-CNN network for CT image processing [37]. We executed the TV method with 40 iterations on the Tomographic Iterative GPU-based Reconstruction (TIGRE) toolbox with the default parameters [38]. RED-CNN is a DL-based denoiser that operates in the spatial domain. In this work, the RED-CNN network was applied to further improve the obtained FBP reconstruction results. The training settings of RED-CNN were set according to the original paper [37].

IV. RESULTS

A. RESULTS FOR THE 360-DEGREE VIEWS

To show the effectiveness of the proposed network, we performed CT reconstruction under different sparse-view scanning conditions. Fig. 2 demonstrates the reconstruction performances of different methods under the 360-degree view condition. The three rows show three reconstruction cases. The first column shows the corresponding ground truths. The second, third, and fourth columns show the results yielded by the FBP, TV, and RED-CNN methods, respectively, for comparison purposes. The fifth column shows the iBP-Net results. In Fig. 2, the images in the regions of interest (ROIs) specified by the blue rectangles are zoomed in and displayed

TABLE 1. Quantitative Results (Mean ± STD) Associated with Different Methods Regarding the Reconstructed images.

Views	Metrics	FBP	TV	RED-CNN	iBP-Net
360	PSNR	15.1679±0.8019	**35.9543±1.2075**	33.5626±0.8677	34.9063±1.1199
	SSIM	0.2770±0.0160	**0.8895±0.0136**	0.8117±0.0182	0.8716±0.0174
	NMSE	0.7219±0.2187	**0.0059±0.0014**	0.0104±0.0037	0.0077±0.0026
180	PSNR	12.1267±0.6012	32.2591±1.4086	31.3470±0.6287	**33.9521±0.9449**
	SSIM	0.2002±0.0144	**0.8352±0.0185**	0.7383±0.0131	0.8334±0.0220
	NMSE	1.4419±0.3970	0.0139±0.0042	0.0173±0.0048	**0.0094±0.0026**
90	PSNR	9.4686±0.5390	26.2686±1.3527	28.9445±1.8008	**32.6016±0.8296**
	SSIM	0.1472±0.0153	0.7324±0.0215	0.6619±0.0455	**0.7934±0.0244**
	NMSE	2.6436±0.7031	0.0558±0.0178	0.0344±0.0347	**0.0130±0.0038**

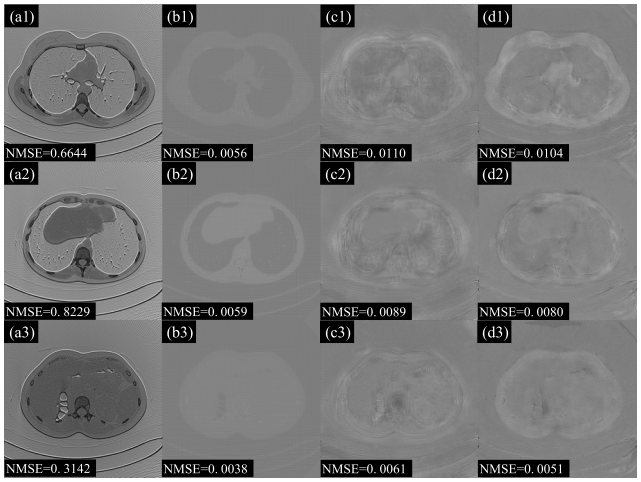


FIGURE 3. Residual images yielded by different methods under the 360-degree view scanning condition. The three rows show three reconstruction cases in Fig. 2. From (a) to (d): FBP, TV, RED-CNN and iBP-Net. The display window is [−512, 512] HU.

in the lower left corners of the corresponding full images. The PSNR and SSIM indexes of the full images are listed in the lower right corners of the corresponding images. In addition, the residual images between the reconstructed images and the references are presented in Fig. 3. The NMSE metrics are listed in the lower left corners of the corresponding residual images. It can be observed in Fig. 2 and Fig. 3 that the traditional FBP reconstruction approach leads to obviously amplified noise and artifacts, which may severely affect the diagnostic accuracy of method. The IR-type TV reconstruction algorithm works well in terms effective noise suppression and structural feature preservation. However, some tiny streak artifacts can be observed in the reconstructed images, as indicated by the red arrow in Fig. 2(c1). The network-based RED-CNN method can suppress noise and remove artifacts to some extent. The resulting images still suffer from blurred details and fake structures (as indicated in Fig. 2(d1)-(d3)). The proposed iBP-Net method exhibits better anatomical feature preservation, as shown in Fig. 2(e1)-(e3). Upon visual inspection, the proposed iBP-Net method outperforms the other three state-of-the-art methods due to its robust noise suppression, artifact removal, and texture preservation performances. However, the evaluation metrics of the proposed iBP-Net method are slightly inferior to those of the TV method, as shown in Fig. 2 and Fig. 3.

B. RESULTS FOR THE 180-DEGREE VIEWS

As we know, the evaluation metrics and the visual quality of the reconstructed images may be degraded with a decrease in the number of projections. Fig. 4 demonstrates the reconstruction performances obtained under the 180-degree view condition. The corresponding residual images are displayed in Fig. 5. As observed in Fig. 4 and Fig. 5, the traditional FBP algorithm obviously increases the noise and streak artifacts in the reconstructed images, resulting in the loss of projection data. In contrast, while the TV algorithm can effectively reduce noise and artifacts, the resulting images suffer greatly from blurred details and streak artifacts (as indicated in Fig. 4(c1) and Fig. 5(b1)). As indicated in Fig. 4(d1)-(c3), the RED-CNN method severely smooths out some of the anatomical details and introduces some fake structures. Referring to the ground truth, the proposed method effectively preserves the structural and textural information in the reconstructed images. The proposed method outperforms competing methods in terms of image quality metrics (as shown in Fig. 4 and Fig. 5) and visual sensitivity.

C. RESULTS FOR THE 90-DEGREE VIEWS

With the further reduction in the amount of projection data, the quality of the reconstructed images is severely degraded. Fig. 6 demonstrates the reconstruction performances obtained under the 90-degree view condition. The corresponding residual images are displayed in Fig. 7. As displayed in Fig. 6 and Fig. 7, the reconstructed images produced by the three competing methods suffer from the loss of subtle structural characteristics, streak artifacts, and noise features. The proposed iBP-Net method also blurs some content and leads to the loss of structural details under the 90-degree view scanning condition. However, the iBP-Net method better preserves the underlying structural information and achieves higher scores on the evaluation metrics than those of the competing methods (in Fig. 6 and Fig. 7).

D. QUANTITATIVE ANALYSIS

In Table 1, the PSNR, NMSE, and SSIM indexes are given in terms of the Means ± SDs (average scores ± standard

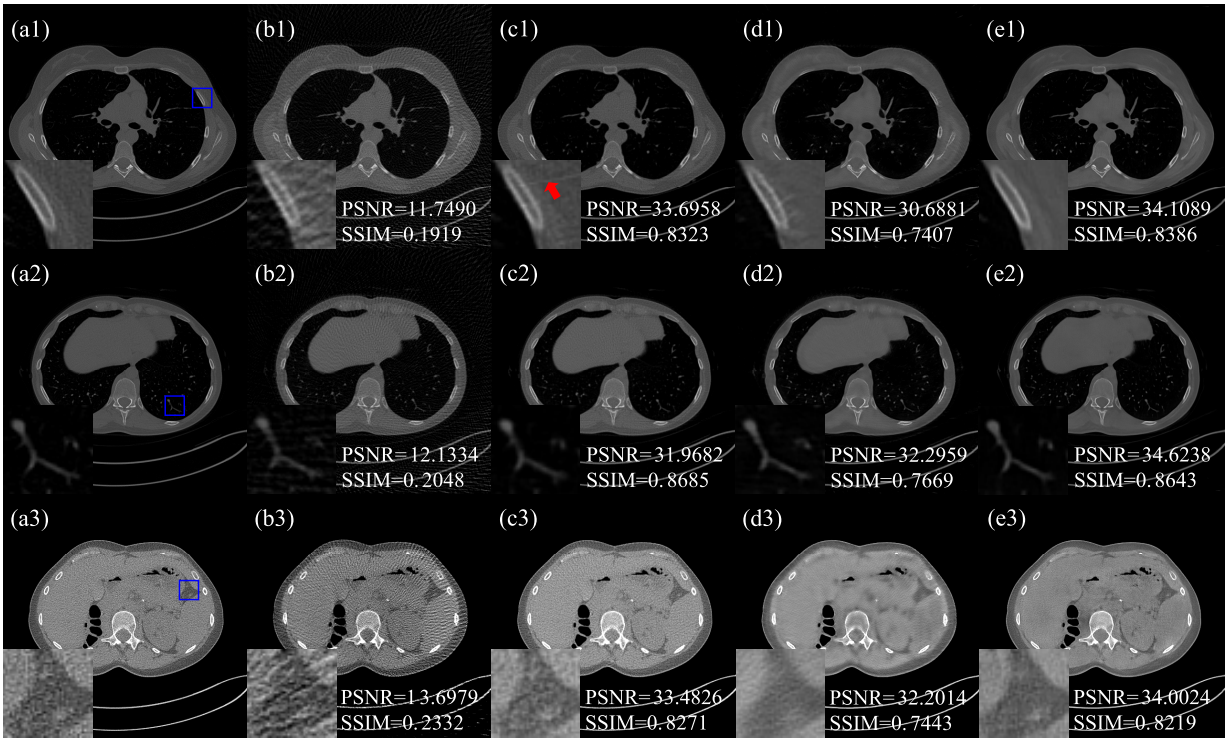


FIGURE 4. Reconstruction results obtained on the AAPM challenge data (chest subdata) under the 180-degree view scanning condition. (a1)-(a3) Ground truths; (b1)-(b3) images obtained with FBP method; (c1)-(c3) images obtained with the TV method; (d1)-(d3) images obtained with the RED-CNN method; (e1)-(e3) images obtained with the iBP-Net method. The display window is [−900, 2100] HU for the images in the first two rows and [−360 440] HU for the images in the last row.

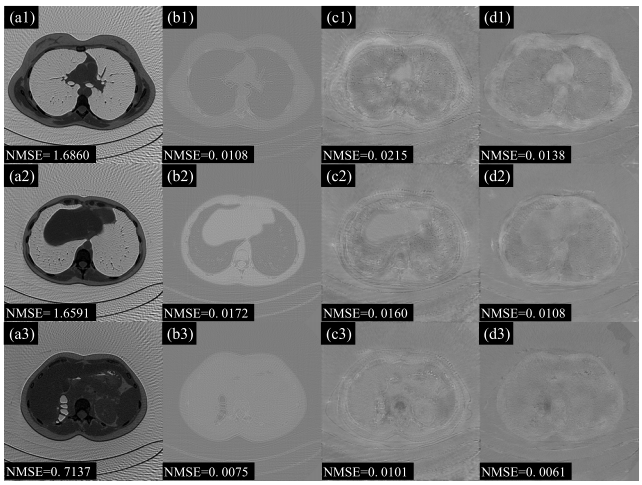


FIGURE 5. Residual images yielded by different methods under the 180-degree view scanning condition. The three rows show three reconstruction cases in Fig. 4. From (a) to (d): FBP, TV, RED-CNN and iBP-Net. The display window is [−512, 512] HU.

deviations) for the images reconstructed by the different methods mentioned in this work. The total number of test images obtained from one patient was 195. It can be seen in Table 1 that the FBP method obtains the lowest SSIM and PSNR scores. The TV method achieves the highest scores under the 360-degree view scanning condition. However, while the performance of the proposed iBP-Net method slowly worsens as the number of projections decreases, this method achieves higher scores than those of the other three competing methods under both the 180-degree view and 90-degree view scanning conditions. To some extent, this quantitative analysis is consistent with the visual performance results in Figs. 2-7.

E. EVALUATING AND ANALYZING THE NETWORKS

To evaluate the CT reconstruction performance of the proposed network, iBP-Net method was applied for CT reconstruction under different view conditions, i.e., 360, 180, and 90 degrees. The convergence curves of the iBP-Net training process are presented in Fig. 8, and the resulting network after 160000 training iterations is used in this work. It can be observed that the energy of the objective function decreases with the increase in the number of training iterations. A runtime comparison between the tested reconstruction methods under three different scanning conditions is given in the Table 2. As observed, the runtime is shortest for the FBP method. The RED-CNN method is slightly faster than the proposed iBP-Net method because fewer convolutional layers are involved in RED-CNN. The TV method is the slowest among the four reconstruction methods.

The proposed iBP-Net consists of three components: a pre-CNN, a back-projection layer, and a post-CNN. If any of the components are removed or replaced, the network may

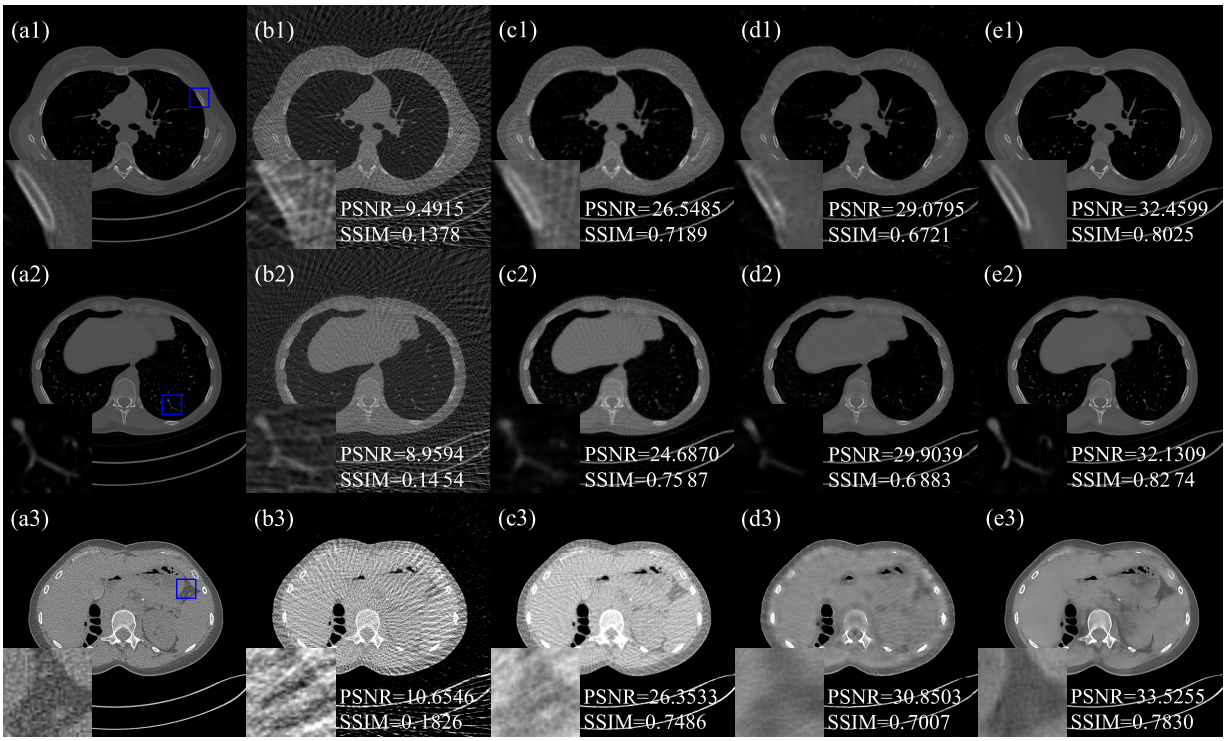


FIGURE 6. Reconstruction results obtained on the AAPM challenge data (chest subdata) under the 90-degree view scanning condition. (a1)-(a3) Ground truths; (b1)-(b3) images obtained with the FBP method; (c1)-(c3) images obtained with the TV method; (d1)-(d3) images obtained with the RED-CNN method; (e1)-(e3) images obtained with the iBP-Net method. The display window is [−900, 2100] HU for the images in the first two rows and [−360 440] HU for the images in the last row.

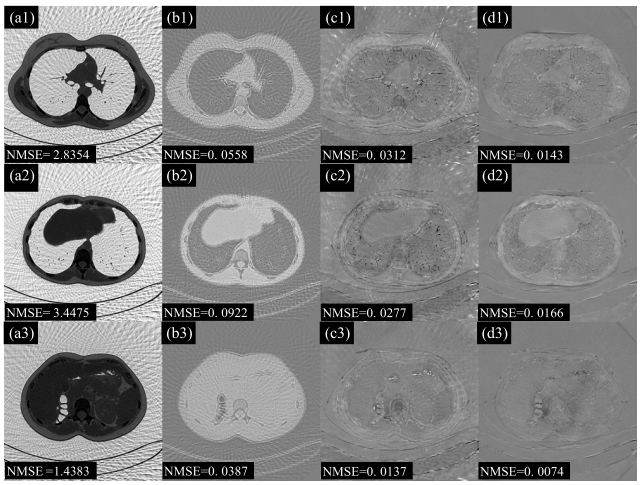


FIGURE 7. Residual images yielded by different methods under the 180-degree view scanning condition. The three rows show three reconstruction cases in Fig. 6. From (a) to (d): FBP, TV, RED-CNN and iBP-Net. The display window is [−512, 512] HU.

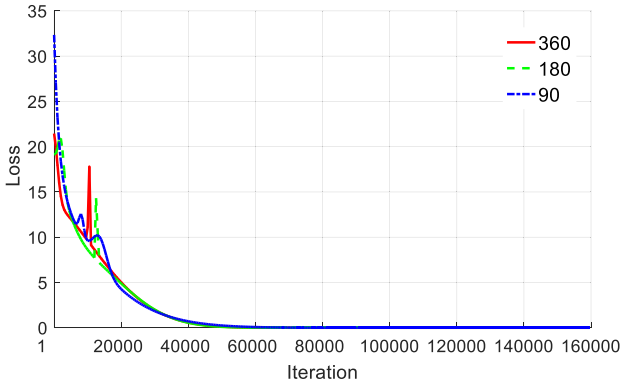


FIGURE 8. Training losses incurred under different scanning conditions.

TABLE 2. Comparison of reconstruction methods in terms of runtime (seconds).

Views	FBP	TV	RED-CNN	iBP-Net
360	0.0206	1963	0.0517	0.0564
180	0.0147	967	0.0302	0.0374
90	0.0093	495	0.0245	0.0288

perform poorly in terms of reconstruction and lead to severe image quality degradation. By reassembling these three components, we can obtain the following new reconstruction methods. The 'FBP' method is obtained by replacing the pre-CNN with a Ram-Lak filter and removing the post-CNN. The 'pre-CNN+BP' method is produced by removing the post-CNN. The 'FBP+post-CNN' method is obtained by replacing the pre-CNN with a Ram-Lak filter. Fig. 9 presents the reconstruction results of these new methods under the 360-degree view scanning condition. As observed, the results

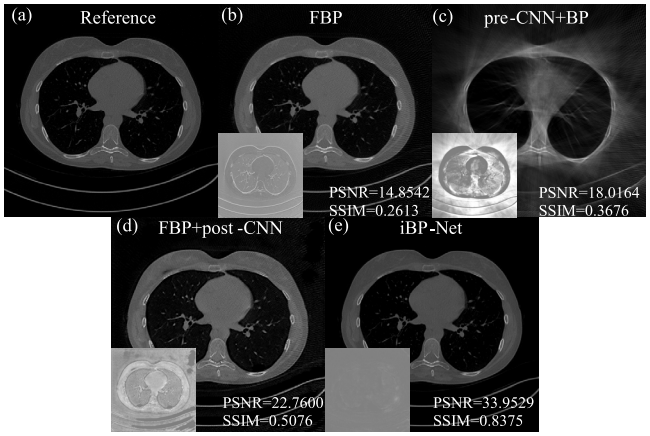


FIGURE 9. Reconstruction performances of different combinations of the three components of the proposed iBP-Net framework. The display window is [−900, 2100] HU.

of the 'FBP' method are contaminated by streak artifacts. The 'pre-CNN+BP' method yields the worst results in terms of severe streak artifacts and fake structures in the reconstructed images. The reconstruction results of 'FBP+post-CNN' method suffer from severe fake structures. The proposed iBP-Net method outperforms the other three new reconstruction methods. The reconstruction performances of the above new reconstruction methods indicate the feasibility of the combination chosen for the proposed iBP-Net framework.

V. DISCUSSION AND CONCLUSION

In this work, we proposed a novel CT reconstruction framework based on CNN, named iBP-Net. Inspired by the data-dependent FBP method [25] and DL-based FBP reconstruction paradigms [26]–[29], the proposed iBP-Net method adopted the CNN technique to model a filtering operation in the projection domain. Between-manifold projection was conducted by a back-projection operation, thereby avoiding the application of fully connected layers. By extracting detailed information from the outputs of the back-projection layer, the post-CNN yielded promising results with high scores in terms of quantitative metrics and good visual quality for the resulting images. The experimental results with clinical patient data demonstrated the feasibility of iBP-Net for CT reconstruction tasks.

For the data-dependent FBP method [25], its wide application in general CT reconstruction tasks is limited by the application of a data-dependent filter with the same experimental setup. In this work, to address this limitation, CNN-based techniques were applied to learn the optimal filter for a given situation from big data. The learned filter is no longer limited by the chosen experimental setup. Compared to the filter learned from a neural network based on a fully connected layer [26], [27], the proposed iBP-Net method greatly reduces the difficulty of the neural network training process. To effectively learn optimal filters by using DL techniques, the proposed iBP-Net method fused a pre-CNN architecture, a back-projection layer, and a post-CNN architecture into an end-to-end network. Compared to non-end-to-end networks [28], [29], the proposed iBP-Net method provides a strict mathematical derivation and learns the optimal filter directly from big data.

Compared to IR algorithms [4] (as shown in Figs. 2-7), the proposed iBP-Net method is capable of preserving more underlying structural details and achieving images with better visual quality. This advantage is increasingly obvious as the amount of projection data decreases. However, some tiny structures in the reconstructed images are also smoothed, due to the loss of projection data. DL approaches using neural networks with large receptive fields, such as U-Net, have demonstrated impressive performance in sparse-data CT reconstruction [39]. However, the existing U-Net architecture does not satisfy the frame condition, and it overly emphasizes the low-frequency component of a given signal [40]. Further work will be devoted to investigating a more sophisticated network and the perceptual loss incurred by the neural network [41]; this superior network can maintain more detailed information in reconstructed images. Additionally, compared with IR algorithms, the proposed iBP-Net method does not adopt an iterative process to solve the reconstruction problem, and this significantly reduces the required computational time.

In the proposed iBP-Net method, both the pre-CNN and post-CNN adopted similar U-Net architectures. As described in section II.C, due to the characteristics of X-ray imaging, the U-Net architecture was employed to enlarge the receptive field of the pre-CNN. U-Net is the most widely used CNN architecture for combining deep learning with a postprocessing approach in sparse-data CT reconstruction [42]–[44]. It has many properties that make it well suited for artifact removal, especially when the number of views is extremely small [44]. Due to the enlarged receptive field, the U-Net can easily capture globally distributed streaking artifacts [42]. Because it involves the mapping of a back-projection operator, the proposed method, to some extent, exhibits the characteristics of FBP algorithms. Thus, in this work, the post-CNN leveraged the U-Net architecture to solve globally distributed few-view artifacts.

The proposed iBP-Net method was also tested on a dataset with distributions that were obviously different from those of the training set. The reconstruction results were contaminated by severe artifacts. This indicates that the proposed DL method relies heavily on the chosen training set. By enlarging the training datasets, the proposed method can also yield promising results.

Many groups have explored directly learning a tomographic mapping from sensor data to an underlying image, such as by using AUTOMAP [13] and FAR-Net [45]. In these methods, multiple fully connected layers are applied to learn the entire reconstruction process through a manifold encoding-decoding process. The main limitations of a neural network based on multiple fully connected layers are its critical dependency on big data and its expensive computational cost [14], [46]. In this work, to address this

problem, we proposed a CNN-based network for CT reconstruction, in which between-manifold projection was performed by a back-projection operation. Compared to a neural network based on fully connected layers, a CNN-based neural network is much easier to train [14].

For a fully 3D CT scan, because data in the projection domain contain more inner correlations, working only in the image domain might not be optimal [47]. Recently, some authors have focused on incorporating an analytical reconstruction-based operator to link data from the projection domain to the image domain and thereby conduct end-to-end network learning in a dual domain [47]–[50]. As mentioned in [24] and [25], a modified ramp filter can further improve the results of the conventional FBP method [2], especially for sparse-view CT reconstruction. Compared to the methods of the studies in [47]-[49], the proposed method is devoted to modifying the ramp filter and replacing it with a CNN. Compared to hdNet [50], the proposed method further improves the back-projection layer and the CNN applied on the projection data. Moreover, the theoretical derivation for replacing the ramp filter with a CNN is described in detail (in sections II.A and II.B). The network architecture of iBP-Net is similar to that of one unroll of an unrolling strategy [18]–[23]. However, the proposed method derives its reconstruction network architecture from the data-dependent FBP method rather than from an IR method. The proposed method significantly reduces the complexity of the reconstruction network compared to that of an unrolling strategy.

As mentioned in [27], a common CNN structure is added after the back-projection layer to further improve the overall performance of the system. However, in this work, the post-CNN plays a more important role in the proposed iBP-Net method. As observed in Fig. 9(c), without the post-CNN architecture, the reconstructed results would be contaminated seriously by severe streak artifacts and fake structures. As described in [51], good reconstruction results can be achieved by the 'pre-CNN+BP' method with a large receptive field and large training datasets. In this work, to decrease the complexity of the pre-CNN, the proposed method reduces the depth of the pre-CNN, which results in a reduction of the receptive field. Thus, for a small receptive field and small training datasets, it is difficult to obtain good reconstruction results (in Fig. 9(c)) using the 'pre-CNN+BP' method mentioned in this work. However, the pre-CNN and post-CNN in the proposed method work together to generate feasible reconstructed images.

ACKNOWLEDGMENT

The authors would like to thank Dr. C. McCollough, the Mayo Clinic, the American Association of Physicists in Medicine for providing the data used in this study.

REFERENCES

[1] D. J. Brenner and E. J. Hall, "Computed tomography—An increasing source of radiation exposure," *New England J. Med.*, vol. 357, no. 22, pp. 2277–2284, Nov. 2007.

[2] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. NS-21, no. 3, pp. 218–227, Jun. 1974.

[3] K. J. Batenburg and L. Plantagie, "Fast approximation of algebraic reconstruction methods for tomography," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3648–3658, Aug. 2012.

[4] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, no. 17, pp. 4777–4807, Aug. 2008.

[5] S. Niu, Y. Gao, Z. Bian, J. Huang, W. Chen, G. Yu, Z. Liang, and J. Ma, "Sparse-view X-ray CT reconstruction via total generalized variation regularization," *Phys. Med. Biol.*, vol. 59, no. 12, pp. 2997–3017, Jun. 2014.

[6] H. Zhang, J. Huang, J. Ma, Z. Bian, Q. Feng, H. Lu, Z. Liang, and W. Chen, "Iterative reconstruction for X-ray computed tomography using prior-image induced nonlocal regularization," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 9, pp. 2367–2378, Sep. 2014.

[7] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1682–1697, Sep. 2012.

[8] S. Ye, S. Ravishankar, Y. Long, and J. A. Fessler, "SPULTRA: Low-dose CT image reconstruction with joint statistical and learned image models," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 729–741, Mar. 2020.

[9] Z. Li, S. Ravishankar, Y. Long, and J. A. Fessler, "DECT-MULTRA: Dual-energy CT image decomposition with learned mixed material models and efficient clustering," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1223–1234, Apr. 2020.

[10] F. Xu and K. Mueller, "Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 3, pp. 654–663, Jun. 2005.

[11] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)," *J. Struct. Biol.*, vol. 176, no. 2, pp. 250–253, Nov. 2011.

[12] X. Zhou, W. Gong, W. Fu, and F. Du, "Application of deep learning in object detection," in *Proc. IEEE/ACIS 16th Int. Conf. Comput. Inf. Sci. (ICIS)*, May 2017, pp. 631–634.

[13] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017.

[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[15] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1289–1296, Jun. 2018.

[16] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018.

[17] Y. Li, K. Li, C. Zhang, J. Montoya, and G.-H. Chen, "Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2469–2481, Oct. 2019.

[18] D. Wu, K. Kim, G. E. Fakhri, and Q. Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2479–2486, Dec. 2017.

[19] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1333–1347, Jun. 2018.

[20] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018.

[21] J. Wang, L. Zeng, C. Wang, and Y. Guo, "ADMM-based deep reconstruction for limited-angle CT," *Phys. Med. Biol.*, vol. 64, no. 11, May 2019, Art. no. 115011.

[22] J. He, Y. Yang, Y. Wang, D. Zeng, Z. Bian, H. Zhang, J. Sun, Z. Xu, and J. Ma, "Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 371–382, Feb. 2019.

[23] Q. Ding, G. Chen, X. Zhang, Q. Huang, H. Ji, and H. Gao, "Low-dose CT with deep learning regularization via proximal forward–backward splitting," *Phys. Med. Biol.*, vol. 65, no. 12, Jun. 2020, Art. no. 125009.

[24] G. L. Zeng, "A filtered backprojection MAP algorithm with nonuniform sampling and noise modeling," *Med. Phys.*, vol. 39, no. 4, pp. 2170–2178, Mar. 2012.

[25] D. M. Pelt and K. J. Batenburg, "Improving filtered backprojection reconstruction by data-dependent filtering," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4750–4762, Nov. 2014.
[26] T. Wurfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, "Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1454–1463, Jun. 2018.
[27] J. He, Y. Wang, and J. Ma, "Radon inversion via deep learning," *IEEE Trans. Med. Imag.*, vol. 39, no. 6, pp. 2076–2087, Jun. 2020.
[28] X. Yin, J.-L. Coatrieux, Q. Zhao, J. Liu, W. Yang, J. Yang, G. Quan, Y. Chen, H. Shu, and L. Luo, "Domain progressive 3D residual convolution network to improve low-dose CT imaging," *IEEE Trans. Med. Imag.*, vol. 38, no. 12, pp. 2903–2913, Dec. 2019.
[29] W. Wang, X.-G. Xia, C. He, Z. Ren, J. Lu, T. Wang, and B. Lei, "A deep network for sinogram and CT image reconstruction," 2020, *arXiv:2001.07150*. [Online]. Available: http://arxiv.org/abs/2001.07150
[30] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, May 2015, pp. 234–241.
[31] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4809–4817.
[32] G. L. Zeng and G. T. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE Trans. Med. Imag.*, vol. 19, no. 5, pp. 548–555, May 2000.
[33] C. McCollough, "TU-FG-207A-04: Overview of the low dose CT grand challenge," *Med. Phys.*, vol. 43, no. 6, pp. 3759–3760, Jun. 2016.
[34] A. Paszke, S. Gross, S. Chintala, G. Chanan, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *Proc. NIPSW*, 2017, pp. 1–4.
[35] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Dec. 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[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] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.
[38] A. Biguri, M. Dosanjh, S. Hancock, and M. Soleimani, "TIGRE: A MATLAB-GPU toolbox for CBCT image reconstruction," *Biomed. Phys. Eng. Exp.*, vol. 2, no. 5, Sep. 2016, Art. no. 055010.
[39] Y. Han and J. C. Ye, "Framing U-Net via deep convolutional framelets: Application to sparse-view CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1418–1429, Jun. 2018.
[40] J. C. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *SIAM J. Imag. Sci.*, vol. 11, no. 2, pp. 991–1048, Jan. 2018.
[41] C. You, W. Cong, G. Wang, Q. Yang, H. Shan, L. Gjesteby, G. Li, S. Ju, Z. Zhang, Z. Zhao, and Y. Zhang, "Structurally-sensitive multi-scale deep neural network for low-dose CT denoising," *IEEE Access*, vol. 6, pp. 41839–41855, Jul. 2018.
[42] Y. S. Han, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis," Nov. 2016, *arXiv:1611.06391*. [Online]. Available: http://arxiv.org/abs/1611.06391
[43] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
[44] S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse Problems Sci. Eng.*, vol. 27, no. 7, pp. 987–1005, Jul. 2019.
[45] G. Ma, Y. Zhu, and X. Zhao, "Learning image from projection: A full-automatic reconstruction (FAR) net for computed tomography," *IEEE Access*, vol. 8, pp. 219400–219414, Nov. 2020.
[46] G. Wang, J. C. Ye, and B. D. Man, "Deep learning for tomographic image reconstruction," *Nature Mach. Intell.*, vol. 2, no. 12, pp. 737–748, Dec. 2020.
[47] A. Zheng, H. Gao, L. Zhang, and Y. Xing, "A dual-domain deep learning-based reconstruction method for fully 3D sparse data helical CT," *Phys. Med. Biol.*, vol. 65, no. 24, Dec. 2020, Art. no. 245030.
[48] Y. Ge, T. Su, J. Zhu, X. Deng, Q. Zhang, J. Chen, Z. Hu, H. Zheng, and D. Liang, "ADAPTIVE-NET: Deep computed tomography reconstruction network with analytical domain transformation knowledge," *Quant. Imag. Med. Surg.*, vol. 10, no. 2, pp. 415–427, Feb. 2020.
[49] W. Lin, H. Liao, C. Peng, X. Sun, J. Zhang, J. Luo, R. Chellappa, and S. K. Zhou, "DuDoNet: Dual domain network for CT metal artifact reduction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10512–10521.
[50] Q. Zhang, Z. Hu, C. Jiang, H. Zheng, Y. Ge, and D. Liang, "Artifact removal using a hybrid-domain convolutional neural network for limited-angle computed tomography imaging," *Phys. Med. Biol.*, vol. 65, no. 15, Aug. 2020, Art. no. 155010.
[51] A. Yamaev, M. Chukalina, D. Nikolaev, A. Sheshkus, and A. Chulichkov, "Lightweight denoising filtering neural network for FBP algorithm," *Proc. SPIE*, vol. 11605, Jan. 2021, Art. 116050L.

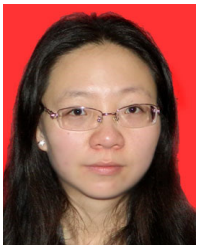
FENGYUAN JIAO received the M.S. degree in software engineering from the North University of China, Taiyuan, China, in 2017, where she is currently pursuing the Ph.D. degree. Her research interest includes image denoising and enhancement.

ZHIGUO GUI received the Ph.D. degree in signal and information processing from the North University of China, in 2004. He is currently working as a Professor with the North University of China. His research interests include image processing and image reconstruction.

KUNPENG LI is currently pursuing the M.S. degree in electronics and communication engineering with the North University of China, Taiyuan, China. His research interests include medical image processing and pattern recognition.

HONG SHANGGUANG received the B.E. degree in biomedical engineering and the Ph.D. degree in signal and information processing from the North University of China, in 2011 and 2016, respectively. Since 2016, she has been with the Taiyuan University of Science and Technology, where she is currently an Associate Professor. Her research interests include medical image processing and pattern recognition.

YANLING WANG received the Ph.D. degree in signal and information processing with the North University of China, in 2018. In 2018, she joined the School of Information Management, Shanxi University of Finance and Economics. Her research interests include image processing and CT image reconstruction.



YI LIU received the Ph.D. degree in signal and information processing from the North University of China, in 2014. In 2015, she joined the School of Information and Communication Engineering, North University of China. Her research interests include CT imaging and image processing.

PENGCHENG ZHANG received the Ph.D. degree in computer science and technology from Southeast University, Nanjing, China, and the Ph.D. degree in traitement du signal et télécommunications from Université de Rennes 1, Rennes, France, in 2014. He currently engages in teaching and research with the North University of China. His research interests include medical image reconstruction, medical image analysis, dose calculation, and planning optimization.